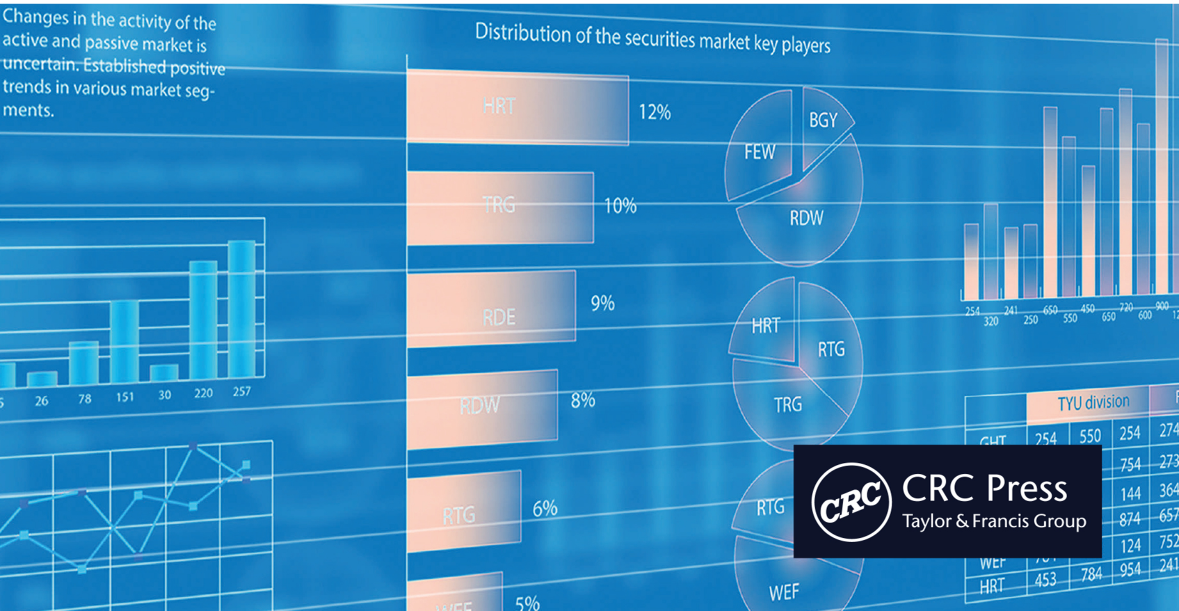
Information Technology, Management and Operations Research Practices

ANALYTICS IN FINANCE AND RISK MANAGEMENT

Edited by
Nga Thi Hong Nguyen, Shivani Agarwal
and Ewa Ziemba



CRC Press
Taylor & Francis Group

Analytics in Finance and Risk Management

This book presents contemporary issues and challenges in finance and risk management in a time of rapid transformation due to technological advancements. It includes research articles based on financial and economic data and intends to cover the emerging role of analytics in financial management, asset management, and risk management.

Analytics in Finance and Risk Management covers statistical techniques for data analysis in finance. It explores applications in finance and risk management, covering empirical properties of financial systems. It addresses data science involving the study of statistical and computational models and includes basic and advanced concepts. The chapters incorporate the latest methodologies and challenges facing financial and risk management and illustrate related issues and their implications in the real world.

The primary users of this book will include researchers, academicians, postgraduate students, professionals in engineering and business analytics, managers, consultants, and advisors in IT firms, financial markets, and services domains.

Information Technology, Management, and Operations Research Practices

Series Editors:
Vijender Kumar Solanki, Sandhya Makkar, and Shivani Agarwal

Operations Research Calculations Handbook, Second Edition
Dennis Blumenfeld

Introduction to Linear Optimization and Extensions with MATLAB®
Roy H. Kwon

Multiple Criteria Decision Making in Supply Chain Management
Edited By A. Ravi Ravindran

Multiple Criteria Decision Analysis for Industrial Engineering
Methodology and Applications
Gerald William Evans

Supply Chain Engineering
Models and Applications
A. Ravi Ravindran, Donald P. Warsing, Jr.

Analysis of Queues
Methods and Applications
Natarajan Gautam

Operations Planning
Mixed Integer Optimization Models
Joseph Geunes

Big Data Analytics Using Multiple Criteria
Decision-Making Models
Edited By Ramakrishnan Ramanathan,
Muthu Mathirajan, A. Ravi Ravindran

Service Systems Engineering and Management
A. Ravi Ravindran, Paul M. Griffin, Vittaldas V. Prabhu

Probability Models in Operations Research
C. Richard Cassady, Joel A. Nachlas

Fundamentals of Stochastic Models
Zhe George Zhang

Supply Chain Engineering
Models and Applications, Second Edition
A. Ravi Ravindran, Donald P. Warsing, Jr., Paul M. Griffin

Analytics in Finance and Risk Management

Edited by Nga Thi Hong Nguyen,
Shivani Agarwal, and Ewa Ziemba

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Cover image: © Shutterstock

First edition published 2024
by CRC Press
2385 NW Executive Center Drive, Suite 320, Boca Raton FL 33431

and by CRC Press
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

ISBN: 978-0-367-41686-7 (hbk)
ISBN: 978-0-367-74840-1 (pbk)
ISBN: 978-0-367-85469-0 (ebk)

DOI: 10.1201/9780367854690

Typeset in Times
by codeMantra

Contents

Editors

Nga Thi Hong Nguyen is an associate professor and vice director of Centre for Post-Graduate Studies, Hanoi University of Industry, Vietnam (HaUI). She has more than 20 years of academic experience in accounting, auditing, and finance. She has authored and co-authored many research articles published in journals, books, and conference proceedings. She has published 12 books. She teaches graduate- and post-graduate-level courses at HaUI, Vietnam. She received a PhD in accounting from Academy of Finance, Hanoi, Vietnam, in 2011; a master's degree in accounting from the Commerce University, Hanoi, Vietnam in 2006; and a bachelor's degree in accounting from National Economic University, Hanoi, Vietnam in 2001. She is a proceeding editor for 2nd International Conference on Management and Technovation (ICRMAT 2021), and 3rd ICRMAT 2022. She was an active organization chair and program chair for the 5th International Conference on Business Management and Accounting (IBSM), Hanoi University of Industry, Hanoi, Vietnam, 2018. She was an organization chair and a program chair for the 2nd ICRMAT, Hanoi University of Industry, Hanoi, Vietnam, 2021; and a program chair for the 3rd ICRMAT, Swinburne University, Danang, Vietnam, 4th ICRMAT, Hanoi University of Industry, Hanoi, Vietnam. She has chaired many Vietnamese national conferences and events at different universities in Vietnam. She is an editor of the *Journal of Accounting and Auditing*, Vietnam, ISSN 1859-1914, and an editor of the *Journal of Science and Technology*, Vietnam, ISSN 1859-3585.

Shivani Agarwal is an assistant professor, Galgotias University, Greater Noida, India. She earned her PhD from Indian Institute of Technology (IIT, Roorkee) in Management. She is engaged in teaching, research, and consultancy assignments. She has more than ten years of experience in teaching and in handling various administrative as well as academic positions. She also presented several research papers at national and international conferences. Dr. Agarwal has contributed chapters in different books published by Taylor and Francis, Springer, and IGI Global. She has conducted and attended various workshops, FDPs, and MDPs. She is the series editor of *Information Technology, Management & Operations Research Practices*, CRC Press, Taylor & Francis Group, USA. She is Guest Editor with IGI-Global, USA. Her research interests include quality of work life, trust, subjective well-being, knowledge management, employer branding innovation, and human resource management.

Ewa Ziemba is a full professor of Management Information Systems at the University of Economics in Katowice, Poland. Her research focuses on information systems and technologies for business, public administration, society, and sustainable development. She is one of the leaders in developing a multi-dimensional approach to a sustainable information society.

Ewa Ziemba has published over 240 peer-reviewed papers, three books, and 24 edited volumes, including a few published with Springer and Cambridge Scholars Publishing. She has played an instrumental role as a coordinator and principal investigator in over 40 prestigious domestic and international research projects. She has an extensive experience in training, guest lectures, and keynote speeches. She is an expert for the National Centre for Research and Development in Poland, the Academy of Finland, the Malta Council for Science and Technology, and the Science Fund of the Republic of Serbia. She serves on the editorial boards of international journals as an editor and reviewer. She is the editor-in-chief of the *Journal of Economics and Management* and the *Interdisciplinary Journal of Information, Knowledge, and Management*, and an associate editor of the *Journal of Computer Information Systems*. She is engaged in organizing prestige international and domestic scientific conferences, e.g., *Conference on Computer Science and Intelligence Systems FedCSIS*.

The academic world values and recognizes Ewa Ziemba's work and valuable contribution to the academic community. She was elected an ordinary member of the European Academy of Sciences and Arts (EASA) in Salzburg (Class V, Social Sciences, Law, and Economics). She received numerous awards for excellent research achievements and young academic staff development and works for the scientific community, including the *Prize of Polish Minister of Science* and *Higher Education for Excellent Research Achievement*, *The Excellence in Research & Scholarship Award*, and *Fellow & Distinguished Scholar Award* from International Institute for Applied Knowledge Management, six *Silver and Bronze Journal Editor Award*s, 26 *Awards of the Rector of the University of Economics in Katowice*, and ten *Best Research Paper Awards.*

Contributors

Shivani Agarwal
Galgotias University
Greater Noida, India

Subhash Chander Arora
Gurukula Kangri Deemed to be University
Haridwar, Uttarakhand, India

Semra Aydoğdu Bağci
Ankara Yıldırım Beyazıt University
Ankara, Turkey

Haşim Bağci
Aksaray University
Aksaray, Turkey

Divya Bansal
Amity University
Noida, Uttar Pradesh, India

Abhay Bhardwaj
KIET School of Pharmacy
KIET Group of Institutions
Delhi-NCR
Ghaziabad, Uttar Pradesh, India

Linh Khanh Bui
FPT University
Hanoi, Vietnam

Beata Czarnacka-Chrobot
Warsaw School of Economics
Warsaw, Poland

Maria Czech
University of Economics in Katowice
Katowice, Poland

Cuong Tri Dam
Industrial University of Ho Chi Minh City
Ho Chi Minh City, Vietnam

Beata Dratwińska-Kania
University of Economics in Katowice
Katowice, Poland

Aleksandra Ferens
University of Economics in Katowice
Katowice, Poland

Michal Gajda
Migamake Pte Ltd
Singapore

Vikas Gupta
GIBS, Delhi
Rohini, Delhi, India

Piotr Kania
University of Economics in Katowice
Katowice, Poland

Anna Karmańska
University of Economics in Katowice
Katowice, Poland

Ceyda Yerdelen Kaygin
Kafkas University
Kars, Turkey

Ewa Wanda Maruszewska
University of Economics in Katowice
Katowice, Poland

Binh Giang Nguyen
Vietnam Institute of Economics
Hanoi, Vietnam

Ha Thi Thu Nguyen
Electric Power University
Hanoi, Vietnam

Nga Thi Hong Nguyen
Hanoi University of Industry
Hanoi, Vietnam

Tam To Nguyen
Electric Power University
Hanoi, Vietnam

Xuan Trung Nguyen
Vietnam Institute of Americas Studies
Hanoi, Vietnam

Huy Quang Pham
University of Economics Ho Chi Minh City (UEH)
Ho Chi Minh City, Vietnam

Blandyna Puszer
University of Economics in Katowice
Katowice, Poland

Puja Roshani
Center of Management Studies
Jain University, Banglore, India

Vinod Kumar Singh
Gurukula Kangri Deemed to be University
Haridwar, Uttarakhand, India

Vijender Kumar Solanki
CMR Institute of Technology
Hyderabad, Telangana, India

Sripal Srivastava
Galgotias University
Greater Noida, Uttar Pradesh, India

Marlena Stanek
Warsaw School of Economics
Warsaw, Poland

Tuan Minh Tran
Ho Chi Minh National Academy of Politics
Hanoi, Vietnam

Maciej Andrzej Tuszkiewicz
University of Economics in Katowice
Katowice, Poland

Phuc Kien Vu
University of Economics Ho Chi Minh City (UEH)
Ho Chi Minh City, Vietnam

Filip Wójcik
Wroclaw University of Economics and Business
Wrocław, Poland

1 Does the effectiveness of accounting information system intensify sustainability risk management? An insight into the enabling role of the Blockchain-enabled Intelligent Internet of Things Architecture with Artificial Intelligence

Huy Quang Pham and Phuc Kien Vu

1 INTRODUCTION

Based on the perspectives of Septriadi et al. (2020), accounting information should be of flawless quality, appropriate, and helpful to enable the organization to make decision-making successful. The accounting information system (AIS) has been fruitful to various stakeholders because it provided information for organizational management and governance practices (Neogy, 2014). Thus, the EAIS has been pondered to be a prerequisite as it would warrant that all degrees of management would acquire adequate, proper, accurate, and timely information for planning and controlling organizational operations (Khassawneh, 2014; Monteiro et al., 2021). The industrial revolution 4.0 (Industrie 4.0) has engendered a significant impact not only on businesses but also on AIS (Salem et al., 2021). More particularly, it has converted the AIS into a more advanced digitalized system (Salem et al., 2021). The Internet-related technologies, namely artificial intelligence (AI), big data, cloud, and

DOI: 10.1201/9780367854690-1

blockchain (BC), would become the preconditions for accounting practices (Moll & Yigitbasioglu, 2019). Numerous academics' notes have deepened the analyses on the advantages of implementing these Internet-related technologies in accounting, namely the influence of Big Data, BC, and AI on Cloud-based AIS (Ionescu, 2019); the effect of BC and AI on Cloud-based AIS (Alkan, 2022); and the effect of AI on AIS (Hashem & Alqatamin, 2021). The advantages of BC on AIS have been considered as the issues that drew much more concerns from the academician and practitioner communities (i.e., ALSaqa et al., 2019; Faccia & Petratos, 2021; Fullana & Ruiz, 2021; Nugraha et al., 2021; Sarwar et al., 2021). Strikingly, the implementation of BC in AIS would achieve an important adoption mass in the near future and would become mainstream in 2025 (Karajovic et al., 2019). As such, Ølnes et al. (2017) advocated for much more investigation into the benefits of BC in ameliorating and developing public service to tackle the governance issues. Of these, risk management has been well acknowledged as the foremost concern in PSOs.

Due to multiple targets as well as numerous stakeholders (McAdam et al., 2005), PSOs must highly recognize and operate on their responsibilities to a variety of stakeholders. For this rationale, PSO risk management has been viewed as being much more difficult and having societal impacts as its primary focus (Ahmeti & Vladi, 2017). The fundamental duties of these organizations, however, have been to assure the public that no current or foreseeable danger would threaten the public value due to the complexity and heterogeneity of risks these organizations had to deal with in their daily operations (Ahmeti & Vladi, 2017). Additionally, according to the United Nations General Assembly, PSOs are not an exception to the consensus for a shared focus on economic, environmental, and social dimensions in the global community. This is because of their contribution to sustainable development, which is reflected in the sustainable development goals. To put it differently, risk management in PSOs should focus on the achievement of sustainable development and the generation of public value. Against this backdrop, SRM has been well recognized as one of the most vigorous solutions made available to PSOs to push forward these goals.

Unfortunately, despite the fact that a lack of security in Internet of Things (IoT) systems has increased the likelihood of intrusion and hacking attempts against sensitive data and critical infrastructure, the absence of an IoT cybersecurity risk management framework would make it extremely difficult to make informed decisions regarding IoT cyber risk management (Lee, 2020).

With respect to this, the architecture for the convergence of BC and AI for IoT introduced by Singh et al. (2020) was supposed to be efficient and effective measures for the above-mentioned issues. Building on the recommendations of Singh et al. (2020), this sort of architecture concentrated on reducing issues with decentralization, digitally signed, distributed, authenticated, public digital ledger, smart contracts, safe data sharing, and understandable AI (Singh et al., 2020). It could transform any environment, namely hospitals, schools, or other types of organizations, into a smart environment in which decision-making capacities as well as learning capacities were employed for all operations automatically rapidly and securely (Singh et al., 2020). With the support of this architecture, all the valuable information in relation to risk would be efficiently and effectively ascertained, analysed on both qualitative and quantitative facets, and further responded to, which were coupled with generating fruitful policies and strategies to prevent and mitigate the occurrence of risk incidents.

This book attempts to address the theoretical gaps in the body of literature concerning research on and implementations of risk management and digital technologies in PSO. This study started with a fascinating research question as follows.

RQ1. *What is the effect of BlockIoTIntelligence on SRM?*
RQ2. *Does EAIS act as a mediator on the interconnection between BlockIoTIntelligence on SRM?*

Aside from the introduction, Section 2 deals with the knowledge and background needed to enable the discussions of the study hypotheses in the section that followed. Then, in Section 4, the methodology of this study is described. This includes the research design and the techniques for gathering and analysing data. Section 5 is the primary episode that presents an overview of the statistical analysis. Theoretical contributions, practical implications, and a suggested research plan are all presented in Section 6.

2 THEORETICAL UNDERSTANDING AND FOUNDATION

2.1 Adoption model

Contingency theory (CT). CT was first introduced in 1950 (Donaldson, 2001; Nohria & Khurana, 2010) and was exploited in voluminous investigations related to the issues of organizations (Donaldson, 2001; Sauser et al., 2009). One of the first and foremost scientific ventures on AIS employed by the lens of this theory was research performed by Gordon and Miller (1976), who demonstrated the fundamental framework for pondering AIS from the contingency points of view. Subsequently, Otley (1980) enlarged this paradigm by advocating that the structure of AISs should be heavily based on the effectiveness of the organization, technology, environment, as well as organizational configuration. As such, AIS was recommended to be formulated in a flexible manner and related to certain identified situations (Otley, 2016). In addition, CT was also employed in several studies in terms of risk management (i.e., Grötsch et al., 2013; Teller et al., 2014).

Survival-based theory (SBT). As stated by Khairuddin (2005), SBT was promoted by Herbert Spencer, who propounded that organizations should ceaselessly roll with the punches to changes in internal and external environment for their survival. There has been a great consensus among organizational ecologists on the significance of a good fit between organizational idiosyncrasies and a changing atmosphere. In other words, there was much more likelihood that an entity which possessed the adaptability with changing surroundings would survive in the long term (Witteloostuijn et al., 2018). In addition, effective strategy has been considered as the most important factor for organizational survival. Therefore, businesses would continue to exist if they could deliver higher-quality goods and services in the shortest possible period and with the fewest resources (Khairuddin, 2005). Simultaneously, this strategy also reiterated that an entity should make a selection on a wide range of strategies which resulted in gaining adaptability to the present environment rather than concentrating on only one strategy (Lynch, 2000). This theory was employed in the present research to illuminate the urgent demand for a course of actions, which the PSOs were supposed to undertake for their survival. In this regard, implementation of strategic planning and adoption of digital information technologies would enable PSO to achieve sustainable development and generate public value.

2.2 Conceptual respect

BlockIoTIntelligence. With incessant changes in information technologies, BC, AI, and IoT devices have turned out to be the most contributing technologies which have been catalyzing the pace of innovation ideas in all areas. Of these, IoT was supposed to be connected or interrelated with numerous devices through the internet (Tang et al., 2018) and comprised three key elements such as hardware, middleware, and presentation (Gubbi et al., 2013). Besides, the application of IoT could create a linkage between heterogeneous and massively decentralized devices (Fazackerley et al., 2015). While AI was identified as the simulation of human intelligence in machines that were programmed to mimic human intelligence (Vaishya et al., 2020), BC was a set of computing nodes linked in a peer-to-peer way with mutually substantiated transactions undertaken in the network. Each block cryptographically sealed a collection of transactions and was connected with the prior block to establish a hash-based chain of blocks. Admittedly, BC has been touted as a powerful technology to improve business performance, particularly in the accounting and auditing field due to the enhancement of transparency and accountability of information (Rîndaşu, 2019). BC plays the role of public ledger of all transactions or digital events based on consensus mechanism.

The convergence of BC and AI for IoT focuses on minimizing the problem in terms of decentralization, distributed, validated, secure share data, and so on, expanding on the perspective of Singh et al. (2020). This suggested paradigm, which used the terms cloud intelligence, fog intelligence, edge intelligence, and device intelligence to show how to combine BC and AI to address large data analysis, security, and centralized challenges of IoT applications, was divided into four intelligences (Singh et al., 2020)

The effectiveness of accounting information system. AIS is widely recognized as an important computerized system that has handled both financial and non-financial activities to produce high-quality information for managing and governing company processes, planning, controlling, supervising, coordinating, as well as monitoring performance (Ibrahim et al., 2020). Additionally, it was the harmonious integration of the financial statement system, data processing system, data storage system, and data input system that increased the productivity of accounting work and produced useful information for decision-making (Huy & Phuc, 2020).

On how to evaluate EAIS, there is still ongoing disagreement and a dearth of empirical studies (Huy & Phuc, 2020). In contrast, Ernawatiningsih and Kepramareni (2019) argued that an AIS may be considered effective provided it was able to produce information in a fast, accurate, and reliable manner. Based on Huy and Phuc's (2020) viewpoints, EAIS could only be achieved if each component—the data input system, data processing system, data storage system, and financial statement system—performed efficiently (Huy & Phuc, 2020). Al-Okaily (2021) asserted, however, that EAIS took into account how well systems, information, processes, collaboration, and services were executed.

Sustainability risk management. Risk was defined as being exposed to risk, hazard, or the volatility of unexpected results (Howells & Bain, 1999; Jorion & Khoury, 1996). The term "sustainability risk" alluded to dangers associated with social justice or environmental issues, and Anderson's writings outlined a number of strategies for how people and organizations could be affected (2005). Risk management considered the crucial responsibilities that firms must carry out if they want to accomplish their business goals (Lark, 2015). In order to quickly make wise decisions and take steps that would result in positive outcomes, it was also determined that the process was codified in terms of defining distinct risks, rating them, and prioritizing them (Barbosa et al., 2022). Because of this, pursuing a sustainable risk management environment would require an effective risk management system (Al-Tamimi, 2002). According to Hofmann et al. (2013) and Schulte and Knuts (2022), SRM was defined in this research as risk management that focused on the effects of internal and external stakeholder value development.

3 SUBSTANTIATION OF RESEARCH HYPOTHESES

The integration of BC into AIS would become the potential for efficient and effective accounting processes (Shyshkova, 2018). As such, the time delays would be completely eradicated with the support of BC (Potekhina & Riumkin, 2017). Based on the arguments of Andersen (2016), the newly produced BC would be a reliable, distributed, and freely accessible ledger with low operating costs. Additionally, since BC was clear about any software modifications, it would be difficult to change or eliminate written accounting records (Potekhina & Riumkin, 2017). On the other hand, it has been confirmed that AI techniques have significantly increased the efficiency and efficacy of AIS by focusing on the understandability, reliability, and comparability of outputs (Hashem & Alqatamin, 2021). Since the accuracy of AIS's outputs allowed organizational leaders to make efficient accounting and financial decisions, integrating AI into the system would significantly reduce the percentage of losses associated with giving incorrect and inaccurate accounting information (Askary et al., 2018). Despite the notion that IoT will make the AIS much easier to use (Lee, 2020), a lack of security in the IoT systems has unfortunately created multiple opportunities for intrusions and hackers to access critical infrastructure and sensitive data (Cao & Zhu, 2012). In light of this, BlockIoTIntelligence would enable distributed cloud storage to create a cutting-edge solution for a database problem that increased the size of data from IoT applications (Singh et al., 2020). Additionally, BlockIoTIntelligence would provide the decentralized AI architecture that was used for autonomous transactions in a secure, genuine manner that was also supported by miners (Sharma et al., 2017). Additionally, by incorporating BC into IoT, devices, and gateways could protect the data that was processed and stored at the node (Singh et al., 2020). By enabling real-time accounting and reporting systems, the BlockIoTIntelligence would help the organizational AIS become effective and efficient and provide better decision-making tools. These studies served as the impetus for the first hypothesis that was developed in the current study (Figure 1.1).

Hypothesis 1 (H1). *BlockIoTIntelligence evinces a substantially positive influence on EAIS.*

The use of BC would make risk management processes much more proactive, integrated, and capable of identifying intangible hazards and supplying several layers of defense (Kouhizadeh et al., 2020). Furthermore, its enhanced visibility made it possible to increase openness and guarantee the security and privacy of donations for several operations (Khan et al., 2021). By stepping up information security (Kodym et al., 2020), reducing information uncertainty in credit choices (Dashottar & Srivastava, 2021), and enhancing cyber threat intelligence sharing platforms for risk management, BC could obtain the role of risk management (Riesco et al., 2019). AI algorithms could provide analytical capabilities for organizations to understand the effects of risks (Bechtsis et al., 2021), release automated suggestions to minimize and manage these risks (Larkin et al., 2021), respond quickly to the changing environment (Yang et al., 2021), and ascertain trends to inform policies (Johnson et al., 2021). Moving on to IoT applications, these devices were used to collect a significant amount of data in a centralized manner, which led to security and space issues (Jeong & Park, 2019), despite the fact that the lack of an IoT cybersecurity risk management paradigm would make it difficult for organizations to make wise decisions (Lee, 2020). Regarding this, the BlockIoTIntelligence would produce the distribution and decentralization method by using consensus protocols for scalability and security (Singh et al., 2020). Alternately, the distributed cloud and intelligent storage, micro-server, and smart contracts produced in the BlockIoTIntelligence would be utilized to achieve secure authentication and validation (Singh et al., 2020). The assistance of BlockIoTIntelligence would simultaneously provide scripting code for encryption, a hash function, a micro-server, and digital identification (Singh et al., 2020). By doing this, the BlockIoTIntelligence would enhance and intensify risk management productivity in a long-term way. These studies served as the impetus for the second hypothesis that was developed in the current study.

Hypothesis 2 (H2). *BlockIoTIntelligence evinces a substantially positive impact on SRM.*

The risk management process has been thought to require the consolidation and translation of various datasets into heterogeneous formulations in order to perform complex financial supervision, manage poor visibility, and traceability during operations, reduce discrepancies resulting from manual reporting, and govern conflicting information in the absence of a central shared database. The success of these methods has shown that they have a big impact on risk forecasting, planning, and crisis management. Automated decision reinforcement mechanisms should link stakeholders and management phases in order to reap the benefits of improved information management. This will enable leaders to understand the information and use it to support their decision-making (Comes et al., 2020). In certain cases,

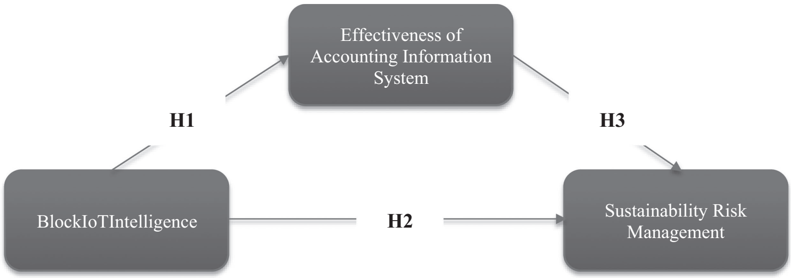


FIGURE 1.1 The hypothesized model.

gathering and storing data is no longer the problem; rather, it's translating the knowledge gleaned from digital data strings into insights and practical applications (Günther et al., 2017). Numerous academic works have emphasized the importance of risk management, which has been spread through a variety of accounting technologies and timelines, which together provide the basis for the others (Hall et al., 2015). As a result, it affected how risk management was understood and used (Rodríguez-Espíndola et al., 2022). The methods and tools used in risk management and accounting shape the temporal directions of factors. Surprisingly, it was confirmed that the National Bank's credit risk evaluation has been significantly impacted by electronic AIS (AL-Masharfi & Matriano, 2022; Qudah, 2021). These studies served as the impetus for the third hypothesis that was developed in the current study.

Hypothesis 3 (H3). *EAIS evinces a substantially positive impact on SRM.*

4 METHODOLOGICAL APPROACH

4.1 Research design

This study made use of semi-structured interviews because, according to DeJonckheere and Vaughn (2019), they enable open-ended data collection and the discovery of new themes (Gray, 2014). To find the right groups of field informants, purposive sampling was chosen. The focus of the current study was on PSO leaders who could offer insightful commentary and in-depth knowledge about the pertinent topics from the perspectives of their individual organizations. The criteria used to determine the interviewees' options included managerial seniority, which was intended to ensure that they had the necessary amount of decision-making power, and at least ten years of work experience in the organizations they represented, which was intended to verify that they were actively involved in the design of AIS. The current study solicited involvement from two overlapping groups of specialists. In addition, the group of BlockIoTIntelligence experts included people who were aware of the benefits of both SRM and BlockIoTIntelligence, while those who made up the AIS experts had an awareness of both AIS and SRM's advantages. Eisenhardt (1989) suggested that the number of cases should vary between four and ten or until a suitable level of saturation was reached, at which point the data became redundant. In relation to this, eight leaders of PSOs in South Vietnam were interviewed. Remote phone calls were used to conduct these interviews between November 2021 and February 2022. The many questions of this study tool were formulated around all of the axes of the proposed model, building on the findings of semi-structured interviews.

To ensure that the translation of the scale items was accurate, the English version of the questionnaire was created, translated into Vietnamese, and then back into English. This process could make it possible to confirm the consistency of the true meaning of each item scale in the original questionnaires (Saunders et al., 2009). Additionally, each item scale was contextualized based on expert opinion and suited to the inherent traits of the target group. In order to eliminate ambiguous questions from the questionnaire, pilot studies have been considered as an essential step before the major data collection (Tay et al., 2020). As a result, 30 people with characteristics resembling those of the survey group were asked to participate in the small-scale pilot test. With the assistance of SPSS version 26.0, the collected data were examined to determine the reliability of the items' scale, and the results showed greater reliability outcomes with Cronbach's Alpha coefficients exceeding 0.7 (Mahmoud et al., 2022). The final structured questionnaire was created and distributed because there were no items that needed to be refined.

4.2 Operationalization of variables for measurement

BlockIoTIntelligence. The first construct of BlockIoTIntelligence encompassed six components, namely Analytics Intelligence; Digital Identity; Distributed Cloud Storage; Decentralization and Distribution; Authentication and Verification; and Chain Structure which stemmed from the contribution of Singh et al. (2020).

Effectiveness of accounting information system. The first construct of EAIS encompassed five ingredients, namely System quality; Information quality; Service quality; and Process quality and Collaboration quality. Of these, System quality; Information quality; and Service quality emanated from the findings of Ifinedo et al. (2010). In the meanwhile, Process quality and Collaboration quality proceed from the works of Urbach et al. (2010).

Sustainability risk management. The first construct of SRM included four elements, namely Sustainability risk assessment, Sustainability risk identification, Sustainability risk analysis and evaluation, Sustainability risk treatment and communication which arose from the contribution of Schulte and Knuts (2022).

Participants were required to indicate a pertinent choice on a seven-Likert scale ranging from 1 for "vigorously disagree" to 7 for "vigorously agree" in order to evaluate their thoughts on each measure.

4.3 Sampling procedure and data collection

The participants in this study, which had PSOs as its main focus, were accountants. One argument in favor of this method stemmed from the fact that all organizational information and decision-making procedures were measured, disclosed, and warranted by accountants. A prerequisite for the favorable results would also be accountants due to the progressive diffusion of digital technologies (Zybery & Rova, 2014). To ensure high survey participation, the questionnaire was also distributed in person (Calvo-Porral & Pesqueira-Sanchez, 2022). The sample of the current study was established on the basis of convenience and snowball sampling. The sample size identified rested on the recommendations of Hair et al. (2019), which was 200 and higher as the volume of variables and the expected volume of variables gained. However, Hinkin (1995) argued that the optimal sample size should range from 1:4 to 1:10 whilst Urbach and Ahlemann (2010) suggested on a sample size with 200–800 responses. The data collection process took place in the Southern regions of Vietnam during the months of March 2022 and September 2022. All the responses suffered from 10% of missing data would be eradicated. The final sample size left for analysis covered with 723, with an 11.83% data loss rate.

4.4 Statistical analysis and calculations

With the aid of SPSS version 26.0 and AMOS version 26.0, the proposed model was examined. Using a two-stage structural equation modeling approach, this work was built on the advice of Anderson and Gerbing (1988). As a result, the initial phase of the inquiry concentrated on establishing the accuracy and validity of the measurement model. In the second stage, the structural model as a whole was examined in order to assess the overall model fit and hypothesized relationships using standardized regression coefficients (β) and p-values.

5 RESULT ANALYSIS

5.1 Statistics for demographic variables

Predominantly, female respondents made up 73.44% of the total respondents, while male respondents made up 36.56%. In terms of lifespan, those who were young to mature (under 40) made up 80.22% of the total respondents, while those who were middle-aged (40–50) and elderly (beyond 50) made up 17.70% and 2.07%, respectively. When it comes to academic proficiency, practically all of the responders have earned a graduate degree at the very least. The respondents had over ten years of experience working as PSO accountants.

5.2 Validity test of the model

Based on the recommendations of Hair et al. (2022), the criteria employed for the analyses included internal consistency, reliability of indicator, convergent validity as well as discriminant validity.

Convergent validity substantially depicted how an indicator correlated with the other of the same construct in a positive manner (Hair et al., 2022). The indicators with the minimum criteria or threshold for factor loading were above the value of 0.6 and served its goal (Hair et al., 2021; Oke et al., 2022). The Average Variance Extracted (AVE) of 0.5 or greater satisfies the demands for convergent validity (Hair et al., 2022).

The internal consistency of the model was evaluated through composite reliability (CR) and Cronbach's alpha. Both of these indicators could perform at a similar degree for corroboration, fluctuating from 0 to 1, in which the scores closer to 1 illustrated that the construct would become much more internally consistent. Accordingly, the values over 0.7 would be the most fitting; nevertheless, for exploratory studies, the values ranging from 0.6 to 0.7 were pondered acceptable (Hair et al., 2022).

On the basis of the outputs in Table 1.1, the measurement model in the current research reached the perfect reliability, internal consistency, and convergent validity.

When there are differences between the two constructs or variables, discriminant validity is mentioned (Hair et al., 2021). The Fornell–Larcker criterion application was determined to be the best method for evaluating discriminant validity. Fornell and Larcker (1981) suggested that discriminant validity could only be reached when the square root of AVE for each construct was greater than its greatest correlation with any other constructs. Additionally, according to Kline (2015), all inter-construct correlation values should be less than 0.85 or significantly different from 1.0 (Philips & Bagozzi, 1986). All the constructs in the hypothesized model exposed discriminant validity for the empirical data, because the correlation matrix shown in Table 1.2 satisfied these requirements.

Model Evaluation Criteria. The fitness between the hypothesized model and the statistical data collected was commonly evaluated through a set of indexes, namely the ratio of χ^2 to its degree of freedom (χ^2/df), root mean square error of approximation (RMSEA), comparative fit index (CFI) and Tucker–Lewis index (TLI), goodness-of-fit (GFI). On the basis of the outputs in Table 1.3, the measurement and structural models in the current research were authenticated to impeccably fit the procured data when all of the obtained indices evidently complied with the threshold suggested by previous researchers.

5.3 Correlations among the constructs

Direct effect. The parameter estimates and outcomes of the model hypotheses were presented in Table 1.4. More instrumentally, the impact of BlockIoTIntelligence (**H1**: $\beta = 0.467$; $p < 0.001$) illustrated a positive association with EAIS, while the effect of BlockIoTIntelligence (**H2**: $\beta = 0.344$; $p < 0.05$) underscored a markedly positive interconnection with SRM. The interlink between EAIS and SRM (**H3**: $\beta = 0.577$; $p < 0.01$) was staunchly supported. Consequently, **H1**, **H2**, **H3**, were buttressed.

Mediating effect. According to Memon et al. (2018), a variable may be considered a mediator when both the significance of the connections between the independent variable and the mediating variable and the connections between the mediating variable and the dependent variable have been established. In light of this, both the direct and indirect effects were confirmed to have a significant impact, leading to partial mediation (Cheung & Lau, 2007). Instead, where the indirect effect was strong compared to the insignificant direct effect, the complete mediation was recorded. Based on the results in Table 1.5, it was proven that EAIS served as a part-mediator in the relationship between BlockIoTIntelligence and SRM.

TABLE 1.1
Results summary of convergent validity and construct reliability

		Convergent validity		Construct reliability		
Constructs and operationalization	**Items (abbreviation)**	**Factor loadings ranges**	**AVE**	**Cronbach's alpha**	**Composite reliability**	**Discriminant validity**
		BlockIoTIntelligence				
Analytics intelligence	ANI	0.743–0.846	0.607	0.854	0.857	Yes
Digital identity	DII	0.805–0.878	0.707	0.874	0.875	Yes
Distributed cloud storage	DCS	0.713–0.797	0.588	0.845	0.848	Yes
Decentralization and distribution	DAD	0.818–0.852	0.692	0.864	0.867	Yes
Authentication and verification	AAV	0.802–0.866	0.704	0.871	0.874	Yes
Chain structure	CHS	0.819–0.879	0.727	0.885	0.822	Yes
		Effectiveness of accounting information systems				
System quality	SYQ	0.789–0.883	0.690	0.865	0.866	Yes
Information quality	INQ	0.825–0.904	0.752	0.855	0.856	Yes
Process quality	PRQ	0.808–0.866	0.716	0.877	0.88	Yes
Collaboration quality	COQ	0.827–0.876	0.738	0.886	0.891	Yes
Service quality	SEQ	0.803–0.881	0.705	0.87	0.874	Yes
		Sustainability risk management				
Sustainability risk assessment	SRA	0.697–0.815	0.615	0.859	0.861	Yes
Sustainability risk identification	SRI	0.780–0.829	0.662	0.851	0.851	Yes
Sustainability risk analysis and evaluation	SRAE	0.705–0.791	0.555	0.857	0.858	Yes
Sustainability risk treatment and communication	SRTC	0.704–0.852	0.600	0.852	0.853	Yes

TABLE 1.2
Results summary of discriminant validity

	SRAE	SRA	ANI	SRTC	DCS	COQ	CHS	DII	PRQ	AAV	SEQ	SYQ	DAD	SRA	INQ
SRAE	1														
SRA	0.034	1													
ANI	0.255	0.105	1												
SRTC	0.128	0.171	0.106	1											
DCS	0.054	0.064	0.053	0.184	1										
COQ	0.097	0.071	0.087	−0.005	−0.003	1									
CHS	0.146	−0.010	0.044	0.076	0.123	0.065	1								
DII	0.017	−0.025	0.145	−0.008	0.034	0.076	0.021	1							
PRQ	0.318	0.145	0.124	0.108	0.034	0.150	0.078	0.002	1						
AAV	−0.003	0.001	0.047	0.041	0.136	0.047	0.226	0.161	0.059	1					
SEQ	0.146	0.191	0.125	0.047	0.013	0.228	0.124	0.103	0.013	0.113	1				
SYQ	0.123	0.051	0.119	0.070	0.045	0.136	−0.024	0.033	0.209	0.133	0.110	1			
DAD	0.185	0.039	0.089	0.011	0.227	0.025	0.182	−0.155	0.069	0.152	0.112	0.079	1		
SRA	0.121	0.211	0.140	0.147	0.075	0.043	0.065	0.066	0.067	−0.014	−0.043	0.089	0.011	1	
INQ	0.062	0.018	0.122	0.070	0.015	0.095	0.072	0.062	0.160	0.061	0.126	0.232	−0.020	0.067	1

TABLE 1.3
Results of measurement and structural model analysis

The goodness of fit measures	Minimum cutoff	Parameter estimates of measurement model	Parameter estimates of structural model	Recommended by
Chi-square/df	<2	1.822	1.929	Schumacker and Lomax (2016)
TLI	≥0.95	0.948	0.941	Schumacker and Lomax (2016)
CFI	≥0.95	0.945	0.944	Schumacker and Lomax (2016)
GFI	0.774 ≤ x ≤ 0.923	0.905	0.893	Motawa and Oladokun (2015)
RMSEA	≤0.05	0.034	0.036	Schumacker and Lomax (2016)

TABLE 1.4
Results summary of hypotheses acceptance

Hypothesis No.	Hypothesized path			Standardized	S.E.	C.R.	Status
H1	BlockIoTIntelligence	→	EAIS	0.467***	0.183	3.384	Buttressed
H2	BlockIoTIntelligence	→	SRM	0.344*	0.139	2.143	Buttressed
H3	EAIS	→	SRM	0.577**	0.122	3.079	Buttressed

Notes: *p < 0.05; **p < 0.01; ***p < 0.001.

TABLE 1.5
The outcomes of indirect effect analysis

Route of paths	Direct effect	Indirect effect	Mediation
BlockIoTIntelligence → EAIS → SRM	0.344*	0.269**	Partial mediation

Notes: *p < 0.05; **p < 0.01.

6 CONCLUSION

The EAIS has been considered a requirement because it would ensure that all management levels would have access to sufficient, appropriate, accurate, and timely information for organizing and managing organizational operations (Khassawneh, 2014; Monteiro et al., 2021). The accounting system has been significantly impacted by Industrie 4.0, in addition to industries and manufacturing firms (Salem et al., 2021). More specifically, it has upgraded the AIS to a more sophisticated digital system (Salem et al., 2021). Building on the perspectives of Moll and Yigitbasioglu (2019), AI, big data, cloud computing, and BC are all Internet-related technologies that will be prerequisites for accounting methods. Unfortunately, despite the fact that the IoT systems' lack of security has given hackers and intruders multiple opportunities to access sensitive data and critical infrastructure, making informed decisions about IoT cybersecurity risk management would be incredibly challenging in the scarcity of a framework (Lee, 2020). In light of this, Singh et al. (2020) argued that the architecture for the convergence of BC and AI for IoT is the effective and efficient solution to these problems. The target of this study is to offer an understandable picture of how BlockIoTIntelligence influences SRM. Additionally, it makes an effort to provide in-depth insights into how the EAIS influences the connection between BlockIoT Intelligence and SRM by acting as a mediating factor. Through statistically large-scale response data obtained from convenient and snowball samples of 723 informants within the PSOs, located in the Southern regions of Vietnam, structural equation modeling was used to mathematically investigate the theoretical model hypothesizing the connections between the aforementioned components. The findings gathered throughout this experiment cast light on the potential use of BlockIoTIntelligence in improving and intensifying the EAIS, which would ultimately improve SRM. As a result, both theoretical value and practical significance were produced by the current research.

6.1 Theoretical contribution

Steered by the burgeoning concerns of risk management in PSO (Ahmeti & Vladi, 2017), the current research set its sight on investigating the role of BlockIoTIntelligence on SRM in PSO with EAIS as a mediator. The geographical focus of this work was set using PSOs in the Southern region of Vietnam as a cue.

Due to the IoT's infancy and the dearth of a robust IoT literature (Gil-Garcia et al., 2020), several academics have expanded their studies of the unresolved research questions of BC for AI and AI for BC on IoT (Singh et al., 2020). This work made a theoretically original addition by providing comprehensive insights into the benefits of BlockIoTIntelligence and the nature of the examined interdependencies between BlockIoTIntelligence and the other elements in the proposed model. Future debates could make use of the relationship between BlockIoTIntelligence, SRM, and EAIS as a jumping-off point. With the rapid advancement of technology and the digital age, research on AIS has become increasingly important, building on the ideas of Monteiro & Cepêda (2021). The findings of this empirical study could provide academic researchers with significant theoretical suggestions for additional research on blockchain intelligence and EAIS in PSO in developing nations. Although there have been several papers that have focused on the benefit of advanced technologies, namely BC and AI on AIS, these investigations placed their concerns only on the individual impacts of each advanced technology integrated into the AIS (i.e., Fullana & Ruiz, 2021; Hashem & Alqatamin, 2021; Sarwar et al., 2021). As this has been the first research that linked BC, AI, and IoT in one frame and delved into this convergence on EAIS, empirical findings were firmly to enrich the academia and AIS-related literature. Secondly, the current study gave rise to an empirical structural model, which substantiated that BlockIoTIntelligence positively, significantly, and directly influences SRM. In doing so, the obtained findings in this research bridged the gaps in the literature concerning risk management in PSOs (Osborne & Brown, 2011). Thirdly, this research was also deemed an endeavor to dig into the direct impact of EAIS on SRM. The positive and direct interconnection between EAIS and SRM in the current work totally contradicted the perceptions of Suzan et al. (2019) who scrutinized that operational risk management illustrated a significant positive influence on the EAIS. Last but not least, the current manuscript was a pioneer in the study of EAIS's mediating role in the relationship between BlockIoTIntelligence and SRM. The qualities of decentralization, authentication, smart contracts, immutability, and safe data sharing (Singh et al., 2020) would lead to an improvement in the AIS's quality. These characteristics also created an incredibly vital change in organizational risk management. In a nutshell, the implementation of BlockIoTIntelligence in AIS could enable the PSO to improve and develop SRM.

6.2 Practical implication

The empirics of the current research would bring great benefit to policymakers and practitioners. More instrumentally, the obtained findings would enable prioritizing the practices suggesting that the leaders in PSO should place additional concentrations on the implementation of BlockIotIntelligence to achieve enhanced AIS, and ultimately succeed in SRM. Accordingly, leaders in PSO should be aware of the advantages of modern information technologies and seek for efficient and effective solutions to facilitate the implementation of BlockIotIntelligence within all internal functions as well as across organizational boundaries. Additionally, the leaders in PSOs were encouraged to spend all endeavors to equally distribute the essential resources for improving and developing information technologies infrastructure of the organization as well as carry out the proper human resource management practices.

The outcomes of this study were also valuable for the leaders in PSOs to sense and seize the role of BlockIotIntelligence in formulating the strategies to efficiently and effectively revamp the risk management within their organizations and ultimately achieve the SRM. Intriguingly, the current manuscript cast light on the role of EAIS as a mediator between BlockIotIntelligence and SRM, and this observation could generate a deeper insight for those who were in search of a newer solution for addressing issues related to SRM. Succinctly put, the most paramount catalyst in the organizational success of SRM by leveraging advanced information technologies would be EAIS.

The results of this study were also anticipated to help policymakers and other responsible organizations establish guidelines for the use of digital technologies that would promote their widespread acceptance within PSOs. These findings also enabled policymakers and governmental influencers to facilitate policies and strategies in terms of risk management in PSO.

6.3 Boundaries and further opportunities

The observations in this research were limited by a number of caveats, which constructed a potential orientation for follow-up research. In the first place, the notable shortcoming associated with literature restraint in the face of the SEM analytical method. Theory and literature have been critical facets of the SEM approach.

Under this circumstance, the available literature has not been sufficient for the investigation of this magnitude. Literature on BlockIoTIntelligence, EAIS, and SRM was found to be relatively restricted. Simultaneously, empirical investigations in these areas were comparatively sparse. In the second place, the sample comprised exclusively accountants in PSOs within the southern regions of Vietnam which might seldom stand for the Vietnam context and become a complicated generalization across various nations. In order to enhance additional valuable insights, follow-up studies were recommended to be conducted on a national scale or to gather much more evidence from other emerging markets and developed countries. In the third place, statistical information was also obtained from a single respondent in each PSO, which may have compromised the validity of the findings. To that end, individuals from the other organizational departments were recommended to be included to generalize the findings. In the fourth place, stratified and quota sampling could take the place of the convenience and snowball sampling strategy that was used in this study's data collection in order to produce results that are more understandable and well-supported. Last but not least, the application of a cross-sectional design restrained this research to snapshot perceptions. As such, longitudinal study designs should be taken into consideration to implement in replicated studies

ACKNOWLEDGEMENT

This chapter was funded by University of Economics Ho Chi Minh City (UEH).

REFERENCES

Ahmeti, R., & Vladi, B. (2017). Risk management in public sector: a literature review. *European Journal of Multidisciplinary Studies, 2*(5), 323–329.

Alkan, B. Ş. (2022). How blockchain and artificial intelligence will effect the cloud-based accounting information systems? In: Bozkuş Kahyaoğlu, S. (eds) *The Impact of Artificial Intelligence on Governance, Economics and Finance* (2nd ed., pp. 107–119). Accounting, Finance, Sustainability, Governance & Fraud: Theory and Application. Singapore: Springer.

AL-Masharfi, M. H. R., & Matriano, M. T. (2022). The role of electronic accounting information system in measuring credit risk in National Bank of Oman. *Global Scientific Journals, 10*(1), 1789–1800.

Al-Okaily, M. (2021). Assessing the effectiveness of accounting information systems in the era of COVID-19 pandemic. *VINE Journal of Information and Knowledge Management Systems*, 1–20. doi:10.1108/VJIKMS-08-2021-0148.

ALSaqa, Z. H., Hussein, A. I., & Mahmood, S. M. (2019). The impact of blockchain on accounting information systems. *Journal of Information Technology Management, 11*(3), 62–80.

Al-Tamimi, H. (2002). Risk management practices: an empirical analysis of the UAE commercial banks. *Finance India, 51*(3), 1045–1057.

Anderson, D. R. (2005). *Corporate Survival: The Critical Importance of Sustainability Risk Management*. Bloomington, IN: iUniverse.

Andersen, N. (2016). Blockchain technology: a game-changer in accounting? *Deloitte & Touche GmbH, 3*, 1–4.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: a review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411–423. doi:10.1037/0033-2909.103.3.411.

Askary, S., Abu-Ghazaleh, N., & Tahat, Y. A. (2018, October). Artificial intelligence and reliability of accounting information. In *Conference on e-Business, e-Services and e-Society* (pp. 315–324). New York: Springer.

Barbosa, M. W., Carrasco, S. I. M., & Abarca, P. C. R. (2022). The effect of enterprise risk management competencies on students' perceptions of their work readiness. *The International Journal of Management Education, 20*(2), 1–14. doi:10.1016/j.ijme.2022.100638.

Bechtsis, D., Tsolakis, N., Iakovou, E., & Vlachos, D. (2021). Data-driven secure, resilient and sustainable supply chains: gaps, opportunities, and a new generalised data sharing and data monetisation framework. *International Journal of Production Research, 60*(14), 1–21.

Calvo-Porral, C., & Pesqueira-Sanchez, R. (2022). Does the use of technology create technology engagement? Comparing three structural models. *Spanish Journal of Marketing - ESIC*, 1–20. doi:10.1108/SJME-03-2022-0033.

Cao, H., & Zhu, Z. (2012). Research on future accounting information system in the Internet of Things era. In *2012 IEEE International Conference on Computer Science and Automation Engineering* (pp. 1–4). doi:10.1109/icsess.2012.6269573.

Cheung, G. W., & Lau, R. S. (2007). Testing mediation and suppression effects of latent variables. *Organizational Research Methods, 11*(2), 296–325. doi:10.1177/1094428107300343.

Comes, T., Van de Walle, B., & Van Wassenhove, L. (2020). The coordination-information bubble in humanitarian response: theoretical foundations and empirical investigations. *Production and Operations Management*, 1–42. doi:10.1111/poms.13236.

Dashottar, S., & Srivastava, V. (2021). Corporate banking—risk management, regulatory and reporting framework in India: a blockchain application-based approach. *Journal of Banking Regulation, 22*, 39–51. doi:10.1057/s41261-020-00127-z.

DeJonckheere, M., & Vaughn, L. M. (2019). Semi-structured interviewing in primary care research: a balance of relationship and rigour. *Family Medicine and Community Health, 7*(2), 1–8. doi:10.1136/fmch-2018-000057.

Donaldson, L. (2001). *The Contingency Theory of Organizations*. New York: Sage Publications, Inc.

Ernawatiningsih, N. P. L., & Kepramareni, P. (2019). Effectiveness of accounting information systems and the affecting factors. *International Journal of Applied Business & International Management, 4*(2), 33–40.

Eisenhardt, K. M. (1989). *Building theories from case study research. The Academy of Management Review, 14*(4), 532–550. doi:10.2307/258557

Faccia, A., & Petratos, P. (2021). Blockchain, enterprise resource planning (ERP) and accounting information systems (AIS): research on e-procurement and system integration. *Applied Sciences, 11*(15), 6792. doi:10.3390/app11156792.

Fazackerley, S., Huang, E., Douglas, G., Kudlac, R., & Lawrence, R. (2015). Key-value store implementations for Arduino microcontrollers. In *IEEE 28th Canadian Conference on Electrical and Computer Engineering* (pp. 158–164), Halifax, NS, Canada.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. doi:10.2307/3151312.

Fullana, O., & Ruiz, J. (2021). Accounting information systems in the blockchain era. *International Journal of Intellectual Property Management, 11*(1), 63–80.

Gil-Garcia, J. R., Pardo, T. A., & Gasco-Hernandez, M. (2020). Internet of Things and the public sector. In: Gil-Garcia, J., Pardo, T., & Gasco-Hernandez, M. (eds) *Beyond Smart and Connected Governments. Public Administration and Information Technology* (30th ed., pp. 3–24). New York: Springer.

Gordon, L. A., & Miller, D. (1976). A contingency framework for the design of accounting information systems. *Accounting, Organizations and Society, 1*(1), 59–69. doi:10.1016/0361-3682(76)90007-6.

Gray, D. (2014). *Doing Research in the Real World*. London: Sage Publications, Inc.

Grötsch, V. M., Blome, C., & Schleper, M. C. (2013). Antecedents of proactive supply chain risk management – a contingency theory perspective. *International Journal of Production Research, 51*(10), 2842–2867. doi:10.1080/00207543.2012.746796.

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): a vision, architectural elements, and future directions. *Future Generation Computer Systems, 29*(7), 1645–1660. doi:10.1016/j.future.2013.01.010.

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: a literature review on realizing value from big data. *The Journal of Strategic Information Systems, 26*(3), 191–209. doi:10.1016/j.jsis.2017.07.003.

Hair, J., Black, W., Barry, B., & Anderson, R. (2019). *Multivariate Data Analysis*. Boston, MA: Cengage, Andover.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R*. New York: Springer.

Hall, M., Mikes, A., & Millo, Y. (2015). How do risk managers become influential? A field study of toolmaking in two financial institutions. *Management Accounting Research, 26*, 3–22. doi:10.1016/j.mar.2014.12.001.

Hashem, F., & Alqatamin, R. (2021). Role of artificial intelligence in enhancing efficiency of accounting information system and non-financial performance of the manufacturing companies. *International Business Research, 14*(12), 65–75.

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*(5), 967–988. doi:10.1177/014920639502100509.

Hofmann, H., Busse, C., Bode, C., & Henke, M. (2013). Sustainability-related supply chain risks: conceptualization and management. *Business Strategy and the Environment, 23*(3), 160–172. doi:10.1002/bse.1778.

Howells, P., & Bain, K. (1999). *The Economics of Money, Banking and Finance*. London: Prentice-Hall.

Huy, P. Q., & Phuc, V. K. (2020). The impact of public sector scorecard adoption on the effectiveness of accounting information systems towards the sustainable performance in public sector. *Cogent Business & Management, 7*(1). doi:10.1080/23311975.2020.1717718.

Ibrahim, F., Ali, D. N. H., & Besar, N. S. A. (2020). Accounting information systems (AIS) in SMEs. *International Journal of Asian Business and Information Management, 11*(2), 51–67. doi:10.4018/ijabim.2020040104.

Ifinedo, P., Rapp, B., Ifinedo, A., & Sundberg, K. (2010). Relationships among ERP post-implementation success constructs: an analysis at the organizational level. *Computers in Human Behavior, 26*(5), 1136–1148. doi:10.1016/j.chb.2010.03.020.

Ionescu, L. (2019). Big data, blockchain, and artificial intelligence in cloud-based accounting information systems. *Analysis and Metaphysics, 18*, 44–49. doi:10.22381/AM1820196.

Jeong, Y.-S., & Park, J.-H. (2019). IoT, and smart city technology: challenges, opportunities, and solutions. *Journal of Information Processing Systems, 15*(2), 233–238.

Johnson, M., Albizri, A., Harfouche, A., & Tutun, S. (2021). Digital transformation to mitigate emergency situations: increasing opioid overdose survival rates through explainable artificial intelligence. *Industrial Management & Data Systems*, 1–21. doi:10.1108/IMDS-04-2021-0248.

Jorion, P., & Khoury, S. (1996). *Financial Risk Management: Domestic and International Dimensions*. Cambridge: Blackwell.

Karajovic, M., Kim, H. M., & Laskowski, M. (2019). Thinking outside the block: projected phases of blockchain integration in the accounting industry. *Australian Accounting Review*, 1–12. doi:10.1111/auar.12280.

Khairuddin, H. M. (2005). *Strategic Management*. Singapore: Thompson Learning.

Khan, M., Imtiaz, S., Parvaiz, G. S., Hussain, A., & Bae, J. (2021). Integration of Internet-of-Things with blockchain technology to enhance humanitarian logistics performance. *IEEE Access, 9*, 25422–25436.

Khassawneh, A. A. L. (2014). The influence of organizational factors on accounting information systems (AIS) effectiveness: a study of Jordanian SMEs. *International Journal of Marketing and Technology, 5*(4), 265–279.

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling*. New York: Guildford Press.

Kodym, O., Kubáč, L., & Kavka, L. (2020). Risks associated with Logistics 4.0 and their minimization using Blockchain. *Open Engineering, 10*(1), 74–85. doi:10.1515/eng-2020-0017.

Kouhizadeh, M., Zhu, Q., & Sarkis, J. (2020). Blockchain and the circular economy: potential tensions and critical reflections from practice. *Production Planning & Control, 31*, 950–966. doi:10.1080/09537287.2019.1695925.

Lark, J. (2015). *ISO 31000 Risk Management: A Practical Guide for SMEs*. Geneva: ISO.

Larkin, C., Drummond Otten, C., & Árvai, J. (2021). Paging Dr. JARVIS! Will people accept advice from artificial intelligence for consequential risk management decisions? *Journal of Risk Research*, 1–16. doi:10.1080/13669877.2021.1958047.

Lee, I. (2020). Internet of Things (IoT) cybersecurity: literature review and IoT cyber risk management. *Future Internet, 12*(9), 1–21. doi:10.3390/fi12090157.

Lynch, R. (2000). *Corporate Strategy* (2nd ed.). New York: Macmillan.

Mahmoud, M. A., Ahmad, S. b., & Poespowidjojo, D. A. L. (2022). Validation of the psychological safety, psychological empowerment, intrapreneurial behaviour and individual performance measurements. *RAUSP Management Journal, 57*(3), 219–234. doi:10.1108/RAUSP-11-2020-0252.

McAdam, R., Hazlett, S., & Casey, C. (2005). Performance management in the UK public sector. *International Journal of Public Sector Management, 18*(3), 256–273. doi:10.1108/09513550510591542.

Memon, M. A., Cheah, J., Ramayah, T., Ting, H., & Chua, F. (2018). Mediation analysis issues and recommendations. *Journal of Applied Structural Equation Modelling, 2*(1), i–ix.

Moll, J., & Yigitbasioglu, O. (2019). The role of internet-related technologies in shaping the work of accountants: new directions for accounting research. *The British Accounting Review, 51*(6), 1–20. doi:10.1016/j.bar.2019.04.002.

Monteiro, A. P., Vale, J., Silva, A., & Pereira, C. (2021). Impact of the internal control and accounting systems on the financial information usefulness: the role of the financial information quality. *Academy of Strategic Management Journal, 20*(3), 1–13.

Monteiro, A., & Cepêda, C. (2021). Accounting information systems: scientific production and trends in research. *Systems, 9*, 1–25. doi:10.3390/systems90300.

Motawa, I., & Oladokun, M. G. (2015). Structural equation modelling of energy consumption in buildings. *International Journal of Energy Sector Management, 9*(4), 435–450. doi:10.1108/ijesm-11-2014-0004.

Neogy, D. (2014). Evaluation of efficiency of accounting information systems: a study on mobile telecommunication companies in Bangladesh. *Global Disclosure of Economics and Business, 3*(1), 40–55.

Nohria, N., & Khurana, R. (2010). *Handbook of Leadership Theory and Practice*. Boston, MA: Harvard Business School Press.

Nugraha, A., Daniel, D. R., & Utama, A. A. G. S. (2021). Improving multi-sport event ticketing accounting information system design through implementing RFID and blockchain technologies within COVID-19 health protocols. *Heliyon, 7*, 1–6. doi:10.1016/j.heliyon.2021.e08167.

Oke, A. E., Kineber, A. F., Alsolami, B., & Kingsley, C. (2022). Adoption of cloud computing tools for sustainable construction: a structural equation modelling approach. *Journal of Facilities Management*, 1–18. doi:10.1108/JFM-09-2021-0095.

Ølnes, S., Ubacht, J., & Janssen, M. (2017). Blockchain in government: benefits and implications of distributed ledger technology for information sharing. *Government Information Quarterly, 34*(3), 355–364. doi:10.1016/j.giq.2017.09.007.

Osborne, S. P., & Brown, L. (2011). Innovation in public services: engaging with risk. *Public Money & Management, 31*(1), 4–6. doi:10.1080/09540962.2011.545532.

Otley, D. T. (1980). The contingency theory of management accounting: achievement and prognosis. *Accounting, Organizations and Society, 5*(4), 413–428. doi:10.1016/0361-3682(80)90040-9.

Otley, D. (2016). The contingency theory of management accounting and control: 1980–2014. *Management Accounting Research, 31*, 45–62. doi:10.1016/j.mar.2016.02.001.

Philips, L. W., & Bagozzi, R. P. (1986). On measuring organizational properties of distribution channels: methodological issues in the use of key informants. In: Sheth, J. N. (eds) *Research in Marketing* (Vol. 8, pp. 313–369). Greenwich: JAI Press.

Potekhina, A., & Riumkin, I. (2017). *Blockchain – A New Accounting Paradigm: Implications for Credit Risk Management*. Master degree thesis, Umeå School of Business and Economics.

Qudah, H. A. A. (2021). Credit risks measurement in Islamic banks study model. *Journal of Sustainable Finance & Investment*, 1–19. doi:10.1080/20430795.2021.1891783.

Riesco, R., Larriva-Novo, X., & Villagra, V. A. (2019). Cybersecurity threat intelligence knowledge exchange based on blockchain. *Telecommunication Systems, 73*(2), 259–288. doi:10.1007/s11235-019-00613-4.

Rîndaşu, S. M. (2019). Blockchain in accounting: trick or treat? *Quality - Access to Success, 20*(170), 143–147.

Rodríguez-Espíndola, O., Chowdhury, S., Dey, P. K., Albores, P., & Emrouznejad, A. (2022). Analysis of the adoption of emergent technologies for risk management in the era of digital manufacturing. *Technological Forecasting & Social Change, 178*, 1–25. doi:10.1016/j.techfore.2022.121562.

Salem, S., Nurdayadi, N., & Alfiandri, A. (2021). Cloud accounting: the development of accounting information system in Industry 4.0 in Indonesia. *Conference Series, 3*(2), 282–294. doi:10.34306/conferenceseries.v3i2.597.

Sarwar, M. I., Iqbal, M. W., Alyas, T., Namoun, A., Alrehaili, A., Tufail, A., & Tabassum, N. (2021). Data vaults for blockchain-empowered accounting information systems. *IEEE Access, 9*, 117306–117324. doi:10.1109/access.2021.3107484.

Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research Methods for Business Students*. New York: Pearson.

Sauser, B. J., Reilly, R. R., & Shenhar, A. J. (2009). Why projects fail? How contingency theory can provide new insights – a comparative analysis of NASA's Mars Climate Orbiter loss. *International Journal of Project Management, 27*(7), 665–679. doi:10.1016/j.ijproman.2009.01.004.

Schulte, J., & Knuts, S. (2022). Sustainability impact and effects analysis - a risk management tool for sustainable product development. *Sustainable Production and Consumption, 30*, 737–751. doi:10.1016/j.spc.2022.01.004.

Schumacker, R. E., & Lomax, R. G. (2016). *A Beginner's Guide to Structural Equation Modeling* (4th ed.). New York: Routledge.

Septriadi, D., Zarkasyi, W., Mulyani, S., & Sukmadilaga, C. (2020). Management accounting information system in gas station business. *Utopía Y Praxis Latinoamericana, 25*(1), 244–254.

Sharma, P. K., Rathore, S., & Park, J. H. (2017). Multilevel learning based modeling for link prediction and users' consumption preference in online social networks. *Future Generation Computer Systems*, 1–10. doi:10.1016/j.future.2017.08.031.

Shyshkova, N. (2018). Prospects for the implementation of blockchain in accounting. *Accounting and Finance, 2*, 61–68.

Singh, S. K., Rathore, S., & Park, J. H. (2020). BlockIoTIntelligence: a blockchain-enabled intelligent IoT architecture with artificial intelligence. *Future Generation Computer Systems, 110*, 721–743. doi:10.1016/j.future.2019.09.002.

Suzan, L., Mulyani, S., Sukmadilaga, C., & Farida, I. (2019). Empirical testing of the implementation of supply chain management and successful supporting factors of management accounting information systems. *International Journal of Supply Chain Management, 8*(4), 629–641.

Tang, C.-P., Huang, T. C.-K., & Wang, S.-T. (2018). The impact of Internet of Things implementation on firm performance. *Telematics and Informatics*, 1–50. doi:10.1016/j.tele.2018.07.007.

Tay, L. C., Tan, F. Y., Yahya, K. K., & Rasli, A. (2020). Validation of corporate environmental citizenship measurement in Malaysia. *European Business Review, 33*(2), 345–360, doi:10.1108/EBR-11-2018-0196.

Teller, J., Kock, A., & Gemünden, H. G. (2014). Risk management in project portfolios is more than managing project risks: a contingency perspective on risk management. *Project Management Journal, 45*(4), 67–80. doi:10.1002/pmj.21431.

Urbach, N., & Ahlemann, F. (2010). Structural equation modeling in information systems research using partial least squares. *Journal of Information Technology Theory and Application, 11*(2), 5–40.

Urbach, N., Smolnik, S., & Riempp, G. (2010). An empirical investigation of employee portal success. *The Journal of Strategic Information Systems, 19*(3), 184–206. doi:10.1016/j.jsis.2010.06.002.

Vaishya, R., Javaid, M., Khan, I. H., & Haleem, A. (2020). Artificial intelligence (AI) applications for COVID-19 pandemic. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 14*(4), 337–339. doi:10.1016/j.dsx.2020.04.012.

Witteloostuijn, A. V., Boin, A., Kofman, C., Kuilman, J., & Kuipers, S. (2018). Explaining the survival of public organizations: applying density dependence theory to a population of US federal agencies. *Public Administration, 96*(4), 1–18. doi:10.1111/padm.12524

Yang, Y. C., Islam, S. U., Noor, A., Khan, S., Afsar, W., & Nazir, S. (2021). Influential usage of big data and artificial intelligence in healthcare. *Computational and Mathematical Methods in Medicine*, 1–13. doi:10.1155/2021/5812499

Zybery, I., & Rova, L. (2014). The role of the accountants in the framework of the modern technological developments and digital accounting systems. *European Scientific Journal, 24*, 30–48.

2 Determining the liquidity level of businesses registered on the Polish Stock exchange

Haşim Bağci and Ceyda Yerdelen Kaygin

1 INTRODUCTION

Stocks are financial instruments that assume a crucial task in the management, valuation, and performance of companies besides providing the funds that companies need (Fang et al., 2009). For this reason, determining the liquidity levels of companies' stocks is very important in measuring financial performance (Kimondo et al., 2016). As it is known, companies must have sufficient liquidity levels to fulfill their short-term financial obligations without default. This requirement can only be achieved with the correct management of companies' resources (Loncan and Caldeira, 2014; Robinson et al., 2015).

Successful liquidity management of companies makes their activities sustainable and is seen as a prerequisite for companies' survival (Niresh, 2012; Madushanka and Jathurika, 2018). Exchanges: They have two types of customer structures, companies listing their shares and investors trading on the stock market (Macey and O'Hara, 1999). Liquidity in terms of stocks: Ease of converting investment into cash, meaning that, represents the ability to purchase and sell stocks quickly and easily (Kaliski, 2007). In this context, the concept of liquidity can also be described as the ability of a stock to become convertible into cash at the market price (Salehi et al., 2011). The excess of liquidity is considered an indication that companies' funds are limited to liquid assets. While a low liquidity level is often considered a threat due to the decrease in a company's profitability and solvency, an adequate liquidity level is seen as a margin of safety (Che et al., 2013; Ehiedu, 2014; Loncan and Caldeira, 2014; Alshatti, 2015; Orshi, 2016; Akenga, 2017). The capability of a financial asset to be converted from one form into another (Amihud et al., 2005) also provides ease of trading securities (Ivanović, 1997). Quick and easy purchasing and selling of securities make liquidity one of the main factors that investors take into account, as it is effective in deciding whether or not to invest in securities.

The most important of the basic questions that investors face today is how to measure liquidity, especially in emerging markets. In this context, four dimensions of liquidity come to the fore: trading time, tightness, depth, and resiliency. Trading

 DOI: 10.1201/9780367854690-2

time: the capability to instantly execute the transaction at the current price; tightness: the capability of an asset to be bought and sold at the exact price at the exact time; depth: the capability of purchasing or selling the stock without affecting the predetermined price; while resiliency is known as the ability of the stock to purchase or sell the asset with a few change over the predetermined price (Wyss, 2004).

The liquidity levels of companies are determined by their liquidity ratios. Liquidity ratios are used to calculate companies' ability to fulfill their short-term obligations. Liquidity ratios are of great importance for both internal and external analysts, as they are the financial ratios that are effective in determining the financial performance of companies (Bhunia, 2010; Niresh, 2012). An optimal liquidity level is thought to prevent financial problems such as financial distress or bankruptcy of companies by preventing excessive borrowing. This study aims to assume that having sufficient liquidity will increase financial performance for companies to evaluate alternative business opportunities, reliability, and fulfill their obligations, and this situation will positively affect profitability.

This research aims to detect the liquidity level of seven companies with high trading volume registered in the Polish Stock Exchange in accordance with the liquidity ratios, and then to measure the liquidity performance of the enterprises with the help of the weights of the three ratios.

2 LITERATURE REVIEW

Financial decision makers aim to use the company's resources effectively and efficiently, as well as to provide access to foreign resources with optimum cost. Insufficient liquidity means insufficient capital (Vavrek et al., 2021). Working capital (WC) is described as the difference between current assets and short-term liabilities which indicates liquidity with its effects on both macroeconomic and microeconomic performance (Fulford, 2015; Cretu et al., 2019). For this reason, it is accepted as a crucial financial indicator that has an impact on the profitability of companies.

To ensure sustainable growth, companies should always determine the right WC strategy (Zimon and Dankiewicz, 2020). WC management is a crucial part of corporate finance as it directly influences the profitability and liquidity of the company (Khan, 2017). Companies with a high level of WC are more prone to attain better financial security (Deloof, 2003; Madhou et al., 2015), while the ones with a low level of WC attain weaker financial security, while financial institutions and banks lend more easily to companies with higher levels of WC they tend to give (Ivashina and Scharfstein, 2010). Since even the smallest mistake made while determining the WC may cause liquidity loss in companies, WC management should be meticulously dealt with, especially during crises (Chang et al., 2019).

Liquidity is an indicator of a company's access to cash and current assets (Zeller et al., 1997). Liquidity assesses the company's ability to offset cash outflows with sufficient cash inflows and its capacity to cope with unexpected circumstances (Doina and Mircea, 2008). Lack of liquidity has both indirect and direct effects on capital costs (Benson et al., 2015). Effective liquidity management not only ensures the survival of companies but also enables companies to attain higher profits by mitigating input needs. In addition, effective liquidity management offers strategic advantages to companies in times of economic distress (Veronika et al., 2014). In many studies in the literature, it has been emphasized that determining the optimal liquidity level of companies in financial planning plays a vital role for companies (Afrifa and Padachi, 2016; Vuković et al., 2017; Martinho, 2022). In order to achieve optimal liquidity, it is mandatory to maintain an appropriate liquidity level (Zainudin et al., 2019). The role of liquidity ratios is very important in calculating financial distress forecast performance (Kapounek et al., 2022).

A positive impact of liquidity on expected returns is also effective in portfolio creation, diversification, and investment strategies. For this reason, liquidity is an issue that should be considered as a systematic factor, not as a feature in pricing and risk models (Bradrania and Peat, 2014). Investors learn about the liquidity management of companies by estimating future cash flows and returns to determine the value of stocks and investments and to decide when to purchase and sell (Zamanpour and Bozorgmehrian, 2012). Research on corporate liquidity management generally aims to compare companies with low high and liquidity. However, determining the relationships between liquidity management and real variables allows the significance of the subject to be better understood (Almeida et al., 2014). Dittmar and Mahrt-Smith (2007) found that companies with weaker corporate governance (CG) utilize cash faster than the ones with more substantial governance. Harford et al., (2008) focused on CG impacts on companies' cash assets. It was found that companies with frailer CG structures had a lower amount of cash reserves. Huang (2003) emphasized that investors may need a high liquidity premium when faced with liquidity shocks and forced to borrow against their future income. Companies are motivated for several reasons to keep a certain amount of liquid balance. Information asymmetries among companies and capital markets are essential precautionary reasons for the demand for corporate liquidity (Bruinshoofd and Kool, 2004; Isshaq and Bokpin, 2009).

As it is known, businesses generally tend to hold cash for transactional, speculative, and prudential purposes. Transactional cash holding companies guarantee daily purchases. In other words, the transaction motive is the situation in which companies hold cash for their daily operational activities (Amess et al., 2015). In such a case, the higher the company's revenues, the more companies can benefit from these revenues (Reilly and Brown, 2011). Dittmar and Mahrt-Smith (2007) suggested that optimal cash holding for transactions can support companies undertaking these activities without incurring outsourcing costs and liquidating assets. Speculative cash handling is the tendency of companies to maintain liquidity with the expectation of financial benefits in some periods. Denis and Sibilkov (2009) suggested that the speculative motive allowed companies to exploit future investments that may be omitted due to cash shortages. As companies' investments get even more speculative, they are expected to assume higher credit risk, and financial performance will be affected in case of default (Elnahas et al., 2017). According to Akhtar et al. (2018), precautionary cash holdings are companies' self-insurance in the presence of unexpected expenses or adverse circumstances in the future. In other words, the biggest incentive to hold cash for need is increasingly perceived uncertainty and insecurity. Al-Najjar (2013) observed that liquidation costs were quite lower compared to other assets. Therefore, companies with higher amounts of liquid assets are less prone to holding cash. Palazzo (2012) developed a model suggesting that riskier companies were prone to holding higher amounts of cash due to precautionary saving purposes, as they attain higher correlations between aggregate shocks and cash flows.

Gopalan et al. (2012) stated that a liquid asset could quickly become convertible into cash at a lower cost. Companies with high liquidity levels could quickly convert their assets into cash and fulfill short-term liabilities. Liquidity determined their capability to offset cash outflows with a sufficient level of cash inflows and their capability to withstand unexpected circumstances (Doina and Mircea, 2008). Nejadmalayeri (2021) found that companies with higher amounts of fixed assets encounter lower risk premiums, given the company's asset structure, asset liquidity, and cash holding ratio, which are closely related to business risk. Nikolaou (2009) emphasized that the main reason for liquidity risk involved the presence of incomplete markets and information asymmetry. Diamond and Verrecchia (1991) stated that to mitigate information asymmetry, a company could increase liquidity for its securities, resulting in a lower cost of capital. According to Chen et al. (2021) liquidity risk is not only a symptom of bankruptcy problems of banks; they concluded that banking crises also have an effect on bank performance.

Umar and Sun (2016) found that capital was not a unique determinant for liquidity, but stock liquidity was also effective. A company's illiquidity can cause financial distress if the company lacks external financing (Davydenko, 2013; Amoa-Gyarteng, 2019). Ng et al. (2013) found that liquidity was a driver of returns made in the short and long run for companies. Zamanpour and Bozorgmehrian (2012) determined that liquidity management had an affirmative and significant association with ROA and ROE, which are also called profitability indices, and they revealed that companies' liquidity management could considerably have a great impact on stock returns. Increasing the WC (Alifiah, 2013) and reducing the total debt (Parker et al., 2011) is an important issue in order to create a buffer in case of bankruptcy and declining profits that may put the company in financial distress. Amoa-Gyarteng determined that a company in financial distress in 2019 has a tendency to decrease in retained earnings, profits before interest and tax, and current assets (Davydenko, 2013; Amoa-Gyarteng, 2019).

A company is likely to default on its debts when its loans become due, triggering the company's liquidity problems (Nomani and Azam, 2020). Rating triggers are a kind of debt agreement that obliges the borrower to keep his or her credit rating higher than a definite rating threshold and aims to secure the lender's receivables from the creditor in the event of a downgrade. Although ratings lower the borrowing costs, once activated, they aggravate borrowers' liquidity needs when credit risk is high and increase the probability of borrower default (Parmeggiani, 2013). Lartey et al. (2013) detected a significant association between liquidity and profitability. Decreased liquidity levels can be seen as a harbinger of an increased probability of financial distress. Saleem and Rehman, in their 2011 study, found that companies' profitability and liquidity ratios would have been highly related and that liquidity affects profitability. The increase in total assets, profit before interest and tax, and WC lead to low-profit margins, low asset efficiency, and illiquidity (Amoa-Gyarteng, 2019). It was emphasized that liquidity has positive statistical effects on financial performance (Amal et al., 2012). Corporate bankruptcy not only harms the company entirely but also imposes a consequential cost on company managers (Eckbo, et al., 2016). A largely debt-financed company provides less financial security to creditors in the case of bankruptcy (Moyer et al., 2005). When the literature on liquidity is explicated, it is noticed that it is discussed under various headings such as WC, optimal liquidity level, corporate liquidity management, liquidity risk premium, liquidity shocks, and financial distress. In this context, the determination of the liquidity levels of companies on a company-by-year basis in the research can be considered as an evaluation criterion by financial information users.

3 PROPOSED WORK

The purpose, scope, limitations, data collection techniques, and methods of the research are discussed under separate headings in this part.

3.1 Purpose, scope, and limitations of the study

The purpose of the study is to detect the liquidity levels of seven companies with high trading volume registered in the Polish Stock Exchange and to measure the liquidity performance in line with the determined levels. To determine the liquidity level; liquidity ratios are used. Although ten years of data were reached in the research, since the year interval in which the liquidity ratios are regularly published within the last five years, the time frame of the sample used consists of five years covering the years 2018–2022. In addition, although there were 12 companies whose data could be accessed in the study, three companies that did not publish their liquidity ratios and two companies with missing data were excluded from the scope and seven companies were analyzed. Seven companies constituting the sample of the research; Projekt, Polsat, Jastrzebska, Orange Polska, PGE Polska, Polskie, and Powszechny Zaklad enterprises. The liquidity levels of these enterprises were determined and their five-year liquidity performance rankings were made. Apart from the main purpose of the study, there are also secondary objectives. These are the calculated liquidity performances to measure its effect on the Altman Z Score, which shows ROA, ROE, ROC, and financial failure.

3.2 Data collection techniques

The data of the research were accessed from the website investing.com (InvestingPro, 2022: https://www.investing.com/pro/watchlist). Since the research data are ready-made, the secondary data set was used.

3.3 Research method

In the research, multi-criteria decision-making methods were used to calculate the liquidity level. First, the SD method, also known as “Standard Deviation”, was used to figure out the degree of significance of liquidity ratios. Since the SD method is an objective method, it eliminates any subjective interpretation. The weight of liquidity ratios was determined in the research using the SD method. Then, the MAIRCA method was employed to measure liquidity performance. MAIRCA is also an objective method, and as a result of this method, companies are ranked in periods determined according to their liquidity performance.

After the measurements were made for the main purpose of the research, the liquidity performance levels were determined. Panel data analysis was conducted to measure the effect of liquidity performance on ROA, ROE, ROC, and financial failure. Dependent and independent variables were determined by panel data analysis and the effect of liquidity performance on four determined variables was measured using panel regression analysis.

3.3.1 SD method

SD Method is a technique known as standard deviation. In this technique, by detecting the deviation of the variables from the mean, it is determined how much each variable deviates from the mean of the whole dataset. It is an objective evaluation method that is completely objective. In the SD method, the criteria weights are found with the help of their standard deviations.

The SD Method consists of three stages. First, a decision matrix is created from the dataset. In the second stage, the decision matrix is standardized. The purpose of standardization is to bring together every data with different values in a common value range. The standardization process is shown with formulas (2.1) and (2.2).

$$r_{ij} = \frac{x_{ij} - x_j^{min}}{x_j^{max} - x_j^{min}}, i = 1, 2, \ldots, m; j = 1, 2, \ldots, n \text{ (for benefit criteria)} \tag{2.1}$$

$$r_{ij} = \frac{x_j^{max} - x_{ij}}{x_j^{max} - x_j^{min}}, \quad i = 1, 2, \ldots, m; j = 1, 2, \ldots, n \text{ (for cost criteria)} \tag{2.2}$$

The standardization process seen in formula (2.1) is formulated for benefit items, while that shown in formula (2.2) is formulated for cost items. In addition, in the equation shown with formulas (2.1) and (2.2), "m" indicates alternatives, and "n" indicates evaluation criteria. The third stage is to estimate the criteria weights. At that stage, according to the standardized data, the process of determining which variable is important and how much weight is applied. This process is done with the formula number (2.3).

$$w_j = \frac{\sigma_j}{\sum_{k=1}^{n} \sigma_k} (j = 1, 2, \ldots, m) \tag{2.3}$$

In formula (2.3), the weight of each variable is estimated by dividing the standard deviation of each variable by the standard deviation of the total dataset. These calculated weights were used in the next method (Diakoulaki et al., 1995).

3.3.2 MAIRCA method

Multi Attributive Ideal-Real Comparative Analysis (MAIRCA) Technique is one of the current methods developed by Gigovic et al. (2016). The main purpose of the method involves calculating the gap values between the criteria and determining the most ideal alternative in the final stage, and the society prefers the alternative with the lowest gap value (Pamucar et al., 2018).

The MAIRCA method consists of seven stages (Pamucar et al., 2018):

Step 1: Creating the decision matrix

It is seen in matrix number (2.4).

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \tag{2.4}$$

Step 2: Identifying alternative priorities

Prioritizing alternatives is up to the decision maker. If the decision maker is neutral, there is no priority among the alternatives, and thus the priority is calculated by dividing the number of alternatives, m. This calculation is seen in formula number (2.5). Also, the sum of all priorities must equal 1.

$$PAi = 1/m \tag{2.5}$$

Step 3: Creating a theoretical rating matrix

Weights and priorities are used in this calculation. The matrix is created by multiplying the weights with the priorities and if the priorities are equal, the matrix consists of a single row.

$$\begin{array}{cc} & \begin{array}{cccc} w_1 & w_2 & \cdots & w_n \end{array} \\ T_p = & \begin{bmatrix} P_{A1}w_1 & P_{A1}w_2 & \cdots & P_{A1}w_n \\ P_{A2}w_1 & P_{A2}w_2 & \cdots & P_{A2}w_n \\ \vdots & \vdots & & \vdots \\ P_{Am}w_1 & P_{Am}w_2 & & P_{Am}w_n \end{bmatrix} \end{array} \tag{2.6}$$

Step 4: Generating the actual rating matrix

At this stage, formula (2.7) is utilized for positive criteria and formula (2.8) is utilized for negative criteria, and a new matrix is created.

$$t_{rij} = t_{pij}\left(\frac{x_{ij} - x_{ij}^{-}}{x_{ij}^{+} - x_{ij}^{-}}\right) \tag{2.7}$$

$$t_{rij} = t_{pij}\left(\frac{x_{ij} - x_{ij}^{+}}{x_{ij}^{-} - x_{ij}^{+}}\right) \tag{2.8}$$

Step 5: Creating the total gap matrix

At this stage, a new matrix is established by taking the difference between the matrices used in the 3rd and 4th stages, and this matrix is called the total gap matrix which is found by subtracting the actual rating matrix from the theoretical rating matrix and is seen in matrix (2.10).

$$g_{ij} = t_{pij} - t_{rij},\ g_{ij} \in \left[0, \infty\right) \tag{2.9}$$

$$G = T_p - T_r = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & & \vdots \\ g_{m1} & g_{m2} & & g_{mn} \end{bmatrix} \tag{2.10}$$

Step 6: Calculation of the final criterion functions of the alternatives (Qi)

The total gap matrix of the alternatives is taken as a basis and the scores are found by summing the rows of each alternative in the gap matrix.

$$Q_i = \sum_{j=1}^{n} g_{ij},\ i = 1, 2, \ldots, m \tag{2.11}$$

Step 7: Identifying the best alternative

In the last stage, according to the Qi scores determined by the formula number (2.11), the lowest final criterion is designated as the best alternative and the ranking is made from smallest to largest.

4 RESULTS AND DISCUSSION

The study consists of seven companies covering the five-year time frame between 2018 and 2022. First of all, the liquidity levels of seven companies were determined using the SD method and are shown in Table 2.1.

According to the SD scores seen in Table 2.1;

- In 2018, the most important liquidity ratio was the cash ratio with approximately 34%, whereas other indicators followed it with 33% and 32%.
- In 2019, the most important liquidity level was the cash ratio with 36.5%, whereas other indicators were listed at 32% and 31%.
- In 2020, unlike other years, the current ratio ranked first at 33.6%, the cash ratio ranked second at 33.5%, and the acid-test ratio ranked last at 32.7%.
- Similar to 2018 and 2019, the cash ratio ranked first at 34.8% in 2021, whereas other indicators ranked at 33% and 32%.
- In 2022, the cash ratio ranked first at 35%, followed by the acid-test ratio and current ratio at 33% and 31%, respectively.

The most important liquidity indicator for the seven companies registered in the Polish Stock Exchange is the cash ratio, which is determined by the SD method

TABLE 2.1
SD scores

Years/indicators	Current ratio	Quick ratio	Cash ratio
2018	0.32401586	0.333390001	0.342594139
2019	0.310321327	0.323900511	0.365778163
2020	0.336881824	0.32760137	0.335516807
2021	0.320237381	0.331309159	0.34845346
2022	0.316529495	0.331017171	0.352453334

scores. Only in 2020, the current rate ranked first. The main reason for this is the harsh impact of the COVID-19 pandemic on the world financial markets in 2020. The reason why the cash ratio is the most important indicator in other years is that the determined companies attach great importance to the level of cash they have in order to experience the liquidity problem at a minimum level.

In the second part of the study, using the liquidity levels determined, the liquidity scores of seven companies were created in a five-year period and the companies were ranked according to their liquidity performances according to these scores.

In order to make the ranking, the MAIRCA method was preferred.

Table 2.2 contains the MAIRCA scores (Qi) for the years 2018–2022. While determining these scores, the criteria weights determined by the SD method were used. While sorting in the MAIRCA technique, the operation of the system is from small to large. Therefore, the best-performing business is the one with the lowest score. According to the results; in the five years between 2018 and 2022, the company with the best liquidity performance is Projekt, while the company with the lowest liquidity performance is Powszechny Zaklad. This result is due to the fact that Projekt's cash level is higher than other companies.

After the analyses were conducted according to the main aims of the study, the situation of the sub-purpose of the research was determined and the five-year observation value of seven enterprises was examined while the liquidity performance scores of five years were created and the research model number (2.12) was established to observe the effects of ROA, ROE, ROC, and financial failure on the liquidity level.

$$\text{Liquidity Scores}\left(\text{LS}\right) = \beta 0\text{it} + \beta 1\text{it}\left(\text{ROA}\right) + \beta 2\text{it}\left(\text{ROE}\right) + \beta 3\text{it}\left(\text{ROC}\right) + \beta 4\text{it Altman Z Score}\left(\text{AZS}\right) + \mu\,\text{it} \tag{2.12}$$

The hypotheses of the research model are;

H0: ROA has no effect on liquidity scores.
H1: ROA has an impact on liquidity scores.
H2: ROE has no effect on liquidity scores.
H3: ROE has an impact on liquidity scores.
H4: ROC has no effect on liquidity scores.
H5: ROC has an impact on liquidity scores.

TABLE 2.2
MAIRCA scores

Firms/scores	Qi	Ranking
	Year 2018	
Projekt	0	1
Orange Polska	0.124628372	2
Jastrzebska	0.13073442	3
Polskie	0.131171971	4
Polsat	0.137710941	5
PGE Polska	0.137788907	6
Powszechny Zaklad	0.14161586	7
	Year 2019	
Projekt	0	1
Polskie	0.101184282	2
Jastrzebska	0.11236971	3
PGE Polska	0.115997023	4
Polsat	0.118173179	5
Orange Polska	0.119246475	6
Powszechny Zaklad	0.142857143	7
	Year 2020	
Projekt	0	1
Polskie	0.063989003	2
Polsat	0.11364412	3
Jastrzebska	0.113804322	4
PGE Polska	0.115983441	5
Orange Polska	0.134089877	6
Powszechny Zaklad	0.142324576	7
	Year 2021	
Projekt	0	1
Polsat	0.114676064	2
PGE Polska	0.125228448	3
Jastrzebska	0.126005606	4
Polskie	0.126112916	5
Orange Polska	0.128829081	6
Powszechny Zaklad	0.142857143	7
	Year 2022	
Projekt	0	1
Polsat	0.11605302	2
Jastrzebska	0.116987386	3
Polskie	0.126465837	4
PGE Polska	0.129492572	5
Orange Polska	0.13086951	6
Powszechny Zaklad	0.142857143	7

H6: AZS has no effect on liquidity scores.
H7: AZS has an impact on liquidity scores.

Four of the established hypotheses are negative hypotheses and four are positive hypotheses. Therefore, the effects of ROA, ROE, ROC, and AZS on liquidity scores were measured using panel regression analysis. Descriptive statistics results of the panel dataset used before panel regression analysis are given below.

According to Table 2.3, there are five variables in the research model, namely LS, ROA, ROE, ROC, and AZS, while the LS variable is the dependent variable, and the ROA, ROE, ROC, and AZS variables are the independent variables. In the descriptive statistics table, since seven companies belonging to five years were examined, there are 35 observation values. The panel regression analysis results for the variables summarized in the descriptive statistics table are presented below (Table 2.4).

TABLE 2.3
Descriptive statistics

Variables	Observations	Mean	Standard deviation	Minimum	Maximum
LS	35	0.0938267	0.0422908	0	0.1416159
ROA	35	0.0642286	0.0944822	–0.103	0.494
ROE	35	0.1194	0.1349654	–0.19	0.636
ROC	35	0.1115143	0.1062205	–0.14	0.491
AZS	35	6.537143	4.233972	–0.8	16.9

TABLE 2.4
Panel regression results

Group indicator: Firms
Number of observed: 35
Number of groups: 7
Correlate (u_i, Xb): 0.000
Prob: 0.0000
R square: 0.8048

LS	Coefficient	Standard error	Z	P > |z|	95% confidence interval	
					Minimum	Maximum
ROA	–0.2020474	0.1763728	–1.15	0.252	–0.5747316	0.1436369
ROE	0.1129496	0.1118556	1.01	0.313	–0.1062833	0.3321825
ROC	0.0127539	0.0781124	0.16	0.870	–0.1403435	0.1658514
AZS	–0.0062893	0.0016159	–3.89	0.000	–0.0094563	–0.0031222
Constant	0.1330094	0.0130405	10.20	0.000	0.1074505	0.1585683

sigma_u: 0.1099623
sigma_e: 0.1707359
rho: 0.29318574

Accordingly, the established research model is found to be meaningful. Because the probability value is less than 5%. In addition, when the R^2 result of the model is examined, it is seen that it is about 80%, and this result shows that the explanation percentage of the model is high. However, according to the results showing the effects of liquidity scores on other indicators, no significant association was detected between liquidity scores and ROA, ROE, and ROC. A significant association was detected between liquidity scores and AZS, which denotes financial failure. The direction of this relationship seems negative. However, since the MAIRCA technique, which measures the liquidity score, performs the performance ranking from small to large, regression analysis was performed according to these scores. Therefore, the direction of the relationship between them is actually positive. In other words, the higher the Altman Z Score and the more successful the companies, the higher the liquidity level. In order for businesses to be successful at Altman Z Score, the Z Score value must be greater than 3. For this reason, businesses with an Altman Z Score of more than 3 are both financially successful and have high liquidity levels. Furthermore, it was observed that the business success increased the liquidity level by 0.6% for the determined sample group.

5 CONCLUSION AND FUTURE WORK

Businesses need funds at every stage of their life cycle. It may not always be possible for a business to access the funds it needs from external sources, both in the establishment phase and in its survival and development. For this reason, it is very important for businesses to have both sufficient capital and sufficient liquidity to evaluate business opportunities and grow. The borrowing ratio and repayment power are issues that should be considered in obtaining the funds that businesses need from external sources. As it is known, the tendency to borrow due to lack of liquidity can cause businesses to experience financial difficulties or even bankruptcy. In this context, the research was built on two purposes. Trying to determine the probability of encountering financial distress by using the liquidity scores calculated on the basis of the liquidity ratios of the enterprises and the Altman Z Score can provide an opportunity to predict whether the enterprises are sustainable. Excessive borrowing can be detrimental to a company and its investors. An uncontrolled level of debt can cause the company to go bankrupt. The insufficient liquidity level of a company can have a negative impact on its financial performance due to its failure to implement its operating plans. Financial performance is important because it gives information about the current financial status of enterprises. Financial performance evaluation is a concept that should be considered by users of financial information in terms of making investment plans for the future, apart from revealing the current situation.

The purpose of the study is to determine the liquidity level of publicly traded enterprises in Poland and to rank these enterprises according to their liquidity performance. In addition, the sub-purpose of the study is the aim is to determine the effect of the determined liquidity scores on ROA, ROE, ROC, and Altman Z Score. In order to make these determinations, five-year data of seven companies traded in the Polish Stock Exchange, between 2018 and 2022, was used. Three different methods were used. In the research, the SD method was used to find the weight of the liquidity ratios, and the MAIRCA technique was used to rank the liquidity performance. Panel regression analysis was performed to measure the impact of other variables on liquidity.

The findings obtained from the study: according to SD scores, the most important liquidity indicator is usually the cash ratio. The main reason for this is to keep the cash level high in order to minimize the liquidity problem of the enterprises. In the MAIRCA technique, which is another method, seven enterprises are ranked according to their liquidity performances as per the importance levels of their liquidity ratios, and Projekt firm has been determined as the firm with the highest liquidity level. Then, according to the results of the panel regression analysis used to measure the impact of ROA, ROE, ROC, and Altman Z Scores on liquidity scores, the panel regression model established at a 95% confidence interval was significant and the explanatory power of the model was found to be 80%. However, while no significant relationship was found between dependent variable liquidity scores and ROA, ROE, and ROC, a significant association was found between liquidity scores and AZS. The direction of this relationship is positive. Therefore, as the AZS value increases, that is, as the companies become successful, the liquidity level also increases. In other words, it has been suggested that successful businesses have high liquidity levels.

The importance of the financial success of enterprises in the development of sustainable economic growth policies of countries is an issue that draws attention day by day in many different aspects, especially in the commercial activity cycle and employment. Considering the financial success and liquidity levels of the enterprises together, the fact that a significant association exists between financial success and liquidity highlights the importance of policymakers' need to focus on this issue sensitively.

The results obtained from the research are the sample that is limited to five years, three liquidity ratios, and seven companies. In order to spread the results more generally, the research should be supported by analyzing different data from Poland with many mathematical, statistical, and econometric methods.

REFERENCES

Afrifa, G. A., & Padachi K. (2016). Working capital level influence on SME profitability. *Journal of Small Business and Enterprise Development*, 23, 44–63.

Akenga, G. (2017). Effect of liquidity on financial performance of firms listed at the Nairobi Securities Exchange, Kenya. *International Journal of Science and Research*, 6(7), 279–285.

Akhtar, T., Tareq, M. A., Sakti, M. R. P., & Khan, A. A. (2018). Corporate governance and cash holdings: the way forward. *Qualitative Research in Financial Markets*, 10(2), 152–170.

Alifiah, M. N. (2013). Prediction of financial distress companies in the consumer products sector in Malaysia. *Sains Humanika*, 64(1), 85–91.

Almeida, H., Campello, M., Cunha, I., & Weisbach, M. S. (2014). Corporate liquidity management: a conceptual framework and survey. *Annual Review of Financial Economics*, 6(1), 135–162.

Al-Najjar, B. (2013). The financial determinants of corporate cash holdings: evidence from some emerging markets. *International Business Review*, 22(1), 77–88.

Alshatti, A. S. (2015). Effect of liquidity management on profitability of Jordanian commercial banks. *International Journal of Business and Management*, 10(1), 62–71.

Amal, Y. A., Sameer, A. A., & Yahya, Z. A. (2012). Factors affecting the financial performance of Jordanian insurance companies listed at Amman stock exchange. *Journal of Management Research*, 4(2), 266–289.

Amess, K., Banerji, S., & Lampousis, A. (2015). Corporate cash holdings: causes and consequences. *International Review of Financial Analysis*, 42, 421–433.

Amihud, Y., Mendelson, H., & Pedersen, L. H. (2005). *Liquidity and Asset Prices* (Vol. 1). Hanover: Now Publishers Inc.

Amoa-Gyarteng, K. (2019). Financial characteristics of distressed firms: an application of the Altman algorithm model. *Journal of Corporate Accounting & Finance*, 30(1), 63–76.

Benson, K., Faff, R., & Smith, T. (2015). Injecting liquidity into liquidity research. *Pacific-Basin Finance Journal*, 35, 533–540.

Bhunia, A. (2010). A trend analysis of liquidity management efficiency in selected private sector Indian steel industry. *International Journal of Research in Commerce and Management*, 1(5), 9–21.

Bradrania, M. R., & Peat, M. (2014). Characteristic liquidity, systematic liquidity and expected returns. *Journal of International Financial Markets, Institutions and Money*, 33, 78–98.

Bruinshoofd, W. A., & Kool, C. J. M. (2004). Dutch corporate liquidity management: new evidence on aggregation. *Journal of Applied Economics*, 7(2), 195–230.

Chang, C.-C., Kam, T.-Y., Chang, Y.-C., & Liu, C.-C. (2019). Effects of the 2008 financial crisis on the working capital management policy of US enterprises. *International Journal of Business & Economics*, 18, 121–140.

Che, Y. K., Gale, I., & Kim, J. (2013). Efficient assignment mechanisms for liquidity-constrained agents. *International Journal of Industrial Organization*, 31(5), 659–665.

Chen, W. D., Chen, Y., & Huang, S. C. (2021). Liquidity risk and bank performance during financial crises. *Journal of Financial Stability*, 56, 100906.

Cretu, D., Iova, A. R., & Nastase, M. (2019). Financial diagnosis of the company based on the information derived from the balance sheet. Case study. *Scientific Papers Series Management, Economic Engineering and Rural Development*, 19, 99–105.

Davydenko, A. S. (2013). *Insolvency, Illiquidity and the Risk of Default*. Toronto: Joseph L. Rotman School of Management, University of Toronto.

Deloof, M. (2003). Does working capital management affect profitability of Belgian firms? *Journal of Business Finance & Accounting*, 30(3–4), 573–588.

Denis, D. J., & Sibilkov, V. (2009). Financial constraints, investment, and the value of cash holdings. *The Review of Financial Studies*, 23(1), 247–269.

Diakoulaki, D., Mavrotas, G., & Papayannakis, L. (1995). Determining objective weights in multiple criteria problems: the critic method. *Computers & Operations Research*, 22(7), 763–770.

Diamond, D. W., & Verrecchia, R. E. (1991). Disclosure, liquidity, and the cost of capital. *The Journal of Finance*, 46(4), 1325–1359.

Dittmar, A., & Mahrt-Smith, J. (2007). Corporate governance and the value of cash holdings. *Journal of Financial Economics*, 83(3), 599–634.

Doina P., & Mircea, M. (2008). Analysis of a company's liquidity based on its financial statements. *Annals of the University of Oradea, Economic Science Series*, 17(3), 1366–1371.

Eckbo, B. E., Thorburn, K. S., & Wang, W. (2016). How costly is corporate bankruptcy for the CEO? *Journal of Financial Economics*, 121(1), 210–229.

Ehiedu, V. C. (2014). The impact of liquidity on profitability of some selected companies: the financial statement analysis (FSA) approach. *Research Journal of Finance and Accounting*, 5(5), 72–80.

Elnahas, A. M., Hassan, M. K., & Ismail, G. M. (2017). Religion and ratio analysis: towards an Islamic corporate liquidity measure. *Emerging Markets Review*, 30, 42–65.

Fang, V. W., Noe, T. H., & Tice, S. (2009). Stock market liquidity and firm value. *Journal of Financial Economics*, 94(1), 150–169.

Fulford, S. L. (2015). How important are banks for development? National Banks in the United States, 1870–1900. *The Review of Economics and Statistics*, 97, 921–938.

Gigovic, L., Pamucar, D., Bajic, Z., & Milicevic, M. (2016). The combination of expert judgment and GIS-MAIRCA analysis for the selection of sites for ammunition depots. *Sustainability*, 8(4), 1–30.

Gopalan, R., Kadan, O., & Pevzner, M. (2012). Asset liquidity and stock liquidity. *Journal of Financial and Quantitative Analysis*, 47(2), 333–364.

Harford, J., Mansi, S. A., & Maxwell, W. F. (2008). Corporate governance and firm cash holdings in the US. *Journal of Financial Economics*, 87(3), 535–555.

Huang, M. (2003). Liquidity shocks and equilibrium liquidity premia. *Journal of Economic Theory*, 109(1), 104–129.

InvestingPro (2022). Data Explorer, 30.06.2022. Review from: https://www.investing.com/pro/watchlist.

Isshaq, Z., & Bokpin, G. A. (2009). Corporate liquidity management of listed firms in Ghana. *Asia-Pacific Journal of Business Administration*, 1(2), 189–198.

Ivanović, Z. (1997). *Financijski menadžment, drugo izmijenjeno i dopunjeno izdanje*. Rijeka: Sveučilište u Rijeci. https://urn.nsk.hr/urn:nbn:hr:191:069288.

Ivashina, V., & Scharfstein, D. (2010). Bank lending during the financial crisis of 2008. *Journal of Financial Economics*, 97(3), 319–338.

Kaliski, B. S. (2007). *Encyclopedia of Business and Finance* (Vol. 1: AI). New York: Macmillan Reference USA.

Kapounek, S., Hanousek, J., & Bily, F. (2022). Predictive ability of Altman Z-score of European private companies. *Politicka Ekonomie*, 70(3), 265–287.

Khan, M. (2017). Z-Score analysis on efficiency of working capital management: an evidence of selected Indian food processing companies listed in NSE. *Pacific Business Review International*, 10, 106–113.

Kimondo, C. N., Irungu, M., & Obanda, M. (2016). The impact of liquidity on the financial performance of the nonfinancial firms quoted on the Nairobi Securities Exchange. *Researchjournali's Journal of Accounting*, 4(2), 1–12.

Lartey, V., Antwi, S., & Boadi, E. (2013). The relationship between liquidity and profitability of listed banks in Ghana. *International Journal of Business and Social Science*, 4(3), 12–34.

Loncan, T. R., & Caldeira, J. F. (2014). Capital structure, cash holdings and firm value: a study of Brazilian listed firms. *Revista Contabilidade & Finanças*, 25, 46–59.

Macey, J. R., & O'Hara, M. (1999). *Globalization, Exchange Governance and the Future of Exchanges*. Ed. R. E. Litan and A. M. Santomero. Brookings-Wharton Papers on Financial Services. Washington, DC: Brookings Institution.

Madhou, A., Moosa, I., & Ramiah, V. (2015). Working capital as a determinant of corporate profitability. *Review of Pacific Basin Financial Markets and Policies*, 18(4), 1550024.

Madushanka, K. H., & Jathurika, M. (2018). The impact of liquidity ratios on profitability. *International Research Journal of Advanced Engineering and Science*, 3(4), 157–161.

Martinho, V. J. P. D. (2022). Profitability and financial performance of European Union farms: an analysis at both regional and national levels. *Open Agriculture*, 7(1), 529–540.

Moyer, C. R., McGuigan, J., & Kretlow, W. (2005). *Contemporary Financial Management*. Boston, MA: Cengage Learning.

Nejadmalayeri, A. (2021). Asset liquidity, business risk, and beta. *Global Finance Journal*, 48, 100560.

Ng, Y. H., Yong, H. H. A., & Faff, R. (2013). The long-and short-run financial impacts of cross listing on Australian firms. *Australian Journal of Management*, 38(1), 81–98.

Nikolaou, K. (2009). *Liquidity (Risk) Concepts: Definitions and Interactions*. Working Paper (1008/February 2009). ECB.

Niresh, J. A. (2012). Trade-off between liquidity & profitability: a study of selected manufacturing firms in Sri Lanka. *Researchers World*, 3(4), 34.

Nomani, A., & Azam, M. K. (2020). Sharīʿah compliant working capital financing – a case-study of Indian sugar industry. *Journal of Islamic Accounting and Business Research*, 11(3), 674–693.

Orshi, T. S. (2016). *Impact of Liquidity Management on the Financial Performance of Listed Food and Beverages Companies in Nigeria*. Master's thesis, Federal University Dutsin-Ma.

Palazzo, B. (2012). Cash holdings, risk, and expected returns. *Journal of Financial Economics*, 104, 162–185.

Pamucar, D. S., Tarle, S. P., & Parezanovic, T. (2018). New hybrid multi-criteria decision-making DEMATEL-MAIRCA model: sustainable selection of a location for the development of multimodal logistics centre. *Economic Research Ekonomska istraživanja*, 31(1), 1641–1665.

Parker, S., Peters, G., & Turetsky, H. (2011). Corporate governance and corporate failure: a survival analysis. *Corporate Governance*, 2(2), 4–12.

Parmeggiani, F. (2013). Rating triggers, market risk and the need for more regulation. *European Business Organization Law Review (EBOR)*, 14(3), 425–463.

Reilly, F. K., & Brown, K. C. (2011). *Investment Analysis and Portfolio Management* (11th edn). Boston, MA: Cengage Learning.

Robinson, T., Henry, E., Pirie, W., & Broihahn, M. (2015). *International Financial Statement Analysis* (3rd edn). Denver, CO: John Wiley &Sons, Inc.

Saleem, Q., & Rehman, R. (2011). Impacts of liquidity ratios on profitability, interdisciplinary. *Journal of Research in Business*, 1(7), 78–91.

Salehi, M., Talebnia, G., & Ghorbani, B. (2011). A study of the relationship between liquidity and stock returns of companies listed in Tehran Stock Exchange. *World Applied Sciences Journal*, 12(9), 1403–1408.

Umar, M., & Sun, G. (2016). Interaction among funding liquidity, liquidity creation and stock liquidity of banks: evidence from BRICS countries. *Journal of Financial Regulation and Compliance*, 24(4), 430–452.

Vavrek, R., Gundová, P., Kravčáková Vozárová, I., & Kotulič, R. (2021). Altman model verification using a multi-criteria approach for Slovakian agricultural enterprises. *Ekonomie a Management*, 24(1), 146–164.

Veronika, F., Tibor, T., & Péter, V. (2014). Financial indicators in managerial decision-making. *Annals of the University of Oradea, Economic Science Series*, 23(1), 893–904.

Vuković, B, Andrić, M., & Jakšić, D. (2017). The impact of working capital management on company's profitability: empirical evidence from Serbia. *Custos e Agronegocio*, 13(1), 333–349.

Wyss, R. (2004). *Measuring and Predicting Liquidity*. PhD diss., Universitat St. Gallen, Zurich.

Zainudin, Z., Kantakji, M. H., Thabet, O. B., Ani, N. S., & Rahman, N. A. (2019). An investigation of the moderating effect of liquidity on the relationship between debt and financial performance of REITs in Malaysia: an optimal liquidity estimation. *Contemporary Economics*, 13(3), 225–238.

Zamanpour, A., & Bozorgmehrian, S. (2012). The study of the effect of liquidity management on return on assets and return on rights of the shareholders of the firms listed on Tehran stock exchange. *Life Science Journal*, 9(4), 1831–1835.

Zeller, T. L., Stanko B. B., & Cleverley, W. O. (1997). A new perspective on hospital financial ratio analysis. *Healthc Financ Manage*, 51(11), 62–6. PubMed.

Zimon, G., & Dankiewicz, R. (2020). Trade credit management strategies in SMEs and the COVID-19 pandemic - a case of Poland. *Sustainability*, 12(15), 6114.

3 The reporting comprehensiveness of financial asset risk and company value

Beata Dratwińska-Kania, Aleksandra Ferens, and Piotr Kania

1 INTRODUCTION

1.1 Theoretical background

Reporting is the main tool for presenting information to stakeholders. The preparation of the annual report is subject to various accounting rules, aimed at generating reports of adequate quality. Determining the quality and quantitative scope of the presented information in corporate reporting should be considered as inspiring public confidence and is not easy. However, it is very much needed because access to such information is recognized worldwide as the basis for value creation (DiPiazza, Eccles, 2002, p. 3) and the basis for gaining trust in reliability and comprehensibility. Stakeholders assess the risk of the firm's activities by referencing financial information, and thus having high-quality information is valuable to them, as it reduces the risk of wrong decisions and should lower the information asymmetry component of the firm's cost of capital (Leuz, Verrecchia, 2000, p. 91). Risk information appears to be useful to stakeholders, and there is no financial report dedicated to this issue alone. In the study, the quality category, in particular reporting comprehensiveness, is examined concerning the risk of financial assets.

The research is therefore organized around the need for information and the originality of the research model.

The goals of this study are:

- The analysis of the quality components of financial statements in the literature,
- Developing a reporting comprehensiveness concept about the disclosures on the risk of financial assets as a quality component of financial statements,
- Examining the reporting comprehensiveness about the risk of financial assets on a sample of Polish companies listed on the Warsaw Stock Exchange,

 DOI: 10.1201/9780367854690-3

- Examining the correlation between the reporting comprehensiveness about the risk of financial assets and the value of the company calculated with accounting methods.

The methodology of examining a reporting comprehensiveness about the risk of financial assets will be based on an original model, which will then be used to audit the reporting presented by Polish listed companies in 2018–2021. The reporting comprehensiveness model about the risk of financial assets will include three main components: The complexity of reporting about the risk of financial assets; the transparency of reporting about the risk of financial assets; and the volume of reporting about the risk of financial assets.

Scoring for the audited enterprises will be awarded based on the individual judgment of the researchers. The scoring on the risk of financial assets created, thanks to the original model, will be used to test the correlation relationships (Spearman's rank correlation) with the value of the company established with accounting measures.

1.2 Comprehensiveness as a qualitative characteristic of financial statements

The literature has defined the quality of financial reporting as the extent to which financial statements provide information useful in making investment decisions (Schipper, Vincent, 2003, p. 98). Plato defined quality as a "degree of perfection," Aristotle, on the other hand, referred to a universal cognitive category that distinguishes objects from each other or determines their similarity (Hamrol, 2005, p. 17). Therefore, from the point of view of accounting, which is the concept of a universal information system as well as control, quality can be perceived as a degree of utility through the financial and non-financial benefits it brings. Appropriate information presented in financial statements should cover what users need to know to assess a company's value and make correct decisions (Tibiletti et al., 2021, p. 57). Furthermore, according to Hairston and Brooks (2019, p. 84), the quality of financial reporting requires financial information to be appropriate and be faithfully presented in financial statements to improve the decision-making usefulness of this information for investors and creditors. Schipper and Vincent (2003) considered that the most important characteristic of the quality of financial information is its usefulness for stakeholder decision-making. They also presumed that utility, in line with the main objective of the requirements of accounting standards, should be supported by relevance, reliability, and comparability (Schipper, Vincent, 2003, p. 103). Relevance and accurate presentation are also cited as important qualities of the quality of financial information under the Conceptual Framework for Financial Reporting (www2, p. 14), as they help investors make optimal decisions. Four characteristics that are normative preconditions required for the government's financial statements to meet the desired quality – relevance, reliability, comparability, and understandability – are also highlighted by Dewi et al. (2019). An organization's financial statements are of high quality when they have all the characteristics resulting from fundamental accounting principles, i.e., the true and fair view principle, accrual principle, materiality, prohibition of offsetting, matching revenues and expenses, prudent valuation, going concern, continuity, and substance over form. According to IFRS, qualitative characteristics of financial statements are such elements that characterize them that make the data included in them useful for external recipients. Fundamental qualitative characteristics are therefore usefulness and faithful presentation, while auxiliary characteristics include relevance, completeness, neutrality, faultlessness, predictive value, confirmatory value, as well as comparability, timeliness, verifiability, and understandability. Similar features should also be characteristic of non-financial statements (Ferens, 2019). The qualitative characteristics of financial statements are presented in Figure 3.1.

The authors assume that, in addition to the characteristics mentioned making up the substance of the presented information, the report quality category should include characteristics that make up the proper reception of information (supporting the comprehensibility characteristics), which can be generally called non-substantive and is examined in this study. Therefore, the study does not examine the content of the information on the risk of financial assets, and consequently, also the issues of whether the presented information on the risk of financial assets is beneficial for the enterprise or not (whether it reflects well on it or not). The authors introduce a model to assess the reporting comprehensiveness of the financial asset risk, which will include three main components: The complexity of reporting on the risk of financial assets; the transparency of reporting on the risk of financial assets; and the volume of reporting on the risk of financial assets.

The model is described in detail in Section 3 of this study. This model was employed to analyze the reports of companies listed on the Warsaw Stock Exchange in

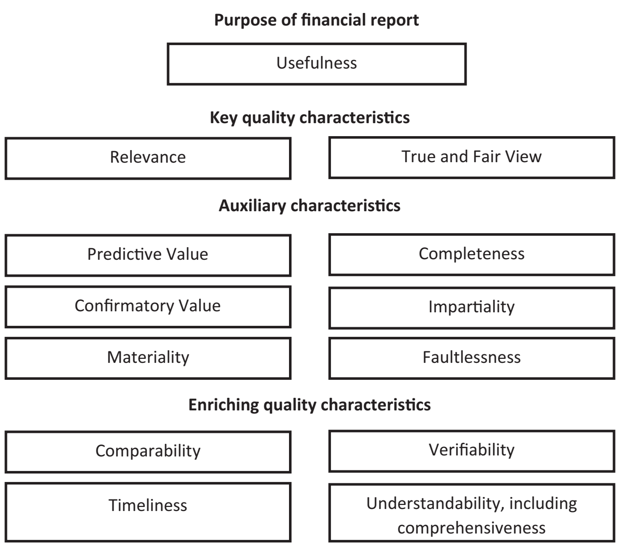


FIGURE 3.1 Qualitative characteristics of financial statements.

Source: Own study based on: International Financial Reporting Standards (IFRS 2013), Accountants Association in Poland, Warsaw, 2014, pp. 42–46; Piosik, A. (ed.), Kształtowanie zysków podmiotów sprawozdawczych w Polsce. MSR/MSSF a ustawa o rachunkowości, C.H. Beck, Warszawa, 2013, p. 37.

2018–2021 and to test the correlation with the company value (Section 4), calculated with accounting measures.

1.3 The concept of financial asset risk in reporting

The risk of financial assets is the probability of a change in the value of financial assets, considered in terms of threats (sources) of these circumstances (financial assets are understood as investments held, resulting from a contracted financial instrument). There are many classifications of risk, the study adopts the classification contained in IFRS 7, which is presented in Table 3.1.

The risk of financial assets is reflected in the balance sheet by changing the value of assets (write-downs) or creating/releasing provisions (Dratwińska-Kania, 2016). The effects of the financial asset risk are reflected in the comprehensive income statement as financial costs or financial income. Moreover, the risk should be described in explanatory notes and the non-financial report (Dratwińska-Kania, 2015, pp. 41–48).

IFRS 7 requires disclosure of the nature and extent of the risks arising from financial instruments to which the entity is exposed during the period involved and at the end of the reporting period, as well as how to manage these risks. This standard requires the so-called sensitivity analysis of the enterprise's exposure to market risk, related to the financial instruments held. Risk disclosures should be both qualitative and quantitative. Qualitative disclosures relate to the description of risk management

TABLE 3.1
IFRS 7 risk classification

Credit risk	The risk that one party will default, and the other party will suffer a loss as a result – also applies to financial assets, which are examined.
Liquidity risk	The risk that the company will encounter difficulties in meeting its obligations – does not apply directly to financial assets, it is not examined.
Market risk	The risk of changes in the fair value of a financial instrument caused by market factors – which also applies to financial assets, is examined.
- currency risk	The risk that the value or future cash flows of a financial instrument will fluctuate due to changes in currency exchange rates.
- interest rate risk	The risk that the value of a financial instrument or the related future cash flows will fluctuate due to changes in market interest rates.
- other price risks	The risk that the value of a financial instrument or the related future cash flows will fluctuate due to changes in market prices caused by factors specific to individual instruments or their issuers, or by factors affecting all instruments traded on the market.
Other risks, e.g., operational, strategic, business, and reputational risk	Of the mentioned types of risk, the operational risk may relate to financial assets – this is the risk of loss resulting from inadequate or failed internal processes, people, systems, or external events. This risk is examined.

Source: own study based on IFRS 7.

objectives, policies, and processes. Quantitative disclosures are intended to provide the necessary information on the extent of an enterprise's risk exposure.

There are no specific rules on how to report risks. Businesses approach this differently. Risk information is extensive or short, is in one place in the report, or is scattered across different parts of the report; it is a coherent whole or incomplete. Therefore, financial asset risk reporting research appears to be an important and urgent issue.

Addressing the topics mentioned also has a practical dimension in the form of an incentive to improve corporate risk reports. The authors assume that an interest in risk reporting will encourage entrepreneurs to improve their reports and policymakers to develop more detailed risk reporting guidelines. Thus, the connection of the reporting comprehensiveness of the financial asset risk with the company value, calculated with accounting measures, is an incentive and stimulus for keeping up with the top runners.

2 LITERATURE REVIEW

The authors have established that the literature only examined the comprehensiveness characteristics concerning Corporate Social Responsibility (CSR) information. For example, Bouten et al. (2011) investigated whether firms comprehensively report CSR information. Comprehensive reporting, as interpreted by them, requires companies to disclose three types of information for each disclosed CSR item: (1) vision and goals, (2) management approach (specific activities), and (3) performance indicators (actual results presented in a valuable manner). The results of the research confirmed the low level of comprehensiveness of CSR reporting. 27% of Belgian companies do not publish CSR data, 73% of companies provide very little information, only 36% of items indicate at least one PI (performance indicators), and over 59% of reporting companies do not provide all three types of information. Research by other authors (Bouten et al., p. 187) suggested that enterprises should comprehensively report, providing information about their (1) goals and intentions, (2) actions, and (3) subsequent results on various CSR issues. For example, for the CSR "emissions" item, this approach means that the company discloses (1) a target to reduce emissions by a certain amount, (2) specific actions to achieve that intention, and (3) the actual reduction achieved. In this way, it is possible to obtain a clearer and more precise picture of the social and environmental responsibility implemented, improve corporate image and transparency, and increase investors' confidence in investment decisions (Yusoff et al., 2013, pp. 213–220).

In the authors' view, the multi-criteria analysis of comprehensiveness was not investigated in the literature. On the other hand, the statement by Reis et al. (2019) is true in that the techniques used for measurement (in our case the comprehensiveness characteristics, as for any other object) should reflect the tested reality as accurately as possible (Wątróbski et al., 2022, p. 2). Therefore, the analysis of the comprehensiveness characteristics is based on three components: complexity, transparency, and volume of reporting about the financial asset risk. These components, according to the authors, reflect the non-substantive dimension of the presented information on the risk of financial assets and constitute the adopted research model, which is presented in Section 3.

The model to assess the reporting comprehensiveness of the financial asset risk proposed here is an original proposition for creating company value, a relationship that has been widely discussed in the literature. For example, Healy et al. (1999) examined the consistency of book values and financial results between two groups of enterprises, using Ohlson's valuation model. They proved that higher-quality disclosures are associated with greater consistency between book values and financial results. According to the management accounting organization (CIMA, 2002), shareholders come first, while recognizing that long-term sustainable value can only be realized when the needs of all stakeholders are considered. Firms should share all information – unless they have a good reason not to. The literature research indicated that clearer disclosures lead to greater detectability of earnings management (Hunton et al., 2006, p. 135). In their view, the increased transparency of reporting by companies significantly reduces, but does not eliminate, the attempts of earnings management. These results suggest that the state or other legislative bodies mainly responsible for the financial reporting of enterprises should require transparency. This will reduce attempts to manipulate the results in more transparent areas or shift such attempts to less transparent ones (Hunton et al., 2006, p. 151).

Research conducted based on data from the analysis of the Environmental, Social, and Corporate Governance (ESG) of Polish companies, investigating whether investing in the most transparent companies is profitable also confirms the thesis that investing in transparent companies is less risky and more profitable (Analiza ESG spółek w Polsce, www1). As emphasized by Marcinkowska (2008, pp. 47–49), the research results indicated that with the increase in the information transparency of the company and the quality of disclosures, the cost of financing the entity decreases. The quality of disclosures affects the assessment of an entity's credibility and creditworthiness, and the ability to obtain certain types of debt financing. Disclosure of information about debt financial instruments also affects the assessment of the effectiveness of shaping the capital structure, choosing the level of debt, cost of capital, or profitability of using debt (Łukasik, 2009). There are also studies conducted on a group of investment funds in Poland that confirm the correlation between the transparency level of the profit and loss account on the operations of investment funds and selected features characteristic for a designated group of funds, including changes in the value of participation units in investment funds (Dratwińska-Kania, 2018).

The broadly understood quality of information disclosed in the financial statements – concerning selected aspects of business entities and a comprehensive presentation of their financial situation and performance – has been researched by many authors (Biddle et al., 2009; Bushman, Smith, 2001, Leuz, Verrecchia, 2000). Bushman and Smith (2001) drew attention in their theoretical considerations, among others, to the importance of financial information to reduce asymmetry among investors. Leuz and Verrecchia (2000) focused on such issues as increasing the level and quality of disclosures as an important component of reducing information asymmetry and a key aspect in protecting the interests of investors and creditors (Leuz, Verrecchia, 2000). Studies of German companies on the impact of increased reporting, qualitatively adjusted to the international reporting strategy, proved that the level of disclosures brings benefits that are economically and statistically significant. It should be noted that the authors of this research interpreted the term "increased level of disclosures" on a par with increasing the quality of disclosures. The degree of financial reporting quality and investment efficiency was studied by Biddle et al. (2009). The results of their research indicated that higher quality of investments supplied by financial reporting is connected with lower investment among cash-rich and unlevered companies, and higher investment among cash-tight and highly leveraged companies. Furthermore, companies providing high-quality financial reports invest less when aggregate investment is high, and invest more when aggregate investment level is low.

Following this observation, Chang, Dasgupta, and Hilary (Biddle et al., 2009, p. 8) proposed a model of dynamic adverse selection and proved empirically that companies have more flexibility to issue capital when their financial reporting is better. They posited that when financial reporting quality downscales adverse selection costs, it can be connected with investment efficiency through the reduction in external financing costs and through the reduction in the likelihood that a firm obtains excess funds because of temporary mispricing. They advocate that high-quality financial reporting also drives the decrease in adverse selection. Another model, where information asymmetry between the company and investors triggered the company's underinvestment, was developed by Myers and Majluf. They noted that when managers take measures favorable toward existing shareholders and the company needs to seek financial support for an impending positive NPV project, managers may object to raising funds at a discounted price even if that results in turning down good investment opportunities (Verdi, 2006, pp. 5–6). Szewieczek, Dratwińska-Kania, and Ferens (2021) examined a number of disclosures on the business model by listed companies and also investigated the correlation between the number of disclosures and selected economic parameters characterizing enterprises, such as total assets, performance, board, EBIT, fixed assets, current assets, equity, liabilities. Significant correlation relationships were indicated.

The second important aspect of reporting comprehensiveness analysis is the impact of the SARS-CoV-2 pandemic on the information presented. Similar problems have been discussed in the literature concerning various issues. For example, Tibiletti et al. (2021) examined the information in financial statements and related documents (accounting notes, comments) against the requirements that public companies are expected to provide for disclosure due to SARS-CoV-2 pandemic, specifically about the going concern principle. The study concerned Italian companies. It has been investigated whether or not the companies showed consistency in their reporting of future performance by comparing what was reported in the financial statements for the year ended December 31, 2019, and the results achieved as of June 30, 2020. The analysis showed that the information that the companies declared as of December 31, 2019, differed significantly from the information obtained on June 30, 2020. The companies recorded mainly a decrease in revenues compared to what was declared, which means that the companies underestimated the impact of the SARS-CoV-2 pandemic. The research also revealed that the scope of risk information presented in the financial statements has expanded in the companies that have suffered the most from the SARS-CoV-2 pandemic. Information on liquidity risk, credit risk, and sections "letter to shareholders" was particularly extensive. The scope of information on the impact of the SARS-CoV-2 pandemic on the company's results was included in most reports, though not all companies experienced a negative impact.

Lassoued and Khanchel (2021) studied the impact of the SARS-CoV-2 pandemic on earnings management. The results of empirical research obtained by them on 2,031 companies from 15 European countries suggest reduced credibility of financial statements generated during the pandemic and a significant profit manipulation in 2020. Their findings indicated that during the pandemic, companies managed earnings upward by alleviating the level of reported losses to rebuild investors and stakeholders. Other studies also showed that in times of crisis, companies tend to manage their profits upward to survive the economic downturn, mitigate the effects of the crisis, maintain positive relationships with stakeholders, or avoid falling stock prices (Arthur et al., 2015; Lisboa, Kacharava, 2018). Similar studies were conducted in the case of natural disasters. In such circumstances, managers managed their profits upward, wanting to increase their remuneration or secure their position in the company (Lassoued, Elmir, 2012; Morck et al., 1990).

The second topic in the literature indicates that companies in times of crisis demonstrate a decrease in profit, which is justified by the existing situation and becomes a reason, for example, to renegotiate repayments and debt conditions, employee concessions, obtain government support, or other benefits (Asquith et al., 1994; Filip, Raffournier, 2014). Similarly, Chen et al. (2021),found that the economic concerns caused by SARS-CoV-2 pandemic outweighed other opportunistic considerations. Bugshan et al. (2020) indicated that companies in the Gulf Cooperation Council (GCC) countries showed a tendency to manipulate profits downward during the oil price crisis in 2014.

The third topic in the literature argues that periods of economic downturn (crisis) are less favorable for profit management than periods of expansion, as these periods are much more closely monitored by auditors, and showing poor performance during these periods is justified and tolerated (Ahmad-Zaluki et al., 2011; Chia et al., 2007). Therefore, in these periods, companies decide to manage profits much less frequently.

3 PROPOSED WORK

The methodology of examining the reporting comprehensiveness of financial asset risk is based on an original model, which is then used to audit the reporting presented by the listed Polish companies in 2018–2021. The model to assess the reporting comprehensiveness of the financial asset risk includes three main components: The complexity of reporting about the risk of financial assets; the transparency of reporting about the risk of financial assets; and the volume of reporting about the risk of financial assets.

The risk complexity component is designed to examine whether the financial statements contain risk information and whether this information covers all types of risk (equity risk, currency risk, interest rate risk, credit risk, operational risk), and the way of managing the risk. The sensitivity analysis establishes the market risk associated with financial assets. 1 point is awarded for each type of risk reported.

The risk reporting transparency component, i.e., clarity of the information on the risk of financial assets, examines two characteristics: The consistency in the reporting rhetoric about the risk of financial assets (whether the same economic categories and other terms are used in the risk reporting of financial assets, whether risk communication is uniform). 1 point will be awarded for positive consistency) and the cohesion or dispersion of information about the risk of financial assets in different parts of financial statements. If the information on the risk of financial assets was in one place, 1 point was awarded. If patterns of financial asset risk were dispersed across different parts of the report, no point was awarded.

The last component of the model is the reporting volume on the risk of financial assets. Here, a score is given for each risk information page in the annual financial statements. Only the total number of points is awarded. No point is awarded if the reporting covers less than half a page. In general, the points awarded here reflect the amount of reported information on the risk of financial assets.

Scoring for the audited enterprises is awarded based on the individual judgment of the researchers. The scoring on the risk of financial assets created thanks to the original model is used to test the correlation relationships (Spearman's rank correlation) with the value of the company established with accounting measures (total assets, fixed assets, current assets, equity capital, comprehensive income, ROA, and ROE).

There is a stream in the literature to prove a positive relationship between the generation of high-quality financial statements, the economic and financial condition of the company, and its ability to create company value. Based on the above, the authors put forward hypotheses:

H1: It is assumed that the reporting comprehensiveness of the financial asset risk varies depending on the company value, calculated with accounting methods.

It also seems important to investigate the impact of the SARS-CoV-2 pandemic on the reporting comprehensiveness of financial asset risk. The authors put forward the following hypothesis:

H2: It is assumed that the SARS-CoV-2 pandemic affected the reporting comprehensiveness of financial asset risk.

The authors demonstrate that the reporting comprehensiveness of financial asset risk, i.e., complexity, transparency, and volume, differed during the pre-pandemic and the period of the pandemic.

4 RESULTS AND DISCUSSION

4.1 Analysis of the reporting comprehensiveness of the risk of financial assets accounted by Polish enterprises

The empirical study was conducted on large Polish companies that are not financial institutions, apply the accounting principles contained in IFRS, have financial assets in their portfolio, and report them. It was important for the authors that the selected entities prepared separate financial statements under the International Accounting Standards, and reported there on the risk of financial assets.

24 companies were selected for the study – such a sample was found to be representative of Poland. Four years were selected for the study: 2018 and 2019, in which

there was no SARS-CoV-2 pandemic, and 2020 and 2021 with the SARS-CoV-2 pandemic.

Since the pandemic prevailed for two years, the two years preceding it were also taken into account.

The empirical verification of the model of reporting on the risk of financial assets is presented in Table 3.2.

When analyzing Table 3.2, it should be noted that the points awarded to a given company for specific components in the analyzed years did not differ much. This was especially true for components such as complexity and transparency. This was because the information on the risk of financial assets was often copied from year to year. The content was very similar in the analyzed years. Most often, the risk reports

TABLE 3.2

Empirical verification of the model of reporting comprehensiveness model of financial asset risk

	Complexity				Transparency				Volume			
Company	**2018**	**2019**	**2020**	**2021**	**2018**	**2019**	**2020**	**2021**	**2018**	**2019**	**2020**	**2021**
Lotos s.a.	6	6	6	6	2	2	2	2	7	7	7	11
PKN Orlen s.a.	5	5	5	6	2	2	1	1	4	5	9	8
PGE s.a.	5	5	5	5	1	1	1	1	7	7	8	10
Tauron s.a.	6	6	6	6	2	2	2	2	10	10	10	10
Enea s.a.	4	4	4	4	2	1	1	1	10	10	10	10
PGNIG s.a.	2	4	4	4	1	1	1	1	7	5	4	4
JSW s.a.	6	6	6	6	2	2	2	2	7	7	10	9
KGHM s.a.	6	6	6	6	2	2	2	2	13	18	22	25
Unimot s.a.	5	6	6	6	1	1	1	1	5	5	5	6
LPP s.a.	4	4	4	4	2	2	2	2	4	2	2	2
Orange s.a.	7	7	7	7	2	2	2	2	7	7	7	7
Cyfrowy Polsat s.a.	6	5	6	6	1	1	1	1	16	16	16	14
Kogeneracja s.a.	5	5	5	5	2	2	2	2	8	7	8	8
CCC s.a.	5	4	5	5	1	1	1	1	6	7	8	6
Asseco s.a.	5	5	5	5	2	2	2	2	7	7	7	7
Grupa Azoty s.a.	6	6	6	6	2	2	2	2	8	10	10	9
Żywiec s.a.	5	5	5	5	2	2	2	2	3	3	3	3
PKP Cargo s.a.	5	5	5	5	2	1	1	1	6	7	7	6
Rafako s.a.	5	5	5	5	1	1	1	1	8	8	11	8
Wawel s.a.	4	4	4	4	2	2	1	1	1	1	3	2
Agora s.a.	5	5	5	5	1	1	1	1	9	8	7	8
Asseco see s.a.	4	5	5	5	2	2	2	2	2	5	5	5
Grupa Kęty s.a.	6	6	6	6	2	2	2	2	6	6	5	6
Kruk s.a.	6	6	6	6	2	2	2	2	10	11	10	10

Source: own study.

differed in the sensitivity analysis performed. It should also be noted that very few companies were awarded the maximum 7 points for the complexity component. The companies rarely informed others of risk in sharing prices and the operational risk. Most often, points were awarded for reporting on market risk. The range of information on the risk of financial assets (complexity component) was 5 points, and the awarded score was in the range <2:7>. Most of the audited companies did not report much on risk management, while market risk and sensitivity analysis were discussed more fully.

In the transparency component, the first parameter considered, i.e., rhetorical consistency, was assessed as the same for all audited companies. All the companies were awarded 1 point for this parameter. Thus, it can be concluded that economic categories are used coherently and consistently in financial reporting. The spread for the transparency component is small, 1 point. The score given is in the range <1:2>. When analyzing the reports on the risk of financial assets, there were also some inaccuracies – some companies treated trade receivables as commercial receivables and classified them as financial instruments. Nevertheless, this parameter has not been tested and assessed.

When examining the volume component, a broad range of published information on the risk of financial assets was noticed – from 1 page to 25 pages. Nevertheless, the companies most often reported on the medium level risk – the mode, and similar to this – the median – are presented in Table 3.3. Moreover, the most extensive risk reports were published by companies such as KGHM SA and Cyfrowy Polsat SA. The least information about risk was published by Wawel SA and LPP SA.

The period of the SARS-CoV-2 pandemic (2020–2021) did not represent much diversity in the complexity parameter compared to the period before the outbreak of the pandemic (2018–2019). In the years 2020–2021, the awarded scores for the complexity component were the same for many companies. The scoring in 2018–2019, compared to the period of the pandemic was different for four companies: PGNiG SA, Cyfrowy Polsat SA., CCC SA, and Asseco See SA.

The period of the pandemic did not have a significant impact on the scores for the transparency component. Differences of 1 point were recorded in the case of four companies: PKN Orlen SA, Enea SA, PKP Cargo SA, and Wawel SA. It should be noted, however, that for this component, 1 point is 50% of possible points. Therefore, when analyzing from this angle, it should be stated that the differences in scores for these companies were significant. The changed score was related to the parameter of

TABLE 3.3
Mode and median for the volume component

Measurement/year	2018	2019	2020	2021
Mode	7	7	7	8
Median	7	7	7.5	8

Source: own study.

cohesion or dispersion of information on the risk of financial assets. In addition, in all the cases, the pandemic score shifted downward compared to the pre-pandemic period. Some companies published information on the risk impact of the SARS-CoV-2 pandemic, which was included in another part of the report, and was, therefore, awarded fewer points in the pandemic years, compared to the pre-pandemic period, due to the spread of risk information.

For the volume component, the amount of information on the risk of financial assets in 2020–2021, compared to 2018–2019, changed quite often. In 11 cases, it was a change from a smaller to a larger volume of information on the risk of financial assets. In four cases, it was a change from more to less information. In five cases, the changes were varied – in one case (in one year) the changes were decreasing and in the second year, they were increasing the amount of information about the risk. The most common change in the amount of information presented on the risk of financial assets was the sensitivity analysis. To sum up, for the volume component, the period 2020–2021 (pandemic period) was rather more comprehensive than the period 2018–2019 (pre-pandemic period).

4.2 Examination of correlation relationships between the reporting comprehensiveness of financial asset risk and the size of the company

In the next stage of the research, the relationship between the points awarded for the reporting model of the financial asset risk for all its components (complexity, transparency, and volume) and the accounting measures of the company's value, among which the following parameters were selected: total assets, fixed assets, current assets, equity capital, comprehensive income, ROA, and ROE. The Shapiro–Wilk test was carried out to investigate the distribution of the studied variables. Since the variables do not have a normal distribution, Spearman's rank correlation method was chosen to study the relationship between the variables. The significance level of the study was set at 0.1. The results of the correlation calculations for all model components for each year are presented in Table 3.4.

When analyzing the results of the study, it should be noted that there are significant correlation relationships (marked in Table 3.4), mainly for the volume component, and they are not too large. In addition, for the ROA and ROE parameters and the volume component, the correlations have a negative sign. This means that there is a slight relationship between the amount of reported information on the risk of financial assets and the accounting measures of the size of the company – for the total assets, fixed assets, current assets, and comprehensive income in selected years, this relationship is positive. Therefore, for the growing amount of risk information, it increases also parameters measuring the company value (in an accounting manner). For the ROA and ROE parameters in 2018, 2019, and 2020, the correlations are negative, which means that the most information about risk was presented by the company with the lowest ROA and ROE parameters. It should also be noted that the correlations with ROA and ROE were also relatively greater than for other examined parameters, except for the correlation with fixed assets in 2021 and comprehensive income in 2021.

For the model component: transparency, a significant correlation was obtained only in 2018 and 2019 with the ROA and ROE parameters, and these correlations are positive. The transparency component consists of two characteristics, but in practice, the company's reporting differed only in the spread of risk information across the various parts of the financial statements. A positive correlation means that companies that consistently reported risk (in one place) had higher ROA and ROE parameters than companies whose risk information was scattered across different parts of the financial statements.

For the complexity component, significant and positive correlations were obtained in 2021, with the parameters such as comprehensive income, ROA, and ROE. This means that in this case, the companies that reported more disclosures (risk types, risk management, sensitivity analysis) also showed greater comprehensive income, ROA, and ROE.

The study contributes to the discussion on disclosing the risk of financial assets. The conducted research proves that all the audited companies go into sufficient detail reporting on the risk of their financial assets. No company received a total of 0 points for any of the analyzed components. Nevertheless, the companies report on risks differently – in a stakeholder-friendly manner. This is evidenced by the awarded scores, which are presented in Table 3.2 and discussed in Section 4.

In the analyzed companies, the range of reporting comprehensiveness of the financial asset risk was quite large for the volume component and relatively small for the transparency component (however, 1 point difference in this component equals 50% of the score, so when analyzing this component separately, it can also be said that the differences in the scores were large), and the mean for the complexity component. The research findings indicate that companies mainly report on individual types of risk and are very economical with reporting on risk management – it is most often information on the adopted hedging instruments and sensitivity analysis. On the other hand, not much information could be found in the separate financial statement on the objectives or policy of risk management.

The authors assume that risk reporting needs to be improved. In the current situation, reports of Polish companies tend to fail to inform potential stakeholders about risk. Therefore, measures such as complexity and volume of risk information were adopted in the model used. Nevertheless, the opposite problem must be borne in mind. The opposing category for increasing disclosures in financial statements is their brevity. Slimming down financial statements has also been one of the dominant narratives in accounting in recent years. Comprehensive reporting is not only rich in additional disclosures but also concise, so comprehensiveness is looking for a compromise between these two categories.

As regards the reporting comprehensiveness of the financial asset risk before and after the SARS-CoV-2 pandemic, this study is part of the trend of research confirming the medium, but significant impact of the pandemic on corporate reporting on the risk of financial assets. It can be stated that the companies reported more risks during the SARS-CoV-2 pandemic, but the transparency of this information and content showed moderate variability compared to the pre-pandemic period.

The conducted empirical research confirms the theoretical assumption that reporting comprehensiveness as a component of the quality of the report affects the value

TABLE 3.4
Spearman's rank correlation – study results

Accounting measures of company size	Complexity 2018	Complexity 2019	Complexity 2020	Complexity 2021	Transparency 2018	Transparency 2019	Transparency 2020	Transparency 2021	Volume 2018	Volume 2019	Volume 2020	Volume 2021
Total assets	0.25893	0.08745	0.11735	0.25193	0.09932	0.08082	–0.01812	0.00604	0.36338	0.29378	0.36829	0.46465
Fixed assets	0.26123	0.10966	0.14343	0.25845	0.12581	0.01865	0.03020	0.01812	0.39062	0.32445	0.38675	0.50582
Current assets	0.14308	0.00139	0.04424	0.14715	0.05959	–0.06838	–0.11477	–0.09060	0.12567	0.05067	0.25227	0.34903
Equity	0.14723	0.03331	0.10617	0.25007	0.03311	–0.03108	–0.04228	0.00604	0.36645	0.30801	0.37708	0.54435
Comprehensive income	–0.14769	0.00000	–0.06752	0.45543	0.27148	0.30462	0.13893	–0.00604	–0.08085	–0.13689	–0.27776	0.54567
ROA	–0.23077	0.02221	–0.11362	0.32131	0.40391	0.41652	0.25973	0.04228	–0.48157	–0.49023	–0.63550	0.09372
ROE	–0.20308	0.06663	–0.10804	0.37114	0.41715	0.39165	0.11477	0.07852	–0.42050	–0.45290	–0.42059	0.02146

Source: own study.

of the company, calculated with accounting measures. This is one more affirming voice in the discussion that concerns issues such as:

- quality and profit shaping (Hunton et al., 2006),
- information transparency and reducing the financing cost (i.e., the impact on the financial result) (Marcinkowska, 2008, pp. 47–49),
- reporting quality and increasing investment efficiency, i.e., the impact on the financial result (Bushman, Smith, 2001),
- reporting quality and stock price (Myers, Majluf, 1984).

In this thematic area, the authors found only one study in which the opposite trend was demonstrated, i.e., studies in which the quality of the financial statements would not improve the functioning of the enterprise (measured by various measures). Lundholm and Myers (2002) in their studies showed no clear evidence that the consistency of accounting information on current earnings increases as the quality of disclosures increases. According to their research findings, corporate disclosures provide reliable and consistent information not reflected in current earnings but reflected in current share prices.

The obtained research results are consistent with the research conducted by Szewieczek, Dratwińska-Kania, and Ferens (2021), in which the amount of information about the business model was positively correlated with selected economic parameters characterizing the company's value, i.e., assets, performance, EBIT, equity, or liabilities. In our case, the conducted research shows that there is a significant, although weak, correlation between the amount of risk information and parameters such as total assets, fixed assets, current assets, equity, comprehensive income, ROA, and ROE. We can therefore conclude that the amount of risk information presented is greater for larger enterprises that perform better.

5 CONCLUSION

The study presents a model to assess the reporting comprehensiveness of the risk of financial assets as a quality component of financial statements and its practical verification, based on Polish companies listed on the Warsaw Stock Exchange in the period 2018–2021. This model is an informative supplement to the basic qualitative characteristics of financial statements with features that make up the proper reception of information (supporting the understandability characteristics), which are addressed in this study as non-substantive, meaning that the reporting content was not examined, but complexity, transparency, and volume of information on the risk of financial assets.

The study concluded that, during the SARS-CoV-2 pandemic, companies reported a relatively greater amount of information about the risk of financial assets than in the pre-pandemic period, while the content of the reported categories (risk types, risk management, sensitivity analysis), i.e., the complexity component of the model, was only slightly more varied. In the case of the transparency component, the period of the pandemic did not significantly affect the score in the model.

The results of the practical model verification of the reporting comprehensiveness of the financial asset risk were used to examine the correlation relationships (Spearman's rank correlation, significance level 0.1) with accounting parameters characterizing the company value, such as total assets, fixed assets, current assets, equity, comprehensive income, ROA, and ROE. Significant correlations were obtained, the most numerous for the volume component of the model. It should be noted that the volume component of the model reached results that confirmed the adopted research hypotheses to the greatest extent.

REFERENCES

Ahmad-Zaluki, N. A., Campbell, K., Goodacre, A. (2011). Earnings management in Malaysian IPOs: The East Asian crisis, ownership control, and post-IPO performance. *International Journal of Accounting*, 46(2), pp. 111–137.

Arthur, N., Tang, Q., Lin, Z. (2015). Corporate accruals quality during the 2008–2010 global financial crisis. *Journal of International Accounting Auditing and Taxation*, 25, pp. 1–15.

Asquith, P., Gertner, R., Scharfstein, D. (1994). Anatomy of financial distress: An examination of junk-bond issuers. *The Quarterly Journal of Economics*, 109, pp. 625–658.

Biddle, G. C., Hilary, G., Verdi, R. S. (2009). How does financial reporting quality relate to investment efficiency? *Journal of Accounting and Economics*, 48(2–3), pp. 112–131.

Bouten, L., Everaert, P., Van Liedekerke, L., De Moor, L., Christiaens, J. (2011). Corporate social responsibility reporting: A comprehensive picture? *Accounting Forum*, 35(3), pp. 187–204.

Bugshan, A., Lafferty, G., Bakry, W., Li, Y. (2020). Earnings management during the oil price crisis. *Journal of Applied Economic Sciences*, 68(2), pp. 297–309.

Bushman, R., Smith, A. (2001). Financial accounting information and corporate governance. *Journal of Accounting and Economics*, 31, pp. 237–333.

Chen, H., Liu, S., Liu, X., Wang, J. (2021). Opportunistic timing of management earnings forecasts during the COVID-19 crisis in China. *Accounting and Finance forthcoming*. https://doi.org/10.1111/acfi.12830.

Chia, Y. M., Lapsley, I., Lee, H.-W. (2007). Choice of auditors and earnings management during the Asian financial crisis. *Managerial Auditing Journal*, 22, pp. 177–196.

CIMA Executive Briefing (2002). Business transparency in past-Enron world, restoring trust in capital markets, August. http://www.cimaglobal.com/Documents/ ImportedDocuments/ BustransparencypostEnron_techrpt_0802.pdf (dostęp: 10.02.2010).

Dewi, N., Azam, S., Yusoff, S. (2019). Factors influencing the information quality of local government financial statement and financial accountability. *Management Science Letters*, 9(9), pp. 1373–1384.

DiPiazza Jr, S. A., Eccles, R. G. (2002). *Building public trust: The future of corporate reporting*. Denver, CO: John Wiley & Sons.

Dratwińska-Kania, B. (2015). *Teoretyczne problemy sprawozdawania o ryzyku*. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu Nr 388, Wrocław, pp. 41–48.

Dratwińska-Kania B. (2016). *Rachunkowość instrumentów finansowych*. Wydawnictwo UE w Katowicach, Katowice.

Dratwińska-Kania B. (2018). *Transparentność rachunku wyniku z operacji funduszu inwestycyjnego – ujęcie teoretyczne i propozycja pomiaru*. Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice.

Filip, A., Raffournier, B. (2014). Financial crisis and earnings management: The European evidence. *International Journal of Accounting*, 49(4), pp. 455–478.

Hamrol, A. (2005). *Quality management with examples*. PWN, Warsaw, p. 17.

Healy, P., Hutton, A., Palepu, K. (1999). Stock performance and intermediation changes surrounding sustained increase in disclosure. *Contemporary Accounting Research*, 16(3), pp. 485–520.

Fijałkowska, J., Hadro, D. (2022). Risk information in non-financial disclosure. *Risks*, 10(1), p. 11.

Ferens, A. (2019). Informacje niefinansowe w sprawozdawczości spółek branży energetycznej. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 63(11), pp. 7–18.

Hairston, S. A., Brooks, M. R. (2019). Derivative accounting and financial reporting quality: A review of the literature. *Advances in Accounting*, 44, pp. 81–94.

Hunton, J. E., Libby, R., Mazza, C. L. (2006). Financial reporting transparency and earnings management (retracted). *The Accounting Review*, 81(1), pp. 135–157.

Lassoued, N., Khanchel, I. (2021). Impact of COVID-19 pandemic on earnings management: An evidence from financial reporting in european firms. https://doi.org/10.1177/09721509211053491.

Lassoued, N., Elmir, A. (2012). Portfolio selection: Does corporate governance matter? *Corporate Governance: The International Journal of Business in Society*, 12(5), pp. 701–713.

Leuz, C. Verrecchia, R. E., (2000). The economic consequences of increased disclosure. *Journal of Accounting Research*, 38, pp. 91–124, https://doi.org/10.2307/2672910.

Lisboa, I., Kacharava, A. (2018). Does financial crisis impact earnings management evidence from Portuguese and UK. *European Journal of Applied Business and Management*, 4(1), pp. 80–100.

Łukasik, G. (2009). Podstawy nowoczesnego zarządzania finansami przedsiębiorstwa. W: A. Bielawska (red.), *Nowoczesne zarządzanie finansami przedsiębiorstwa*. Wydawnictwo C.H. Beck, Warszawa, pp.14–29.

Lundholm, R. L., Myers, C. A. (2002). Bridging the future forward: The effect of disclosure on the returns – Earnings relation. *Journal of Accounting Research*, 40(3), pp.809–839.

Marcinkowska, M. (2008). Wpływ władztwa korporacyjnego na tworzenie wartości przedsiębiorstwa. W: St. Rudolf (red.), *Rola nadzoru korporacyjnego w kreowaniu wartości przedsiębiorstwa*. Wydawnictwo Uniwersytetu Łódzkiego, Łódź, pp. 35–58.

Morck, R., Shleifer, A., Vishny, R. W. (1990). Do managerial objectives drive bad acquisitions? *The Journal of Finance*, 45(1), 31–48. https://doi.org/10.2307/2328808.

Piosik, A. (2013). *Kształtowanie zysków podmiotów sprawozdawczych w Polsce. MSR/MSSF a ustawa o rachunkowości*. C.H. Beck, Warszawa.

Reis, I. F., Ferreira, F. A., Meidute-Kavaliauskiene, I., Govindan, K., Fang, W., Falcão, P. F. (2019). An evaluation thermometer for assessing city sustainability and liability. *Sustainable Cities and Society*, 47, Article 101449. http://doi.org/10.1016/j.scs.2019.101449.

Schipper, K., Vincent, L. (2003). Earnings quality. *Accounting horizons*, 17, pp. 97–110.

Szewieczek, A., Dratwińska-Kania, B., Ferens, A. (2021). Business model disclosure in the reporting of public companies – An empirical study. *Sustainability*, 13, Article 10088.

Tibiletti, V., Marchini, P. L., Gamba, V., Todaro, D. L. (2021). The impact of COVID-19 on financial statements results and disclosure: First insights from Italian listed companies. *Universal Journal of Accounting and Finance*, 9(1), pp. 54–64.

Verdi, R. S. (2006). Financial reporting quality and investment efficiency. Available at SSRN 930922.

Wątróbski, J., Bączkiewicz, A., Ziemba, E., Sałabun, W. (2022). Sustainable cities and communities assessment using the DARIA-TOPSIS method. *Sustainable Cities and Society*, 83, pp. 1–24.

Yusoff, H., Mohamad, S. S., Darus, F. (2013). The influence of CSR disclosure structure on corporate financial performance: Evidence from stakeholders' perspectives. *Procedia Economics and Finance*, 7, pp. 213–220.

(www1) https://www.slideshare.net/EYPoland/analiza-esg-spek-w-polsce-prezentacja-wynikw-badania.

(www2) https://www.ifrs.org/content/dam/ifrs/publications/pdf-standards/english/2021/issued/part-a/conceptual-framework-for-financial-reporting.pdf.

4 Gold as an alternative investment in times of turmoil

Blandyna Puszer and Maria Czech

1 INTRODUCTION

With the development of financial markets, the concept of alternative investments has emerged, and investors have become keen to use alternative forms of financing. Investors' decisions were influenced, among other things, by studies that indicated a negative correlation between equities and alternative investments. The interest of investors increased rapidly, as the purchase of alternative financial market instruments contributed to the reduction of risk and allowed them to properly diversify their investment portfolio (Aspadarec, 2013, pp. 9–10). Alternative investments are defined as investments outside the realm of well-known financial instruments such as shares, bonds, other debt instruments and banking instruments that include, among others, certificates of deposit (Swedroe and Kizer, 2008, p. 23). Accordingly, a whole catalogue of alternative investments can be distinguished, which include hedge funds, fund of funds (FOF), private equity funds, venture capital funds, structured products, credit derivatives, currencies, commodity markets, real estate markets, collector coin markets, emotional investments, and infrastructure investments.

One form of alternative investment is investment in the commodities' market, which includes energy commodities, agricultural commodities, industrial metals, and precious metals (Gierałtowska, 2013, p. 89). The most popular precious metal is gold, which is treated as the global currency of the world. According to capital market analysts, gold should be the foundation of any investment, as an investment in gold is virtually risk-free and is considered to be timeless and the safest in the world. Gold is a metal that has been an indicator of wealth and affluence for centuries. It played a key role in the development of the international monetary system – until the 20th century it served as the main means of payment. Currently, gold is widely used in jewellery, electronics, industry, and medicine. It is seen as a hedge against inflation and declining purchasing power (Borowski, 2016, p. 27), and in times of increased uncertainty and economic downturn, gold becomes a safe instrument and an indispensable insurance policy; thus, the demand for it is increasing. Today, there are also strong disturbances in the economic and social spheres, caused, among others, by the Covid-19 pandemic and the armed conflict (the Ukrainian–Russian war). Meanwhile, in the long term gold gains in value and is an asset desired by individual and institutional investors.

 DOI: 10.1201/9780367854690-4

The research subject in this chapter is the examination and evaluation of gold demand and supply, as well as fluctuations in world gold prices as a result of exogenous factors (Covid-19 and the Ukrainian–Russian war). There are both theoretical and practical reasons for adopting such a formulated research subject. The current state of knowledge regarding the situation on the global gold market is incomplete. The issue of the impact of pandemics and armed conflict on the situation on the gold market is important not only from a cognitive point of view, but also for economic practice. The situation and development of this market during turmoil is important from the point of view of individual and institutional investors, business entities, as well as the economy as a whole. Meanwhile, the results of the research allow for the conclusion that the Covid-19 pandemic and the armed conflict have determined the situation on the global gold market. This state of affairs prompted the study of the relationship between the demand for investment gold and the development of the pandemic and the armed conflict.

Consequently, the main objective of the study is to examine the impact of the Covid-19 pandemic and the armed conflict on the level of world gold prices. The specific objective is to examine the gold market in terms of supply and demand during the Covid-19 pandemic and the armed conflict. The implementation of the main objective consists of theoretical-cognitive and application objectives. The implementation of the theoretical-cognitive objective in terms of presenting existing knowledge required the identification of detailed areas regarding the situation in the global gold market (primarily during the pandemic). In turn, the implementation of the theoretical and cognitive goal required the study of the size and dynamics of the gold price and the development of a model of the relationship between the determinants of gold demand during a pandemic. The application (practical) objective was to draw conclusions, as well as postulates for various groups of stakeholders interested in the issue of gold investment in times of uncertainty, including investors, institutional entities, practitioners, and students from various fields of study.

With reference to the research subject and the objectives adopted, the hypothesis is formulated that the Covid-19 pandemic and the armed conflict have contributed to the increase in gold prices, but have negatively affected the level of gold supply. The considerations undertaken in the study fall within the framework of economic sciences in the discipline of finance, with particular emphasis on the sub-disciplines of financial markets, international finance, and alternative investments. The issues addressed in the study refer to current problems of contemporary finance, also in the aspect of dilemmas related to the development of the science of finance.

2 LITERATURE REVIEW

As one of the most valuable metals, gold has played an important role in the development of monetary systems around the world. It is now one of the reserve assets of central banks, and investors are keen to use it as a form of investment. Demand for this precious metal is also expressed in the jewellery industry and other industries. On the other hand, the unstable situation in global financial markets in recent years is one of the factors behind the growing interest in alternative investments, including investments in the gold market.

There are two basic forms of investing in gold in the source literature: direct or indirect investment. Direct investment involves the physical purchase of gold. Surveys such as that conducted by Borowski (2016), Gierałtowska (2016), Walczak (2012), and Wang (2012) have shown that direct gold investment can take a variety of forms:

- investing in the spot market – the investor becomes the owner of bullion stored in a licensed storage house; the bullion is subject to specific standardisation rules, applying the principle that the gold supplier has the right to choose between the different standards; the gold delivered is characterised by the following elements: quantity, delivery date, location, physical composition and form (Borowski, 2016, p. 10);
- investing in collectible bars / investment bars – this requires a significant capital expenditure and consideration of storage, transportation, and insurance costs; the weight of gold bars is quoted in ounces, and popular weights are: 1 kg (32 ounces), 10 ounces or 400 ounces; these bars can be purchased from mints or specialist financial institutions; an element that affects the liquidity of physical bars is their manufacturer – the best-known manufacturers include Pamp Suisse, Perth Mint, and Royal Canadian Mint; an investor wanting to find out whether it is worth buying a particular bar should check whether the mint is on the Good Delivery list, which is prepared by the LBMA (London Bullion Market Association) – the most important accrediting body for precious metals producers (Gierałtowska, 2016, p. 129);
- investing in bullion coins that carry bullion value – coins are minted in large numbers and therefore their production costs slightly increase the price of the coins over the value of the bullion they contain; these coins most often do not have a denomination because their current valuation depends on the price of the bullion they contain; they have high liquidity due to the fact that they are widely known around the world and their authenticity can be more easily verified (e.g. through the use of a mint ruler) (Gierałtowska, 2016, p. 129);
- investing in collector coins which, apart from the appropriate gold content, also have a unique character, and sometimes even a historical origin, and whose prices may significantly exceed their bullion value;
- investing in gold medals mainly for the needs of governments and government organisations;
- investing by means of certificates for a specified amount of gold offered by financial institutions or private brokers – these are securities that evidence ownership of a specified amount of gold, allowing the sale of the metal represented by them without the need to physically transfer it; acquisition of the certificate means the purchase of investment bars, which are deposited in the safe of the offering institution or one selected by the intermediary; these securities have a specific validity, but can be resold ahead of schedule, and the prices are calculated on the basis of the current gold price, increased by the cost of production, storage and insurance costs, as well as the dealer margin (Walczak, 2012, p. 386).

Forms of direct investment can also include so-called gold accounts. Financial institutions (primarily banks/depositories) offer two types of gold accounts – allocated and unallocated. Allocated accounts are considered the safest form of investment in physical gold. The gold is stored in a vault owned and managed by a recognised bullion dealer or depositary, while the gold bars or coins are numbered and identified according to characteristics, weight, and sample – they are then allocated to the investor (who pays the depositary a custody and insurance commission in addition to the price of the gold). The investor in the allocated account has full ownership of that gold, and the depositary cannot trade, lease, or lend the gold – except under specific instructions from the account holder. The second type is allocated accounts, where investors do not have specific bars allocated to them. The advantage of these accounts is that there are no storage or insurance fees, as the bank reserves the right to lease the gold. Banks offering gold accounts do not generally handle transactions of less than 1,000 ounces, so their customers are institutional investors, private banks acting on behalf of their clients, central banks, and gold market participants wishing to buy or borrow large quantities of gold. Nowadays, financial institutions also offer so-called alternative accounts for smaller investors wishing to invest below 1,000 ounces. An example of such an account is Gold Pool account, where one can invest in as little as one ounce. Various electronic 'currencies' linked to gold in allocated storage are also available. These offer a simple and cost-effective way to buy and sell gold and use it as money. Any amount of gold can be purchased, and these currencies allow gold to be used to send online payments around the world (Wang, 2012, pp. 80–81).

Gold Accumulation Plans (GAPs) are an alternative to gold accounts; they are similar to traditional savings plans (so-called systematic savings) and based on the principle of putting aside a fixed amount of money at regular intervals, e.g., monthly. For a given amount in a given month, gold is purchased on the market. The amounts deposited may be small in value and there is no charge for purchases of small bars or coins. As small amounts of gold are purchased over a long period of time, exposure to short-term price fluctuations is limited. At any time during the term of the contract (usually a minimum of one year) or after the GAP closes, investors can receive the gold in the form of bars or coins and sometimes even jewellery, and in case they decide to sell their gold, they can also receive cash (Wang, 2012, pp. 81–82).

Much of the current literature on gold investment pays particular attention to indirect investment in gold. The most important indirect forms include futures, options, CFDs (Contract for Difference), investing in the shares of gold mining companies, structured products linked to the gold market, as well as units/certificates of investment funds operating on the gold market, for which a distinction is made between index funds and ETFs.

One form of indirect investment in the gold market is investment in forward contracts, where the underlying instrument is the price of gold. They are a kind of contract between the buyer and the seller, which represents a commitment by both parties to buy and sell gold at a forward price set on the date of conclusion of the contract. In addition, other terms of the transaction are included in the contract, concerning quantity, delivery, and settlement date: forward contracts with delivery (actual contracts), which oblige sellers to deliver the investment object on the expiry date, or cash-settled forward contracts (unfunded contracts), in which there is no physical delivery of gold, but only on the expiry date the parties to the transaction settle with each other through the difference between the contracted price and the current market price (Krężołek, 2020, pp. 38–40).

The second type of contract is futures contracts, which in theory are contracts for the delivery of a specified quantity and quality of gold for a fixed price at a specified date, while in practice no physical delivery takes place and the contracts are settled in cash based on profit or losses relative to the gold price on the market (Cai et al., 2008, p. 717). These contracts are traded on the stock exchange and their standards are regulated by the market. The seller and buyer cannot negotiate the terms of the contract as they are standardised, and the futures price is dependent on the gold price. In this type of contracts, one party undertakes to buy, i.e., to take a short position, and the other party, by taking a long position, undertakes to sell the asset specified in the contract. Taking a short position involves a commitment to sell the object of the transaction at a specified price and in a specified future to the person taking a long position. By taking a long position, the investor purchases the object of the transaction, which he or she undertakes to sell at a specified future. If the value of the underlying instrument increases, the investor, selling it after a specified time at a higher price than when they purchased it, makes a profit, and if the value of the purchased object decreases, the investor makes a loss on the transaction. Taking a short position, on the other hand, obliges the investor to repurchase the object of the transaction in the future, which makes it possible to make a profit if the price at the time of concluding the contract is higher than the future sale price (Rembisz, 2009, pp. 31–33). Participants in the futures market include investors who wish to profit from a transaction, as well as gold producers. The purpose of futures transactions for gold producers is to hedge against commodity price volatility. Using such instruments, producers operating in the commodity market improve their financial stability by neutralising negative price movements in the real market (Cai et al., 2008, p. 717).

A form of indirect investment in gold similar to futures contracts is gold-based contracts for differences (CFDs.) CFDs are based on a kind of bet between two parties on the future price of gold. This allows investors to make money on both a rise and a fall in the price of the commodity. These contracts are over-the-counter instruments, so they are not subject to as much regulation as futures contracts, which appear on the stock exchange. The way CFDs work and are structured is similar to futures, but CFDs have greater flexibility and transparency. CFDs are settled in real time, which means that the trader can monitor their profits and losses in real time (Borowski, 2016, p. 13).

Investing in gold can be based on investing in option contracts. Gold options give the holder the right, but not the obligation, to buy (call option) or sell (put option) a specified quantity of gold at a predetermined price on a fixed date. The price of a gold option depends on a number of key factors, for example, the spot price of gold, the strike price, the interest rate, the estimated volatility of the gold price and the time remaining until expiry of the option (Cai et al., 2008, p. 718).

An instrument similar to options are gold warrants, which give the buyer the right to buy gold at a specific price on a specific date in the future. The buyer pays a premium for this right. Like futures contracts, warrants are generally leveraged against the price of the underlying asset (in this case gold). In the past, gold warrants have mainly involved shares in gold mining companies (Wang, 2012, p. 80).

Indirect investing also includes investing in the shares of gold mining companies, but it should be emphasised that the share price is not compressed with the price of gold, but reflects fundamental factors (Gierałtowska, 2016, p. 129). Investing in gold mine shares is a logical substitute and complement to investing in other forms of physical gold. Mines' share prices are also influenced by the companies themselves, their projects, reserves of unmined gold underground or revenue streams from mining royalties. A number of factors can influence share valuation, such as the maturity and geographic scope of mining projects, gold reserves, ore grades, costs, margins, company profitability, profile, debt size, and management quality (Wang, 2012, p. 82).

Investing in structured products linked to the gold market is another form of indirect investment. Structured products are most often issued by banks and gold dealers and often come in the form of structured bonds with varying structures and pay-out profiles (gold-linked bonds). These investments are important because they provide some exposure to changes in the gold price, rates of return, and offer varying degrees of capital protection. They can be issued depending on whether the investor is optimistic about the gold price or disagrees with the gold price. Depending on this view, part of the investment will be placed in put or call options. The remainder of the investment is then placed in the money market to generate a return. These products may also provide capital protection depending on the product design and the investor's expectations and risk profile (Cai et al., 2008, p. 719). In some markets, an example of this type of instrument is commodity-linked notes, which are medium-term debt instruments whose value at maturity depends on the price of the underlying asset, possibly on the value of a basket of derivatives (Borowski, 2016, p. 15).

Investing in units/certificates of gold mutual funds is becoming a popular form of indirect investment. It should be pointed out that there are many forms of collective investment in gold, including mutual funds, open-ended investment companies, close-ended funds, and unit trusts; and the range of their investments can be quite substantial, some of which will only invest in the shares of mining companies, while in others there may also be exposure to the price of gold through the use of derivatives or direct investment in gold (Cai et al., 2008, p. 719). One type of investment fund is index funds, whose portfolios include gold mine shares, derivatives based on the gold price, as well as other assets. Gold Exchange Traded Funds (ETFs) based on gold (Gold Exchange Traded Fund – GETF) are becoming increasingly popular among investors. These are investment funds that aim to imitate the performance of specific stock market indices. Their participation titles are traded on a regulated or organised market, most often on an exchange (Czech and Puszer, 2021, p. 134). An ETF is a fund that invests in standard gold bullion as the underlying asset, and its certificates are listed on exchanges and can therefore be bought and sold, just like shares. A gold ETF is a fund based on the value of the underlying asset, including gold. These funds issue certificates representing physical gold and can be in paper or dematerialised form and are traded on exchanges like individual company shares. Gold ETFs give investors the opportunity to participate in gold bullion markets without the need for physical delivery, and to buy and sell this participation through trading in exchange-listed securities (Saleem and Khan, 2013). Among ETFs, a distinction is made between funds whose investment portfolios consist mainly of gold, held in storage houses – these are called commodity-related ETFs; and commodity-linked ETFs, whose portfolios are predominantly gold and derivatives, rather than shares in gold-producing companies. A further criterion is used to divide ETFs according to whether they purchase gold on the spot market (bullion-backed ETF) or whether their investment portfolio is based on futures contracts (future-backed ETF) (Borowski, 2016, p. 13).

Numerous studies have suggested that gold is one of the most popular forms of alternative investment in the commodities market. This is because the price of gold is treated as one of the most important indicators of the health of the economy. Volatility in the price of gold can signal economic problems, especially when the price of gold rises, as the international price of gold is based on the US dollar. When the gold price rises, a fall in the USD exchange rate takes place. Also, during a financial crisis, many investors invest in gold to protect their investments. Gold can also be used as a hedge against inflation. It maintains its value even in times of economic downturn. Due to its properties, gold can be considered an exceptional investment. Gold is a valuable bullion that can function as a unit of value, a source of wealth, and a highly liquid asset. It also has industrial applications, e.g., in jewellery or dentistry. Another factor that determines the attractiveness of gold investments is the fact that demand for gold is increasing, while resources are decreasing. The value of gold is also determined by its historical significance. Gold has invariably retained its value ever since its first use as a currency (Juras, 2021, p. 83).

Goodboy (2013) suggests that gold is the favourite precious metal of most investors and that it is the most durable investment in the world. It serves as a medium of exchange and a store of value. Besides, he points out that despite its differences and similarities as a commodity, investors prefer to own gold over, for example, silver, the reason for which is that silver is mainly used for industrial purposes, while gold is used in investing and industry (Goodboy, 2013).

On the other hand, Pullen et al., observed that a gold mining company stocks and mutual funds investing in gold, is forms of investment that are safe havens for investors. Consequently, investors who wish to secure a safe haven investment in gold should generally not rely on mining company stocks or mutual funds, but should instead take positions directly in gold or in bullion-based ETFs (Pullen et al., 2014, p. 76).

According to Feldman, gold is an accumulated physical asset taking the form of deferred consumption, the characteristics of which are one of the reasons investors use it in investment strategies. He points out that investors use gold as a 'safe haven' during political disturbances/conflicts, or in times of economic or financial uncertainty. Gold is an investment that provides diversification to an investment portfolio in both the short and long term, in addition to providing a hedge against inflation and currency depreciation, primarily of the dollar. He also believes that there is growing investor demand for gold in the global economy and during times of political uncertainty, so there are various options related to the use of gold within investment strategies, from physical gold to gold-based exchange-traded products. The price of gold is influenced by a number of factors, the most important being central bank reserves held in gold, the USD exchange rate, political uncertainty, economic concerns around the world, hedging transactions by gold producers, and the trading activities of speculators (Feldman, 2010, pp. 12–13).

Baur and Lucey also treat gold as a 'safe haven', understood as a resource that is uncorrelated with another asset or portfolio in times of market stress or turbulence, and as a hedge i.e., a resource that is uncorrelated or negatively correlated with another asset. Safe-haven assets are distinguished from hedging and diversification assets, which provide average diversification benefits, but not necessarily when they are most needed, i.e., in times of market turmoil. They believe that gold functions as a safe haven only for a limited period and that investors buy gold on days of extremely negative returns and sell it when market participants regain confidence and volatility is lower (Baur and Lucey, 2010, p. 228).

Similarly, Baur and McDremott demonstrated that gold is an asset that is a safe haven, offering protection to investors from financial losses in financial markets. Their research suggests that investors respond to short-term and extreme shocks by seeking a safe haven in gold, and they pointed to different responses to shocks in developed and emerging markets. Gold is, at best, a weak safe haven for some emerging markets as investors suffering losses in emerging market equities, rather than looking for an alternative asset, may simply adjust their portfolios, pulling out of emerging markets in favour of developed markets. Gold, on the other hand, is a strong safe haven for most developed markets. In addition, the authors believe that gold has the potential to act as a stabilising force for the global financial system by limiting losses when it is needed most (Baur and McDremot, 2010, p. 1897).

Investment interest in gold is driven by a number of factors, including the volatility of the dollar, inflationary expectations, continued high levels of geopolitical uncertainty, and the increased acceptance of gold's role as an investment portfolio diversifier. Trends currently emerging in the gold market may suggest that the role of gold is changing, with increasing demand for gold in jewellery and industry, and less demand for gold as a reserve asset. However, the important role is that of gold as an investment asset. Whether for physical investment in coins and bars or for hedging or diversification, gold as an asset is viewed very positively. Increased demand from the industry and jewellery sector, and the need to hedge and offset adverse movements in gold prices, means that demand for gold derivative products significantly affects the price of gold (Cai et al., 2008, p. 735).

Kovinska sees gold as a safe investment that provides diversification benefits to an investment portfolio. According to her research, gold is an alternative to a weak US dollar, as investors tend to invest more in gold when the US dollar loses value, leading to higher precious metal prices. However, when the US dollar strengthens, investors trust paper money more than gold. Furthermore, gold cannot provide a stable hedge against inflation, but it is believed that investing in gold will provide investors with the ability to protect their wealth against possible future inflation. Investing in gold ETFs, on the other hand, provides a weak hedge against inflation. The author argues that gold ETFs are an option for investors who seek the benefits of portfolio diversification, despite being a relatively young investment vehicle on the financial market (Kovinska, 2014, p. 30).

Šoja's research shows that gold is an important instrument for investment portfolio diversification. He indicates that the recommended share of this metal in a portfolio should be between 1% and 9%. He thinks that it is reasonable to include gold in investment portfolios if the portfolio is combined with European bonds and equities. Gold is a good basis for portfolio diversification, both from the point of view of a risk-averse investor and an investor prepared to take more risk (Šoja, 2019, p. 51).

Pule highlights that gold has a low correlation with other assets (i.e., equities), this confirms the use of gold as an instrument for diversifying an investment portfolio. Gold is a useful asset in times of stock market weakness, but equities outperform gold over the long term when the market recovers. Investing in gold has become popular due to the development of gold ETFs, which are readily available (Pule, 2013, p. 56).

According to Waghmare and Shukla, investment in gold in its various forms (e.g. mining company shares, gold coins, or ETFs) is responsible for meeting two investment needs: functional and protection. They indicate that investors consider gold as a leading alternative investment in their portfolios in order to achieve high returns over the long term, provide stable wealth, mitigate risk during market turmoil, increase liquidity, and improve portfolio quality. Gold is also treated by investors as a buffer against financial risk, adverse currency market movements, and inflation. Gold consistently increases investment risk-adjusted returns, provides stable growth, and generates liquidity during periods of market turmoil to cover liabilities. Gold has long been considered a valuable commodity to return to in both good and bad times. The long-term return on gold is competitive with returns on equities or government bonds (Waghmare and Shukla, 2021, p. 115).

In times of heightened uncertainty and economic downturn, gold becomes a safe instrument and an indispensable insurance policy, and the demand for it therefore increases. At present, there is a strong economic and social disturbance caused by the Covid-19 pandemic and the Ukrainian–Russian war, where, under conditions of strong restrictions, gold is gaining in value and is an asset desired by individual and institutional investors. According to Wagmare and Shukla, gold is a good asset class to invest in during the Covid-19 pandemic, and there is no significant gender difference in investment preferences. Investors are influenced by behavioural biases that affect their preference to invest in gold during Covid-19; they are more likely to buy gold ETFs than during a normal economic situation (Waghmare and Shukla, 2021, p. 119).

Akhtaruzzaman et al. in their study of the use of gold as a safe haven demonstrated that gold fared relatively better than many other assets during the pandemic, and even though its price fell in March, the decline was smoothed by the fall in stock prices and other commodities such as oil. This demonstrates gold's role as a safe haven asset and portfolio diversifier in the very early stages of shocks that could affect the global economy and financial markets. Secondly, the rapid and radical response of the Federal Reserve (and later other central banks) provided ample liquidity to financial markets, which in turn quickly restored investor confidence and increased appetite for risky assets, in addition to the demand for alternative investments. Consequently, correlations between gold and other financial assets increased, leading to higher hedging costs. Therefore, investors should be wary of the potential for policymakers to take decisive action and implement effective measures during economic downturns, and closely monitor markets so as to adjust their portfolios according to the prevailing market conditions (Akhtaruzzaman et al., 2021).

Juras emphasises that gold as an alternative investment is a good hedge in times of crisis and therefore also in times of pandemic. When financial markets experience declines in the price of financial instruments, gold behaves inversely – its price has been steadily increasing since 2000, and therefore investors looking for an alternative to increase capital during a crisis can definitely choose gold (Juras, 2021, p. 87).

Salisu et al. argue that during a pandemic the gold market is a safe investment alternative compared to other financial assets. Gold is considered the most useful hedging financial asset that protects investors' portfolios. In other words, investors are better able to protect their investments by diversifying their portfolios to include the acquisition of gold (Salisu et al., 2021).

According to Berry and Syal, any type of crisis, whether financial or health-related, such as Covid-19, has a negative impact on the economic, professional, and personal lives of individuals. Economic growth declines and people lose their jobs, resulting in a drop in disposable income and a lack of surplus. As incomes fall, people stop investing cash and use their income to meet current needs. During a crisis, investors are risk-averse and prefer to shift their investments from risky e.g., equities, to safe instruments such as gold. With the onset of a financial crisis and pandemic, share prices start to fall and the value of gold starts to rise. If the crisis is health-related, people start to focus on investing their money in health-related instruments and products, so there is a very low correlation between the returns generated by gold and stock indices. During a crisis, there is a difference in the behaviour of gold and indices, and in order to maximise returns investors start to sell equities, and with an increase in risk on the stock market they start to buy gold (Berry and Syal, 2022, p. 211).

At the time of the pandemic, in countries where savings are based on Shariah (India, among others), investors put their money in an investment instrument in the form of gold. This is because when the JCI (Composite Stock Price Index) and the rupee's exchange rate against the US dollar falls, the price of gold rises. Then, activities that have been carefully designed and analysed beforehand involve investors putting their money in gold-based savings, because essentially Shariah-based products have the feature of no interest in any form, as it is usury. Consequently, customers pledge gold in Shariah pawn shops because the process of pawning gold is easy, quick, and safe. In Shariah-based countries, investing in gold by buying, storing, and then selling it through pawn services, i.e., buying, saving, and then pawning, is a gold investment strategy and also a way to maintain gold investment during the Covid-19 pandemic (Dhuha et al., 2021, p. 82).

The worldwide crisis triggered by the outbreak of the SARS-CoV-2 coronavirus pandemic contributed to an increased interest in gold investments. Following the outbreak of the pandemic, investors were interested in investing directly in gold as well as investing indirectly. The research showed that as time went on and the Covid-19 pandemic unfolded, physical gold gained relative to ETFs and was a more desirable asset relative to ETFs. In doing so, it should be noted that despite the decline in demand for ETFs, their holdings of gold increased (Czech and Puszer, 2021).

Madhavan and Sreejith studied the relationship between gold, alternative gold assets, and international capital during Covid-19. They also identified linkages between gold and gold-backed investments to assess the phenomenon of flight to quality, contagion, and disconnect in the gold market during Covid-19. They found that gold was the safest haven during Covid-19. Assets with returns focused solely on gold, such as gold ETFs, the gold volatility index, and gold-backed cryptocurrencies, were safe havens only during higher market extremes. The most dependent gold markets were gold futures, the gold volatility index, and the gold market, which provides less diversification benefits to the portfolio. Gold futures, the gold volatility index, and mining stocks were close but not perfect substitutes for gold bullion in a portfolio. At the same time, gold ETFs and gold-backed cryptocurrencies became a complementary product in a portfolio along with direct gold during Covid-19. They also point out that gold ETFs and gold bullion provided diversification benefits in both normal and extreme market conditions during Covid-19. Investors could add gold direct, gold volatility index, or gold futures to their portfolio to gain diversification benefits during Covid-19, as these are close substitutes for gold bullion, while the choice of these assets depends on market conditions. In the event of a severe stock market decline, adding the gold volatility index or gold bullion to a portfolio is better; otherwise, gold futures are preferred (Madhavan and Sreejith, 2022, p. 431).

3 PROPOSED WORK

The implementation of the study objective was based on a critical analysis of the source literature and on quantitative research. The literature analysis made it possible to identify the role of gold in a period of turbulence and turmoil in financial markets. The quantitative research used dynamic analysis and multiple linear regression analysis. The dynamics analysis examined the direction and rate of pandemic development and the dynamics of change in the gold market. In addition, within the framework of multiple linear regression analysis, a model was built on the basis of which the correlation between the gold price and selected indicators determining the development of the Covid-19 pandemic was investigated. In the estimated model, the gold price expressed in ounces (USD/Oz) was taken as the dependent variable, while the following variables were taken as independent variables (predictor variables):

Total Covid-19 cases (b_1)
New Covid-19 cases (b_2)
Total Covid-19 deaths (b_3)
New Covid-19 deaths (b_4)
Reproduction rate (R number) (b_5)
Total Covid-19 vaccinations (b_6)
People vaccinated with at least one dose against Covid-19 (b_7)
People fully vaccinated against Covid-19 (b_8)
New Covid-19 vaccinations (b_9)

The estimated linear multivariate regression model was described by the equation:

$$\hat{Y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7 + b_8x_8 + b_9x_9 \pm \zeta$$

Where: b_i – partial regression factors, model parameters representing independent variables affecting gold prices.

Multiple linear regression analysis was conducted on a sample of 697 observations of gold price changes (dependent variable) and 6,273 observations of the independent variables (697 observations each of the selected indicators characterising the global pandemic situation).

In the multiple linear regression analysis, the stepwise progressive regression method was used. In building the model, all the indicators listed above were considered, but only those that were statistically significant were entered into the model. Statistical significance was assessed using a t-test, assuming a maximum 5% probability of error in inference. Thus, those variables whose value was higher than the critical value resulting from the Student's t-distribution at an alpha level of <0.05 were considered statistically significant. Then, once all statistically significant variables were included in the model, linear significance was tested for the entire constructed model using the F-test statistic.

The use of a multivariate linear regression model allowed us to answer the question of whether gold price levels changed during the development of the Covid-19 pandemic.

The research covered the period between January 23, 2020, and September 30, 2022. This period was dependent on the availability of statistical data, as the first global data on the Covid-19 pandemic was published on January 23, 2020. The upper limit of the period, on the other hand, was set at the end of the study and includes the most recent statistical data. The statistical data was taken from the Stooq.com database, Our World in Data, Statista.com, and the World Gold Council.

4 RESULTS AND DISCUSSION

4.1 Development of the Covid-19 pandemic

The Covid-19 pandemic was caused by a new, hitherto unknown virus, SarsCov2. Covid-19 is an infectious disease spread by water droplets. Its most common symptoms include loss of smell and taste, dry cough, shortness of breath, fatigue, and fever (Kucharska, 2022). The first confirmed case of Covid-19 was reported on November 17, 2019 in China, in the city of Wuhan (Ma, 2020). Since then, the disease has spread very rapidly worldwide and the number of Covid-19 cases has increased significantly. On January 30, 2020, the World Health Organisation (WHO) declared Covid-19 disease a public health emergency of international concern (Kucharska, 2022), and on March 11, 2020, WHO declared the disease as a pandemic, which has continued until today.

The number of Covid-19 cases and the number of deaths due to the disease were used as a measure of the progression of the pandemic. Figures 4.1 and 4.2 show that the course of the pandemic was uneven throughout the study period. In the initial period, both the number of cases and the number of deaths were relatively constant, but as time went on, these figures changed significantly (Figures 4.1 and 4.2).

Both the number of new cases of Covid-19 as well as the number of deaths due to Covid-19 (Figures 4.1 and 4.2 – left axes) varied over time, confirming that the

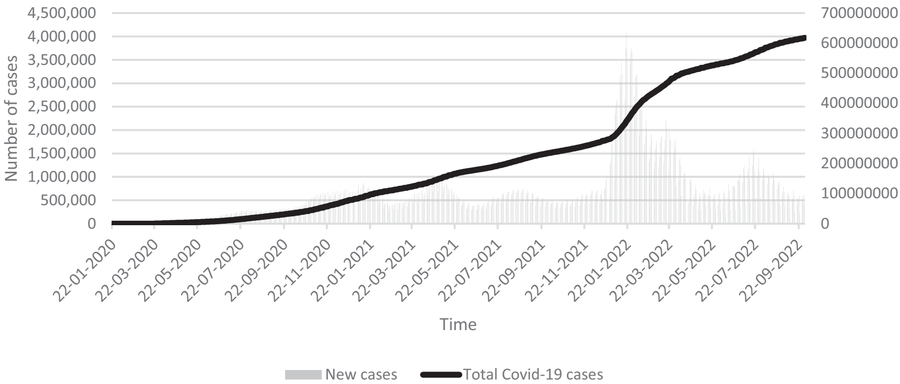


FIGURE 4.1 Covid-19 cases worldwide (new cases – left axis, total cases – right axis).

Source: own calculations based on: Our World in Data.

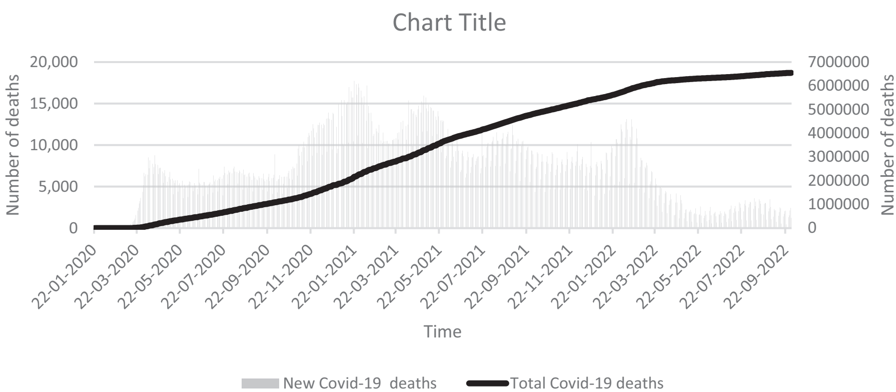


FIGURE 4.2 Covid-19 deaths worldwide (new deaths – left axis, total deaths – right axis).

Source: own calculations based on: Our World in Data.

pandemic occurred in a wave-like manner. The results of the study indicate that five waves of the pandemic can be distinguished throughout the study period (Table 4.1), wherein, with each wave, the number of Covid-19 infections increased relative to the previous waves.

Figure 4.1 clearly shows that the lowest wave formed during the first stage of the pandemic, i.e. from the announcement of the epidemic (later pandemic) until July 2020. Due to the lack of sufficient knowledge regarding the new pathogen, this period was characterised by a high level of uncertainty and information chaos. After this period, the pandemic slowed down somewhat and remained relatively stable until October 2020. Due to the change of seasons, the SarsCov2

TABLE 4.1
Health crisis waves

Health crisis wave	Duration of the health crisis wave (contractual period)
1	January 23, 2020–June 30, 2020
2	October 1, 2020–February 15, 2021
3	March 15, 2021–May 31, 2021
4	July 15, 2021–September 30, 2021
5	December 15, 2021–April 18, 2022
Whole research period	January 23, 2020–September 30, 2022

Source: own calculations based on: Our World in Data.

virus became active again in October 2020. This is when the second wave of the pandemic began – much higher in terms of the number of cases and deaths. This wave lasted until February 2021, after which the number of infections returned to relative stability for a very short period (just one month). The third wave of the pandemic lasted about three months – until the end of May 2021. At this time, the fight against Covid-19 was not complete as an effective drug to prevent the disease had still not been developed, and vaccination, due to difficulties in the production and availability of the vaccine, was not so widespread as to protect the entire population. Therefore, in the second half of July 2021, new outbreaks of Covid-19 and an increase in new infections were reported. This triggered a fourth wave of the pandemic, lasting until the end of September 2021. The SarsCov2 virus has not been constant since the beginning of its activation but has constantly mutated. Since the beginning of the pandemic, new strains of SarsCov2 virus were being discovered, which were no more dangerous than the original pathogen. However, on November 24, 2021, the emergence of a new and very dangerous mutation, which was given the name Omicron, was reported to the WHO. Omicron was responsible for causing the fifth and highest ever wave of Covid-19 cases, lasting until April 2022. Currently, the number of cases and the number of deaths have stabilised, which can be explained by the acquisition of natural immunity in the population, as well as immunity due to the availability of vaccinations.

As previously mentioned, due to the lack of sufficient knowledge on how the Sars-CoV-2 virus spreads and how to prevent Covid-19 infection, the disease has caused significant difficulties for individuals, businesses, and countries. During the period of highest transmission of the virus, a number of restrictions were put in place around the world to prevent new infections (including restrictions on mobility, working and studying remotely, maintaining social distance, significantly limiting tourism, covering mouths and noses, etc.). Despite the measures taken worldwide, the number of cases of Covid-19 and the number of deaths from Covid-19 have steadily increased throughout the study period, i.e. from January 23, 2020 to September 30, 2022. As of September 30, 2022, 617,583,875 people have contracted Covid-19 and 6,545,452 people have died (Figures 4.1 and 4.2 – right axes).

4.2 The gold market during the Covid-19 pandemic

For centuries, gold has been a safe form of capital investment. Referring to historical data, it is noted that during times of heightened uncertainty and socio-economic turmoil, the demand for gold increases. This is due to the fact that gold has come to be regarded as a safe haven, which, due to the fact that it is not positively correlated with other assets, provides protection to investors against financial losses (Baur and Lucey, 2010, p. 228; Baur and McDremot, 2010, p. 1897, et al.). However, the current socio-economic turmoil has a very different cause, one that is not economic, but health-based.

Prior to the outbreak of the Covid-19 pandemic, the demand for gold, as well as its supply, remained on an upward trend, but the outbreak of the pandemic unbalanced this market (Figure 4.3).

As Figure 4.3 shows, in consecutive quarters preceding the outbreak of the Covid-19 pandemic, both the demand and supply of gold, as well as its price, increased steadily. In contrast, immediately after the pandemic outbreak – from January 2020 onwards (the first wave of the pandemic) – both the demand for and supply of gold decreased. In the first wave of the pandemic, central bank demand for gold was the first to decline, as well as jewellery production demand (Figure 4.4). At the same time, despite falling demand, the gold price remained on an upward trend. The likely reason for the increase in the price of gold with the concomitant decrease in demand in the first wave of the pandemic was, on the one hand, the reduction in the supply of gold due to insufficient labour and, on the other hand, the break in supply chains caused by the numerous lockdowns implemented in all the countries of the world. The reduction in gold production compounded by the rupture of supply chains consequently led to shortages of gold on the market in the first wave of the pandemic. In contrast, in the subsequent waves of the pandemic, both gold demand and supply steadily increased from quarter to quarter and the price of gold stabilised.

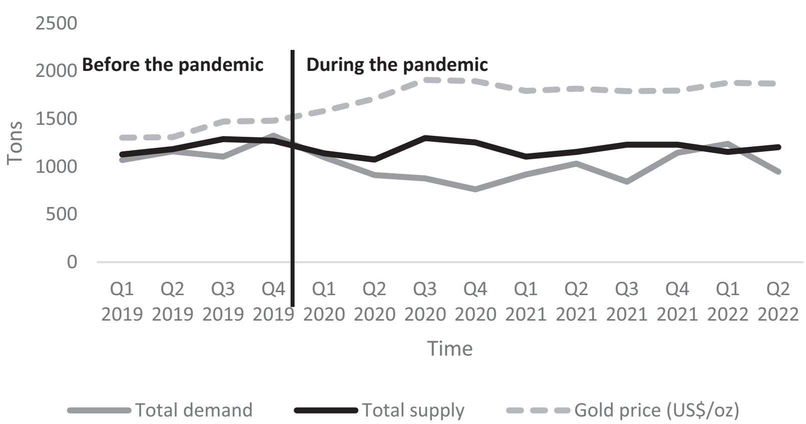


FIGURE 4.3 Gold price, demand, and supply before and during the Covid-19 pandemic.

Source: own work based on World Gold Council.

Ultimately, over the entire Covid-19 pandemic, gold demand increased by an average of 16 tonnes from quarter to quarter, while gold supply increased by an average of 4.5 tonnes from quarter to quarter.

Throughout the pandemic, jewellery manufacturers reported the highest demand for gold (47.34%), followed by investors (36.16%). In contrast, the smallest gold purchases throughout the pandemic were made by the technology industry (8.10%) and also central banks (9.18%) (Figure 4.4).

Figure 4.4 shows that in the first wave of the pandemic, the largest purchases of gold were made by investors, but as the pandemic progressed their demand for gold dropped significantly (on average around 24 tonnes from quarter to quarter). At the same time, there was a dynamic increase in demand for gold for jewellery production (averaging around 34 tonnes from quarter to quarter). This confirms that during the development of the pandemic, gold buyers showed a greater propensity to purchase physical gold relative to other forms of investment in the gold market. As mentioned earlier, investing in gold can take two forms: direct and indirect. As historical data shows, in times of socio-economic turbulence and turmoil, investors move away from indirect investments to physical gold. The results of the analysis confirm that physical gold is an attractive form of capital investment also during non-economic crises. At the same time, the results of the study indicate that during the health crisis triggered by the outbreak of the Covid-19 pandemic, physical gold acted as a safe haven instrument (Figure 4.5).

Figure 4.5 confirms that demand for physical gold, particularly gold bars, increased significantly during the Covid-19 pandemic. Demand for gold bars in the second quarter of 2022 increased by just under 72% (66 tonnes) compared to the first quarter of 2020. During the same period, a similar situation existed in the gold coin market, where a clear increase in demand is noted in the subsequent quarters of the pandemic. The research shows that the demand for official coins at the end of

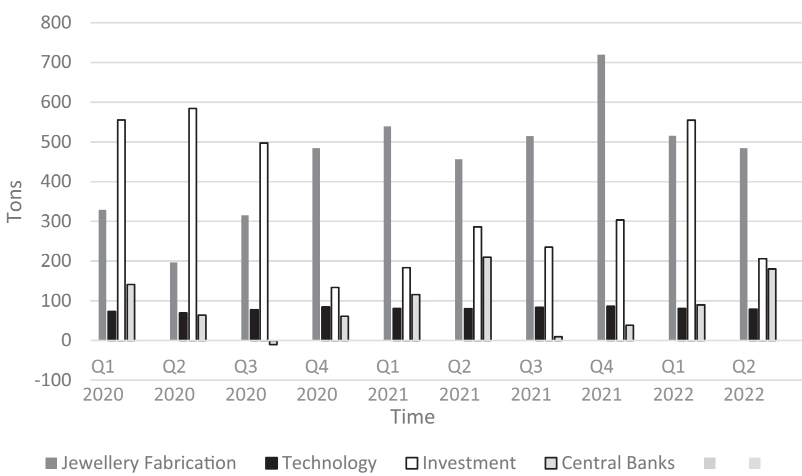


FIGURE 4.4 Gold demand during the Covid-19 pandemic.

Source: own work based on World Gold Council and stooq.com.

the period studied (Q2 2022) increased by 59% compared to the initial period, while the demand for precious metal coins during the same period was 25%. On the other hand, as the research shows, the significant reduction in demand for gold in gold-based ETFs is mainly responsible for the decline in investment demand during the Covid-19 pandemic. The highest ETF demand for gold was observed in the second quarter of 2020 (just under 436 tonnes). Thereafter, this demand declined dynamically from quarter to quarter, and from the third quarter of 2020 onwards, negative gold demand was reported for these funds (Figure 4.5). A temporary increase in demand for gold in ETFs was only reported in the first quarter of 2022, which was related to the outbreak of the Ukrainian–Russian war.

From the second quarter of 2020 to the fourth quarter of 2020, the gold market experienced a divergence between gold demand and supply. During this period, there was a significant decrease in demand for gold (−16.45%), while its supply increased (16.87%). However, as the pandemic progressed, this disparity narrowed. As Figure 4.3 shows, from the first quarter of 2021 onwards, gold demand increased, while gold supply was characterised by a slight increase. Throughout the Covid-19 pandemic, as previously stated, gold supply remained on an upward trend, increasing by an average of 4.5 tonnes from quarter to quarter (Figure 4.6).

Figure 4.7 shows that gold is supplied to the market through mine production and through recycled gold. Comparing the two segments of gold supply, it is concluded that mine production was responsible for the increase in gold supply throughout the Covid-19 pandemic. As can be seen from the figure above, mine production followed a wave-like pattern, wherein it should be noted that the lowest gold production occurred in the second quarter of 2020 (788.9 tonnes), as well as in the first quarter of 2021 (834.9 tonnes) and in the first quarter of 2022 (852.6 tonnes). This coincides with the period of greatest incidence of Covid-19 and

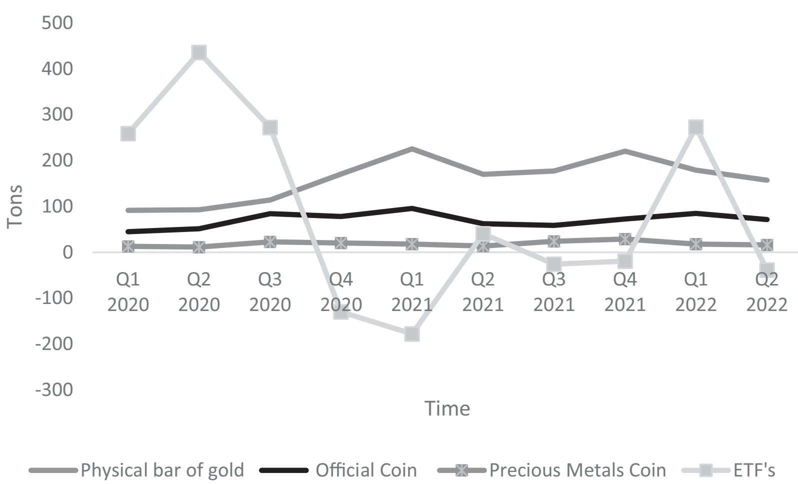


FIGURE 4.5 Gold investment demand during the Covid-19 pandemic.

Source: own work based on World Gold Council.

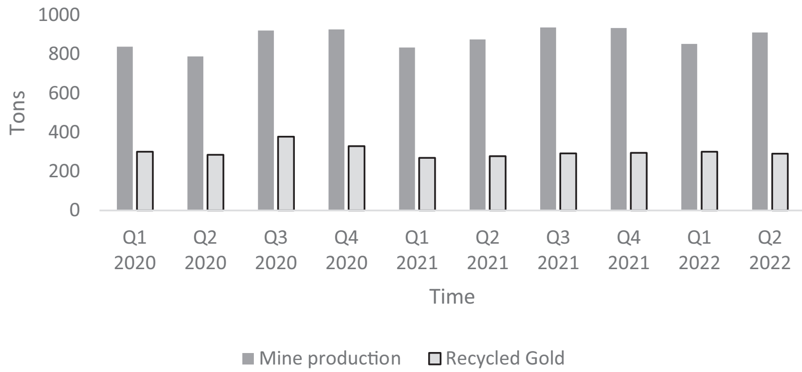


FIGURE 4.6 Gold supply during the Covid-19 pandemic.

Source: own work based on World Gold Council.

indicates that gold mining was dependent on the development of the pandemic. On average, gold production increased by 7.5 tonnes from quarter to quarter throughout the study period, and at the end of the study period (Q2 2022), gold production was 8.72% higher than at the start of the study period (Q1 2020). The average share of mine production in gold supply throughout the study period was 74.49%. Recycled gold, on the other hand, followed a downward trend throughout the study period (–2.98 tonnes from quarter to quarter). With physical gold acting as a safe haven instrument in times of turbulence, gold holders were reluctant to sell their gold holdings and, as a result, the supply of recycled gold decreased. As with mine production, the smallest supply of recycled gold occurred in the second quarter of 2020 and the first quarter of 2021, which also indicates that the supply of the gold price increased dynamically. This was due to the high level of uncertainty regarding the development of the pandemic and its consequences for the economy, as well as the difficult access to the gold market due to the restrictions put in place (restriction of mobility, disrupted supply chains, etc.). In contrast, as the pandemic unfolded, knowledge of the new pathogen increased, and the gold price stabilised from the fourth quarter onwards. Until the end of the study period, the gold price fluctuated slightly. Towards the end of the study period (in the first and second quarters of 2022), the gold price increased, which should be explained by the emergence of a new key risk factor, namely the escalation of the Russia–Ukraine armed conflict.

4.3 The impact of the Covid-19 pandemic on gold prices

In the second part of the quantitative research, multiple linear regression analysis was used. Based on this, a model was developed to investigate the impact of the Covid-19 pandemic on gold prices. Nine independent variables were entered into the model sequentially. The analysis showed that of these, only five were statistically significant (Table 4.2).

TABLE 4.2
Results of multiple linear regression analysis

Dependent variable	Parameters of the independent variables								
	Total Covid-19 cases	**New Covid-19 cases**	**Total Covid-19 deaths**	**New Covid-19 deaths**	**Reproduction rate (R number)**	**Total Covid-19 vaccinations**	**People vaccinated with at least 1 dose against Covid-19**	**People fully vaccinated against Covid-19**	**New Covid-19 vaccinations**
	(bl)	**(b2)**	**(b3)**	**(b4)**	**(b5)**	**(b6)**	**(b7)**	**(b8)**	**(b9)**
Gold price (ozt/USD)		0,000			−292,699	0,000	0,000	0,000	

Source: own calculations.

Based on the results obtained in Table 4.2, the following five variables were found to have an impact on gold ounce prices expressed in USD during the Covid-19 pandemic:

- Reproduction rate (R number)
- People fully vaccinated against Covid-19
- Total Covid-19 cases
- People vaccinated with at least one dose against Covid-19
- Total Covid-19 vaccinations

The results are homogeneous, with the exception of the variable Reproduction rate (R number). The results show a strong negative correlation between this indicator and the gold price. Each 1% change in the R-number rate resulted in a change in the price of gold of just under US$ 293 per troy ounce, with the price of gold falling by US$ 293 as the number of people to whom each Covid-19 patient transmits infection increased. Conversely, as the number of people to whom each patient transmitted Covid-19 infection decreased, the gold price increased by US$ 293. During the period of increased transmission of the SarsCov2 virus, there was an increase in restrictions to contain the spread of the virus (restriction of mobility, implementation of personal protection procedures, lockdowns, self-isolation, etc.), which consequently translated into significant supply chain disruptions in all markets, including the global gold market. Because of this, the supply of gold was reduced and, in turn, the demand for gold was not met. In turn, the reduction in restrictions resulted in the release of labour and, consequently, an increase in gold supply and sales.

At the same time, research results indicate that, in addition to the Reproduction rate variable (R number), the price of gold is also influenced by four other variables, but the strength of this correlation is negligible. The research has therefore shown that there is no direct interrelation between the health situation and the gold price, which is also confirmed by the model fit parameters (Table 4.3).

Table 4.3 shows that the estimated multiple linear regression model is statistically significant. This is confirmed by the values of the F statistic (45.90) and

TABLE 4.3
Results of the evaluation of multiple linear regression model parameter significance

Model adjustment parameters				
Determination coefficient R^2	Estimation error	F-test statistic value	Test probability level ($p < 0.05$)	T-test statistic value (intercept term)
0.3275	56.55	45.90	0.0000	2,287

Source: own calculations.

$p < 0.05$. This is because the value of the F statistic shows the significance of the model components: if $F = 0$, the model components are insignificant, and if $F \neq 0$, the components are significant. The significance of the model is also confirmed by the estimation error and the value of the t-statistic (free expression). The estimation error indicates the average difference between the observed values of the dependent variable and the theoretical values (Rabiej, 2012). In the estimated model, the free expression is different from zero, while the standard error of the estimation of the free expression relative to its value is 2,287. The coefficient of determination R2 indicates the degree to which the constructed model explains changes in price, which means that the model explains 32.75% of changes in the price of gold. This means that more than 67% of the gold price depends on other factors.

4.4 Russia–Ukraine armed conflict and the gold market

The escalation of the Russian–Ukrainian conflict, which has been ongoing since 2014, began on February 24, 2022. Since then, by September 29, 2022, more than 5,990 civilians, including 382 children, have been killed and 8,848 civilians, including 676 children, have been injured in Ukraine (Statista, 2022). With the start of military operations, thousands of Ukrainian citizens (mostly women and children) left their country seeking safety in Central and Eastern European (CEE) countries, mainly Poland (Figure 4.8).

Figure 4.7 illustrates the number of border crossings between Ukraine and other CEE countries after Russia's invasion of Ukraine. As can be seen from the figure, the largest number of Ukrainian citizens (over 6.1 million) crossed into Poland, followed by Russia (over 2.5 million) and Hungary (over 1.3 million). The lowest number of Ukrainian citizens emigrated to Moldova (over 606 thousand) and Belarus (over 16 thousand). In total, more than 12.6 million citizens left Ukraine in search of refuge and security, as the Ukrainian forces were a poor match for the Russian forces (Table 4.4).

As can be seen from the table above, Russia has much greater military potential than Ukraine. For this reason, other countries have directed additional military assistance to Ukraine, including:

- Weapons (United States, United Kingdom, France, Netherlands, Germany, Canada, Sweden, Norway, Denmark, Finland, Belgium, Czechia, Croatia)

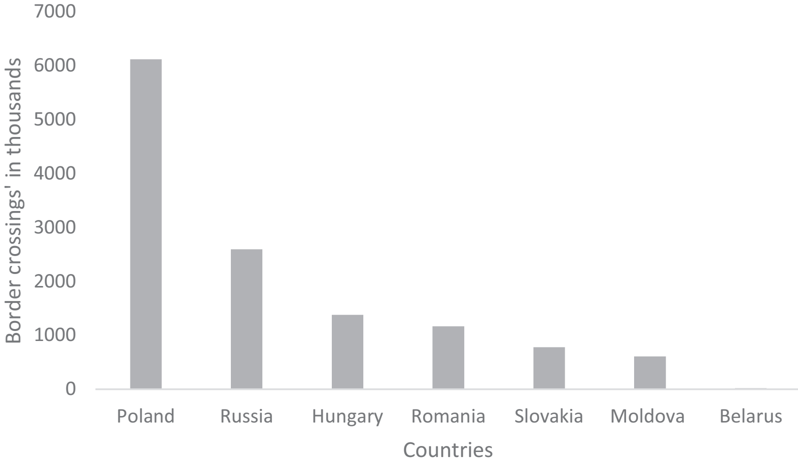


FIGURE 4.7 Number of border crossings from Ukraine to CEE, from February 24 to September 13, 2022.

Source: www.statista.com.

TABLE 4.4
Comparison of the military capabilities of Russia and Ukraine as of 2022

Indicator	Russia	Ukraine
Total military personnel	1,350,000	500,000
Total aircraft and helicopters	5,716	430
Ground combat vehicles	60,078	18,460
Total military ships	605	38
Nuclear warheads	6,255	0

Source: www.statista.com.

- Medical and defensive equipment (Spain)
- Soldier's equipment (Romania, Portugal, Greece)

The outbreak of the war in Ukraine has become a new risk factor reflected in the economy. As a result of the military conflict, the supply of fossil fuels, of which Russia is one of the key producers, was disrupted. The gold market was not indifferent to the new threat either, especially as Russia is among the world's leading gold distributors. Alongside Russia, the world's largest gold producers are China and Australia (Figure 4.8).

Research has shown that China, Russia, and Australia are responsible for less than 30% of global gold production. These countries are followed by Canada and United

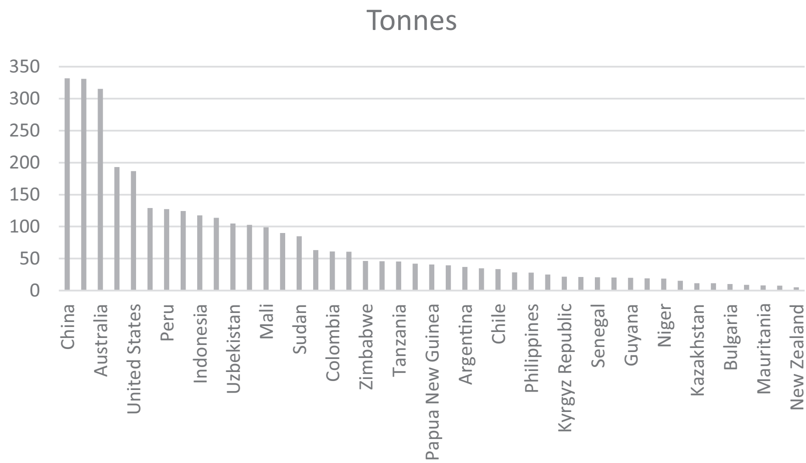


FIGURE 4.8 Global gold mine production. Data as of December 31, 2021.

Source: World Gold Council.

States, which account for 11.5% of global gold production. Russia, which is among the top gold producers, accounts for 10.02% of gold production. Due to the military conflict between Russia and Ukraine and the restrictions imposed on the Russian Federation in this regard, the supply of gold from this side is decreasing. On the other hand, the increase in military spending in Russia may increase the supply of gold from Russia, which will undoubtedly affect the gold price.

From the outbreak of the armed conflict between Russia and Ukraine in 2022 until the end of the research period, the gold price was in a downward trend. Only the timing of the outbreak of the conflict (February 24, 2022) contributed to a short-term increase in the gold price. During this time, the gold price rose by 7.66% over a 12-day period (until March 8, 2022), after which the trend was reversed and the gold price regularly fell by USD 1.43 per ounce. The average rate of decline in the gold price over the entire study period was −0.09% per day (Figure 4.9).

The fall in the price of gold during the military crisis has been driven by declining demand for bullion. Following Russia's invasion of Ukraine, demand for gold fell markedly. In the first quarter of 2022, gold demand was more than 23.5% higher than in the second quarter of 2022. The research shows that as the military conflict unfolded, gold demand decreased by just under 292 tonnes (Figure 4.10). This is a rather surprising finding given that gold has historically been a safe haven during military conflicts.

The research showed that the largest demand for gold during the military crisis was from China, which generated just under 45% of global gold bullion demand (Figure 4.11).

As shown in Figure 4.11, the United States (12.55% of world demand) and Germany (10.55% of world demand) also reported high gold demand during the military crisis.

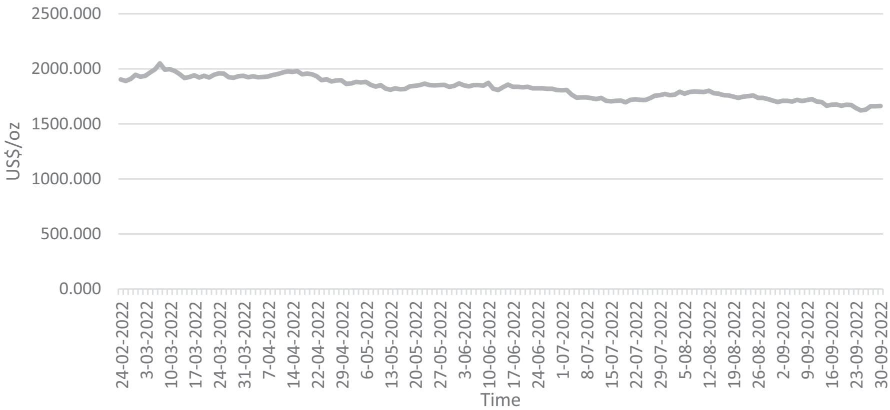


FIGURE 4.9 Gold price after the Russian invasion of Ukraine.

Source: own work based on stooq.com.

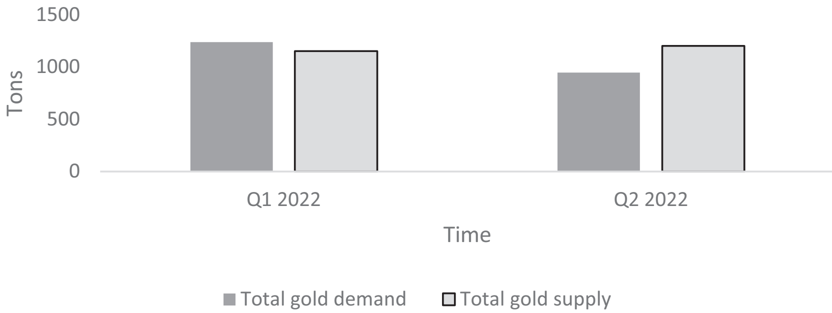


FIGURE 4.10 Gold demand and supply after the Russian invasion of Ukraine.

Source: own work based on World Gold Council.

The research showed that Russia's gold demand during the period under review was moderate at over 46 tonnes, or just 2.57% of world demand. Given that Russia is one of the world's leading gold producers, it is assumed that Russia's gold demand is met by its own production of gold bullion.

The fall in the price of gold during the military crisis, in addition to the decrease in demand for the bullion, is also influenced by an increase in the supply of gold. According to the research, since the Russian invasion of Ukraine, gold supply has increased by 4.26%, from 1,153.7 tonnes in the first quarter of 2022 to 1,202.8 tonnes in the second quarter of 2022. An increase in the supply of gold with a concomitant decrease in the demand for gold bullion has a negative impact on the price of gold bullion, as demonstrated in this research.

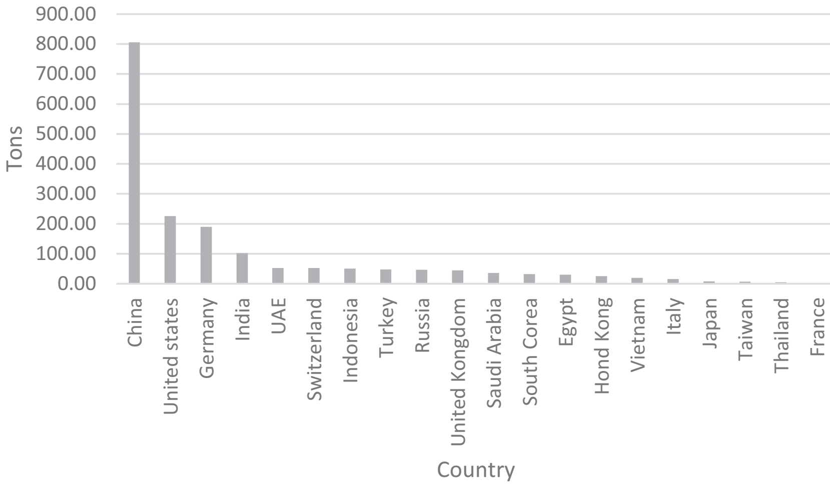


FIGURE 4.11 Gold demand by country after the Russian invasion of Ukraine.

Source: own work based on www.usdebtclock.org.

4 CONCLUSIONS AND FUTURE WORK

Since 2020, the world has been facing a health crisis caused by the new SarsCov2 virus, which has recently (as of February 24, 2022) been compounded by a military crisis. Both the health crisis and the military crisis have caused tremendous socio-economic turmoil and contributed to changes in all markets including commodities, especially gold.

The object of the research in this study was to investigate and assess changes in gold demand and supply and fluctuations in global gold prices as a result of the impact of Covid-19 and the Russian–Ukrainian armed conflict. On the basis of the research, it is concluded that the global crisis triggered by the outbreak of the SARS-CoV-2 coronavirus pandemic contributed to an increased interest in gold investments. Initially, after the outbreak of the pandemic, investors showed moderate interest in the gold market. Research showed that during the first wave of the pandemic, both the demand and supply of gold remained on a downward trend, while the price of gold increased. The decline in the supply of gold bullion was shown to be responsible for the decline in the labour force, triggered by numerous restrictions designed to limit the transmission of the virus, which led, among other things, to the disruption of supply chains (e.g. restriction of mobility and temporary closure of economies). The results were also repeated in the other waves of the pandemic, and were confirmed by regression analysis, which showed that the gold market was largely influenced by the virus's ability to spread (R number). The study proved that a reduction in the supply of gold negatively affected the ability to purchase gold bullion, which had a direct impact on the increase in the price of gold during this period. The analysis of the gold market during the subsequent periods of the pandemic showed that, over time, demand and supply in the gold market increased and the price stabilised.

The results of the analysis confirmed that during the health crisis turmoil, investors moved away from intermediate investments in the gold market to investments in physical gold. Investments in gold bars and coins proved particularly attractive to investors. Particularly noteworthy, however, was the increased demand for jewellery production. During the period under review, this demand increased by 47%. This result, combined with the increase in demand for physical bars of gold (72%) and coins (84%), confirms that gold is a safe haven for investors in times of turmoil. This conclusion is consistent with previous studies (e.g. Baur and Lucey, 2010, p. 228; Baur and McDremot, 2010, p. 1897, et al.).

Slightly different results were obtained when analysing the impact of the military crisis on the gold market. The results of the study indicate that after Russia's invasion of Ukraine, the supply of gold increased, but at the same time the demand for gold bullion decreased, which in turn was reflected in a decrease in the price of gold.

Returning to the research hypothesis posed at the beginning of this paper, it is concluded that the research results only partially support this hypothesis. In doing so, it should be emphasised that the superimposition of the military crisis on the pandemic crisis in such a short period of time makes it difficult to assess the determinants of the demand and supply of gold, as well as its price. The health crisis, due to its duration, provides much more statistical data, allowing an in-depth study of its impact on the gold market. The military crisis, on the other hand, has lasted only a few months and does not provide much statistical data. Therefore, the results should be interpreted with some caution. This study is only a prelude to a broader study into the impact of the military crisis on the gold market. Future research focusing on the correlation between the gold price and the military invasion in the context of the upcoming digital revolution may prove interesting. This would contribute to a deeper understanding of the gold market in a changing socio-political, military, and health context.

REFERENCES

Akhtaruzzaman M., Boubaker S., Lucey B., Sensoy A. (2021), Is gold a hedge or a safe-haven asset in the COVID–19 crisis?, *Economic Modelling*, 102(2021), 105588.

Aspadarec W. (2013), Atrakcyjność inwestycyjna inwestycji alternatywnych, *Finanse*, 1(6), 9–26.

Baur D. G., Lucey B. M. (2010), Is gold a hedge or a safe haven? An analysis of stocks, bonds and gol, *The Financial Review*, 45(2), 217–226. DOI: 10.1111/j.1540-6288.2010.00244.x .

Baur D. G., McDremot T. K. (2010), Is gold a safe haven? International evidence, *Journal of Banking & Finance*, 34(8), 1886–1898.

Berry R., Syal S. (2022), Comparative study on performance of selective indices and gold during the time of crisis, [in:] *Emerging Trends in Management Sciences*, Kuldeep C., Vinay S., Amar R. (red.), Proceedings of Yogananda International Conference on Contemporary Issues in Social Sciences-2021, House of Journals, Pune, 198–213.

Borowski K. (2016), *Rynek inwestycji alternatywnych. Od inwestowania w metale i kamienie szlachetne do numizmatyki, skrypofilistyki i falerystyki*, CeDEWu, Warszawa.

Cai C., Clacher I., Faff R., Hillier D. (2008), A practical guide to gold as an investment asset, [in:] *The Handbook of Commodity Investing*, Fabozzi F., Füss R., Kaiser D. (red.), John Wiley and Sons, Hoboken, NJ, 712–735. DOI: 10.1002/9781118267004.ch31.

Czech M., Puszer B. (2021), Wpływ pandemii COVID-19 na zasoby złota w funduszach ETF, [w:] *Inwestycje alternatywne. Nowe spojrzenie*, Królik-Kołtunik K., Skibińska-Fabrowska I. (red.), CeDeWu, Warszawa, 131–146.

Dhuha S., Cahyadi H., Atourrahman M. (2021), Gold investment in a Syariah portgage during the COVID-19 pandemic time, *Cross-Border Journal of Business Management*, 1(1), 74–84.

Feldman K. (2010), Insight into investing in gold, *Alchemist The London Bullion Market Association*, (59), 74–84, https://cdn.lbma.org.uk/downloads/Alch59.pdf.

Gierałtowska U. (2013), Inwestowanie w metale szlachetne jako alternatywna forma lokowania kapitału, [in:] *Inwestycje finansowe i ubezpieczenia – tendencje światowe a rynek polski*, Jajuga K., Ronka-Chmielowiec W. (red.), Wydawnictwo Uniwersytetu we Wrocławiu, Wrocław, 88–100.

Gierałtowska U. (2016), Bezpośrednie i pośrednie inwestycje w metale szlachetne, *Annales Universitatis Mariae Curie-Skłodowska, sectio H – Oeconomia*, (4), 125–137.

Goodboy D. (2013), Gold or silver: Which is better for your portfolio?, *Yahoo Finance*, September 12. https://finance.yahoo.com/news/gold-silver-better-portfolio-030000278.html.

Juras A. (2021), Is gold as an alternative investment a good solution during pandemic?, *Finanse I Prawo Finansowe, Journal of Finance and Financial Law*, 3(31). DOI: 10.18778/2391-6478.3.31.05.

Kovinska I. (2014), *Gold as Safe Haven or Speculative Investment*, ISCTE Business School, Lisboa.

Krężołek D. (2020), *Modelowanie ryzyka na rynku metali*, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice.

Kucharska, I. (2022), *Rządowe Centrum Bezpieczeństwa*, Pobrano z lokalizacji Koronawirus - wirus SARS-CoV-2. https://www.gov.pl/web/rcb/koronawirus---wirus-sars-cov-2.

Ma, J. (2020). Coronavirus: China's first confirmed Covid-19 case traced back to November 17, *Pobrano z lokalizacji South China Morning Post*. https://www.scmp.com/news/china/society/article/3074991/coronavirus-chinas-first-confirmed-covid-19-case-traced-back.

Madhavan S., Sreejith S. (2022), A comparative analysis on the role and market linkages of gold backed assets during COVID-19 pandemic, *Scientific Annals of Economics and Business*, 69(3). DOI: 10.47743/saeb-2022-0019.

Pule B. (2013), *Evaluation of Gold as an Investment Asset: The South African Context*. https://core.ac.uk/download/pdf/39671663.pdf.

Pullen T., Benson K. L., Faff R. W. (2014), A comparative analysis of the investment characteristics of alternative gold assets, *ABACUS A Journal of Accounting, Finance and Business Studies*, 50(1). DOI: 10.1111/abac.12023.

Rabiej, M. (2012). *Statystyka z programem Statistica*. Gliwice: Helion

Rembisz W. (2009), *Instrumenty rynku rolnego*, Vizja Press & IT, Warszawa.

Saleem M., Khan M. (2013), The overview of gold ETFs and its various positive features, *International Journal of Marketing, Financial Services & Management Research*, 2(4), 125–135.

Salisu A., Raheem I., Vinh Vo X. (2021), Assessing the safe haven property of the gold market during COVID-19 pandemic, *International Review of Financial Analysis*, (74), 101666, DOI: 10.1016/j.irfa.2021.101666.

Šoja T. (2019), Gold in investment port folio from perspective of European inwestor, *EJAE*, 16(1), DOI: 10.5937/EJAE15–19652.

Statista. (2022), www.statista.com.

Swedroe L. E., Kizer J. (2008), *The Only Giude to Alternative Investments You'll Ever Need*, Bloomberg, New York.

US Debt Clock. (2022). www.usdebtclock.org.

Waghmare P., Shukla S. (2021), The impast of COVID-19 pandemic on gold investors behavior, Anvesh-2020 Doctoral Research Conference in Management Institute of Management, Nirma University. https://www.researchgate.net/publication/349848618.

Walczak M. (2012), Efektywność zdywersyfikowania portfela rynkowego GPW przy użyciu inwestycji w metale szlachetne, *Annales Universitatis Mariae Curie-Skłodowska, sectio H – Oeconomia*, 46(1).

Wang L. (2012), *Investment in Gold – An Empirical Study of the Gold Return from 90s to 21st*, Copenhagen Business School, Denmark, p. 383-393.

5 Use of artificial neural networks and decision trees for cost estimation of software projects – A model proposal

Marlena Stanek and Beata Czarnacka-Chrobot

1 INTRODUCTION

The aim of this chapter is to propose a model using selected data mining techniques to estimate the functional size of a product of a software development project on the basis of information about the project known at the beginning of its life cycle. The estimated software functional size enables the proper planning of project costs (and its other parameters, e.g., duration). This, in turn, facilitates the reduction of losses on investment in unsuccessful and often very expensive software projects. This is particularly important in the case of initiatives financed from the state budget, which contradicts the idea of sustainable development, as it hampers economic growth and social progress. Due to the high percentage of failures of IT/software projects and related losses, an area has been identified that can be supported with analytical methods, e.g., data mining techniques. The aim was to propose a model that could be used to support software project cost planning in order to increase the number of successful projects.

The chapter presents the process of building a model based on data from the International Software Benchmarking Standards Group (ISBSG) and with the application of selected data mining techniques, i.e., artificial neural networks and decision trees. The best results were obtained for decision trees using the gradient boosting method. Model scoring allowed the assessment of its predictive power. The obtained accuracy of the model is approximately 85%, which is a satisfactory result. The results can be considered good enough to meet the assumed criteria, but the proposed model can still be improved on the condition of access to complete and qualitatively better benchmarking data.

Therefore, the chapter focuses on the following issues: (1) the use of data mining in the management of software projects based on the analysis of the subject literature; (2) data mining techniques used to build the proposed predictive model,

DOI: 10.1201/9780367854690-5

i.e., artificial neural networks and decision trees; (3) a practical part that presents the entire process of the research performed in SAS Enterprise Miner, a tool created by SAS Institute, including a description of the ISBSG benchmarking data used and the subsequent stages of model development: sampling, exploration, modification, modeling and assessment (SEMMA) of the obtained results; (4) conclusions from the study together with the assessment of the proposed model – its strengths and weaknesses, indicating the future direction of the research.

2 THEORETICAL FRAMEWORK OF DATA MINING AND APPLICATION IN SOFTWARE PROJECT MANAGEMENT

Despite a high number of IT project failures (Johnson, 2021), as a result of which enormous amounts invested in software projects are wasted, a considerable increase in market demand for such systems is noticeable. Therefore, the principal issue is to identify the causes of this common undesirable phenomenon. It is normally attributed to a group of factors that can be synthetically defined as improper project management (e.g., Fatima, 2017). Therefore, the possibility to support the software project management process is crucial. This is because knowing what caused failures of software projects proves to be an insufficient method for controlling and planning them. And contracting entities, which are in demand of often very expensive software systems want to have the highest possible certainty that the undertaken project will not bring a financial loss. Contractors are also interested in avoiding failures, since their aim is the simultaneous increase in profits from and quality of projects. To this end, it is worth verifying the techniques permitting them to improve the quality of a process where this quality is dependent on the proper estimation of unknown parameters based on incomplete knowledge in high-risk conditions, with both these factors occurring predominantly at the beginning of a software system project. Such a technique is predictive analytics as one of the possibilities offered by data mining.

In some definitions, data mining is identified with knowledge discovery:

> Data mining, known also as knowledge discovery in databases is the nontrivial attempt at extracting implicit, previously unknown, and potentially useful information from data. This includes a range of techniques, such as clustering, data summarization, classification, correlation identification, change analysis and anomaly detection.
>
> *(Pujari, 2013)*

Other definitions point also to further aspects of the notion: data mining is the discovery of new significant relationships and global correlation patterns or trends existing in large databases, but hidden among enormous quantities of data, by screening them with the use of pattern recognition techniques as well as statistical and mathematical techniques. The purpose of extracting the identified knowledge from a database is to use it in decision-making, prediction, and estimation processes (Pujari, 2013). Data mining is a combination of numerous techniques: ones that permit recognition of the patterns found in data, thus enabling their better understanding, and ones that enable predictive modeling (Witten et al., 2017). In the case of prediction, the unknown attribute values are estimated and generalized based on the identified patterns.

Hence, predictive analytics allows the development and enhancement of predictive models based on the existing large amount of data. As mentioned above, prediction can be particularly helpful in the initial phase of a software system project, but it is useful also during its entire life cycle, since prediction enables, among others, the following (Fatima, 2017): (1) an objective overview of the project-related risks, the areas at risk, so that the identified gaps could be remedied, (2) improvement of the project performance by prioritizing the activities that might facilitate the success, (3) implementation of the values determined for projects by selecting specific implementation measures, (4) minimization of potential financial costs and losses and maximization of efficiency, (5) enhancement of organizational effectiveness by eliminating redundant project attributes and reinforcement of the features that contribute to the success of a project, (6) mitigation of the risk involved in estimating a project by comparing it with projects having similar attributes.

The application of data mining techniques can be exemplified by controlling the project schedule. In the publication by Pospieszny (2017), the author proposed the employment of data mining in estimating the initial values of a project and focused on a model making use of classification techniques. For that purpose, a data set was divided into four subsets, each of which was used for predicting another variable. Target variables included duration, task completion time, number of hours spent by a given contractor on completing the task, and number of tasks completed as of the analysis date. The models were built by means of three decision tree techniques: C4.5 algorithm, random tree, and CART (Classification and Regression Tree). The best results were achieved when using the random tree algorithm. An important statement made by the author of the analysis is his proposal to employ various methods to make the model capable of adapting to diverse organizational cultures, methodologies, and project data. In order to improve schedule management performance of development enterprises, it might be beneficial to employ advanced machine learning models and other data mining techniques. The application of such a solution will permit the use of the aforementioned benefits. The knowledge mined from data could serve as an effective support for managers in reasonable task planning and successive cycles of a project, thus allowing the delivery of complete products within the set deadline. Along with technological progress, the process might become fully controlled by machines (Wei & Rana, 2019).

In order to fully monitor a project, the EVM (Earned Value Method) can be combined with data mining techniques. Decision trees, neural networks, and association rules for predicting the total project duration using the EAC (Estimate at Completion) indicator as an outcome can be employed to this end. The EVM is a method integrating the scope, schedule, and measurement of resources for the purpose of assessing a project in terms of its efficiency and progress. It is an interesting solution for estimating the total effort and duration, since it is combined with the currently applied method. This might result in an easier transfer of the ideas presented in scientific studies to real-life applications. Another possibility is to use the knowledge arising from previous stages of a project for predicting effort in successive phases. It concentrates on the fuzzy set theory and association rules (Pospieszny, 2017).

Another essential aspect in which data mining is capable of supporting a project team is product quality. In response to business needs, researchers proposed several approaches to estimating quality-related problems. One of the proposals is a model estimating the number of defects by their criticality: minor, major, and extreme. As part of that proposal, three models founded on a decision tree were built based on MARS (Multivariate Adaptive Regression Splines). The obtained results differ insignificantly for each model and their estimation accuracy is above 90%. The satisfactory results suggest that the model could be used in practice. Additionally, when combined with the prediction of effort, they could offer valuable information in the project planning process. A different approach is the prediction of bug complexity and repair time. Here, clustering was applied to categorizing them by complexity (simple, medium, complex) and k-means algorithm – to calculating the repair time (Pospieszny, 2017).

The application of various machine learning (ML) algorithms in software engineering was the aim of the Systematic Literature Review (SLR) presented in (Mezouar & El Afia, 2022). It is, to the best of our knowledge, the most up-to-date analysis of this issue. The authors found that ML algorithms were the most commonly applied in the discussed area with respect to software project management to software defect/bug/fault prediction and its quality improvement. This is followed by data classification and software development effort estimation. According to the authors, the latter issue was covered by merely four studies, where the random forest algorithm was applied to predicting teamwork effectiveness, an approach based on neural networks, support vector machines, and generalized linear models was proposed for estimating software project effort and duration, the ensemble model approach to estimating such project parameters based on three ML algorithms and the software size measurement function point method were analyzed (Pospieszny et al., 2018). In addition, attention was paid also to effort estimation methods for agile projects based on methods using the so-called story points, which serve the purpose of software size measurement (they have not been accepted by international standardization organizations ISO/IEC for that purpose; see Czarnacka-Chrobot, 2023). They are followed in the SLR by: recommendations for code structure, detection of code changes, detection of the design pattern, evaluation of the software process, prediction of software reliability, improvement of software testing, and verification of software (Mezouar & El Afia, 2022). In this study, the ML algorithms that are most commonly used in the area of software engineering activities, including software project management support, were also analyzed as part of the SLR. It was found that, among 15 ML algorithms, three of them were the most frequently applied: the decision tree, the support vector machine, and the random forest algorithm. They are followed by naive Bayes algorithm, artificial neural networks (ANN), K-nearest neighbors, extreme learning machine, logistic regression, multilayer perceptron, while others are used in a negligible number of studies. The previously mentioned analysis shows that different ML algorithms are employed to handle different problems, but, in general terms, the usefulness of ML is noticeable at many stages of a software project life cycle: specification of requirements, design, including the capability to accurately predict software development effort, implementation, verification, and maintenance. Moreover, the authors declare that the interest in the subject matter of their examination has been markedly increasing since 2017, and hence it can be considered relatively new (Mezouar & El Afia, 2022).

Access to a database containing the information about the values of initial project parameters and the end results might permit, if employing the proper techniques, the creation of a certain set of rules. This would be an invaluable knowledge for software project managers. Based on the created rules and information about the projects managed by them, they could make better decisions and take more appropriate actions required for their effective execution. The existing systems enable the modeling and simulation of complex behaviors in the software development process as well as the generation of diverse scenarios for the same project. One of the greatest advantages of the process is the capability to check, almost cost-free, their influence on the project and the effect of various management rules or lack of such rules, and to allow for the individual attributes in the scenarios. The combination of the dynamic simulation of system development and data mining permits the obtainment of automatic management rules, which might make the decision-making process easier to project managers. The created rules might be useful for controlling the work, introducing the necessary changes in order to achieve the assumed goals, and also presenting the possible changes during the project and monitoring the aspects that give rise to objections (Alvares et al., 2004).

Hence, various data mining techniques can be employed as tools supporting the software project management process, including estimation of project attributes, such as software system size, which is supposed to be a product of a given project. The knowledge of the software size to be delivered at early development stages is a necessary condition for the correct (reliable and objective) estimation of effort, personnel costs, duration of the project, and – as a consequence – pricing of its product. This enables an increase in the number of software projects that will succeed, and a reduction of the huge financial losses incurred by contracting entities as a result of project failures. For contractors, in turn, data mining techniques might be a way to convert performance into measurable business benefits, such as a higher number of satisfied customers, financial profitability, more effective business processes, and greater market share. However, the multitude of benefits coming from the use of the information contained in data has not led to their broader application yet (Pospieszny et al., 2018). The high number of scientific studies, most of which are conceptual, does not translate into a high percentage of implementations. The effect is a gap between the worlds of science and business, which should be bridged.

3 THEORETICAL FRAMEWORK OF PREDICTIVE MODELING: ARTIFICIAL NEURAL NETWORKS AND DECISION TREES

This subchapter presents synthetically, due to the abundant relevant literature, two data mining techniques used for the predictive model proposed in this chapter: (1) artificial neural networks, and (2) decision trees.

Artificial neural networks are systems dedicated to processing information and modeling complex non-linear dependencies, the creation of which was inspired by a human nervous system. Artificial neurons composing a network and its structure are constructed based on biological functioning. The foundation of a neural network is a neuron, which is responsible for processing information. The components are connected with each other through links with weights that are modifiable while learning.

The basic network components are characterized by many input signals and one output signal (Tadeusiewicz & Szaleniec, 2015). The fundamental neuron components, in turn, are weights with their values, the internal processing function, and the activation function. The output signal value calculation stage is divided into two substages. First, input signals are multiplied by appropriate weights and subjected to processing by the preset function. This stage is called the internal processing function. The next stage is the activation function, which serves the purpose of obtaining the output signal. Activation functions can be any continuous functions generating a value from the relevant scale in the output neuron. These are usually the following functions: linear, unit step, sigmoid, and hyperbolic tangent (Stegowski, 2004). In terms of architecture, the following network types can be distinguished: single-layer – in such a network, neurons form one layer; multi-layer – its characteristic feature is the existence of at least one hidden layer; recurrent – there is a feedback between the input and output layers; hybrid – a mixture of the multi-layer and recurrent networks (Klamra et al., 2020). In the backpropagation algorithm, each neuron in the input layer is connected to all neurons in the hidden layer and each neuron in this layer is connected to each neuron in the output layer. The neuron receives a signal from the connections linking it to other neurons and then aggregates the inhibiting and exciting impulses it receives and sends another signal (Rajola, 2013). When learning, a neural network adapts through an iterative procedure, which modifies weights until it reaches the optimal point, where the discrepancies between the input, output and the historical values are the smallest. The test set is intended to evaluate the consistency between the output data provided by the network and a range of historical data, which were not included in the learning phase. Outputs are largely influenced by not only selected input variables but also a multitude of network configuration factors, such as, above all, network structure, and particularly the number of hidden layers, which affect both the results and time of learning. In reality, a network with an excessive number of hidden neurons and a high number of intermediate layers makes the learning time longer without improving the results. A network with a limited number of hidden neurons, in turn, might lead to a poor analysis (Rajola, 2013).

Another technique is decision trees. They are predictive models designed to select the variables that will best explain the target variable by means of a division of the analyzed data set. A characteristic feature of the method is easy interpretation of the results. A graphic representation of the tree shows the internal relationships between data in a simple and intuitive manner. However, decision trees tend to grow, which might hamper the interpretation of results due to a high number of attributes and their values. Decision trees are successfully implemented in supervised learning systems in order to classify objects into a specified number of classes. Tree nodes are designated with names of attributes, and branches show the values that can be assumed by attributes. At the same time, they are conditions determining the split of a feature. Leaves represent the classes resulting from the split. Depending on the split type, there are binary, tertiary, or multiple clusters. In order to classify the examined case, we must begin with the root of the tree and pass through successive nodes to reach the appropriate leaf. Decision trees are suitable for analysis of both qualitative and quantitative data. They are employed in applications for data mining and, for instance, segmentation, or construction of prognostic models (Rajola, 2013).

The algorithms that are most frequently used for building a decision tree are as follows: C4.5 (an improved version of the ID3 algorithm), CART, and CHAID (Chi-square Automatic Interaction Detector). What is an essential element in constructing a model is the selection of stopping criteria. The ones applied most commonly include the maximum number of steps taken in a process, the minimum group size, the minimum deviation of the parent group, and the minimum capability to explain the best split at each stage. Pruning, in turn, is not a stopping criterion; it is a procedure intended to reduce the size of a tree by cutting off less significant branches. Another essential factor is tree stability. A way to determine that is the random substitution of certain data to observe differences with reference to the original data set. One more method is to divide a data set into several random parts and confront the obtained results. If they are the same in both cases as in the original data set, the tree is stable (Rajola, 2013).

In this study, the gradient-boosting decision tree model was used. Gradient boosting is a powerful ML algorithm, which combines weak prediction models, thus forming a strong and effective classifier. Decision tree models that are trained based on the residuals of the previous tree are constructed iteratively. This enables the achievement of the goal by reducing the residual error in the learning process (Denga et al., 2019).

4 PROCESS OF BUILDING A PREDICTIVE MODEL – RESEARCH AND RESULTS ACCORDING TO THE SEMMA METHODOLOGY

4.1 Description of data and selection of variables

The data used for building the model come from the International Software Benchmarking Standards Group (ISBSG) repository. It is a non-profit institution founded by a group of national organizations dealing with software measurement methodologies in Australia in 1997. They intend to promote the employment of appropriate methods and data in the IT industry in order to improve the software development process. The ISBSG collects data in a few repositories holding information about various types of software projects executed in a variety of industries along with information about all significant parameters. The data come from software enterprises and are gathered in accordance with strictly defined guidelines in order to ensure their high quality (ISBSG, 2020a; ISBSG, 2020b; Czarnacka-Chrobot, 2023).

The data set received by the authors of this study from the ISBSG contained more than 9,000 observations (records) and more than 100 variables describing them. It concerned various projects completed in the period 1989–2018. However, the final data set consisted of more than 5,000 observations, since the study focused only on: (1) new development/enhancement/re-development projects, while omitting maintenance projects; and (2) business applications (data driven), while omitting projects such as infrastructure development or mathematical applications (real-time driven). This arises from the fact that the aim of the research was to construct a predictive model for business software functional size estimation in software development projects (e.g., new development/enhancement/re-development projects) based on the project information available at the beginning of its life cycle. It is worth noting here that the timespan within which the projects were executed is less significant for this aim.

Due to the aim of the research, the target variable selected for modeling was *Functional Size*. The variable describes the number of unadjusted function points (FP) of the software Functional Size Measurement (FSM) methods (ISO/IEC 14143:2002–2012) recognized by ISO/IEC as a standard of such measurement (Czarnacka-Chrobot, 2023). At the initial stage, the following 20 variables were selected, as they might explain the modeled phenomenon and are normally known at the beginning of the project life cycle (ISBSG, 2020c):

1. *Industry Sector* – the type of the industry in which the software implementation enterprise operates;
2. *Development Type* – the type of software development project: new development, enhancement, or re-development;
3. *Development Platform* – it defines the primary development platform: PC, mid-range, mainframe, or multi-platform;
4. *Language Type* – the type of programming language used in a project, e.g., 3GL (3rd Generation Language), 4GL (4th Generation Language), application generator, etc.;
5. *Primary Programming Language* – the primary programming language used in a project;
6. *Count Approach* – a method used for determining the size of the software: for most projects in the ISBSG repository, this is one of FSM methods, and the most popular of them is the IFPUG method devised by the International Function Point Users Group and standardized, in the principal part, by ISO/IEC (ISO/IEC 20926:2009), since over 70% of the products of projects are measured with this method (ISBSG, 2020b);
7. *Project Activity Scope* – the project phases completed during its execution: planning, specify, design, build, test, and implement;
8. *Resource Level* – the level at which data on effort per project are recorded (there are four levels: 1 = development team effort, 2 = development team support effort, 3 = computer operations involvement effort, 4 = end users' or client's effort);
9. *CASE Tool Used* – whether and, if so, what CASE (Computer-Aided Software Engineering) development support tool is used in a project;
10. *Used Methodology* – whether a team of developers apply a methodology to develop software;
11. *1st Operating System* – the primary operating system used during a project;
12. *1st Data Base System* – the primary database used during a project;
13. *Development Methodologies* – methodologies applied during the whole development process (e.g., agile, JAD, waterfall, etc.);
14. *Functional Sizing Technique* – a technique (tool) applied to support the process of measuring the functional size of the product of a project;
15. *Architecture* – the software architecture: stand-alone, multi-tier, client-server, or multi-tier with web public interface;
16. *DBMS Used* – whether a database management system is used while executing a project;
17. *Metrics Program* – whether the development organization has implemented a software process and product measurement program;

18. *Package Customization* – whether a project contains packages requiring customization;
19. *Project Type*;
20. *Upper CASE Used* – whether tools providing support at early stages of the project life cycle, such as requirement analysis and design, are used.

The entire analysis was carried out based on the SEMMA methodology developed by the SAS Institute (SAS, 2017). The SAS Institute is the firm that offers analytics software, which supports data modification and data analysis. The proposed methodology offers a logical list of successive steps taken during data mining. Its individual stages are as follows: Sampling, Exploration, Modification, Modeling, and Assessment (SEMMA). In order to explore large quantities of data, two applications from the SAS Institute were employed: SAS Enterprise Miner Workstation (Kattamuri, 2017) and SAS Enterprise Guide (Delwiche & Slaughter, 2018).

4.2 Data sampling and exploration: selection of variables

The first stage in building a model according to the SEMMA methodology is sampling. At this stage, impute nodes and data partitioning nodes were used. Before modeling, roles of individual variables in the set were specified. Additionally, the variables for which missing data exceeded 80% were rejected. Afterwards, the input data set was divided into three sub-sets: training, validation, and testing. Observations were selected for the sub-sets based on the 6:3:1 proportion.

The next stage was data exploration understood as a selection of variables. Preliminary data were analyzed as part of this point. This served as the basis for selecting the explanatory variables used for modeling.

Due to the high number of levels of the *Functional Size* target variable value, it was decided to categorize the variable. Responses from the ISBSG dataset were grouped into three categories. The division of the target variable into categories and their number of occurrences are presented in Tables 5.1 and 5.2. As follows from Table 5.2, the largest category is the first one, which refers to projects with the product size falling within the range of 0–100 in IFPUG unadjusted function points. The smallest category is the last one, which presents products with the largest sizes.

Next, preliminary data exploration, i.e., selection of variables, was performed. It allowed the verification of the correlation between explanatory variables and the

TABLE 5.1
Target variable categorization

Initial functional size values	Size of the product of a project (software)	Assigned value
0–100 FP	Small	0
101–500 FP	Medium	1
Above 501 FP	Large	2

Source: own work.

target variable as well as the distribution of their values with respect to the response variable (Kattamuri, 2017). The selection of the variables used when modeling the phenomenon was based on two approaches. One of them was automatic and involved choosing the “Selection of variables” node. The method selected for the target variable was chi-square and the upper threshold for missing data was set at 80%.

TABLE 5.2
Description of the target variable after categorization

Variable	Assigned value	Number of occurrences	Percentage
Functional size	0	2,537	46.6360
Functional size	1	2,215	40.7169
Functional size	2	688	12.6471

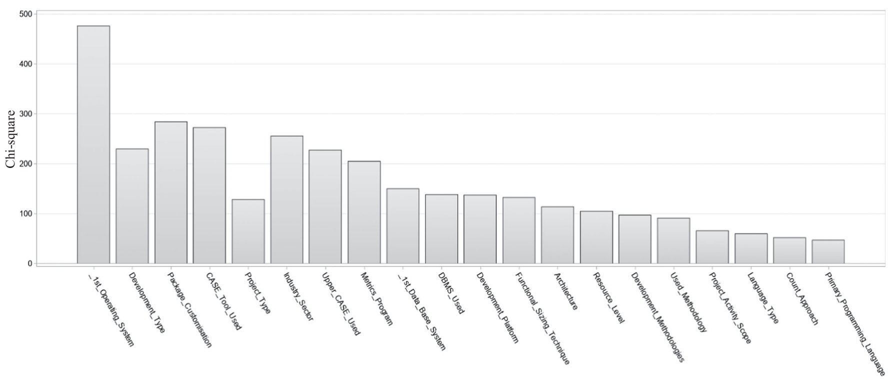


FIGURE 5.1 Chi-square coefficient values for selected dependent variables.

Source: own work.

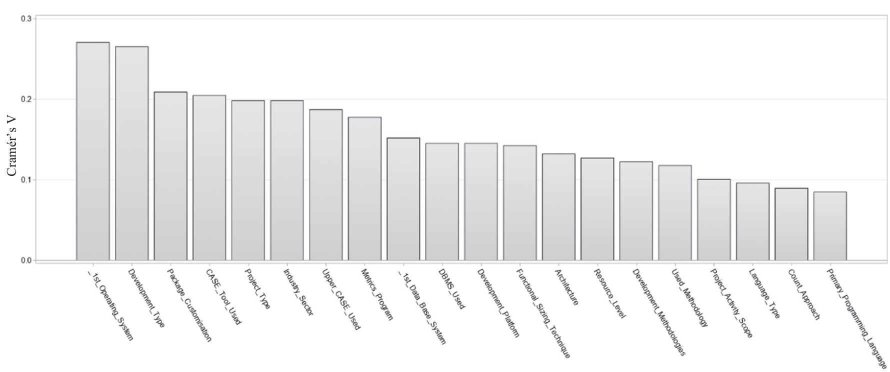


FIGURE 5.2 Cramér’s V coefficient values for selected dependent variables.

Source: own work.

The other approach was the application of the PROC CORR function computing correlation statistics. This permitted the verification of the correlation not only between the response variable and the explanatory variables but also between the explanatory variables themselves. The ultimate explanatory variables used when constructing the model were selected based on the chi-square coefficient. The significance level was determined at 0.05. The next step was the verification of the collinearity of the explanatory variables present in the model. The VIF (Variance Inflation Factor) statistic, which can be determined for each of the predictors, was used to that end. Cramér's V coefficient was applied, as it is one of the correlation measures that determines the level of correlation between two nominal variables, at least one of which assumes more than two values. The effects of employing those methods can be seen in Figures 5.1 and 5.2.

The numerical variables were selected based on the WoE (Weight of Evidence) method. Based on the information criterion indicating the most predictive variables with respect to the modeled variable, it was decided that the *Max Team Size* variable, signifying the maximum number of persons working on a project and being a continuous rather than qualitative variable, would be added to the set. In the case of the variables the individual categories of which explained only one category of the target variable, it was decided to remove them. The variables that were characterized by high correlation with other explanatory variables were also removed. A list of the following explanatory variables, which were used in the modeling process, was created as an effect: *Architecture*, *Development Methodologies*, *Development Platform*, *Development Type*, *Industry Sector*, *Language Type*, *Max Team Size*, *Metrics Program*, *Package Customization*, *1st Operating System*.

4.3 Data modification

Therefore, the model construction process allowed for the ten explanatory variables selected at the previous stage, nine of which were qualitative variables and one was a continuous variable. The contingency tables, i.e., ones showing the distribution of one variable in rows and another in columns, used to study the correlation between the two variables (see Tables 5.4–5.13), and charts of distribution with respect to the target variable after the recategorization process (combination of selected categories of the variable) only for the variables included in the selected model (see Figures 5.3–5.12) are presented below. Table 5.3, in turn, shows variables together with their classes before and after recategorization. The variables were recategorized due to a high number of levels of variable values or a very small size of their specific levels.

Initially, the *Architecture* variable was divided into seven categories. Due to the small size of five classes, they were consolidated into one: Other. Contingency Table 5.4 and Figure 5.3 present the distribution of categories after the recategorization of the variable.

The *Development Methodologies* variable initially assumed many unique values, which could be assigned to 12 categories. Due to the high fragmentation of results, the variable was divided, in accordance with the most common division, into two substantially differing project management methodologies (Waterfall, Agile), and the remaining methodologies were consolidated into one category: Other.

The distribution of categories after recategorization of the variable is demonstrated in Contingency Table 5.5 and Figure 5.4.

Initially, the *Development Platform* variable was divided into seven categories. After categorization, four categories were formed, including three referring to the most popular platforms. Contingency Table 5.6 and Figure 5.5 present the distribution of categories after recategorization of the variable.

The *Development Type* variable was initially divided into three categories, but enhancement and re-development are often considered to be one type of software project and therefore they were combined into a single category. Contingency Table 5.7 and Figure 5.6 present the distribution of the categories of the variable.

TABLE 5.3
Recategorization of explanatory variables

No.	Variable	Categories before recategorization	Categories after recategorization
1.	*Architecture*	• Stand-alone • Multi-tier/client-server • Client-server • Multi-tier with web public interface • Multi-tier • Embedded software • Other	• Stand-alone • Client-server • Other
2.	*Development Methodologies*	• Waterfall • Multifunctional teams • Incremental • Rapid application development (RAD) • Agile development • Spiral • Joint application development (JAD) • Timeboxing • Unified process • Scrum • Lean • Personal software process (PSP)	• Waterfall • Agile development • Other
3.	*Development Platform*	• MR • Multi • MF • PC • Proprietary • Handheld • Laptop	• PC • MF • Multi • Other
4.	*Development Type*	• New development • Enhancement • Re-development	• New development • Enhancement and re-development

(Continued)

TABLE 5.3 (CONTINUED)

No.	Variable	Categories before recategorization	Categories after recategorization
5.	*Industry Sector*	• Service industry • Communication • Construction • Wholesale & retail • Insurance • Banking • Medical & health care • Electronics & computers • Professional services • Manufacturing • Financial • Government • Utilities • Logistics • Mining • Education • Defense • Human resource • Estate • Food and beverage • Other	• Service industry • Communication • Banking and financial • Insurance • Manufacturing • Government • Other
6.	*Language Type*	• 4GL • 3GL • ApG • 2GL • 5GL	• Other • 3GL • 4GL
7.	*Max Team Size*	• Continuous variable	• From 0 to 5 • From 5.5 to 10 • From 10.5 to 20 • Above 20
8.	*Metric Program*	• No • Yes • Don't know	• No • Yes
9.	*Package Customization*	• No • Don't know • Yes • People soft • TIBCO	• No • Yes
10.	*1st Operating System*	• UNIX • Windows • DOS • MVS • IMS/VS • Solaris	

(Continued)

TABLE 5.3 (CONTINUED)

No.	Variable	Categories before recategorization	Categories after recategorization
10.	*1st Operating System*	• SUN • VME • Linux • VMS • SAP • HP Unix • Web client • IMS • Web application • AIX • VSE • Custom • .Net • SINIX • Other	• Windows • Other

Source: own work.

Initially, the *Industry Sector* variable was divided into 21 categories. Due to the high number of categories and a small share of most of them, it was decided to consolidate some classes. The effect was seven categories. Contingency Table 5.8 and Figure 5.7 present the distribution of categories after recategorization of the variable.

The *Language Type* variable was initially divided into five categories. However, the majority of projects were executed using 3GL and 4GL. The remaining categories were combined into one: Other. Contingency Table 5.9 and Figure 5.8 present the distribution of categories after recategorization of the variable.

The *Max Team Size* variable is numerical in source data and, due to the high number of unique values, it was decided to categorize it. Contingency Table 5.10 and Figure 5.9 present the distribution of categories after the recategorization of the variable.

Initially, the *Metric Program* variable had three categories, but, as a result of recategorization, it was decided to merge categories "No" and "Don't know" into one. Contingency Table 5.11 and Figure 5.10 present the distribution of categories after recategorization of the variable.

The *Package Customization* variable was initially divided into five categories. Due to the fact that some of the categories recurred, it was decided to merge them. Categories "No" and "Don't know" were also combined into one. Contingency Table 5.12 and Figure 5.11 present the distribution of categories after recategorization of the variable.

The *1st Operating System* variable in input data assumed a large number of unique values. After their analysis, it was decided that two categories will be created. Contingency Table 5.13 and Figure 5.12 present the distribution of categories after recategorization of the variable.

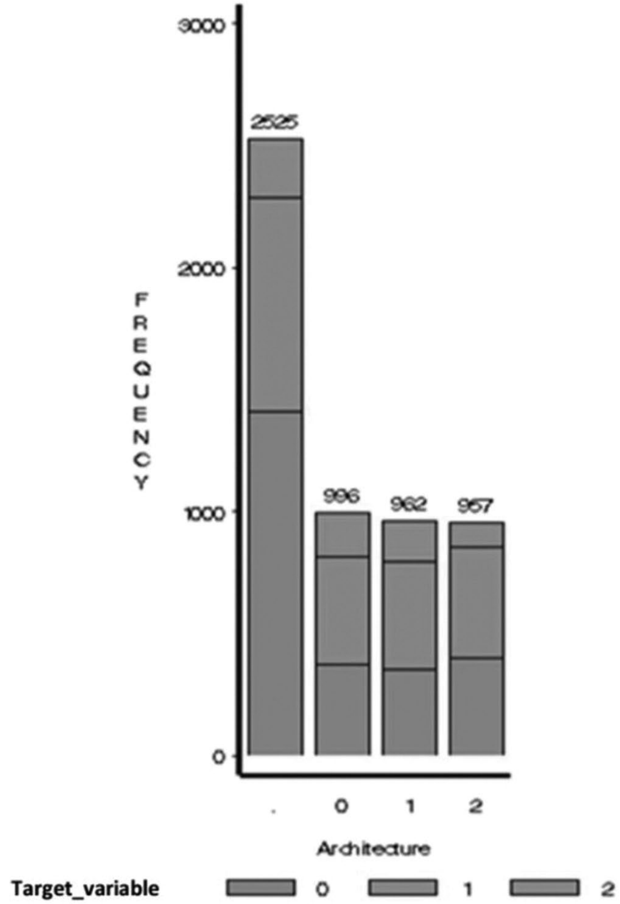


FIGURE 5.3 Distribution of the *Architecture* variable with respect to the target variable.

Source: own work.

TABLE 5.4
Contingency table for the *Architecture* variable

	Target_ variable	Architecture 0	1	2	Sum
Number	0	373	356	399	1,128
Percentage		12.80	12.21	13.69	38.70
Row pct.		33.07	31.56	35.37	
Column pct.		37.45	37.01	41.69	
	1	444	441	455	1,340
		15.23	15.13	15.61	45.97
		33.13	32.91	33.96	
		44.58	45.84	47.54	

(Continued)

TABLE 5.4 (CONTINUED)

Target_ variable	Architecture			
	0	1	2	Sum
2	179	165	103	447
	6.14	5.66	3.53	15.33
	40.04	36.91	23.04	
	17.97	17.15	10.76	
Sum	996	962	957	2,915
	34.17	33.00	32.83	100.00

Number of missing data = 2,525

Source: own work.

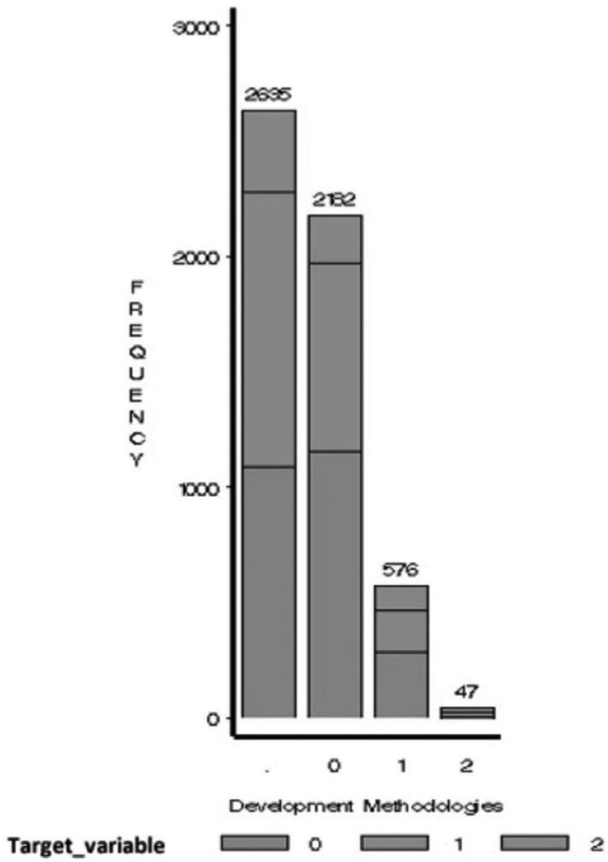


FIGURE 5.4 Distribution of the *Development Methodologies* variable with respect to the target variable.

Source: own work.

TABLE 5.5
Contingency table for the *Development Methodologies* variable

	Development Methodologies			
Target_variable	**0**	**1**	**2**	**Sum**
0	1,154	285	5	1,444
	41.14	10.16	0.18	51.48
	79.92	19.74	0.35	
	52.89	49.48	10.64	
1	818	185	24	1,027
	29.16	6.60	0.86	36 61
	79.65	18.01	2.34	
	37.49	32.12	51.06	
2	210	106	18	334
	7.49	3.78	0.64	11.91
	62.87	31.74	5.39	
	9.62	18.40	38.30	
Sum	2,182	576	47	2,805
	77.79	20.53	1.68	100.00
	Number of missing data = 2,635			

Source: own work.

4.4 Modeling

The next step, in accordance with the SEMMA methodology, was modeling. At that point, the research was divided into two separate processes. The reason was the use of input data with missing data. Decision trees handle the difficulty well and therefore, after selecting variables, the modeling process was commenced. Many models differing in terms of the maximum branching, depth, target variable splitting rules, significance level, and assessment measure were tested in order to select the best decision tree.

In addition to decision trees, models were constructed with the use of neural networks. However, this technique is not robust to missing data, which required the application of imputation before the final modeling. The imputation was carried out based on decision trees. By means of this method, a decision tree is built for a variable; it treats other variables as explanatory ones and the variable as a response variable (Skobel, 2020).

The explanatory variables, which were used in the final model, were selected also at that stage. Diverse variants of independent variables in the model were tested and the ultimately selected variables were the ones for which their distributions with respect to the target variable and contingency tables were presented above, i.e., *Architecture*, *Development Methodologies*, *Development Platform*, *Development Type*, *Industry Sector*, *Language Type*, *Max Team Size*, *Metrics Program*, *Package Customization*, and *1st Operating System*.

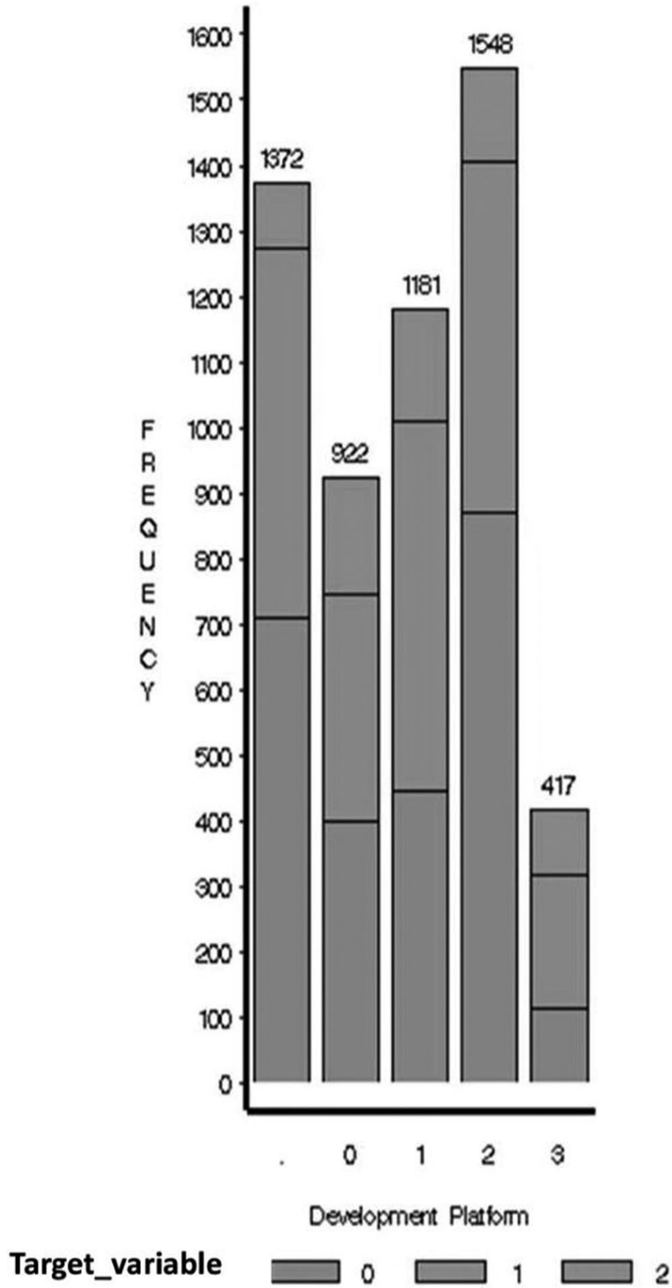


FIGURE 5.5 Distribution of the *Development Platform* variable with respect to the target variable.

Source: own work.

TABLE 5.6
Contingency table for the *Development Platform* variable

Target_ variable	Development Platform 0	1	2	3	Sum
0	398	446	870	113	1,827
	9.78	10.96	21.39	2.78	44.91
	21.78	24.41	47.62	6.19	
	43.17	37.76	56.20	27.10	
1	346	565	536	203	1,650
	8.51	13.89	13.18	4.99	40.56
	20.97	34.24	32.48	12.30	
	37.53	47.84	34.63	48.68	
2	178	170	142	101	4,068
	4.38	4.18	3.49	2.48	100.00
	30.12	28.76	24.03	17.09	
	19.31	14.39	9.17	24.22	
Sum	922	1,181	1,548	417	4,068
	22.66	29.03	38.05	10.25	100.00

Number of missing data = 1,372

Source: own work.

The final model was created with the use of the gradient-boosting algorithm. From among all methods employed when modeling, that one performed best. Its results are presented in Figures 5.13 and 5.14 and Tables 5.14 and 5.15.

Figure 5.13 presents classifications of the target variable for two sets in the selected model. In the case of the training set, the classification effects are better than those for the validation set. Category 0 was classified best for both sets. The highest number of incorrect classifications was obtained for category 1. Considering the category's strength, the highest percentage of incorrect classifications was obtained for category 2.

In Table 5.14, which presents fit statistics, a small difference between the three sets for the mean square error can be seen. In the case of the results obtained for the percentage of incorrect classifications, there are differences between the results for the training set and the validation and test sets. This might mean that the model adapted to the training data set.

Figure 5.14 presents a chart of the mean square error for the training and validation sets along with the increase in the number of leaves. The vertical line indicates the best number of leaves – this is when the model minimizes the fit statistics for the validation set. In the case of this model, the number is 43 leaves.

Table 5.15 demonstrates the obtained statistics for the test set. The highest result was obtained for ROC (Receiver Operating Characteristic). This means that

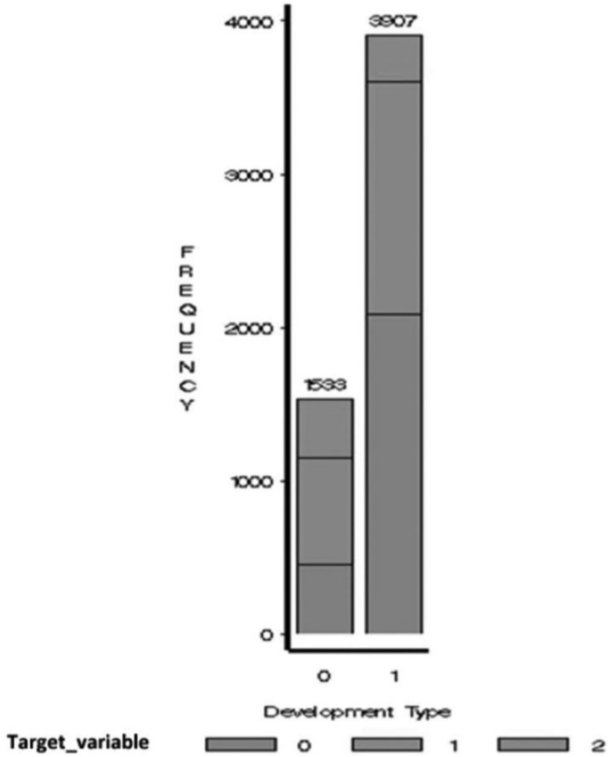


FIGURE 5.6 Distribution of the *Development Type* variable with respect to the target variable.

Source: own work.

TABLE 5.7
Contingency table for the *Development Type* variable

	Development Type		
Target_variable	**0**	**1**	**Sum**
0	452	2,085	2,537
	8.31	38.33	46.64
	17.82	82.18	
	29.48	53.37	
1	697	1,518	2,215
	12.81	27.90	40.72
	31.47	68.53	
	45.47	38.85	
2	384	304	688
	7.06	5.59	12.65
	55.81	44.19	
	25.05	7.78	

(Continued)

TABLE 5.7 (CONTINUED)

	Development Type		
Target_variable	**0**	**1**	**Sum**
Sum	1,533 28.18	3,907 71.82	5,440 100.00

Source: own work.

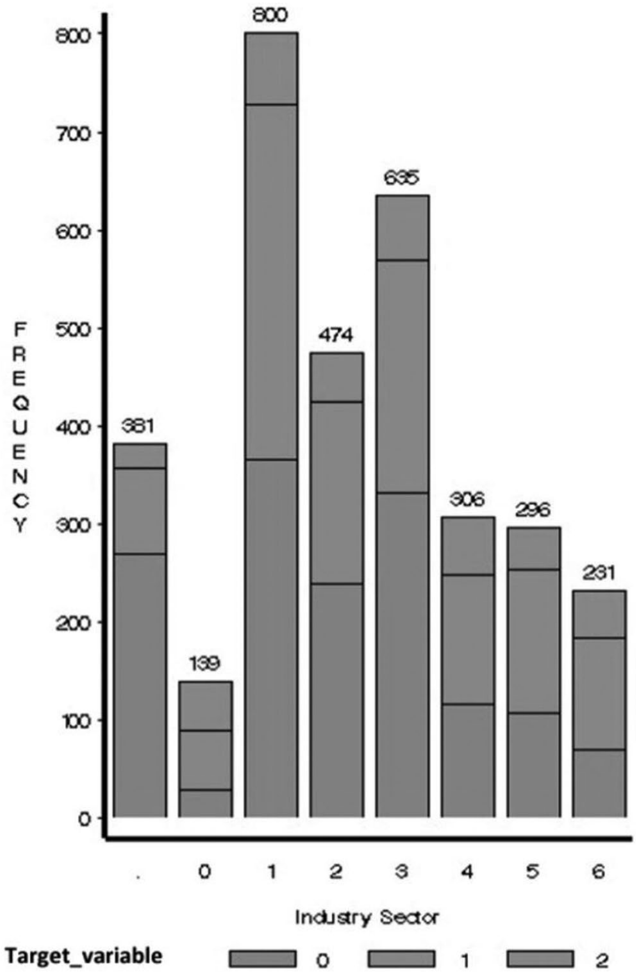


FIGURE 5.7 Distribution of the *Industry Sector* variable with respect to the target variable.

Source: own work.

TABLE 5.8
Contingency table for the *Industry Sector* variable

Target_ variable	Industry Sector							
	0	**1**	**2**	**3**	**4**	**5**	**6**	**Sum**
0	39	613	421	535	189	172	116	2,085
	0.81	12.77	8.77	11.14	3.94	3.58	2.42	43.43
	1.87	29.40	20.19	25.66	9.06	8.25	5.56	
	19.02	47.23	51.28	52.71	34.81	33.14	29.00	
1	92	580	305	383	253	266	191	2,070
	1.92	12.08	6.35	7.98	5.27	5.54	3.98	43.12
	4.44	28.02	14.73	18.50	12.22	12.85	9.23	
	44.88	44.68	37.15	37.73	46.59	51.25	47.75	
2	74	105	95	97	101	81	93	646
	1.54	2.19	1.98	2.02	2.10	1.69	1.94	13.46
	11.46	16.25	14.71	15.02	15.63	12.54	14.40	
	36.10	8.09	11.57	9.56	18.60	15.61	23.25	
Sum	205	1,298	821	1,015	543	519	400	4,801
	4.27	27.04	17.10	21.14	11.31	10.81	8.33	100.00
	Number of missing data = 639							

Source: own work.

the model effectively distinguishes the values of the dependent variable. The Gini coefficient is also at a high level, which permits the ascertainment that the predictive power is satisfactory.

4.5 Model assessment

The last step in the process, in accordance with the SEMMA methodology, was the quality assessment of the final model. For this purpose, the model was scored based on the set of data used for modeling, yet without the dependent variable.

The matrix in Table 5.16 and the results presented in Table 5.17 show that the model identified the best correctly modeled feature (sensitivity, called also recall or true positive rate) for class 0. This arises from the largest number of projects included in the category. The model was capable of learning on a larger quantity of data, which translated into better results for the category. In the case of identifying negative results by the classifier (specificity), the best result was obtained for class 2. It also produces the highest ratio of the predicted positive results to true positive ones (precision, called also positive predictive value). In Table 5.17, the last statistic presents two values: sensitivity and precision. Due to the fact that the ability to correctly identify the positive results of the modeled features is the most significant, the F1 measure statistic was added. Unlike accuracy, it does not take into account true negative values, which are less significant in this case.

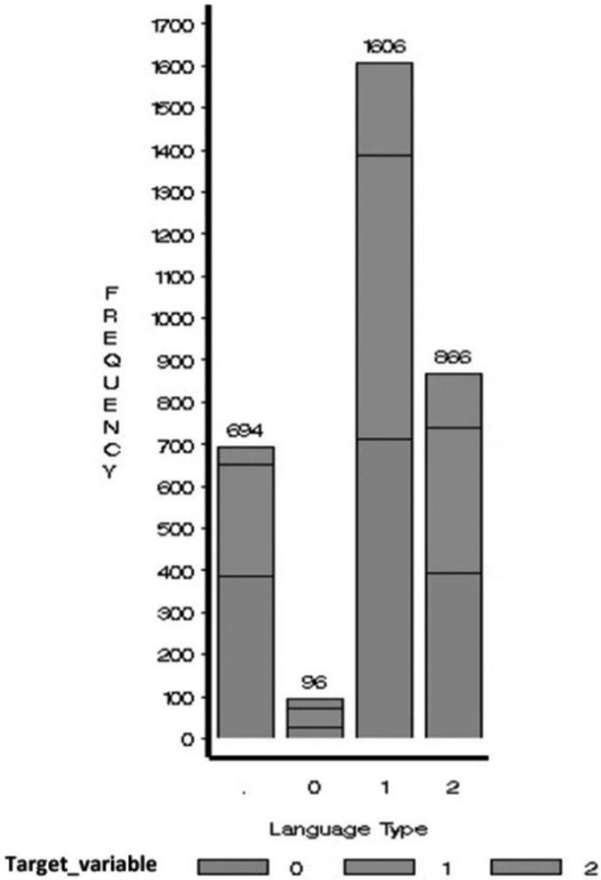


FIGURE 5.8 Distribution of the *Language Type* variable with respect to the target variable.

Source: own work.

Table 5.18 presents the results obtained based on the classification matrix for the whole model. They are compared there with the results demonstrated in the study (Jaekwon et al., 2015).

At first, the sensitivities of both models were compared. When juxtaposed with the model obtained as part of the study described in this chapter, the comparative model performs much better when it comes to identifying positive values of features. This property was considered the most important one when assessing the predictive power. The results show that this is an area for further improvement. In the case of other statistics, better results of the model built as part of this analysis are noticeable, with a particular focus on accuracy, which measures the ratio of the sum of true positive and true negative results to the sum of true positive, false positive, true negative, and false negative results.

TABLE 5.9
Contingency table for the *Language Type* variable

	Language Type			
Target_variable	**0**	**1**	**2**	**Sum**
0	42	1,177	677	1,896
	0.98	27.41	15.77	44.15
	2.22	62.08	35.71	
	27.27	43.89	46.43	
1	79	1,144	555	1,778
	1.84	26.64	12.93	41.41
	4.44	64.34	31.21	
	51.30	42.65	38.07	
2	33	361	226	620
	0.77	8.41	5.26	14.44
	5.32	58.23	36.45	
	21.43	13.46	15.50	
Sum	154	2,682	1,458	4,294
	3.59	62.46	33.95	100.00

Number of missing data = 1,146

Source: own work.

5 CONCLUSIONS AND RECOMMENDATIONS

The aim of this chapter was to propose a model using selected data mining techniques, i.e., neural networks and decision trees, to estimate the functional size of a product of a software development project on the basis of information about the project known at the beginning of its life cycle. The estimated functional size of software is the necessary condition for the proper planning of software project attributes, mainly effort and then costs (and its other parameters, e.g., duration). This, in turn, facilitates the reduction of losses on investment in unsuccessful and often very expensive software projects, especially those financed from state budgets. Therefore, this chapter presents, first of all, the process of building the proposed model based on data from the International Software Benchmarking Standards Group against the synthetic review of the relevant literature. The model was constructed in accordance with the SEMMA methodology: the starting point was sampling, followed by the exploration/selection, modification, and modeling stages, and – finally – assessment of the obtained results.

The best results from the verified approaches, i.e., artificial neural networks and decision trees, were obtained for decision trees by means of the gradient boosting algorithm. The scoring of the model enabled the assessment of its predictive power. The resulting accuracy of the final model is approximately 85%, which is a satisfactory outcome.

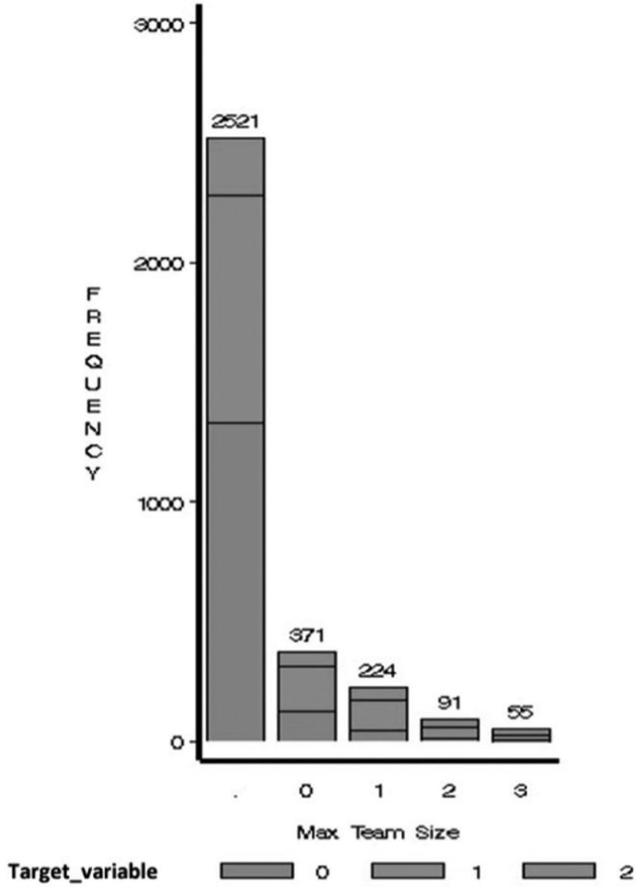


FIGURE 5.9 Distribution of the *Max Team Size* variable with respect to the target variable.

Source: own work.

TABLE 5.10
Contingency table for the *Max Team Size* variable

Target_ variable	Max Team Size				
	0	**1**	**2**	**3**	**Sum**
0	212	67	21	5	305
	17.06	5.39	1.69	0.40	24.54
	69.51	21.97	6.89	1.64	
	33.54	18.56	12.73	5.88	
1	325	208	83	31	647
	26.15	16.73	6.68	2.49	52.05
	50.23	32.15	12.83	4.79	
	51.42	57.62	50.30	36.47	

(Continued)

TABLE 5.10 (CONTINUED)

Target_ variable	Max Team Size 0	1	2	3	Sum
2	95	86	61	49	291
	7.64	6.92	4.91	3.94	23.41
	32.65	29.55	20.96	16.84	
	15.03	23.82	36.97	57.65	
Sum	632	361	165	85	1,243
	50.84	29.04	13.27	6.84	100.00
		Number of missing data = 4,197			

Source: own work.

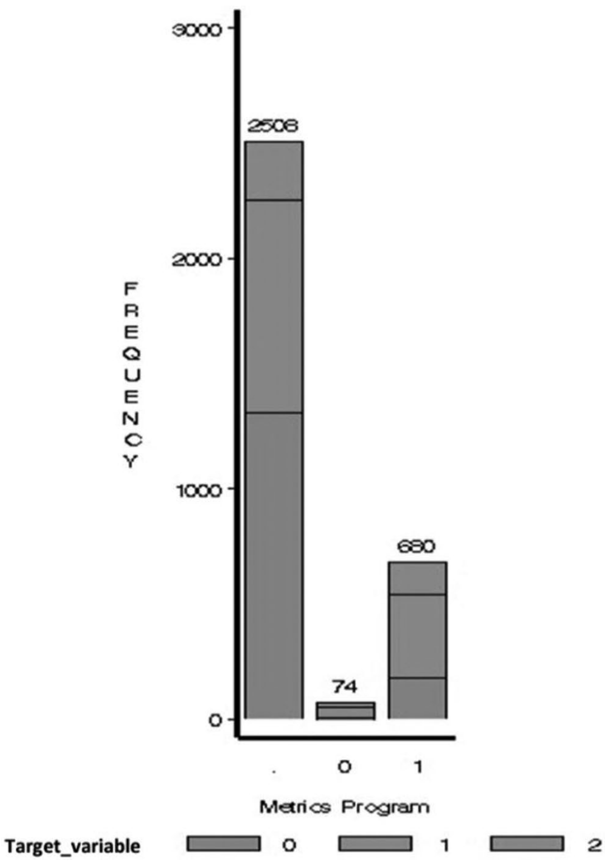


FIGURE 5.10 Distribution of the *Metrics Program* variable with respect to the target variable.

Source: own work.

TABLE 5.11
Contingency table for the *Metrics Program* variable

	Metrics Program		
Target_variable	**0**	**1**	**Sum**
0	19	315	334
	1.45	24.10	25.55
	5.69	94.31	
	13.01	27.13	
1	83	611	694
	6.35	46.75	53.10
	11.96	88.04	
	56.85	52.63	
2	44	235	279
	3.37	17.98	21.35
	15.77	84.23	
	30.14	20.24	
Sum	146	1,161	1,307
	11.17	88.83	100.00

Number of missing data = 4,133

Source: own work.

In order to compare the produced results, they were juxtaposed with the data published in a selected scientific study. The model's strengths and weaknesses can be observed on this basis. The sensitivity ratio of the presented model deserves particular attention. Increasing its value will have a beneficial influence on predictive power, as an effect of which the model could be used for business purposes. At the same time, the model stands out by its high specificity. This means that it performs very well in predicting true negative classes. A drawback is the unequal sizes of the target variable classes. This influences the accuracy of identifying the modeled features and therefore the categories should be equinumerous. Another aspect is missing data. Unfortunately, some variables had over 90% of missing data, which excluded them from the presented process. A better quality of input data permits the obtainment of a larger amount of information and has a positive impact on the prediction.

A limitation that may affect the implementation of the proposed solution is its quite high complexity. It requires experience and knowledge of data mining techniques. The model has also not been tested in real-world conditions on ongoing projects. This means that if such a solution is implemented for business purposes, additional work is likely to be required to improve the usability of the model or to support it by using other techniques. Therefore, the model can be a supportive approach for those with sufficient experience and knowledge.

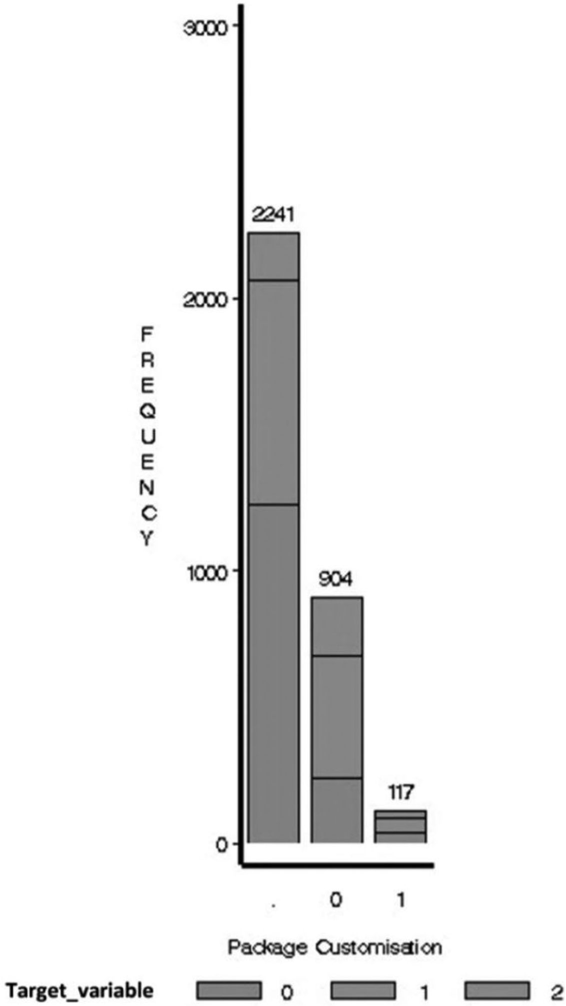


FIGURE 5.11 Distribution of the *Package Customization* variable with respect to the target variable.

Source: own work.

TABLE 5.12
Contingency table for the *Package Customization* variable

	Package Customization		
Target_variable	**0**	**1**	**Sum**
0	391	64	455
	22.76	3.73	26.48
	85.93	14.07	
	25.79	31.68	

(Continued)

TABLE 5.12 (CONTINUED)

	Package Customization		
Target_variable	**0**	**1**	**Sum**
1	768	95	863
	44.70	5.53	50.23
	88.99	11.01	
	50.66	47.03	
2	357	43	400
	20.78	2.50	23.28
	89.25	10.75	
	23.55	21.29	
Sum	1,516	202	1,718
	88.24	11.76	100.00
	Number of missing data = 3,722		

Source: own work.

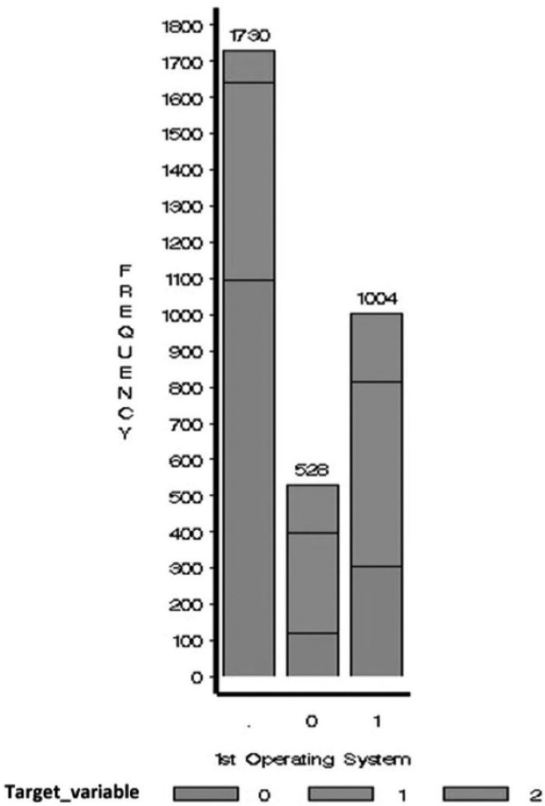


FIGURE 5.12 Distribution of the *1st Operating System* variable with respect to the target variable.

Source: own work.

TABLE 5.13
Contingency table for the *1st Operating System* variable

Target_variable	1st Operating System 0	1	Sum
0	213	517	730
	8.22	19.95	28.17
	29.18	70.82	
	24.01	30.34	
1	458	856	1,314
	17.68	33.04	50.71
	34.86	65.14	
	51.63	50.23	
2	216	331	547
	8.34	12.77	21.11
	39.49	60.51	
	24.35	19.42	
Sum	887	1,704	2,591
	34.23	65.77	100.00

Number of missing data = 2,849

Source: own work.

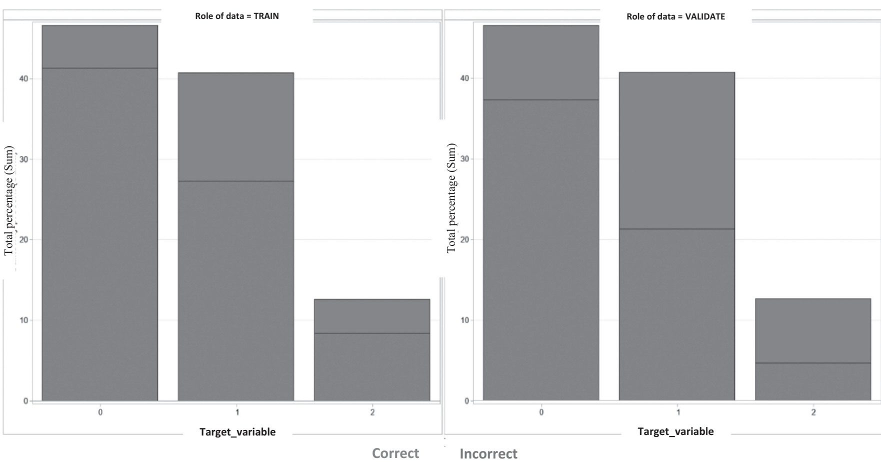


FIGURE 5.13 Chart of target variable classification for training and validation data.

Source: own work.

To recapitulate, although the obtained results can be considered good enough and meet the assumed criteria, the proposed model can still be improved in terms of sensitivity on condition of access to complete and qualitatively better benchmarking data.

TABLE 5.14
Fit statistics

Statistics label	Training	Validation	Testing
Sum of sizes	3,262	1,631	547
Sum of products of weights and numbers of cases	9,786	4,893	1,641
Percentage of incorrect classifications	0.229307	0.366033	0.413163
Maximum absolute error	0.96849	0.991047	0.989136
Sum of square errors	988.6941	775.1904	279.506
Average squared error (ASE)	0.101031	0.158428	0.170327
Root mean square error	0.317855	0.398031	0.412706
Divisor of ASE	9,786	4,893	1,641
Total degrees of freedom	6,524		

Source: own work.

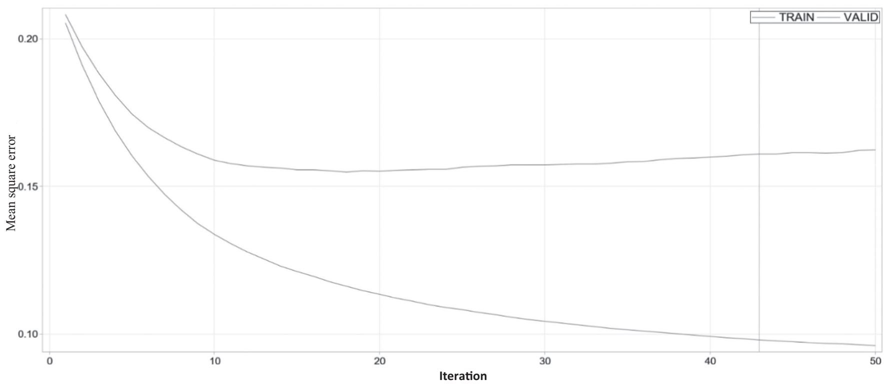


FIGURE 5.14 Chart of the mean square error for the training and validation sets.

Source: own work.

TABLE 5.15
Assessment statistics for the model's predictive power

Model	Kolmogorov–Smirnov statistic	Gini coefficient	ROC	Cumulative lift
Final model	0.545	0.676	0.820	3.694

Source: own work.

TABLE 5.16
Classification matrix

Predicted classes	Actual classes			
	0	**1**	**2**	**Sum**
0	1,348	407	44	1,799
1	165	891	93	1,149
2	9	30	275	314
Sum	1,552	1,328	412	

Source: own work.

TABLE 5.17
Assessment of the model's predictive power for individual classes

Class	Sensitivity	Specificity	Precision	F1 measure
0	0.8839	0.7408	0.7493	0.8110
1	0.6710	0.8666	0.7754	0.7194
2	0.6675	0.9863	0.8758	0.7576

TABLE 5.18
Comparison of models

Model	Accuracy	Sensitivity	Specificity	Precision
Final model	0.8471	0.7716	0.8853	0.7716
Comparative model	0.6951	0.9310	0.2564	0.6995

REFERENCES

Alvares, J.L., Mata, J., & Riquelme, J.C. (2004). Data mining for the management of software development process. *International Journal of Software Engineering and Knowledge Engineering, 14*(06), 665–695.

Czarnacka-Chrobot, B. (2023). Use of machine learning for software projects cost estimation based on ISO/IEC standards. In: Thi Hong Nga, N., Agarwal, S., & Ziemba, E. (eds.) *Analytics in Finance and Risk Management*. CRC Press, Taylor and Francis (in print).

Delwiche, L.D., & Slaughter S.J. (2018). *Little SAS Enterprise Guide Book*. SAS Institute Inc.

Denga, S., Wanga, C., Wanga, M., & Sun, Z. (2019). A gradient boosting decision tree approach for insider trading identification: An empirical model evaluation of China stock market. *Applied Soft Computing, 83*, 1–11.

Fatima, T. (2017). A predictive analytics approach to project management: Reducing project failures in web and software development projects. *World Academy of Science, Engineering and Technology, International Journal of Materials and Metallurgical Engineering, 4,* 10–17.

ISBSG. (2020a). *ISBSG Corporate Presentation*. International Software Benchmarking Standards Group. Retrieved April 7, 2022 from https://www.isbsg.org/wp-content/uploads/2020/02/2020.01.12-ISBSG-Corporate-Presentation.pdf.

ISBSG. (2020b). *ISBSG D&E – Corporate Release Demographics, ISBSG, R1 v1.0*. International Software Benchmarking Standards Group.

ISBSG. (2020c). *Field Descriptions - ISBSG D&E Repository*. International Software Benchmarking Standards Group.

ISO/IEC 14143:2002–2012 Information Technology – Software Measurement – Functional Size Measurement – Part 1–6. ISO/IEC.

ISO/IEC 20926:2009 Software and Systems Engineering – Software Measurement – IFPUG Functional Size Measurement Method. ISO/IEC.

Jaekwon, K., Jongsik, L., & Youngho, L. (2015). Data-mining-based coronary heart disease risk prediction model using fuzzy logic and decision tree. *Healthcare Informatics Research, 21*(3), 167–174.

Johnson, J. (2021). *CHAOS 2020: Beyond Infinity*. The Standish Group International.

Kattamuri, S.S. (2017). *Predictive Modeling with SAS Enterprise Miner Practical Solutions for Business Application, Third Edition*. SAS Institute Inc.

Klamra, M., Kowalski, P., & Krawiec, K. (2020). *Rodzaje sztucznych sieci neuronowych*. Cracow University of Technology. Retrieved March 3, 2020 from https://www.l5.pk.edu.pl/~pkowal/SSN/rodzaje.htm.

Mezouar, H., & El Afia, A. (2022). A systematic literature review of machine learning applications in software engineering. *Lecture Notes in Networks and Systems, 489*, 317–331.

Pospieszny, P. (2017). Application of data mining techniques in project management - An overview. *Collegium of Economic Analysis Annals, 43*, 199–220.

Pospieszny, P., Czarnacka-Chrobot, B., & Kobylinski, A. (2018). An effective approach for software project effort and duration estimation with machine learning algorithms. *Journal of Systems and Software, 137*, 184–196.

Pujari, A.K. (2013). *Data Mining Techniques*. Orient Blackswan.

Rajola, F. (2013). *Customer Relationship Management in the Financial Industry*. Springer.

SAS. (2017). *Introduction to SEMMA*. SAS Institute Inc. Retrieved January 8, 2020 from https://documentation.sas.com/doc/en/emref/14.3/n061bzurmej4j3n1jnj8bbjjm1a2.htm.

Skobel, M. (2020). *Eksploracja danych. Uzupełnianie i transformacja danych wejściowych oraz predykcja z wykorzystaniem sztucznych sieci neuronowych i regresji logistycznej*. University of Zielona Gora. Retrieved January 8, 2020 from http://staff.uz.zgora.pl/mskobel/ED/lab4.pdf.

Stegowski, Z. (2004). *Sztuczne Sieci Neuronowe*. Kernel.

Tadeusiewicz, R., & Szaleniec, M. (2015). *Leksykon sieci neuronowych*. Projekt Nauka (Project "Science").

Wei, W., & Rana, M.E. (2019). Software project schedule management using machine learning & data mining. *International Journal of Scientific & Technology Research, 8*(9), 1385–1389.

Witten, I., Frank, E., Hall, M.A., & Pal, C.J. (2017). *Data Mining: Practical Machine Learning Tools and Techniques, Fourth Edition*. Elsevier Ltd.

6 Data accounting

Michal Gajda

1 INTRODUCTION

Error tracing and error-impact estimation have been a holy grail of practical use of automated approaches to risk reporting (Basel Committee on Banking Supervision 2013), large-scale distributed error tracking (Julian 2017), and automated debugging.

Since the processing of risk reports and data analysis pipelines can be frequently expressed using a sequence relational algebra operations, we propose a replacement of this traditional approach with a data summarisation algebra that helps to determine the impact of errors. It works by defining data analysis of a necessarily complete summarisation of a dataset, possibly in multiple ways along multiple dimensions.

We also present a description to better communicate how the complete summarisations of the input data may facilitate easier debugging and more efficient development of analysis pipelines.

This approach can also be described as a generalisation of axiomatic theories of accounting in data analytics, thus dubbed *data accounting*.

We also propose formal properties that allow for transparent assertions about the impact of individual records on the aggregated data and ease debugging by allowing us to find minimal changes that change the behaviour of data analysis on a per-record basis.

The answer is based on fully bidirectional computation as the first stage of the analytics pipeline, and generalisation of aggregation functions into *complete aggregations* that are taken together to be sensitive to all inputs.

This approach will allow us to make more robust data analytics (Gajda 2020a, 2020b) and accounting analytics, based on a formal mathematical underpinning.

2 LITERATURE REVIEW

Basel Committee Best Practices (BCBS) 239 (Basel Committee on Banking Supervision 2013) is a set of recommendations for financial institutions on how to organise their risk reporting pipelines. These recommendations also give careful considerations to general principles of data aggregation and reporting, and thus became the reference for other work on best practices in aggregated reporting.

We also are inspired by practical rules coming from recent opinions on data pipeline construction methodology (Gajda 2020a, 2020b).

Reversible computing (Bennett 1973) aims to reverse computation from outputs to inputs. Bidirectional computing (Bancilhon and Spyratos 1985) attempts to generalise programs to allow picking a subset of either outputs and inputs and computing the remaining.

 DOI: 10.1201/9780367854690-6

Automated debugging aims to automatically indicate possible places in the program that may be buggy. Automated program derivation aims to propose valid programs given both inputs and outputs.

Differentiable computing (Levitt 1983; Abadi and Plotkin 2019) aims to find a minimum of one function by computing its value and differential at the pivot point. This differential is then used to correct the inputs towards the more desirable outputs.

2.1 Database view update problem

A database view is a computation from input relations to output relations. Database community has long tried to resolve the view update problem (Chen and Liao 2011), which tries to reverse the computation of the view so that the update on results is transformed into an update on source relations.

While this approach was limited in scope to relations, it has a great impact on data analytics since big datasets are often modelled directly as relations or translatable to relations (for example, graphs translated into an incidence relation).

Analysis of the problem has led to the identification of view complement, which contains a relation missing from the view, but essential for computing the update.

In this sense, the database community identified views as non-injective. To achieve updatable views, they attempted to restore the injectiveness of the views by computing view complement.

2.2 Event sourcing

Event sourcing replaces the database as a source of truth with a history of events that happened. This event log corresponds to the general ledger in accounting in the sense that it can only be appended to, and the state becomes just a view or balance on the ledger.

2.3 Automated debugging

Automated debugging approaches aim to find a location of a bug in the program from the fault in the output. This may be realised with an approach akin to a bidirectional program but applied to program source instead of inputs.

2.4 Philosophy and axiomatisation of accounting

Accounting is a science that attempts to predict the future performance (Ijiri 2018, 1975) of companies by insight into cash flows and metrics. The mathematical axiomatisation of accounting (Ijiri 2018) is based on a concept of time-invariant measure that partitions money in a way that changes along the time dimension but preserves a total sum.

The above is consistent with a traditional preference of accountants to use a monotonic measure of cash flow instead of recording only the sum of money currently available, and instead of adding and removing money, moving the money between different accounts.

This approach, called double entry accounting (Ellerman 2007), is best presented by presenting a total balance as pair of debit and credit sides. Subtraction of debit from credit yields a total balance. Both sides are non-negative and strictly monotonic over time. That means that we have a strictly monotonic measure that is easy to verify across time.

The next step is a generalisation of double-entry accounting to multiple accounts: for example, we can make separate accounts for cash balance and bank account balance.

To deal with the difficulty of putting all elements of the business on a monetary scale, modern accountants base statements, not just on a ledger of past transactions, but on an infobase defined as an entire database used to define financial statements (Warsono, Ridha, and Darmawan 2009).

Note that a valuation-centred approach (like fair value accounting – FVA) makes it hard to communicate non-monetary information and is subject to currency stability. It is also hard to describe barter transactions where one company gives services directly in exchange for another service, with no money changing hands (Balzer and Mattessich 1991).

As opposed to FVA, traditional accounting is based on "going-on concern", which values assets on a non-liquidation basis, and thus is oriented towards the future performance of the company.

Information content school emphasises that disclosure of business information attempts to reduce uncertainty regarding the entity. It sees accounting information as a proactive management support tool that is not focused solely on fair value.

2.5 Accounting measure

In the work on axioms of accounting measure, Ijiri (2018) notes that all pertinent objects are quantifiable by amount, volume, weight or another measure, and divided into a countable collection of classes. There must be a physical measure for each class, but they are not the same for different classes. The "principle of substitutability" tells that two sets of objects with the same quantity can be mutually substituted for the purpose of exchanges. Non-substitutable objects must be accounted in a class of itself. Measures must be correct mathematical measures that are non-negative, zero if a class is empty, and countably additive (monotonic). Some objects may possess multiple classes that often correspond to different reporting dimensions. Time is a real variable over which some objects change their classes. But the total measure of objects over time must be invariant. Measures may be allocated, imputed and compared based on the current and future value of an asset. The default is to use the cost of replacement at a given point in time.

For example, we can estimate the cost of purchase for a workstation at 1,000$ in 2020. The value of a workstation will decrease due to wear and obsolescence to 800$ in 2021. But since the total measure must remain 1,000$, we divide this value in 2021 into two classes: The current workstation value is $800, and the written-off value is $200.

Balzer and Mattesich (1991), and Ellerman (2007) made different attempts at axiomatising value-based accounting that partitions resources into disjoint groups, but they go along similar principles.

For example, Ellerman (2007) gives a solution for monotonic tracking of information contained in both cash inflows and outflows by splitting them into pairs of monotonically increasing values of debit and credit used in traditional double-entry accounting systems.

Some accounting philosophers add that the goal of accounting is more than, to sum up, the numbers; it is to predict the future state of business and reduce uncertainty about this state (Warsono, Ridha, and Darmawan 2009). Thus accounting becomes a science of data analysis and summarisation where accounting measure plays just a small part within a big edifice of information summarisation.

2.6 Requirements for error impact reporting

Here we summarise the basic requirements for error impact reporting in bulk:

The analytics pipeline should not crash for a single failing record, but instead should run as a whole for a given algebra expression, and provide correct relation(s) (that contain correct and relevant results) and error-trace relations that complement correct results with the information needed for error reporting. That means functions must be total.

No information is lost or discarded. Whenever a piece of relevant information should not affect the output, we sort it into a different category of information provided, but irrelevant so it can be put into an error trace relation. We may convert between types of information, but non-null information must be converted to entities of non-null information. We treat information as monotonic (it can only increase) and additive.

Every correct relation that contains data that is valid and relevant at the current stage of the pipeline, we also keep an error trace with records that were discarded, and thus they could not impact the processing. This allows us to estimate the possible impact of erroneous records if they were provided on the input. This means that we preserve erroneous inputs in the form of error outputs.

2.7 Railroad-oriented programming

At the end of the analytics pipeline, we should always be able to provide an automatic error dashboard that summarises how much data was correctly processed and summarised in the output, and how much data was discarded and for what reasons. This again means that the dashboard that gives a final summary of the output must be sensitive to all inputs and all forms of errors.

We should automatically tag the error trace with all information necessary to discover at which stage of the pipeline the data was discarded. This means that we should divide the analytics pipeline into summarisation stages, and transformation stages, and transformation stages should preserve sufficient information about bits of input to evaluate their impact until final summarisation.

We conclude that the best approach to fulfill these requirements is to attempt to keep all errored records along with the correct records, and summarise them in parallel.

Railroad-oriented programming methodology (Wlaschin 2012) suggests merging all error records together in a separate output stream that is also summarised, just like the data on the so-called happy path (the main path that yields validated records for summarisation). This leads to two parallel streams of processing, one forwarding the valid data, the other forwarding errors and anomalies.

3 PROPOSED WORK

This work attempts to unify approaches of bidirectional programming, railroad-oriented programming, and accounting to get the best features of all three. That is, we aim to provide both injectiveness of debugging data pipelines (understood as finding differences to the program that fixes it), and in order to do so, we maintain strict injectiveness of view/filter computations and monotonic data aggregations. In this process, we discover that a unifying theory for bidirectional computation and debugging can be found on the basis of a mathematical theory of accounting.

To our knowledge, bidirectional computing was so far not applied to monotonic data aggregations. We also find that this is the first work exploring the connection between axiomatic accounting (Ijiri 2018) and both bidirectional computing and monotonic summarisations (Ross and Sagiv 1997).

3.1 Motivating example

We propose the following running example of the container ship contents.

Container ship example table "items":

Inventory table

Description	Purchase price	Commodity	Quantity	Unit	Insurance
Sailors	Priceless		17	Person	2,000,000$
Nutella	10$		20,000,000	Jar	
Grain		Wheat Grain	200	Tonne	
Milk		Milk	5,000	Litre	
Cat			1	Animal	
Nut oil	Unknown	Nut oil	1,000	Litre	
Ship	2,000,000,000$		1		1,000,000,000$

Commodity prices reference table

Market	Commodity	Price	Currency	Per unit	Quote date	Future date
Forward	Wheat grain	6.0981	USD	Bushel	2021-12-01	2022-02-01
Spot	Wheat grain	6.0575	USD	Bushel	2021-09-01	
Spot	Wheat grain	Closed	USD	Bushel	2021-12-01	
Spot	Milk	33.98	EUR	100 kg	2021-12-01	

Table of reports

Name	Description	Accounted	Unaccounted
Insured value	Total amount paid by the insurer to all parties	Insured value	Cat, nut oil
Replacement cost	Cost of replacement for the owner in case of loss	Purchase price or hiring cost	Cat, nut oil
Weight	Total mass transported	Weight of commodities	Sailors, cat
Weight	Total mass transported	Weight of commodities	Sailors, cat

The point is to show how different values may be computed depending on whether we consider a total payment by the insurer for the ship loss, the cost of replacing the ship by a ship owner or purchasing commodities on the market upon unreasonable delay. Since we will be interested only in complete summarisations of the dataset, we will only build queries that summarise relevant data and summarise omitted or ignored data. This is considered the best practice (Basel Committee on Banking Supervision 2013).

3.2 Outline

The similarity between best practices in impact reporting (Basel Committee on Banking Supervision 2013) and the mathematical theory of accounting measure (Ijiri 2018) is not accidental. We may formalise complete summarisations of an input data set that generalises both the notion of error impact report and the accounting measure defined above.

But instead of measure, which is a rather narrow numeric value, we generalise to use a notion of information that is an additive monoid over a given set. The notion of ownership is translated into partitions of the set. Invariance of a total measure over time is replaced by invariance of total information monoid during processing.

Data space represents a complete summarisation of a dataset. The language that can only formulate complete summarisations would work differently from traditional relational algebra, and thus to define it, we need a mathematical theory.

3.3 Definitions

Partition of a set we call a family of disjoint sets that cover the space.

Reals $\mathbb{R}$ and positive reals $\mathbb{R}_+ = \{x \,||\, x \geq 0 \wedge x \in \mathbb{R}\}$.

We use symbol $\leq$ or any partial ordering, $\cup$ for tagged disjoint union, with in_r for tagging right and in_l for tagging left.

We use $\cup$ as a classical set union, and $\cup$ for a union of sets labelled by right label $\mathrm{in}_r()$ and left label $\mathrm{in}_l()$.

3.4 Data space

As mentioned before, we will focus on complete summarisations of a given input. By complete, we mean that they do not lose information.

We formalise the above notions as data space:

Data space is a tuple with the following $\mathbb{D}, \leq_{\mathbb{D}}, \left\|-\right\|_{\mathbb{D}}, \left(\mathbb{M}, e, \Diamond\right)$:

data carrier set $\mathbb{D}$;
information set $\mathbb{M}$;
partial ordering representing information growth $''_{\mathbb{D}}$ 1;
measure of data by elements of set $\mathbb{M}$: $\left\|-\right\|: 2^{\mathbb{D}} \rightarrow \mathbb{M}$.
information fusion monoid $\Diamond_{\mathbb{D}} : \mathbb{M} \rightarrow \mathbb{M} \rightarrow \mathbb{M}$, that is additive in m; For a set $A = \left\{a_0, \ldots, a_n\right\} \in 2^{\mathbb{D}}$, we use shorthand notation $\Sigma_{\Diamond\mathbb{D}}(A) = a_0 \Diamond_{\mathbb{D}} \cdots \Diamond_{\mathbb{D}} a_n$.
neutral element e of fusion monoid contains is the neutral element of $\mathbb{M}$ that is equal to $\left\|\emptyset\right\|_{\mathbb{D}}$
we expect information fusion to be invariant2 over any partition of the carrier set $\mathbb{D}$:

$$\forall x, y \in \mathbf{2}^{\mathbb{D}} \cdot \Diamond_{\mathbb{D}}()$$

monotonicity of partial ordering with respect to information fusion:

$$a \leq b \wedge c \leq d \Rightarrow a \Diamond c \leq b \Diamond d$$

We note that our partial ordering can often be derived from well-behaved information monoid by equational laws: $a \Diamond b \leq a$ and $a \Diamond b \leq b$.

It describes an additive monoid with partial ordering consistent with this monoid. That is because we are interested only in complete summarisations of data on the given carrier set $\mathbb{D}$.

3.4.1 Examples of data spaces

accounting measure (Ijiri 2018)
any mereological space with a finitely additive metric, which is a family of sets equipped with operations of Boole's algebra, and additive metric over these sets (Arntzenius 2004; Barbieri and Gerla 2021);
typelikes used for recording union types (Gajda 2022) with a partial ordering derived from the information fusion monoid as $a \Diamond b \leq a$ and $a \Diamond b \leq b$, and metric $\left|-\right\|$ as type inference operation.

Banach space (metric space) with set union as information fusion.

Each set also enjoys its identity data space:

$$\begin{array}{lcl} \left\|-\right\|_{id} & := & id \\ \mathbb{D}_{id} & := & \mathbb{M}_{id} \\ \leq_{id} & := & \subset \\ \Diamond_{id} & := & \cup \end{array}$$

Another exciting example is Paccioli algebra which describes double-entry accounting:

$$
\begin{aligned}
\mathbb{D}_p &:= \mathbb{R}_+ \times \mathbb{R}_+ \\
\left(d_1, c_1\right) \Diamond_p \left(d_1, c_2\right) &:= \left(d_1 + d_2, c_1 + c_2\right)
\end{aligned}
$$

$$\text{partial ordering } \left(d_1, c_1\right) \leq_p \left(d_2, c_2\right), \text{ when } d_1 \leq d_2 \text{ and } c_1 \leq c_2$$

This is an example of how we create an additive data space over account balances, by converting each balance into a pair of strictly positive credit and debit.

4 DATA SPACE SUMMARISATION

Whenever the summarisation of data is performed, we should attempt to perform similar summarisation on the fields in the error path, or another measure of impact counting when these fields are absent.

For each table, we indicate possible summarisations by assigning fitting monoid to a column and indicating other columns that may be used for grouping the results.

4.1 Example summarisation

Column	Type	Unique	Result	Aggregation
Order_id	Integer	Unique	Set of identifiers	Union
Product_id	Integer	–	Set of identifiers	Union
Unit_price	Number	–	Number	Min, max, avg
Quantity	Number	–	Number	Sum (for each unit type)
Unit_price * quantity	Number	Derived	Number	Sum

4.2 For example the table order _ details from Table 6.4

Now that we read this example, we also see that reporting monoids exhibit certain trends:

Sets usually summarise OIDs, unique identifiers, and dimensions;
sum, min, max, average summarise numbers.

Using the same summarisation on the same datatype (unit) is most common. Summarising together different datatypes (units) is usually a mistake, in which case we use a containing category, for example, by counting different types of faulty records instead of summing their values.

Automatic summarisation can perform a summarisation on all permitted aggregations in order to detect common patterns.

We can infer that we may assign a heuristic that for every summarisation of correct data would also summarise incorrect records on those monoids that can still be used in case of incomplete data.

That should make a reasonable default summarisation for the data outside the main path (or "happy path") of the analysis.

4.2.1 Information-preserving operations

Information-preserving operations can be formulated in a general way as bidirectional and thus reversible, or as safeguarding the information measure.

Conjecture: Reversible operations are always preserving any specific information measure as defined above

4.3 Partition of data space

Instead of selecting or projecting data from the input, the lossless operation is to partition the data set into subsets of desired properties.

For example, for a predicate p, we can mark it as

$$\begin{array}{lcl} \left(\mathbb{D}_{\text{accepted}}, \mathbb{D}_{\text{rejected}}\right) & := & \text{partition}(p, \mathbb{D}) \\ \mathbb{D}_{\text{accepted}} & = & \{x \mid x \in \mathbb{D} \wedge p(x)\} \\ \mathbb{D}_{\text{rejected}} & := & \{x \mid x \in \mathbb{D} \wedge \text{‹} \, p(x)\} \end{array}$$

4.3.1 Information-preserving union

The union of two sets is information-preserving if members of each set are clearly labelled in a way that allows inferring from which input set they come from. This may be either because they come from disjoint sets, or because they are explicitly tagged:

$$\mathbb{D}_u = \mathbb{D}_1 \cup \mathbb{D}_2 = \left\{ \text{in}_l \left(d_1 \right) \middle\| d_1 \in \mathbb{D}_1 \right\} \cup \left\{ \text{in}_r \left(d_1 \right) \middle\| d_1 \in \mathbb{D}_1 \right\}$$

In some cases, we already have tags that serve the same purpose, that is, there is an origin function f:

$$\exists f \cdot \forall x \cdot f(x) = \begin{cases} \text{in}_1() & x \in \mathbb{D}_1 \\ \text{in}_r() & x \in \mathbb{D}_2 \end{cases}$$

4.3.2 Information preserving function

In case we are interested only in information measures $\mathbb{M}_1, \supset, \mathbb{M}_n$ and $\mathbb{M}_1', \ldots, \mathbb{M}_n'$ we can narrow ourselves to the functions that preserve information for these functions, that is:

Given function $f : \mathbb{D} \to \mathbb{D}'$ we can compute a corresponding mapping of $g : \mathbb{M}_i \to \mathbb{M}_i'$ such that $g(m(x)) = m'(f(x))$.

4.3.3 Examples of information-preserving functions

In the simplest example, it may occur that $\mathbb{D} \to \mathbb{D}'$ and $\mathbb{M}_i \to \mathbb{M}_i'$. When $\mathbb{M}_i \neq \mathbb{M}_i'$, we tell that the function f changes the dimension.

Another example would be re-partitioning the dataset with a predicate, or computing additional quantities for each record.

Note that aggregation generally only preserves information that corresponds to their dimension (for example, summing up quantities only maintains this dimension, but not the number of records). Later we will discuss how to construct aggregations preserving multiple dimensions by using tensor products of aggregations on individual dimensions.

Derivations of complex data spaces from simpler ones.

Now we address a derivation of complex data spaces from simpler ones.

4.4 Tensor products

We already saw that data metric space can be derived from a well-behaved information fusion monoid.

Below we attempt to assemble data spaces from subspaces that conform to different data metrics

4.4.1 Disjoint tensor product

Disjoint tensor product of any two data spaces $\mathbb{D}_1 \otimes \mathbb{D}_2$ is a data space such that:

data carrier set is	$\mathbb{D}$	:=	$\mathbb{D}_1 \times \mathbb{D}_2$
information measure	$\|\|(a, b)\|\|_{\mathbb{D}}$	:=	$\left(\|\| a \|\|_{\mathbb{D}_1}, \|\| b \|\|_{\mathbb{D}_2}\right)$
partial ordering	$(a, b) \leq_{\mathbb{D}} (c, d)$	when	$a \leq_{\mathbb{D}_1} c \wedge b \leq_{\mathbb{D}_2} d$
information fusion	$(a, b) \Diamond_{\mathbb{D}} (c, d)$	:=	$\left(a \Diamond_{\mathbb{D}_1} c, b \Diamond_{\mathbb{D}_2} d\right)$
neutral element	$e_{\mathbb{D}}$	:=	$\left(e_{\mathbb{D}_1}, e_{\mathbb{D}_2}\right)$

4.4.2 Parallel tensor product

Given two data spaces over the same carrier set, we may want to use composite data metric that represents information conveyed by different data metrics over the same carrier set:

Parallel product of two data spaces $\mathbb{D}_1 \Theta \mathbb{D}_2$ is a data space such that:

data carrier set is	$\mathbb{D}$	$= \mathbb{D}_1 = \mathbb{D}_2$
information measure	$\left\Vert x \right\Vert_{\mathbb{D}}$	$= \left(\left\Vert x \right\Vert_{\mathbb{D}_1}, \left\Vert x \right\Vert_{\mathbb{D}_2}\right)$
partial ordering	$a \leq_{\mathbb{D}_1} c \wedge b \leq_{\mathbb{D}_2} d$	$\Rightarrow (a, b) \leq_{\mathbb{D}} (c, d)$
information fusion	$(a, b) \Diamond_{\mathbb{D}} (c, d)$	$= \left(a \Diamond_{\mathbb{D}_1} c, b \Diamond_{\mathbb{D}_2} d\right)$
neutral element	$e_{\mathbb{D}}$	$= \left(e_{\mathbb{D}_1}, e_{\mathbb{D}_2}\right)$

Information reconstruction and information projection

Given two dataspaces $\mathbb{D}_1$ and $\mathbb{D}_2$ such that the data carrier is the same, but different information carriers $\mathbb{M}_1$ and $\mathbb{M}_2$, we can make a $\mathbb{D} = \mathbb{D}_1 \oplus \mathbb{D}_2$ by making a tensor product of information carriers $\mathbb{M} = \mathbb{M}_1 \Theta\, \mathbb{M}_2$ and using induced partial order and information fusion monoid:

$$a \leq \mathbb{D}_1 \oplus \mathbb{D}_2 b \text{ when both } \|a\|_{\mathbb{D}_1} \leq \mathbb{D}_1 \|b\| \text{ and } \|a\|_{\mathbb{D}_2} \leq \mathbb{D}_2 \|b\|_{\mathbb{D}_2}$$

$$m\Diamond_{\mathbb{D}_1 \oplus \mathbb{D}_2} n = \left(m\Diamond_{\mathbb{D}_1} n,\, m\Diamond_{\mathbb{D}_2} n\right)$$

For disjoint information carriers, we may call the component dataspaces $\mathbb{D}_1$ and $\mathbb{D}_2$ projections:

$$\mathrm{Proj}_{\mathbb{D}_1}(\mathbb{D}) = \mathbb{D}_1$$
$$\mathrm{Proj}_{\mathbb{D}_2}(\mathbb{D}) = \mathbb{D}_2$$

4.4.3 Products of functions

Since we are interested in transformations of data spaces, we denote function products that correspond to the above products on data spaces:

1. **Disjoint product of functions:**

$$(f \otimes g)(x)\begin{cases} in_r(x) \to f(x) \\ in_l(y) \to g(y) \end{cases}$$

1. **Parallel product of functions**:

$$(f\Theta g)(x) = (f(x), g(x))$$

4.4.4 Partial functions

Each partial function $f : \mathbb{D} \mapsto \mathbb{M}_1$ can be extended to a total function $\mathbb{T}(f) : \mathbb{D} \to \mathbb{M}$ by using the partition of the data space $\mathbb{D}$ into two different fragments $\mathbb{D}_1$ and $\mathbb{D}_2$:

$$\mathbb{T}(f)(a)\begin{cases} in_r(f(a)), & \text{if } x \in \mathbb{D}_1 \\ in_l(a), & \text{otherwise} \end{cases}$$

That means that $f : \mathbb{D} \mapsto \mathbb{M}_1 \cup \overline{D_1}$

4.4.5 Lookups and joins

Joins may represent lookups of information that may change the impact of processed records. But it may also be true that even when computing the inner join on correct data, some error stream data will have absent key fields necessary for the join.

Because of this, for every join of correct data, we perform a similar outer join on error path data. Because of this, we need to always perform an outer join on error path data, with both records that are on the correct path and error records.

We limit information-preserving operations to outer join with a tagging of results to the left, right, and inner output. The inner output consists of tuples matching by the join condition c, left part consists of tuples from $\mathbb{D}_1$ that were not matched by any of $\mathbb{D}_2$, and right part consists of tuples from $\mathbb{D}_2$ that were not matched by any of $\mathbb{D}_1$.

$$\mathbb{D}_1 \underset{c}{\rhd\lhd} \mathbb{D}_2 := \left(\mathbb{D}_{\text{left}}, \mathbb{D}_{\text{inner}}, \mathbb{D}_{\text{right}}\right)$$

For a predicate P, we can mark it as:

$$\begin{aligned}
&\mathbb{D}_{\text{inner}} := \{(x, y) \mid x \in \mathbb{D}_1 \wedge y \in \mathbb{D}_2 \wedge p(x, y)\} \\
&\mathbb{D}_{\text{left}} := \{xx \mid\in \mathbb{D}_1 \wedge \langle\, \exists_{y \in \mathbb{D}_2} p(x, y)\big\} \\
&\mathbb{D}_{\text{right}} := \{y \mid x \in \mathbb{D}_2 \wedge \langle\, \exists_{y \in \mathbb{D}_1} p(x, y)\big\}
\end{aligned}$$

Note that join condition c is a filter on a Cartesian product of $\mathbb{D}_1$, and $\mathbb{D}_2$ that only depends on the join columns (usually keys).

4.4.6 Value lookup example

We take an example of looking up the current price of the commodity in the table including product and quantity of the position. The reference table is keyed by product and has description, and current_price. Normal lookup is inner join when we have a guaranteed match of product_id, which we convert to an outer join, where the left side corresponds to product_id entries that are missing from the reference table, and the right side corresponds to product_id entries in reference tables that are unused.

Either the left and the right side are aggregated into error feedback for the left table, and unused feedback for the right table to ensure that records are never lost.

This allows us to easily find outdated references, or misspellings like "oli" instead of "oil" and report them correctly as a summary row with a count of 14 rows, and five units with an unknown product name of "Nut oli" that was a result of a typographical error.

4.4.7 Value lookup example

We take an example of looking up the current price of the commodity in the table including product and quantity of the position. The reference table is keyed by product and has description, and current_price. Normal lookup is inner join when we have a guaranteed match of product_id, which we convert to an outer join, where the left side corresponds to product_id entries that are missing from the reference table, and the right side corresponds to product_id entries in reference tables that are unused.

Either the left and the right side are summarised into error feedback for the left table and unused feedback for the right table to assure that records are never lost. The inner join (middle) contains valid results that are passed on to the summarisation of correct data.

This allows us to easily find outdated references or misspellings like oli instead of oil and report them correctly as a summary row with a count of fourteen rows and five units with an unknown product name of oli.

4.4.8 Columnar projection and row format

Since we enforce the principle of never discarding the data that can be relevant to debugging effort, we cannot simply discard fields with projection operator. Because of this, our records contain two types of fields:

fields relevant to the final output
fields irrelevant to final output and only kept for the debugging purposes

For each record, we have a single value for each relevant field, but possibly a set of values for each irrelevant field to take account of discarding duplicate relevant records. That means that we have a sub-relation for each record that keeps only irrelevant fields.

This allows us to implement a lossless projection operation that moves relevant fields into irrelevant fieldset.

Note that the projection operation is executed in the same way on both correct relation and error tracking relation that pairs with it.

4.4.9 Per-row data enrichment operations

In SQL syntax, it is common to add a computed field to a query from a single table:

select "Purchase price"*"Amount" **as** "Total price" **from** items;

We call these operations data enrichment and can execute them on all rows. However, in order to maximise debugging potential, we propose a methodology of automatically shifting computing additional fields as early in the pipeline of operations as possible. That allows us to provide more information. After the record leaves the correct stream and goes into the error stream, we no longer enrich it with new field values relevant to the main computation in order to prevent cascading faults.

We can only enrich these records with more error-specific information, like adding a description of the error found and estimating the impact of this information.

That is why besides fmap operation on correct data stream, we also allow emap operation on the error stream that enriches error information. We can optionally schedule this information to be computed in advance when the stage of the pipeline is reached and before any records are added to the error stream:

```
pipeline = do
 forM filename -> do
 emap (@"Filename" :-> filename) $ do
 ...
```

4.4.10 Putting algebra together

The final step is to use these operations in a relevant way: that means that all data spaces need to be either included in the final aggregation, or used by data spaces that contribute to it. Just like in linear logic, we are not allowed to use certain variables more than once, in data space algebra, we are not allowed to forget a variable.

4.4.11 Translating relational algebra

Relational algebra is a standard algebra used as a basis for describing database operations and analytics pipelines.

These primitives are usually implemented on sets of rows (records of the same shape) called relations. However, they can equally well be implemented on multisets or lists of records

It uses the basic primitives of:

projection which discards a set of columns and keeps only the columns relevant for the further processing
selection which selects a subset of records
rename, which renames some fields to others
cartesian product (cross product) of two relations, with outputs being a list of all possible pairs of rows from both inputs
set operations of union, difference, and intersection

Note that all joins, intersection, and difference operators are technically implemented as outer joins with a post-processing step.

For convenience, we also provide lambda or map operation that computes a new field (or fields) from the others (logically redundant, but provided as convenience).

Translation of operations from relational algebra and SQL to data space algebra.

Relational algebra	SQL	Data space algebra
Projection	`SELECT field FROM ...`	Extraction
Selection	`WHERE ...` or `HAVING ...`	Partition
Rename	`WHERE ...` or `HAVING ...`	Mapping function
Cross product	`SELECT ... FROM table1, table2`	Disjoint tensor product
Outer join	`OUTER JOIN`	Partition induced by NULL on each table, then parallel tensor product
Natural join	`INNER JOIN`	Parallel tensor product on partition induced by non-NULL for each table
Union	`UNION`	Disjoint tensor product (tagged union), then partition removing duplicates, and mapping removing tags
Multiset union	`UNION ALL`	Tagged union, then information-preserving function that removes tags
Difference	`MINUS`	Partition on the first table that puts shared elements in the error set
Intersection	`INTERSECT`	Partition on each table that puts shared element in the output set

Using these operations, we can also implement joins Additionally, there are joins and aggregation operations provided.

Joins are operations that take two relations and produce rows that are contains bigger rows that are merged from both inputs.

We also add the aggregate operations:

- groupings which produce a relation with records that are sub-relations with fields from the same subgroup
- aggregation operations which take any monoid operation over a given field and produce rows that are monoid sum over entirety of records in a relation

The set of operations above is widely acknowledged as sufficient to provide most of the ETL and data aggregation needs. Whenever a new database or data analytics operation is proposed, it is customary to compare it to basic relational algebra operations. Thus we will propose a set of operations that allow us to get the same results as relational algebra expressions.

5 DISCUSSION

5.1 Error aggregates vs error estimates

The use of error-tracing relational algebra allows us to easily make general assertions on how the individual records impact processing (since each total summarisation must be sensitive to all records). When we expect a given record to be 100% correct, we just assert its fields are included in a correct record set at the end of the correct path.

If we expect a given record to be discarded, we may just assert that it is present in error set.

That also means that we may use a failing assertion to quickly point to all possible code sites or data inputs that may need to be changed in order to fix the issue.

5.2 Replacing relational algebra in analytics

We proposed a replacement of relational algebra with a set of operations that allow for the same computation, but structured in a way that assure all data is summarised and included. We justify it by analogy to accounting theory, and thus coin the name "data accounting" for the process of complete summarisation of a data set in order to assure conformance to statistical and analytical best practices. We described the process of translation of classical relational algebra into this new "data space algebra". This algebra allows for a reversible transformation of data and the preservation of mathematical principles of accounting while aggregating the data. Implementation of "data space algebra" operations on top of traditional SQL is shortly described.

5.3 Limitations

Since the use of data space algebra requires us to use information-preserving functions (bidirectional transformations), processing a large amount of data may lead to overheads. This disadvantage is usually trumped by the much higher reliability of the resulting analyses. Even in case some data needs to be discarded, like in the case of recording events in CERN at rates beyond TB/s, we should use the early aggregation of discarded data to still preserve some dimensions of the discarded part of data space.

Another limitation is the necessity to change traditional thinking in analytics from using separate, error-prone transformations to constructing a complete summarisation of the dataset (a dashboard) that simultaneously takes care of all relevant dimensions and remains sensitive to all inputs. This intellectual barrier may be surpassed by placing emphasis on the completeness and correctness of data analytics pipeline instead of quick reward of partial results.

6 CONCLUSION

We also suggest a common summarisation mechanism for all well-defined columns within the error table to facilitate the finding of errors.

We provide data space algebra as an alternative to conventional relational algebra for the purpose of more robust reporting, and suggest using accounting principles for more disciplined data analysis. We briefly described how to convert traditional relational algebra expressions so that there one of the outputs gives the same information as a relational query. Naturally, for a full accounting of input data, we also require additional information to be summarised in order to ensure that all data is accounted for.

BIBLIOGRAPHY

Abadi, Martin, and Gordon D. Plotkin. 2019. "A Simple Differentiable Programming Language." In *Proceedings of the ACM on Programming Languages* (POPL), 4. https://doi.org/10.1145/3371106.

Arntzenius, Frank. 2004. "Gunk, Topology and Measure." In *Oxford Studies in Metaphysics: Volume 4*, edited by Dean Zimmerman. Oxford University Press. https://link.springer.com/chapter/10.1007/978-94-007-0214-1_16

Balzer, Wolfgang, and Richard Mattessich. 1991. "An Axiomatic Basis of Accounting: A Structuralist Reconstruction." *Theory and Decision* 30 (3): 213–43.

Bancilhon, François, and Nicolas Spyratos. 1985. "Algebraic versus Probabilistic Independence in Data Bases." In *Proceedings of the Fourth ACM SIGACT-SIGMOD Symposium on Principles of Database Systems*, March 25–27, 1985, Portland, Oregon, USA, 149–53. ACM. https://doi.org/10.1145/325405.325424.

Barbieri, Giuseppina, and Giangiacomo Gerla. 2021. "Defining Measures in a Mereological Space (an Exploratory Paper)." *Logic and Logical Philosophy* 31 (1): 57–74. https://doi.org/10.12775/LLP.2021.005.

Basel Committee on Banking Supervision. 2013. "BCBS 239: Principles for Effective Risk Aggregation and Risk Reporting." https://www.bis.org/publ/bcbs239.htm.

Bennett, C. H. 1973. "Logical Reversibility of Computation." *IBM Journal of Research and Development* 17 (6): 525–32. https://doi.org/10.1147/rd.176.0525.

Chen, Haitao, and Husheng Liao. 2011. "A Survey to View Update Problem." *International Journal of Computer Theory and Engineering* 3: 23–31.

Ellerman, David. 2007. "Double-Entry Accounting: The Mathematical Formulation and Generalization." *SSRN Electronic Journal*, December. https://doi.org/10.2139/ssrn.1340619.

Fong, Brendan, and David I. Spivak. 2021. *Symmetric Monoidal Preorders*. Massachusetts Institute of Technology.

Gajda, Michal Jan. 2020a. "Agile Data Pipelines: ETL for 2020." Data Science Europe. https://youtu.be/aHAc8ght9Gw.

———. 2020b. "Agile Generation of Cloud API Bindings with Haskell." Haskell.Love. https://skillsmatter.com/skillscasts/14905-agile-functional-data-pipeline-in-haskell-a-case-study-of-multicloud-api-binding.

———. 2022. "Towards a More Perfect Union Type."

Ijiri, Yuji. 1975. *Theory of Accounting Measurement*. American Accounting Association, 10. https://www.worldcat.org/title/theory-of-accounting-measurement/oclc/1384293

———. 2018. "Axioms and Structures of Conventional Accounting Measurement." *Accounting, Economics, and Law: A Convivium* 8 (1): 20170057. https://doi.org/https://doi.org/10.1515/ael-2017-0057.

Julian, Mike. 2017. "Practical Monitoring."

Levitt, Michael. 1983. "Molecular Dynamics of Native Protein: I. Computer Simulation of Trajectories." *Journal of Molecular Biology* 168 (3): 595–617. https://doi.org/10.1016/S0022-2836(83)80304-0.

Ross, Kenneth A., and Yehoshua Sagiv. 1997. "Monotonic Aggregation in Deductive Databases." *Journal of Computer and System Sciences* 54 (1): 79–97. https://doi.org/10.1006/jcss.1997.1453.

Warsono, Sony, Muhammad Ridha, and Arif Darmawan. 2009. "Mathematics in Accounting as a Big Unanswered Question." *SSRN Electronic Journal*, July. https://doi.org/10.2139/ssrn.1439084.

Wlaschin, Scott. 2012. "F# for Fun and Profit." 2012. https://swlaschin.gitbooks.io/fsharpforfunandprofit/content/posts/recipe-part2.html.

Translation of shared concepts within the literature of different domains.

Accounting	**Double-entry accounting**	**General ledger**	**N/A**	**Sensitive to inputs**
Bidirectional computing	Reversible transformation		Bidirectional computation	Bijective transformation
Databases	Mark for deletion, data partition	Monotonic query	View update problem	Injectivity
Event sourcing	N/A	Event log	N/A	N/A
Data structures	Partition	Append-only or log-structured	N/A	Not invariant, injectivity
Railroad-oriented programming	Switching between happy path and error path	N/A	N/A	N/A

1. *Of course,* $''_{\mathbb{D}}$ *is a monoidal preorder (Fong and Spivak, 2021).*
2. *To yield constant value when summed over any partition of carrier set* $\mathbb{D}$.
3. *Named by analogy to relevant logic that disallows discarding premises from the input.*
4. *In which case, there is usually an additional sorting operation. And extracting a subset of unique, distinct records.*

7 A deep reinforcement learning approach for portfolio optimization and risk management – Case studies

Filip Wójcik

1 INTRODUCTION

Portfolio optimization is a critical tool for modern investors willing to maximize their capital utilization in the market and expected returns. As the number of decision factors and valuable data (produced by multiple automated systems) increases, it becomes more challenging to include such non-standard predictors in the optimization process.

Many portfolio optimization techniques are based on different variants of constrained quadratic programming formalizations, taking as inputs expected asset returns, forms of risk metrics (like covariance matrices), and transaction costs. Next, the goal is selected (typically a maximization of return while keeping the risk level constant or minimizing the risk with varying returns), and the problem is solved either via a closed-form equation or gradient minimizing approach. Additional elements or steps can be included, like periodical rebalancing or the inclusion of transaction costs. However, such an approach can be limiting if an analyst would like to consider more predictive factors or extend the problem beyond a formalized structure.

Recent studies on DRL revealed that this family of algorithms could achieve satisfying performance results in complex, volatile, and stochastic environments. DRL originated from "classic" reinforcement learning (RL) that utilizes tabular calculation or function approximation methods. Portfolio optimization can be perceived as a Multi-Stage Decision Problem, which directly corresponds with the primary goal for DRL utilization.

This work presents a literature review of selected portfolio optimization techniques that use deep reinforcement learning algorithms. Approaches based on this technique are compared to classic methods, highlighting the benefits and drawbacks

DOI: 10.1201/9780367854690-7

of each. The second part of the study presents experiments conducted on stock simulators and different assets in challenging market conditions after the COVID-19 outbreak and following pandemic events. DRL agents were compared with benchmarks (stock market indices and classic optimization strategies) in terms of Sharpe ratio, cumulative returns, and annual volatility, taking as their input multiple non-standard factors, like several technical indicators and summary statistics, beyond a typical quadratic programming setting.

Presented DRL algorithms outperformed benchmarks in terms of cumulative returns and Share ratio, demonstrating the ability to optimize example portfolios properly, even in very demanding market conditions.

Experimental results confirm the hypothesis that DRL is a promising set of tools for complex, stochastic, multi-stage decision problems, with various additional, potentially valuable information available.

The contribution of this study is two-fold. Firstly, it reviews and summarizes selected recent attempts to utilize deep reinforcement learning (DRL) algorithms for portfolio management and stock trading. Secondly, systematic experiments using a simulated stock market environment were conducted to empirically verify DRL's theoretical properties and assess the predictions' quality in challenging post-COVID-19 conditions.

This study is organized as follows:

Part 2 Literature review – briefly summarizes theoretical findings and existing literature in the following contexts:
- 2.1 Overview of portfolio optimization.
- 2.2 Reinforcement learning.
- 2.3 Deep reinforcement learning.
- 2.4 Reinforcement learning in Economics.

Part 3 Proposed work – presents the experimental goal, setup and results.

Part 4 Results and discussion – contains the result analysis, perspectives and possible next steps.

Part 5 Conclusions – concludes the whole study and summarizes all findings.

2 LITERATURE REVIEW

2.1 Overview of portfolio optimization methods

Portfolio optimization is one of the instrumental tools for modern investors operating in the financial markets. This area has a long history of research and study, lasting more than sixty years now (Kolm et al., 2014). The goal of this section is to review the essential concepts and introduce the fundamental mathematical apparatus that will be required to understand the experimental studies. There exist many systematic reviews and surveys of methods developed since 1952 (Fabozzi et al., 2007; Kalayci et al., 2019; Markowitz, 2014; Rubinstein, 2002; Steinbach, 2001; Zanjirdar, 2020), which can now be counted in hundreds (e.g., 34,100 research papers submitted on portfolio allocation in years 1998–2018 and 175 specifically about deterministic the Mean–Variance Optimization approach (Kalayci et al., 2019)). The focus of this work is to review the essential concepts and compare some of the existing methods with innovative Deep Reinforcement-Learning models.

2.1.1 Financial time series

Input to any portfolio optimization algorithm is a matrix of asset prices, where an i-th single asset price p is a time-indexed vector of non-negative real numbers, formalized as (Dees & Sidier, 2019; Feng et al., 2016):

$$\vec{\mathbf{p}}_{i,1:T} = \begin{bmatrix} p_{i,1} \\ p_{i,2} \\ \vdots \\ p_{i,T} \end{bmatrix} \in \mathbb{R}_+^T \tag{7.1}$$

Therefore, for multiple assets $S_1, S_2, \dots S_m$, the price matrix $\vec{\mathbf{P}}$ is formalized as:

$$\vec{\mathbf{P}}_{1:T} = \begin{bmatrix} p_{1,1} & p_{2,1} & \cdots & p_{M,1} \\ p_{1,2} & p_{2,2} & \cdots & p_{M,2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1,T} & p_{2,T} & \cdots & p_{M,T} \end{bmatrix} \in \mathbb{R}_+^{T\times M} \tag{7.2}$$

Raw prices can be volatile and unstable; therefore, most of the time, investors operate on returns (Dees & Sidier, 2019; Kalayci et al., 2019) defined either as gross returns (differenced prices) or simple returns (equivalent to the percentage change in asset price). Formally, asset returns can be presented as (Dees & Sidier, 2019; Kennedy, 2016):

1. Gross returns:

$$R_t \doteq \frac{p_t}{p_{t-1}} \in \mathbb{R} \tag{7.3}$$

2. Simple returns (equivalent to percentage price change)

$$r_t \doteq \frac{p_t - p_{t-1}}{p_{t-1}} = \frac{p_t}{p_{t-1}} - 1 = R_t - 1 \in \mathbb{R} \tag{7.4}$$

3. Log returns to avoid numerical overflows

$$\rho_t \doteq \ln\left(\frac{p_t}{p_{t-1}}\right) = \ln\left(R_t\right) \in \mathbb{R} \tag{7.5}$$

A matrix of N asset returns forms a returns matrix, where each row represents a time step, and each column represents a time series of asset returns (Capinski & Zastawniak, 2011; Dees & Sidier, 2019):

$$\vec{R}_{1:T} = \begin{bmatrix} r_{1,1} & r_{2,1} & \cdots & r_{M,1} \\ r_{1,2} & r_{2,2} & \cdots & r_{M,2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{1,T} & r_{2,T} & \cdots & r_{M,T} \end{bmatrix} \in \mathbb{R}_{+}^{T\times M} \tag{7.6}$$

In this context, a portfolio is defined as a set of asset weights vector such that (Kolm et al., 2014):

$$\omega = \left[\omega_1, \omega_2, \ldots, \omega_N\right]^T, \sum_{i=1}^{N} \omega_i = 1 \tag{7.7}$$

where each ω_i is a weight of i-th asset in portfolio.

Simple portfolio return at time t can be then defined as (Dees & Sidier, 2019; Kolm et al., 2014):

$$r_{t,p}\left(\omega\right) = \sum_{i=1}^{M} \omega_{i,t} r_{i,t} = \omega^T r_t \in \mathbb{R} \tag{7.8}$$

Portfolio return matrix (7.6) multiplied by portfolio weights matrix (7.7) over multiple time steps results in overall portfolio returns time series formalized as (Dees & Sidier, 2019):

$$r_{1:T} = \vec{R}_{1:T}\omega_{1:T} \in \mathbb{R}^T \tag{7.7}$$

An essential concept in later methods for portfolio optimization is an asset returns covariance matrix $\bullet$, where σ_i defines the standard deviation of r_i. Formally (Capinski & Zastawniak, 2011; Kolm et al., 2014):

$$\Sigma = \begin{bmatrix} \sigma_{1,1} & \sigma_{2,1} & \cdots & \sigma_{M,1} \\ \sigma_{1,2} & \sigma_{2,2} & \cdots & \sigma_{M,2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1,T} & \sigma_{2,T} & \cdots & \sigma_{M,T} \end{bmatrix} \tag{7.10}$$

where $\sigma_{ii} = \sigma^2$ (variance of j-th asset).

2.1.2 Markowitz model

Given the elements defined in a previous section, the portfolio Mean–Variance optimization problem can be defined as a constrained linear model as follows (Capinski & Zastawniak, 2011; Kennedy, 2016; Kolm et al., 2014):

1. Formulation as maximization problem: maximize portfolio returns while keeping variance below maximal level σ^2_{max}

$$\max_{\omega\in\Omega} \omega^T\mu$$

Subject to

$$\omega^T \sum \omega \le \sigma^2_{max} \tag{7.11}$$

$$\sum \omega = 1$$

2. Formulation as minimization problem: minimize portfolio variance while achieving at least minimal desired portfolio returns R_{min}

$$\min_{\omega\in\Omega} \omega^T \sum \omega$$

Subject to

$$\mu^T\omega \ge R_{min} \tag{7.12}$$

$$\sum \omega = 1$$

Where ∝ is a vector of expected securities return $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_M \end{bmatrix}$, $\mu_i = \mathbb{E}\left[r_i\right] \in \mathbb{R}$ and Ω is a universe of all possible portfolio weights combinations.

If portfolio weights have negative weights, this is interpreted as allowance for short-selling (Capinski & Zastawniak, 2011; Steinbach, 2001). If short-selling is not allowed (or an analyst does not want to use this possibility), then the optimization problem is defined as (Dees & Sidier, 2019; Gill et al., 1981):

$$\underset{\omega\in\Omega}{\text{minimize}}\ \omega^T\omega$$

Subject to

$$\mu^T\omega = \ge R_{min} \tag{7.13}$$

$$\omega \ge 0$$

$$\sum \omega = 1$$

2.1.3 Markowitz model extensions

One of the typically used modifications to the baseline model is the "risk aversion" parameter, defining the degree to which the investor is willing to balance the risk level $\omega^T \sum \omega$ with expected returns $\omega^T \mu$ (Wilmott, 2007). With this aversion parameter o, the model is formalized as the following optimization problem (Dees & Sidier, 2019; Kolm et al., 2014):

$$\begin{aligned} &\underset{\omega \in \Omega}{\text{maximize}}\ \omega^T \mu - \alpha \omega^T \sum \omega \\ &\text{Subject to} \\ &\omega \geq 0 \\ &\sum \omega = 1 \end{aligned} \tag{7.14}$$

Additionally, several objective functions can be selected as the optimization goal. One of the most interpretable ones is the utilization of the Sharpe Ratio (Sharpe, 1966, 1994), defined as the reward-to-variability ratio. Formally (Dees & Sidier, 2019; Goetzmann et al., 2014):

$$SR = \frac{\mu - R_f}{\sigma_p} \tag{7.15}$$

Where $\propto$ is expected assets returns (as defined above), R_f is risk-free rate (often set to zero) and σ_p is a standard deviation of portfolio returns $\sqrt{\omega^T \sum \omega}$.

When used as the objective function in the Markowitz model, the optimization problem is formalized as (Dees & Sidier, 2019; Wilmott, 2007):

$$\begin{aligned} &\underset{\omega \in \Omega}{\text{maximize}} \frac{\omega^T \mu}{\sqrt{\omega^T \sum \omega}} \\ &\text{Subject to} \\ &\omega \geq 0 \\ &\sum \omega = 1 \end{aligned} \tag{7.16}$$

A popular and very easy-to-understand risk measure is the Calmar Ratio (or Drawdown ratio), which utilizes a maximum drawdown instead of standard deviation to measure portfolio instability (Young, 1991). *Drawdown* in this context is defined as a (potential) loss that can happen over a defined period (Bacon, 2013). It can be calculated as a relative change between the maximum value of the portfolio returns and the selected time point afterward. Formally (Drenovak et al., 2021):

$$DD(t) = \frac{V_{p,t}}{\max_s V_{p,s}}, \; s \in [0, t] \tag{7.17}$$

Where $V_{p,t}$ is the portfolio value at time t.

Therefore a Calmar ratio can be calculated as (Bacon, 2013):

$$\text{Calmar ratio CR} = \frac{\mu - R_f}{DD_{max}} \tag{7.18}$$

Another essential metric is Value at Risk (VAR), which is used as a risk benchmark. It measures the expected maximal loss in a target time horizon, with a defined confidence level, under typical market conditions (Larsen et al., 2002; Zanjirdar, 2020). Formally it can be written as (Zanjirdar, 2020):

$$P(P_0 - P_1 \geq VaR) \leq \alpha \tag{7.19}$$

where P_0 is an initial portfolio value, P_1 is the portfolio value at the end of the time horizon, VaR is the deprecation value, and α is the statistical confidence level. While very useful from the business perspective, this measure lacks some essential properties like sub-additivity, and is non-convex, non-smooth, and can have multiple extremes when applied to a finite number of scenarios (Larsen et al., 2002). For this reason, it is often optimized using heuristic approaches (Larsen et al., 2002; Nguyen et al., 2021; Zanjirdar, 2020) or gradient optimization method (Hiraoka et al., 2019).

Typically, operations on the stock market incur some transaction costs. The inclusion of such a variable in the model makes it impossible to solve using closed-form optimization procedures and requires the usage of various heuristic methods (Brown & Smith, 2011; Peng et al., 2011). With transaction costs included, the optimization problem for M assets is formulated as (Dees & Sidier, 2019):

$$\begin{aligned} &\underset{\omega \in \Omega}{\text{maximize}} \; \mathcal{J} - 1_M^T \beta \left\| \omega_0 - \omega \right\|_1 \\ &\text{Subject to} \\ &\omega \geq 0 \\ &\sum \omega = 1 \end{aligned} \tag{7.20}$$

where $\mathcal{J}$ can be any objective function, like the Sharpe Ratio (7.15), (7.16), β is a transactions cost parameter (that can take different forms – numerical, or percentage), ω_0 is the vector of initial portfolio weights.

In practice, problems, including transaction costs, are approached using complex metaheuristics. Some examples include:

1. particle-swarm optimization PSO (Chen & Zhang, 2010; Coello et al., 2004);
2. different forms of evolutionary algorithms (Deb et al., 2002; Liagkouras, 2019; Rong et al., 2009);
3. fuzzy algorithms (Liagkouras & Metaxiotis, 2018; Liu & Zhang, 2015);
4. deep learning (Aboussalah & Lee, 2020).

The more constraints and elements are added to the portfolio optimization problem, the more complicated it becomes and can be (potentially) unsolvable with classic methods (Neuneier, 1996). In that sense, the portfolio optimization problem with transaction costs becomes a "Multi-Stage Decision Problem" (Neuneier, 1996), where the later steps in the optimization process directly depend on the previous steps and can lead to different results. Such a formulation makes the portfolio optimization task suitable for Markov Decision Processes (MDP) modeling (Neuneier, 1996; Sutton & Barto, 2018) and, in conclusion, for Reinforcement Learning (RL), algorithms described in later sections.

Comprehensive reviews of recent literature reveal multiple additional constraints and bounds that can be added to the Markowitz model to obtain the best-performing portfolios. Some of the possibilities include (Kalayci et al., 2019; Kolm et al., 2014; Markowitz, 2014; Steinbach, 2001):

1. Lower/upper limits can be imposed on the asset weights in the portfolio.
2. Limitation on the number of assets in the portfolio.
3. Limitations connected to the market sector to which particular assets belong.
4. Limitations based on the turnover rate for each asset typically utilized in multi-period portfolio optimization.
5. Short-sale constraints, imposing the weights to be non-negative.

All of the methods mentioned above are ways to enrich the optimization task without changing its very core – a foundation laid by the basic Markowitz Mean–Variance model.

2.2 Reinforcement learning

Reinforcement learning (RL) is one of the approaches to solving a control problem, where an artificial entity (an agent) interacts with the environment to achieve some predefined goal (Goodfellow et al., 2016). As such, it can be considered one of the subfields of machine learning, concentrated on Markov Decision Processes (MDP). Typically it consists of the following main elements:

1. An agent is an autonomous entity that aims to learn how to control and operate in the environment (Dees & Sidier, 2019; Goodfellow et al., 2016). Typically, the role of an agent is performed by a selected algorithm or procedure.

2. An environment is a simulated "world" with which an agent can interact and operate. This simulation should have some internal state (so the description of the current status), possible actions, and *transitions dynamics*, so possibilities to move from one state to another after performing a selected action (Sutton & Barto, 2018). In theory, the environment suitable for RL modeling should possess the so-called Markov Property, when the current state depends only on the previous one and carries sufficient information to describe the present fully (Li, 2017).
3. A set of rewards is assigned each time an agent acts in the environment. The reward can be positive (indicating the correct, beneficial move) or negative (as a penalty for wrong decisions) (Arulkumaran et al., 2017). The definition of the reward, its scale, and magnitude is subject to so-called *reward engineering* – a process of careful design and selection based on expert knowledge and observation of a simulated environment (Dewey, 2014).

Formally, the dynamics of agent-environment interactions are defined as follows (Arulkumaran et al., 2017; Li, 2017; Sutton & Barto, 2018):

1. $\mathcal{S} = \{s_0, s_1, \ldots s_m\}$ – is a set of states in which an environment can be.
2. $\mathcal{A}$ – is a set of actions that an agent can perform in the environment.
3. r_t – is a reward obtained at time step t after performing the action.
4. $\mathcal{T}(s_{t+1}, r | s_t, a_t)$ – is a transition dynamics, an internal environment property that describes (possibly non-deterministic) state changes after performing an action.

While the agent interacts with the environment, it leaves a *trace* – a sequence of quadruplets (state, action, reward, next state – (s_t, a_t, r_t, s_{t+1}) that are a subject for the later learning process (Fenjiro & Benbrahim, 2018; Mousavi et al., 2018).

The interaction between the agent and the environment is presented in the picture below (Figure 7.1):

If the environment has a well-defined terminal state – the state at which no further actions are allowed, then a control problem in such an environment is called *an episodic task* (Sutton & Barto, 2018). Because in the beginning, an agent does not know either the transition dynamics or the future rewards that it will receive, it needs to have a possibility to relate past actions with future rewards. The goal defined in this way is called *discounted future reward* or return formalized as follows (Dees & Sidier, 2019):

$$G_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \ldots = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \tag{7.21}$$

The discounting factor gamma describes the emphasis the agent puts on the future rewards from the current (t) time step.

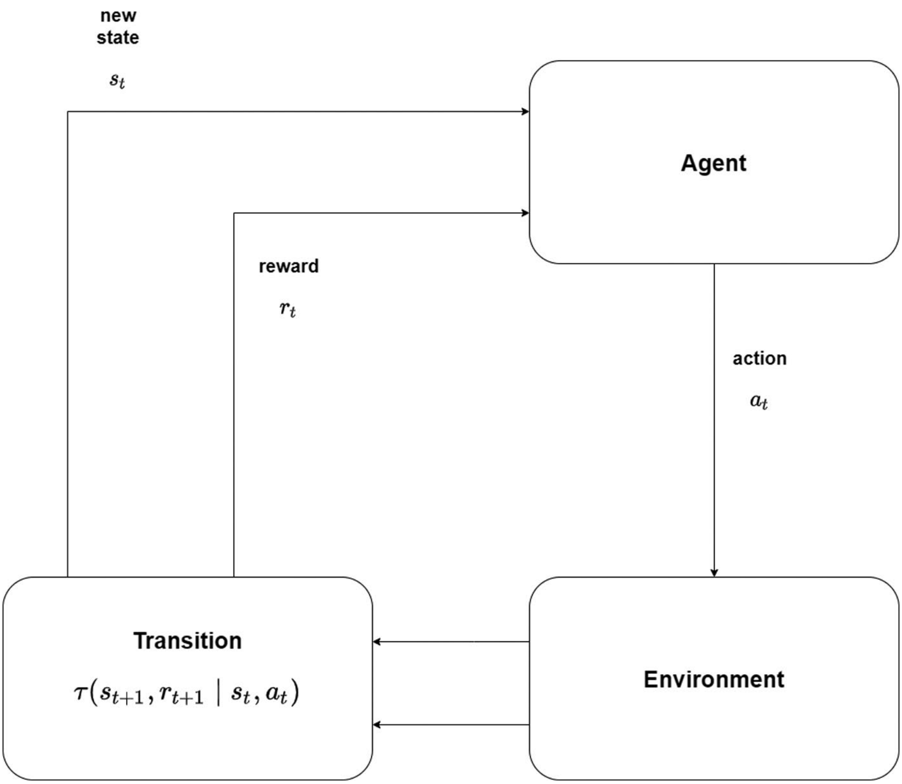


FIGURE 7.1 Interaction between an agent and the environment.

Source: own work based on (Dees & Sidier, 2019; Sutton & Barto, 2018).

The systematic way that describes how an agent acts in the environment is called a "policy" and is formalized as follows (Arulkumaran et al., 2017; Sutton & Barto, 2018):

$$\pi : \mathcal{S} \to \mathrm{p}(\mathcal{A} = \mathrm{a} \mid \mathcal{S}) \tag{7.22}$$

where

$\mathcal{A}$ – set of actions
$\mathcal{S}$ – set of states

It maps the encountered states into actions. Intuitively, a policy is a function that takes a state as an input and outputs the desired action. The optimal policy is the one that maximizes the expected, discounted return for the entire episode (Arulkumaran et al., 2017; Sutton & Barto, 2018):

$$\pi^* = \arg\max_{¿} \mathbb{E}\left[\mathrm{G}_t \middle| \pi \right] \tag{7.23}$$

For obvious reasons, an optimal policy is unknown at the time of the training, and it is the agent's ultimate goal to estimate its value or approximate as closely as possible.

In order to find the best possible policy, that agent utilizes additional functions, out of which two are fundamental for reinforcement learning. Jointly they are called "Bellman Equations" (Fenjiro & Benbrahim, 2018; Sutton & Barto, 2018) and they laid the foundation for most of the theoretical apparatus for future RL algorithms.

1. **State-value function**

   State-value function $V(s)$ estimates the expected future return if the agent follows policy pi from the present state s. Formally (Fenjiro & Benbrahim, 2018; Li, 2017; Sutton & Barto, 2018):

$$V^{\pi}(s) = \mathbb{E}^{\pi}\left[G_t | S_t = s\right] = \mathbb{E}^{\pi}\left[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \middle| S_t = s\right] \tag{7.24}$$

   Intuitively, a state-value function describes "how good" (from the perspective of accumulated rewards) a given state is. In the case of a stock market or portfolio optimization perspective – it might be defined as the profit and loss up to date or expected portfolio returns given particular asset weights (Xiong et al., 2018).

2. **Action-value function**

   Action-value function $Q(s, a)$ describes the expected future return if the agent follows policy pi from the present state s after execution of action a. Formally:

$$Q^{\pi}(s, a) = \mathbb{E}^{\pi}\left[G_t S_t = s, A_t = a\right] = \mathbb{E}^{\pi}\left[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \middle| S_t = s, A_t = a\right]. \tag{7.25}$$

   Its intuitive interpretation is analogous to the interpretation of the (7.24), but additionally considers the action being performed by an agent.

Those two equations can be expanded in a recursive form that emphasizes the probabilistic nature of the simulated environment dynamics and the policy itself (Fenjiro & Benbrahim, 2018; Sutton & Barto, 2018):

$$V^{\pi}(s) = \mathbb{E}^{\pi}\left[G_t | S_t = s\right] = \sum_{a} \pi(a|s) \sum_{s'} \sum_{r} p(s', r|s, a)\left[r + \gamma V^{\pi}(s')\right] \tag{7.26}$$

$$\begin{aligned} Q^{\pi}(s, a) &= \mathbb{E}^{\pi}\left[G_t | S_t = s, A_t = a\right] \\ &= \sum_{s'} \sum_{r} p(s', r|s_t, a_t)\left[r + \gamma \sum_{a'} \sum \pi(s', a') Q^{\pi}(s', a')\right] \end{aligned} \tag{7.27}$$

Bellman equations give theoretical guarantees to find optimal policies, defined by Bellman-optimality equations defined as (Fenjiro & Benbrahim, 2018; Sutton & Barto, 2018):

$$V^*(s) = \max_{\pi}\left[V_\pi(s)\right] = \max_{a \in \mathcal{A}(s)} \sum_{s',r} p(s', r|s, a)\left[r + \gamma V^*(s')\right]. \quad (7.28)$$

$$Q^*(s, a) = \sum_{s',r} p(s', r|s, a)\left[r + \gamma \max_{a'} Q^*(s', a')\right] \quad (7.29)$$

In general, RL algorithms can be divided into two main categories, summarized in the table below (Table 7.1):

Another important taxonomy emphasizes the approach of the RL agent to the learned policy.

1. **On-policy agents** utilize the policy that is learned. Therefore, they follow the same plan of executing gradually improved actions. Such algorithms tend to be more stable but require a longer learning time and are less sample efficient, as they can only utilize the newest samples.
2. **Off-policy agents** are not required to use the same policy being learned. They are more sample-efficient, utilizing samples from any learning phase in exchange for lower stability of results (larger variance).

TABLE 7.1

Comparison of model-based and model-free RL algorithms

Aspect\category	Model-free algorithms	Model-based algorithms
Completeness of the environment	Do not require the complete environment model – prior knowledge of all possible transitions, states, and rewards is not required.	Require a complete description of the operating environment – transitions, states, rewards, and possible actions.
Kind of problems solved	Suitable for complicated, real-world problems, where a simulation is just an approximation and simplified description of reality.	Suitable for technical problems, where strict rules of environment dynamics can be applied.
Algorithms' characteristics	Algorithms of this kind are universal and can work with almost any kind of simulator or problem.	Algorithms of this kind are designed for specific types of problems and dedicated simulators only.
Proliferation	More frequently used in practice.	Less frequently used.

Source: own work, based on (Agostinelli et al., 2018; Arulkumaran et al., 2017; Fenjiro & Benbrahim, 2018; Sutton & Barto, 2018; Trask, 2019; Winder, 2020).

2.3 Deep reinforcement learning

Methods described above are considered to be "classic RL" – utilizing relatively simple, grid-based algorithms, where enumeration of possible states and actions is fairly possible (Agostinelli et al., 2018). As the problems solved by RL became more complex, classic methods started to lack enough functionality in terms of:

1. **Handling high-dimensional spaces**, where states are described using long, complex vectors (Agostinelli et al., 2018; Busoniu et al., 2017).
2. **Handling continuous spaces** – when real numbers describe both states or actions instead of discrete categories (Santamaria et al., 1997).
3. **Stability and convergence** – when the number of iterations required to get stable results increases exponentially, along with the state dimensionality and actions complexity (Boyan & Moore, 1994).

A natural extension to the beforementioned RL approaches was utilizing neural networks to approximate the most complicated functions. The common name given to methods of building multi-layer neural networks is "deep learning" (DL), and deep reinforcement learning (DRL) is only one of its applications (Goodfellow et al., 2016). Neural networks are considered universal approximators, theoretically capable of modeling any function, given the proper training time and data (Cybenko, 1989; Hornik et al., 1989). This property, combined with very flexible architecture (possibility to accept state input vectors of different shapes and multiple inputs), made them especially suitable for solving RL problems (Agostinelli et al., 2018). Initially, two basic approaches were tested.

Value-based DRL – is an approach where a neural network replaces the state-action function Q(s,a) as an approximator (Trask, 2019; Winder, 2020). Q table is no longer needed, and the network computes state-action values for each state-action combination (Agostinelli et al., 2018). Several variations of this model were created – starting from simple Deep Q-networks (DQN) (Mnih et al., 2013), double deep q-networks (DDQN) (van Hasselt et al., 2016), where the over-optimistic learning is slowed down by using two Q-networks and dueling double deep q-networks (DDDQN), utilizing additional action-advantage function (Fujimoto et al., 2018; Wang et al., 2016).

The picture below shows a simplified, intuitive comparison between grid/table-based RL and Deep Q-Networks, where Q-table has been replaced by its approximation by the DL model (Figure 7.2).

Policy-based DRL – emphasizes the role of a policy (defined as in previous chapters) in RL. Policy-based algorithms do not try to search and estimate state and state-action values for all possible environmental conditions but focus on a strategy maximizing the expected reward. This feature makes them especially useful in high-dimensional and continuous action spaces, where agents have the possibility to experience only a small portion of available states (Agostinelli et al., 2018; Fenjiro & Benbrahim, 2018). Especially notable in this context is the policy gradient approach, where the probabilistic policy-calculation function can be used directly to calculate

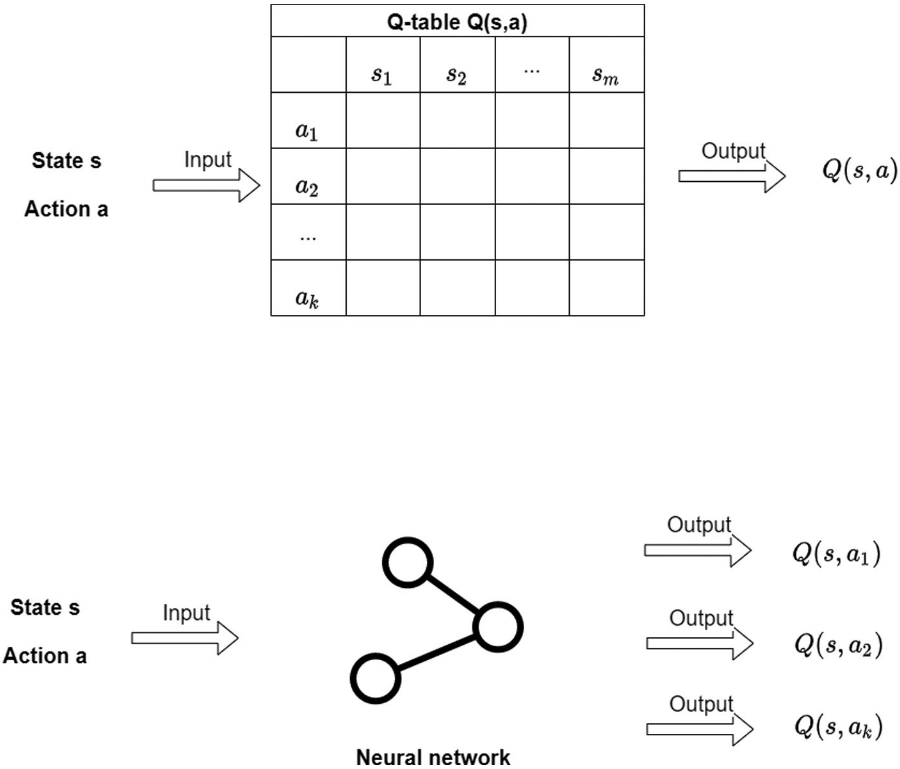


FIGURE 7.2 Simplified comparison between grid-based RL algorithms and DRL.

Source: own work.

neural network gradients during the learning procedure (Sutton et al., 1999). That laid the foundations for the basic REINFORCE algorithms family (Williams, 1992), followed by a Vanilla Policy Gradient (VPG) algorithm (Schulman, 2016; Sutton et al., 1999) and Proximal Policy Optimization (PPO) (Schulman et al., 2017).

One of the significant issues in utilizing policy-based DRL was its variance and instability. A specialized family of algorithms, called actor-critic (AC), was introduced to address this issue, especially in continuous action-space settings (Degris et al., 2012; Konda & Tsitsiklis, 1999) [Degris, 2012; Konda, 1999; Konda, 2003], and later on, it was extended with additional variations. In principle, in AC methods, an agent consists of two sub-estimators. “Critic” is responsible for calculating the value function (it can be state-value or action-value, depending on the algorithm), and the “actor” optimizes the policy using the critic’s approximation (Fenjiro & Benbrahim, 2018). May later implementations utilize this concept as a fundamental baseline, with the most notable being: A3C (Mnih et al., 2016) or A2C (Wu et al., 2017). The picture below shows the simplified concept of AC methods (Figure 7.3).

Comparison of value-based and policy-based DRL architectures is presented in the picture below (Figure 7.4).

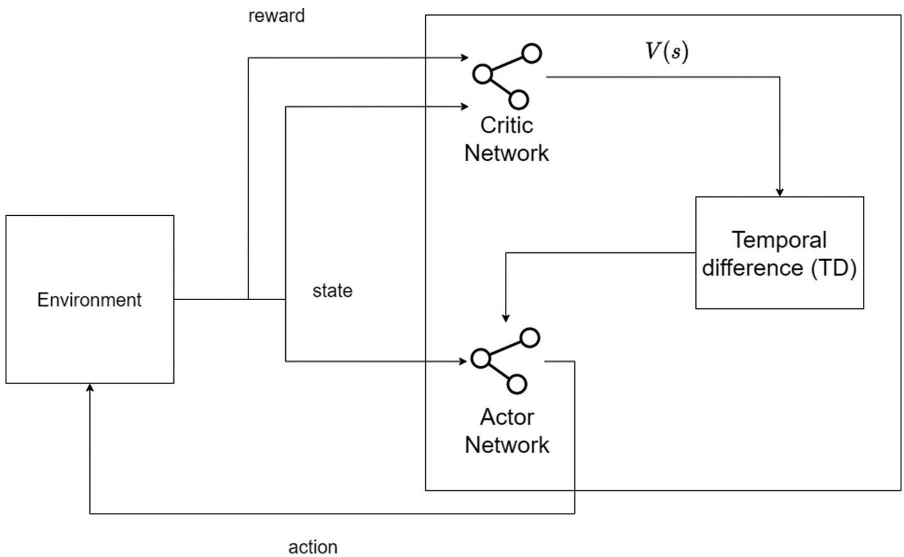


FIGURE 7.3 Actor-critic architecture overview.

Source: own work based on (Fenjiro & Benbrahim, 2018; Fujimoto et al., 2018; Konda & Tsitsiklis, 1999; Schulman, 2016).

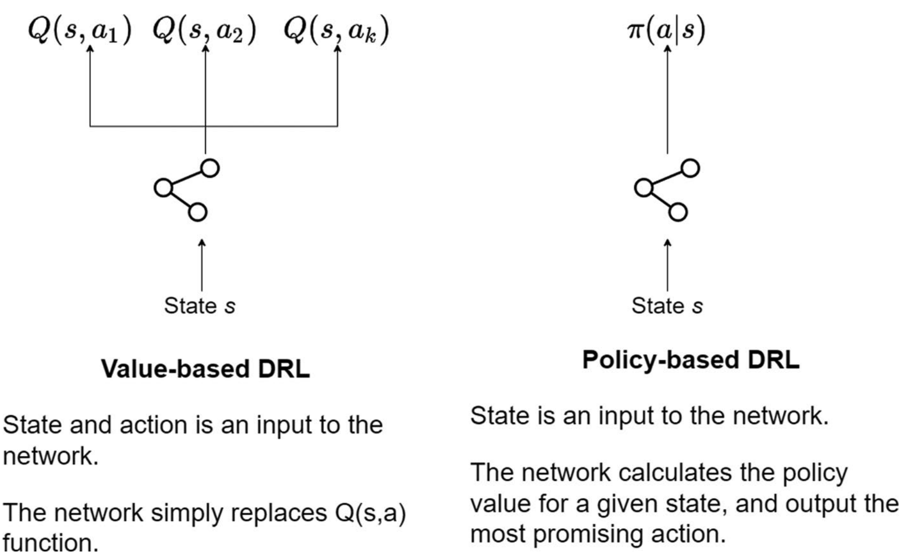


FIGURE 7.4 Value-based and policy-based DRL architecture comparison.

Source: own work, based on (Agostinelli et al., 2018; Sutton & Barto, 2018; Sutton et al., 1999).

Multiple hybrid approaches were invented on the crossing point between the beforementioned approaches. Most of them utilized some elements of on-policy learning combined with off-policy memory buffers. Most notable examples are:

1. **Deep Deterministic Policy Gradient (DDPG)** combines an actor-critic approach with two networks (one for predicting action-value function and one for computing targets during learning) (Silver et al., 2014). In that sense, the algorithm appears to be an extension of Q-learning, but action selection and exploration is much more similar to on-policy procedures.
2. **Twin Delayed Deep Deterministic Policy Gradient (TD3)** is a direct successor of DDPG with multiple additional functionalities that should stabilize results and ensure convergence. Most notable extensions are an additional ("twin") Q-function for estimating Q(s, a), deployed updates of the target network, and policy regularization (by noise injection) during the action selection (Fujimoto et al., 2018).
3. **Soft Actor-Critic (SAC)** is an algorithm published concurrently with TD3, considered another successor of DDPG, bridging the gap between value-based and policy-based DRL approaches. It uses techniques similar to those mentioned in TD3 to stabilize the learning process and online decision-making – like squashing policy values with *tanh* function and entropy-based target values regularization (Haarnoja et al., 2018a; Haarnoja et al., 2018b).

While the approaches mentioned above are modern and efficient implementations of DRL for complex problems, they all suffer from inconveniences, characteristic of both: Deep Learning and Reinforcement Learning (Ding & Dong, 2020). Most importantly:

1. The longer the learning horizon (parametrized by, e.g., n-steps in some algorithms or gamma discounting coefficient), the more unstable and prone to variance become the results. Thus, results might be unstable and hard to reproduce in subsequent studies.
2. As the learning horizon shortens and more emphasis is put on approximation (like in the temporal-difference (TD) (Sutton, 1988) or Generalized Advantage Estimation (GAE) approach (Schulman et al., 2016)) – the prediction bias becomes significant. Obtained results might be suboptimal and away from the target expected values.
3. Sampling inefficiency – as the state description grows in dimensionality or the number of available actions increases, DRL models also grow in size. Therefore, the number of required simulation iterations needs to be increased so that the agent can experience a large enough part of the environment. That might lead to very long and time-consuming learning procedures (Botvinick et al., 2019).
4. Feedback Sparsity and reward attribution – in some simulations, an agent might not receive any reward for a long time, which makes it unable to correlate taken actions with positive or negative feedback (Winder, 2020). Such periods of uncertainty affect the training process negatively and slow it down (Arumugam et al., 2021).

There exist some procedures to mitigate those risks, like one-shot imitation learning (Henderson et al., 2018), meta-learning (Finn et al., 2017), or better reproducibility by enforcing deterministic behavior (Nagarajan et al., 2018) although instability is still considered an issue in real DRL applications (Agostinelli et al., 2018; Dasagi et al., 2019).

2.4 Reinforcement learning in economics

Multiple research studies were published concerning DL, RL, and DRL in Economics (Charpentier et al., 2021; Mousavi et al., 2018). Most promising directions include using DRL agents to replace classic optimization systems, especially when the state description of actions space is beyond the comprehension of classic methods. A selection of studies and use cases is presented below, grouped by general category (Table 7.2).

TABLE 7.2
Review of selected research on DRL utilization in Economics

Category	Use case	Description	References
Sales and revenue optimization in E-commerce	Dynamic pricing	Price adjustment according to the current demand. In such context, RL agents operate in the environment described by customer purchasing behaviors, and they can take different actions to lower/increase prices in response to the demand.	Levina et al. (2009), Schwind (2007) and Zhao et al. (2018)
	Customer recommendation systems	State descriptions include past user behavioral data, item ratings, or sales figures, while actions include the possibility to recommend an item.	Kompan and Bieliková (2010) and Liu et al. (2018)
	Online advert bidding processes	A use case where agents take part in auctions to choose whose online advert will be presented to the end customer, potentially generating revenue and demand.	Cai et al. (2017), Even Dar et al. (2009), Schwind (2007), Zhang et al. (2014) and Zhao et al. (2018)
Operations research	Logistics and planning	DRL systems can help logistics and planning by solving generalized versions of the so-called Traveling Salesman Problem optimization and resource allocation.	Bello et al. (2019) and Deudon et al. (2018)
Portfolio allocation & trading		This family of use cases will be discussed in a separate section below.	

Source: own work.

2.5 Reinforcement learning in trading and portfolio optimization

Classic portfolio optimization procedures (some of which were discussed in Section 2.1), are challenging to implement when the analyst wants to consider additional factors, like transaction costs, market conditions, natural language analysis for sentiment, and similar factors (Mosavi et al., 2020). The application of DL and DRL to such cases has been a subject of intensive studies in recent years. One crucial advantage of DRL over classic RL and other optimization procedures is the fact that neural networks have a very flexible architecture that can utilize the data of mixed types: numerical vectors, textual, correlation matrices, and others at once (Dees & Sidier, 2019). The very design of the portfolio optimization problem – a sequence of actions resulting in changes in portfolio structure (number of stocks held) in response to the current market state (current stock prices and their dynamics) very closely resembles the Markov-Decision Processes of control and simulation that RL and DRL are naturally designed for (Charpentier et al., 2021; Dees & Sidier, 2019; Mosavi et al., 2020).

There is no single solution or the best algorithm to solve a portfolio optimization problem with DRL, as the model needs to be designed especially for every case. Different authors utilized various DRL algorithm families. The table below shows some of the approaches proposed in recent years, with a commentary presenting the most important findings (Table 7.3).

TABLE 7.3
Review of selected research on DRL utilization in trading and portfolio optimization

Publication	Findings & methods
Jin and El-Saawy (2016)	Utilization of an early Deep Q-learning approach for portfolio management using a stock market simulator.
Jiang et al. (2017)	Utilization of various deep reinforcement learning algorithms for cryptocurrency portfolio management.
Xiong et al. (2018)	Utilized a DDPG algorithm capable of handling vast, continuous action spaces, in this case: buy/sell orders for financial instruments.
Li et al. (2019a)	Developed an innovative Adaptive Deep Deterministic RL algorithm that was expected to mitigate the instability of predictions issues and was able to learn from both: positive and negative errors in forecasts.
Azhikodan et al. (2019)	Utilized recurrent neural network (RNN) architecture (capable of learning from the vector of historical and past events) and market news parsing with DRL agents to design a system taking actions in the environment based on stock price predictions.
Li et al. (2019b)	Combined various DRL agents with novel Adversarial Learning to increase the performance of trading agents.
Li et al. (2021), Liu et al. (2020), Liu et al. (2021a) and Liu et al. (2021b)	A series of publications concentrated on building a reliable (in terms of reproducibility, stability of results, ease of use, and transparency) market simulator, combined with various DRL agents and high scalability for complex environments.

Source: own work.

Market simulators are the central point and standard tools used in the mentioned research. While different authors use different toolsets, every RL or DRL model requires an interactive market environment. The main components of such a solution are:

1. Ability to fetch historical stock prices in any format – OHLC (open/high/low/close) or other, that captures the state of the market in a given time.
2. Pre-processing module that cleans and normalizes such data by filling in missing information, removing malformed entries, and parsing dates.
3. Feature calculation module that enriches historical row prices with additional information. It can be a set of technical indicators like correlation/covariance matrix or statistical features (like rolling mean, percentiles, standard deviation).
4. A simulator engine that responds to agent actions. Specifically:
   1. After acting, moves one step (typically a day) forward and reads new stock prices and technical indicators;
   2. Presents the new state description to the agent;
   3. Maintains the internal state – the account balance of the agent, current cash level, and stocks held;
   4. Responds to agent actions – the environment simulates a buy/sell/hold actions and considers transaction fees and additional constraints (like non-negative balance or possibly allowed short selling).

A diagram of a simplified, generic design of such a simulator is presented below (Figure 7.5).

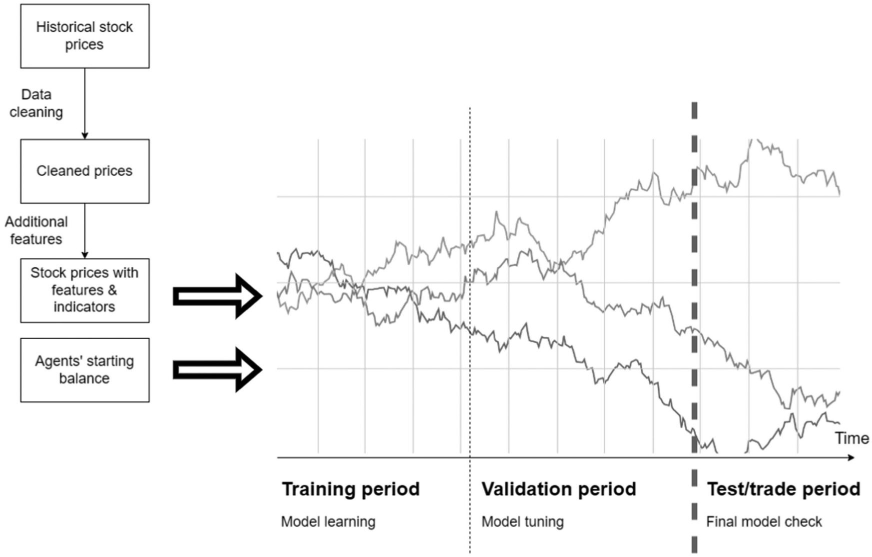


FIGURE 7.5 A simplified overview of a generic DRL stock simulator.

Source: own work.

While implementation details may vary across different research studies and software, the general idea of a simulated DRL market environment remains similar.

3 PROPOSED WORK

This experimental study aims to assess the performance of discussed DRL approaches in challenging market conditions. Subsequent sections show three practical case studies comparing the DRL agents, the standard market index benchmark and the portfolio optimized using the classic Markowitz model.

3.1 Research problem

The research problem in all three case studies has been formulated as follows:

> Is it possible to teach the DRL agent to construct a portfolio that maximizes the total return on a predefined time horizon that includes challenging market conditions (COVID-19 pandemic)?

The following research hypotheses were formulated:

1. The DRL agent will be able to learn the trading strategy during a training period and apply it during the testing period that it has not encountered before.
2. The DRL agent will be able to outperform the market benchmark index and the portfolio, optimized using the Markowitz approach, in terms of at least one of the following metrics: cumulative return, annual return, Sharpe ratio, Calmar ratio, or annual volatility.
3. DRL constructed portfolio should be readable by a human analyst, so its changes can be analyzed over time.

3.2 Experimental setting

The performance of DRL agents has been evaluated using the FinRL simulator (Li et al., 2021; Liu et al., 2020; Liu et al., 2021b). It downloads historical asset prices using Yahoo Finance, and on later stages, these prices are a subject for further rearrangement and enrichment according to the analysts' design. The simulator moves forward by one time-step (trading day) and presents actual stock prices to the agent. For DRL agents, the simulated environment has been configured as follows:

1. In each experimental case, the agent can trade using k predefined stocks (symbols).
2. On each time step (each trading day), the state consists of:
    1. Current balance information $b_t \in \mathbb{R}_+$ – agents current account balance.
    2. Shares vector $h \in \mathbb{Z}_+^k$ – number of k stock holdings in a portfolio.
    3. k stock prices in OHLC format (Open-High-Low-Close) $o_t, h_t, l_t, p_t \in \mathbb{R}_+^k$.
    4. Additional technical indicators for each stock. Especially:
    5. Moving average of stock prices $\mu \in \mathbb{R}^k$ – for 30 and 60 days.
    6. Moving Average Convergence-Divergence (MACD) $\mathrm{MACD} \in \mathbb{R}^k$ (Appel, 1979; Vaidya, 2020).

7. Relative Strength Index (RSI) rsi $\in \mathbb{R}^k$ (Wilder, 1978).
8. Commodity Channel Index (CCI) cci $\in \mathbb{R}^k$ (Lambert, 1983).
9. Bollinger Bands – upper and lower bb $\in \mathbb{R}^{2k}$ (Bollinger, 2002).

Action – for each stock (symbol) in the portfolio, the agent can perform the following actions: buy amount X/sell amount Y/hold, therefore increasing or decreasing the stock holding h by +X, −Y. Therefore $a \in \{-h, -h+1, -h+2, \dots, -1, 0, 1, \dots, h-2, h-1, h, hmax\}^k : h, hmax \in \mathbb{Z}_+$, where h indicates the current holding of each stock, and hmax is the global maximum of stocks to buy (constant number set in the simulator).

Reward function $r(s, a, s') \in \mathbb{R}$ is defined as the overall portfolio value (calculated as the dot product between the stock prices vector and holdings vector plus the account balance b, $p^T h + b$).

A simplified example of the current state, action, and transition is presented in Figure 7.6. The agent is presented with two stocks (AAPL and GOOG) and their technical indicators at time t. The current balance is 1000 USD, current holding is 0 shares of AAPL and ten shares of GOOG. The agent decides to buy five shares of AAPL and sell five shares of GOOG. This operation lowers the balance of the agent. The simulation moves to the next state (next trading day), and the reward is calculated using updated stock holdings times current prices plus the remaining balance (Figure 7.6).

Although the library is supposed to replicate the interaction between the algorithm and the actual stock market, it uses some simplifications and assumptions. Especially:

1. Simulated transaction costs are set on a constant level.
2. No market impact – agent's actions do not affect stock prices.
3. Full market liquidity – each buy–sell operation succeeds and is performed immediately.

	Returns		Prices		Technical indicators				Agent's portfolio		
	AAPL ret	GOOG ret	AAPL price	GOOG price	AAPL ma30	GOOG ma30	AAPL sharpe	GOOG sharpe	AAPL shares	GOOG shares	balance
Date											
2022-02-16	-0.001389	0.007784	172.550003	137.487503	-0.001178	-0.007186	-1.808746	-5.749348	0	10	1000

$a_t = \{5;\ -5\}$

$b_{t+1} = 1000 + [172.55 \quad 137.48]^T [-5 \quad 5] = 1000 - 175.31 = 824.68$

	AAPL ret	GOOG ret	AAPL price	GOOG price	AAPL ma30	GOOG ma30	AAPL sharpe	GOOG sharpe	AAPL shares	GOOG shares	balance
Date											
2022-02-17	-0.021269	-0.037669	168.880005	132.308502	-0.001000	-0.007315	-2.195546	-5.789996	5	5	824.68

$r(s_t, a_t, s_{t+1}) = p^T h + b = 1505.94 + 824.68 = 2330.622$

FIGURE 7.6 Example state – action – balance – next state – reward in financial simulator.

Source: own work.

DRL performance was checked using three simulated environments:

1. Trading with stocks of the top ten S&P companies in terms of market capitalization. Learning and tuning period 2015-01-01 until 2019-12-31.
2. Trading with two different sets of 30 companies' stocks, selected randomly from the top 150 S&P in terms of market capitalization. Learning and tuning period: 2010-01-01 until 2019-12-31.

The "trading period," where the agent makes decisions with full consequences and without a possibility to learn from mistakes, was set to 2020-01-01 until July 2022.

Companies with the largest market capitalization (either top ten or random 30 out of top 150) were selected because of their size, market impact, and fluent liquidity. These factors help mitigate the beforementioned simplifications and assumptions of the simulated environment.

Algorithms selected for the experiment were:

1. **PPO** (Proximal Policy Optimization) – an on-policy algorithm with an actor-critic variant, known for its improved stability due to the policy clipping function (Schulman et al., 2017).
2. **DDPG** (Double Deep Policy Gradient) – an off-policy algorithm used with an environment consisting of continuous state- and action-spaces (Lillicrap et al., 2016).
3. **TD3** (Twin Delayed DDPG) – an off-policy, direct successor to DDPG, extending it with an application of a double Q-learning approach and various learning stability improvements, like a delayed policy update or target smoothing (Fujimoto et al., 2018).

4 RESULTS AND DISCUSSION

The tables below summarize metrics obtained by all of the approaches in each testing case. TD3 algorithm achieved the best results in all experimental cases, outperforming other approaches in terms of all metrics except the annual volatility. The highest advantage has been observed in case 1 (top ten stocks of S&P 500), while in case 3 (random stocks), results were only slightly above those obtained using classic portfolio optimization.

In each case, the DDPG algorithm achieved the lowest annual portfolio volatility (Table 7.4).

Charts below present the visual comparison of the portfolio constructed by the best algorithm TD3 (highlighted as green "Backtest") with the market index benchmark (S&P, highlighted in grey as "Adj Close") and the classic portfolio optimization approach (highlighted in grey as "optimized portfolio") in each case (Figures 7.7–7.9). Careful analysis of both – visualizations and results table reveals the following:

1. In all cases, the best DRL model (TD3) outperformed other approaches in terms of final portfolio value, cumulative returns, annualized return, Shape ratio, and Calmar ratio.

2. The most significant advantage can be observed in cases 1 and 2, while in case 3, the TD3 results are slightly better than those obtained by the optimized portfolio.
3. In cases 1 and 3, volatility of the TD3 algorithm was substantially higher than for the optimized portfolio approach, while in case 2, it was comparable. In general, the TD3 approach had moderate volatility compared to other approaches.
4. Analysis of returns distribution reveals that in all cases DRL approach quickly recovered after the COVID-19 pandemic's initial period (March–May 2020), but suffered from losses after the Russia–Ukraine conflict eruption at the beginning and the mid-2022 period. This relationship is especially visible in case 3, where TD3 outperformed the optimized portfolio approach almost the entire time while losing the advantage at the beginning of 2022.

TABLE 7.4
Comparison of results obtained by different algorithms

	S&P	Optimized portfolio	PPO	DDPG	TD3
			Case 1 (top ten of S&P500)		
Final portfolio value	1,215,495,000	1,975,360,000	1,680,900,000	2,054,000,000	3,068,100,000
Cumulative returns	0.215	0.975	0.681	1.054	2.068
Annualized return	0.080	0.312	0.226	0.326	0.551
Sharpe ratio	0.427	0.954	0.935	0.960	1.287
Calmar ratio	0.235	0.902	1.038	0.822	1.588
Annual volatility	0.258	0.350	0.252	0.363	0.407
			Case 2 (random stocks)		
Final portfolio value	1,272,369,000	1,378,563,437	1,200,352,032	1,216,575,950	1,459,814,492
Cumulative returns	0.272	0.379	0.200	0.217	0.460
Annualized return	0.097	0.130	0.073	0.078	0.157
Sharpe ratio	0.491	0.579	0.432	0.424	0.721
Calmar ratio	0.286	0.336	0.243	0.229	0.526
Annual volatility	0.257	0.284	0.219	0.257	0.244
			Case 3 (random stocks)		
Final portfolio value	1,272,369,000	1,527,874,000	1,402,881,000	1,184,113,000	1,547,890,000
Cumulative returns	0.272	0.528	0.403	0.184	0.548
Annualized return	0.097	0.177	0.139	0.067	0.183
Sharpe ratio	0.491	0.714	0.722	0.388	0.740
Calmar ratio	0.286	0.448	0.467	0.208	0.516
Annual volatility	0.257	0.287	0.212	0.246	0.281

Source: own work.

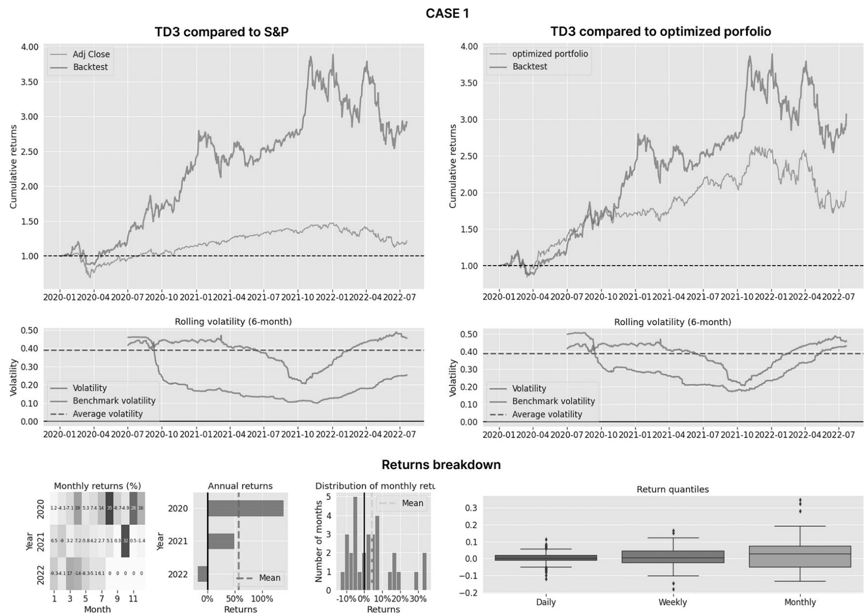


FIGURE 7.7 Results comparison for case 1.

Source: own work.

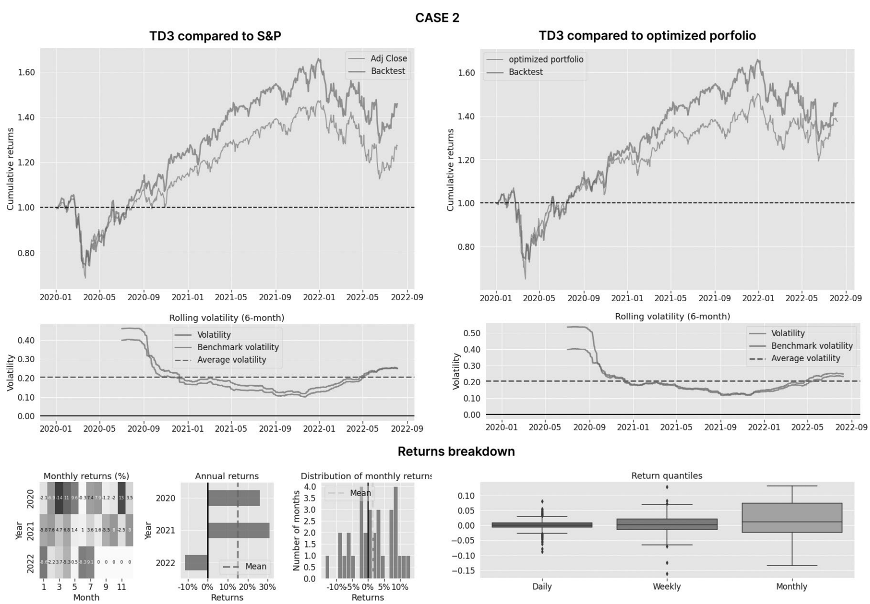


FIGURE 7.8 Results comparison for case 2.

Source: own work.

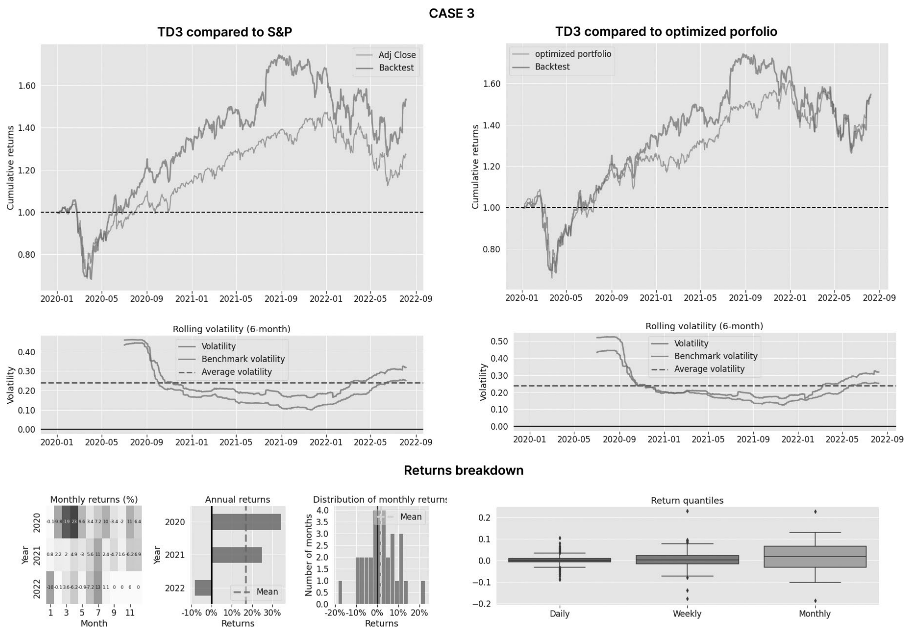


FIGURE 7.9 Results comparison for case 3.

Source: own work.

5. At the same time, TD3 outperformed the market index benchmark in every case, after the short recovery period, following the COVID-19 pandemic start.
6. Agent's portfolio over time takes a form of an activity table, recording the buy/sell/hold orders and the number of stocks held in each time step.

Presented results demonstrate the ability of DRL algorithms to achieve high results in terms of selected metrics using the complex market simulator. Models of this family react to market events and price changes, adjusting their behavior.

Therefore, all research hypotheses that were formulated in Section 3.1 have been confirmed.

At the same time, the experiment results reveal several properties of DRL algorithms that can be considered benefits or drawbacks when applied to portfolio optimization and trading use cases. Specifically:

1. All of the presented DRL can operate on a rich state representation. They were trained on a simulator feeding several technical indicators, current prices, and account information. Additional features could be added to this setting, namely:
   1. Sentiment analysis data from key market-focused new portals.
   2. Calendar information, focused on a time of the year period, working days, holidays, or important financial events (like regular bank status updates, press conferences, etc.).
   3. World geopolitical information about conflicts, wars, or random events.

2. DRL algorithms are highly stochastic and unstable, and moderately high portfolio volatility reflects this.
3. Training time for DRL is long – it can span several hours or minutes, depending on the machine specification. This property makes it inconvenient in some real-world situations.
4. DRL agents leave a “trace of activity” – buy/sell/hold orders and current portfolio composition. They do not produce any solutions to equations or clear explanations of their actions.

5 CONCLUSIONS AND FUTURE WORK

Experiments presented in this work confirmed research hypotheses that DRL could outperform market benchmarks and classic optimization strategies in some contexts, in terms of Sharpe ratio, Calmar ratio, annualized and total portfolio returns. Taking the properties of DRL algorithms described in previous sections into account, they can be perceived as valuable, additional tools for portfolio optimization, potentially very flexible, and capable of utilizing information beyond the reach of classic methods. At the same time, their learning process can be unstable and require significant computational power or time. Additionally, a lack of clear explanation or rationale behind their actions can make them unusable in some contexts, especially law-regulated.

RL and DRL are extensive research fields, and some of the shortcomings described in this work are constantly improved. Examples of recent advancements include the explainability of DRL (Guan & Liu, 2021; Heuillet et al., 2021; Vouros, 2022), reproducibility of results (Nagarajan et al., 2018), or stabilization of variance (Gao et al., 2019; Mao et al., 2019).

Additionally, this work’s future extensions might include utilizing a recurrent data format, where the DLR model processes a predefined time window of past market observations, most probably with some form of recurrent neural network (RNN) policy. Apart from using only market data (like prices and stock indicators), additional data for the model might include sentiment analysis from financial portals and similar natural language processing tools to judge the potential market impact. The flexible structure of neural networks makes such a setup possible.

LITERATURE

Aboussalah, A. M., & Lee, C. G. (2020). Continuous control with Stacked Deep Dynamic Recurrent Reinforcement Learning for portfolio optimization. *Expert Systems with Applications*, *140*. https://doi.org/10.1016/j.eswa.2019.112891.

Agostinelli, F., Hocquet, G., Singh, S., & Baldi, P. (2018). From reinforcement learning to deep reinforcement learning: An overview. In L. Rozonoer, B. Mirkin, & I. Muchnik (Eds.), *Braverman Readings in Machine Learning. Key Ideas from Inception to Current State* (pp. 298–328). Springer.

Appel, G. (1979). *The Stock Option and No-Load Switch Fund Scalper’s Manual*. Windsor Books.

Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep reinforcement learning: A brief survey. In *IEEE Signal Processing Magazine* (Vol. 34, Issue 6). https://doi.org/10.1109/MSP.2017.2743240.

Arumugam, D., Henderson, P., & Bacon, P.-L. (2021). An information-theoretic perspective on credit assignment in reinforcement learning. *ArXiv Preprint ArXiv:2103.06224*.

Azhikodan, A. R., Bhat, A. G. K., & Jadhav, M. v. (2019). Stock trading bot using deep reinforcement learning. In H. S. Saini, R. Sayal, A. Govardhan, & R. Buyya (Eds.), *Innovations in Computer Science and Engineering* (pp. 41–49). Springer.

Bacon, C. R. (2013). Practical risk-adjusted performance measurement. In *Practical Risk-Adjusted Performance Measurement*. https://doi.org/10.1002/9781118673621.

Bello, I., Pham, H., Le, Q. v., Norouzi, M., & Bengio, S. (2019). Neural combinatorial optimization with reinforcement learning. In *5th International Conference on Learning Representations, ICLR 2017 - Workshop Track Proceedings, Toulon, France*.

Bollinger, J. (2002). *Bollinger on Bollinger Bands*. McGraw-Hill.

Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., & Hassabis, D. (2019). Reinforcement learning, fast and slow. In *Trends in Cognitive Sciences* (Vol. 23, Issue 5). https://doi.org/10.1016/j.tics.2019.02.006.

Boyan, J., & Moore, A. (1994). Generalization in reinforcement learning: Safely approximating the value function. *Advances in Neural Information Processing Systems*, *7*, 369-376.

Brown, D. B., & Smith, J. E. (2011). Dynamic portfolio optimization with transaction costs: Heuristics and dual bounds. *Management Science*, *57*(10), 1752–1770.

Busoniu, L., Babuska, R., de Schutter, B., & Ernst, D. (2017). *Reinforcement Learning and Dynamic Programming Using Function Approximators*. CRC Press.

Cai, H., Ren, K., Zhang, W., Malialis, K., Wang, J., Yu, Y., & Guo, D. (2017). Real-time bidding by reinforcement learning in display advertising. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, Cambridge, UK,* (pp. 661–670).

Capinski, M., & Zastawniak, T. (2011). *Mathematics for Finance*. Springer.

Charpentier, A., Elie, R., & Remlinger, C. (2021). Reinforcement learning in economics and finance. *Computational Economics, 62(1)*, 425–462.

Chen, W., & Zhang, W. G. (2010). The admissible portfolio selection problem with transaction costs and an improved PSO algorithm. *Physica A: Statistical Mechanics and Its Applications*, *389*(10), 2070–2076. https://doi.org/10.1016/J.PHYSA.2010.01.016.

Coello, C. A. C., Pulido, G. T., & Lechuga, M. S. (2004). Handling multiple objectives with particle swarm optimization. *IEEE Transactions on Evolutionary Computation*, *8*(3), 256–279.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, *2*(4). https://doi.org/10.1007/BF02551274.

Dasagi, V., Bruce, J., Peynot, T., & Leitner, J. (2019). Ctrl-z: Recovering from instability in reinforcement learning. *ArXiv Preprint ArXiv:1910.03732*.

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, *6*(2), 182–197.

Dees, B. S., & Sidier, G. (2019). *Reinforcement Learning for Portfolio Management*. https://doi.org/10.48550/arxiv.1909.09571.

Degris, T., Pilarski, P. M., & Sutton, R. S. (2012). Model-free reinforcement learning with continuous action in practice. In *2012 American Control Conference (ACC), Montréal, Canada,* (pp. 2177–2182).

Deudon, M., Cournut, P., Lacoste, A., Adulyasak, Y., & Rousseau, L. M. (2018). Learning heuristics for the tsp by policy gradient. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10848 LNCS*. https://doi.org/10.1007/978-3-319-93031-2_12.

Dewey, D. (2014). Reinforcement learning and the reward engineering principle. In *2014 AAAI Spring Symposium Series, Palo Alto, USA*.

Ding, Z., & Dong, H. (2020). Challenges of reinforcement learning. In In H. Dong, Z. Ding, & S. Zhang (Eds.), *Deep Reinforcement Learning* (pp. 249–272). Springer.

Drenovak, M., Ranković, V., Urošević, B., & Jelic, R. (2021). Mean-maximum drawdown optimization of buy-and-hold portfolios using a multi-objective evolutionary algorithm. *Finance Research Letters*. https://doi.org/10.1016/j.frl.2021.102328.

Even Dar, E., Mirrokni, V. S., Muthukrishnan, S., Mansour, Y., & Nadav, U. (2009). Bid optimization for broad match Ad auctions. In *WWW'09 - Proceedings of the 18th International World Wide Web Conference*. https://doi.org/10.1145/1526709.1526741.

Fabozzi, F. J., Focardi, S. M., Kolm, P. N., & Pachamanova, D. A. (2007). *Robust Portfolio Optimization and Management*. John Wiley & Sons.

Feng, Y., Palomar, D. P., & others. (2016). A signal processing perspective on financial engineering. *Foundations and Trends® in Signal Processing*, *9*(1–2), 1–231.

Fenjiro, Y., & Benbrahim, H. (2018). Deep reinforcement learning overview of the state of the art. *Journal of Automation, Mobile Robotics and Intelligent Systems*, 12(3), 20–39.

Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. In *International Conference on Machine Learning, Sydney, Australia,* (pp. 1126–1135).

Fujimoto, S., Hoof, H., & Meger, D. (2018). Addressing function approximation error in actor-critic methods. In *International Conference on Machine Learning, Stockholm, Sweden,* (pp. 1587–1596).

Gao, Q., Hajinezhad, D., Zhang, Y., Kantaros, Y., & Zavlanos, M. M. (2019). Reduced variance deep reinforcement learning with temporal logic specifications. In *ICCPS 2019 - Proceedings of the 2019 ACM/IEEE International Conference on Cyber-Physical Systems*. https://doi.org/10.1145/3302509.3311053.

Gill, P. E., Murray, W., Saunders, M. A., & Wright, M. H. (1981). Aspects of mathematical modelling related to optimization. *Applied Mathematical Modelling*, *5*(2). https://doi.org/10.1016/0307-904X(81)90002-0.

Goetzmann, W. N., Brown, S. J., Gruber, M. J., & Elton, E. J. (2014). *Modern Portfolio Theory and Investment Analysis* (p. 237). John Wiley & Sons.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Guan, M., & Liu, X.-Y. (2021). Explainable deep reinforcement learning for portfolio management: An empirical approach. In *Proceedings of the Second ACM International Conference on AI in Finance, online,* (pp. 1–9).

Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018a). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *35th International Conference on Machine Learning, ICML 2018, Stockholm, Sweden,* 1856-1865.

Haarnoja, T., Zhou, A., Hartikainen, K., Tucker, G., Ha, S., Tan, J., Kumar, V., Zhu, H., Gupta, A., Abbeel, P., & Levine, S. (2018b). Soft actor-critic algorithms and applications. *ArXiv Preprint ArXiv:1812.05905*.

Henderson, P., Islam, R., Bachman, P., Pineau, J., Precup, D., & Meger, D. (2018). Deep reinforcement learning that matters. *Proceedings of the AAAI Conference on Artificial Intelligence*, *32*(1), 3207–3214.

Heuillet, A., Couthouis, F., & Díaz-Rodríguez, N. (2021). Explainability in deep reinforcement learning. *Knowledge-Based Systems*, *214*. https://doi.org/10.1016/j.knosys.2020.106685.

Hiraoka, T., Imagawa, T., Mori, T., Onishi, T., & Tsuruoka, Y. (2019). Learning robust options by conditional value at risk optimization. *Advances in Neural Information Processing Systems*, *32*, 2615-2625.

Hornik, K., Stinchcombe, M., White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366.

Jiang, Z., Xu, D., & Liang, J. (2017). A deep reinforcement learning framework for the financial portfolio management problem. *ArXiv Preprint ArXiv:1706.10059*.

Jin, O., & El-Saawy, H. (2016). *Portfolio Management Using Reinforcement Learning*. Stanford University.

Kalayci, C. B., Ertenlice, O., & Akbay, M. A. (2019). A comprehensive review of deterministic models and applications for mean-variance portfolio optimization. *Expert Systems with Applications*, *125*, 345–368.

Kennedy, D. (2016). *Stochastic Financial Models*. CRC Press.

Kolm, P. N., Tütüncü, R., & Fabozzi, F. J. (2014). 60 Years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, *234*(2). https://doi.org/10.1016/j.ejor.2013.10.060.

Kompan, M., & Bieliková, M. (2010). Content-based news recommendation. In *Lecture Notes in Business Information Processing*, *61 LNBIP*. https://doi.org/10.1007/978-3-642-15208-5_6.

Konda, V., & Tsitsiklis, J. (1999). Actor-critic algorithms. *Advances in Neural Information Processing Systems*, *12*, *1008-1014*.

Lambert, D. R. (1983). Commodity channel index: Tool for trading cyclic trends. *Technical Analysis of Stocks & Commodities*, *1*, 47.

Larsen, N., Mausser, H., & Uryasev, S. (2002). Algorithms for optimization of value-at-risk. In P. M. Pardalos & V. K. Tsitsiringos (Eds.), *Financial Engineering, E-commerce and Supply Chain* (pp. 19–46). Springer.

Levina, T., Levin, Y., McGill, J., & Nediak, M. (2009). Dynamic pricing with online learning and strategic consumers: An application of the aggregating algorithm. *Operations Research*, *57*(2). https://doi.org/10.1287/opre.1080.0577.

Li, X., Li, Y., Zhan, Y., & Liu, X.-Y. (2019a). Optimistic bull or pessimistic bear: Adaptive deep reinforcement learning for stock portfolio allocation. *ArXiv Preprint ArXiv:1907.01503*.

Li, Y. (2017). *Deep Reinforcement Learning: An Overview*. https://doi.org/10.48550/arxiv.1701.07274.

Li, Y., Ni, P., & Chang, V. (2019b). An empirical research on the investment strategy of stock market based on deep reinforcement learning model. In *COMPLEXIS 2019 – Proceedings of the 4th International Conference on Complexity, Future Information Systems and Risk*. https://doi.org/10.5220/0007722000520058.

Li, Z., Liu, X.-Y., Zheng, J., Wang, Z., Walid, A., & Guo, J. (2021). FinRL-Podracer: High performance and scalable deep reinforcement learning for quantitative finance. In *Proceedings of the Second ACM International Conference on AI in Finance, online,* (pp. 1–9).

Liagkouras, K. (2019). A new three-dimensional encoding multiobjective evolutionary algorithm with application to the portfolio optimization problem. *Knowledge-Based Systems*, *163*, 186–203. https://doi.org/10.1016/J.KNOSYS.2018.08.025.

Liagkouras, K., & Metaxiotis, K. (2018). Multi-period mean–variance fuzzy portfolio optimization model with transaction costs. *Engineering Applications of Artificial Intelligence*, *67*, 260–269. https://doi.org/10.1016/J.ENGAPPAI.2017.10.010.

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., & Wierstra, D. (2016). Continuous control with deep reinforcement learning. In *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings, San Juan, Puerto Rico*.

Liu, F., Tang, R., Li, X., Zhang, W., Ye, Y., Chen, H., Guo, H., & Zhang, Y. (2018). Deep reinforcement learning based recommendation with explicit user-item interactions modeling. *ArXiv Preprint ArXiv:1810.12027*.

Liu, X.-Y., Li, Z., Yang, Z., Zheng, J., Wang, Z., Walid, A., Guo, J., & Jordan, M. I. (2021a). ElegantRL-Podracer: Scalable and elastic library for cloud-native deep reinforcement learning. *ArXiv Preprint ArXiv:2112.05923*.

Liu, X.-Y., Rui, J., Gao, J., Yang, L., Yang, H., Wang, Z., Wang, C. D., & Guo, J. (2021b). FinRL-Meta: A universe of near-real market environments for data-driven deep reinforcement learning in quantitative finance. *ArXiv Preprint ArXiv:2112.06753*.

Liu, X.-Y., Yang, H., Chen, Q., Zhang, R., Yang, L., Xiao, B., & Wang, C. D. (2020). FinRL: A deep reinforcement learning library for automated stock trading in quantitative finance. *SSRN Electronic Journal*. https://doi.org/10.48550/arxiv.2011.09607.

Liu, Y. J., & Zhang, W. G. (2015). A multi-period fuzzy portfolio optimization model with minimum transaction lots. *European Journal of Operational Research*, *242*(3), 933–941. https://doi.org/10.1016/J.EJOR.2014.10.061.

Mao, H., Venkatakrishnan, S. B., Schwarzkopf, M., & Alizadeh, M. (2019). Variance reduction for reinforcement learning in input-driven environments. In *7th International Conference on Learning Representations, ICLR 2019*, *New Orleans, USA*.

Markowitz, H. (2014). Mean–variance approximations to expected utility. *European Journal of Operational Research*, *234*(2), 346–355.

Mnih, V., Badia, A. P., Mirza, L., Graves, A., Harley, T., Lillicrap, T. P., Silver, D., & Kavukcuoglu, K. (2016). Asynchronous methods for deep reinforcement learning. In *33rd International Conference on Machine Learning, ICML 2016*, *New York, USA,* 1928-1937.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. *ArXiv Preprint ArXiv:1312.5602*.

Mosavi, A., Faghan, Y., Ghamisi, P., Duan, P., Ardabili, S. F., Salwana, E., & Band, S. S. (2020). Comprehensive review of deep reinforcement learning methods and applications in economics. *Mathematics*, *8*(10). https://doi.org/10.3390/MATH8101640.

Mousavi, S. S., Schukat, M., & Howley, E. (2018). Deep reinforcement learning: An overview. In *Lecture Notes in Networks and Systems* (Vol. 16). https://doi.org/10.1007/978-3-319-56991-8_32.

Nagarajan, P., Warnell, G., & Stone, P. (2018). Deterministic implementations for reproducibility in deep reinforcement learning. *ArXiv Preprint ArXiv:1809.05676*.

Neuneier, R. (1996). Optimal asset allocation using adaptive dynamic programming. *Advances in Neural Information Processing Systems*, *8*, *32*(2).

Nguyen, Q. P., Dai, Z., Low, B. K. H., & Jaillet, P. (2021). Value-at-risk optimization with Gaussian processes. In *International Conference on Machine Learning, online,* (pp. 8063–8072).

Peng, H., Kitagawa, G., Gan, M., & Chen, X. (2011). A new optimal portfolio selection strategy based on a quadratic form mean–variance model with transaction costs. *Optimal Control Applications and Methods*, *32*(2), 127–138.

Rong, X., Lu, M., & Deng, L. (2009). Multi-period model of portfolio investment and adjustment based on hybrid genetic algorithm. *Transactions of Tianjin University*, *15*(6), 415–422.

Rubinstein, M. (2002). Markowitz's "portfolio selection": A fifty-year retrospective. *The Journal of Finance*, *57*(3), 1041–1045.

Santamaria, J. C., Sutton, R. S., & Ram, A. (1997). Experiments with reinforcement learning in problems with continuous state and action spaces. *Adaptive Behavior*, *6*(2), 163–217.

Schulman, J. (2016). *Optimizing Expectations: From Deep Reinforcement Learning to Stochastic Computation Graphs*. UC Berkeley.

Schulman, J., Moritz, P., Levine, S., Jordan, M. I., & Abbeel, P. (2016). High-dimensional continuous control using generalized advantage estimation. In *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*. *San Juan, Puerto Rico*.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. *ArXiv Preprint ArXiv:1707.06347*.

Schwind, M. (2007). *Dynamic pricing and automated resource allocation for complex information services: Reinforcement learning and combinatorial auctions* (Vol, 589). Springer.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, *39*(1), 119–138.
Sharpe, W. F. (1994). The Sharpe Ratio. *The Journal of Portfolio Management*, *21(1), 49-58*.
Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., & Riedmiller, M. (2014). Deterministic policy gradient algorithms. In *31st International Conference on Machine Learning, ICML 2014, Beijing, China*, 387-395.
Steinbach, M. C. (2001). Markowitz revisited: Mean-variance models in financial portfolio analysis. *SIAM Review*, *43*(1), 31–85.
Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*, *3*(1). https://doi.org/10.1007/bf00115009.
Sutton, R. S., McAllester, D., Singh, S., & Mansour, Y. (1999). Policy gradient methods for reinforcement learning with function approximation. *Advances in Neural Information Processing Systems*, *12*, 1057-1063.
Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning : An Introduction.* MIT Press.
Trask, A. (2019). *Grokking Deep Learning* (1st ed.). Manning Publications Co.
Vaidya, R. (2020). Moving average convergence-divergence (MACD) trading rule: An application in Nepalese Stock Market "NEPSE". *Quantitative Economics and Management Studies*, *1*(6), 366–374.
van Hasselt, H., Guez, A., & Silver, D. (2016). Deep reinforcement learning with double q-learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, *30*(1), 2094-2100.
Vouros, G. A. (2022). Explainable deep reinforcement learning: State of the art and challenges. *ACM Computing Surveys (CSUR*), 55(5), 1-39.
Wang, Z., Schaul, T., Hessel, M., Hasselt, H., Lanctot, M., & Freitas, N. (2016). Dueling network architectures for deep reinforcement learning. In *International Conference on Machine Learning*, *New York, USA*,1995–2003.
Wilder, J. W. (1978). *New Concepts in Technical Trading Systems*. Trend Research.
Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, *8*(3), 229–256.
Wilmott, P. (2007). *Paul Wilmott Introduces Quantitative Finance*. John Wiley & Sons.
Winder, P. (2020). *Reinforcement Learning*. O'Reilly Media. https://books.google.pl/books?id=SdcHEAAAQBAJ.
Wu, Y., Mansimov, E., Grosse, R. B., Liao, S., & Ba, J. (2017). Scalable trust-region method for deep reinforcement learning using kronecker-factored approximation. *Advances in Neural Information Processing Systems*, *30*, 5279–5288.
Xiong, Z., Liu, X.-Y., Zhong, S., Yang, H., & Walid, A. (2018). *Practical Deep Reinforcement Learning Approach for Stock Trading*. https://doi.org/10.48550/arxiv.1811.07522.
Young, T. W. (1991). Calmar Ratio: A Smoother Tool. *Future Magazine*, *20*(1), 40.
Zanjirdar, M. (2020). Overview of portfolio optimization models. *Advances in Mathematical Finance and Applications*, *5*(4), 419–435.
Zhang, W., Yuan, S., & Wang, J. (2014). Optimal real-time bidding for display advertising. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/2623330.2623633.
Zhao, J., Qiu, G., Guan, Z., Zhao, W., & He, X. (2018). Deep reinforcement learning for sponsored search real-time bidding. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/3219819.3219918.

8 Leveraging the intelligent internal audit informatization for risk management to pave a route toward circular economy

Does the forensic accounting skill make it different?

Huy Quang Pham and Phuc Kien Vu

1 INTRODUCTION

SMEs have been commonly recognized as the backbone of the economy not only for developed countries (Deen, 2020; Roy et al., 2019) but for also developing countries (Sohal et al., 2022). It is considered a significant component of the manufacturing and production carried out by this kind of business. The CEI in SMEs could be paramount in gaining process amelioration and thus fostering productivity and profitability (Sawe et al., 2021). Admittedly, SMEs have played a main part in transforming from a linear toward circular economy (CE) (Schmidt et al., 2021).

Recently, the definition of a CE has drawn increasing attention and has been seen as an economic framework that might significantly improve the sustainability of firms (Khanra et al., 2021; Reike et al., 2018). CE has thought about how the industrial system may be viewed as a system that integrates economics with ecological design considerations and suggests a completely different strategy for resource use (Awan et al., 2020). In the conclusive economic system, closed-loop material flow could be realized through CE (Yi et al., 2008).

In spite of the many environmental and economic benefits generated from the CE practices, this implementation has been challenging because numerous risks posed hindrances. By adopting the RM principles, it could become much easier to tackle the challenges in terms of risks. In order for CE practices to achieve their benefits to

 DOI: 10.1201/9780367854690-8

the fullest extent, risks must be fully assessed, controlled, and addressed in a proper manner; otherwise, they would continue to hinder the entire framework of the CEI in SMEs and how it was perceived (Rahman et al., 2019). RM primarily focused on achieving the following three objectives: identifying risks, analyzing or evaluating risks, and initiating appropriate responses to address the risks identified (Ahmeti & Vladi, 2017).

Internal auditing enabled an organization to reach its goals by investigating and ameliorating the effectiveness of RM as well as governance systems in a stringent and disciplined manner. As technological advancement would allow SMEs to minimize waste and smoothen the CEI process, governmental influencers and practitioners concentrated on technological innovation to initiate CE practices (Moktadir et al., 2020). The amount and pace of data generation have increased tremendously throughout the fourth industrial revolution, and the wave of informatization has spread to every facet of life (Siggelkow et al., 2018). Due to the age of profoundly informationalized big data, the generation and operation of organizations would create an enormous number of scattered and complicated accounting data. The advancement of big data technology would assist to put forward the constitution of audit informatization in an intelligent manner. Data resources are continually being investigated as big data technology advances and develops. Big data supplied a wider vision for internal audit informatization, enriched audit data sources, strengthened RM functions, and meliorated organizational operation and management which, in turn, enable the SMEs to succeed in the CEI and RM.

Owing to the development of intelligent technologies, integrating forensic accounting technologies with fraud detection has become a vital accounting and management issue (Yang & Lee, 2020) which resulted in the burgeoning demands for forensic accounting (Baldacchino et al., 2020; Borg et al., 2020; Woods et al., 2020). The use of science and technology in forensic accounting has been used to identify dishonest accounting, financial, and corporate practices (Rezaee et al., 2016). In the current economic scenario, forensic accounting is a crucial instrument for detecting financial fraud (Kaur et al., 2022), enhancing the effectiveness of RM, and providing the organization with useful information for decision-making in terms of organizational operations.

A wide range of studies have called for an examination of the CEI in SMEs in developing nations, which are motivated by technological innovation applications and frameworks encouraging the adoption of CE practices, the current manuscript put forward the enthralling research questions (RQs) as follows.

RQ1. What is the effect of IIAI on CEI?
RQ2. Does RM mediate the interconnection between IIAI and CEI?
RQ3. Does FAS moderate the interconnection between IIAI, CEI and RM?

Rested on main observations, the current research generated several contributions both for academician and practitioner communities by bridging many voids. This study's initial theoretical contribution was to further the field of study on CEI in SMEs in developing economies. CE has recently attracted a lot of interest from academics and practitioners (Alhawari et al., 2021). Surprisingly, many professionals working in international development (Nikanorova et al., 2020) and academics have recently started to express concerns about CE practices in poor economies (Schroeder et al., 2019). Startlingly, while there have been many works focusing on larger organizations (Kumar et al., 2019), the study on SMEs' implementation of CE has been deficient(Dey et al., 2020). The obtained findings in this research cast light on the marked effect of IIAI in developing economies on CEI. The advancement in information technologies would enable SMEs to decrease waste and smoothen the CEI process (Moktadir et al., 2020). Simultaneously, the result analyses in this study increased the general findings in precedent works on the association between internal audit and RM (i.e., Coetzee & Lubbe, 2011; Kerazan, 2017; Ojo, 2019). Accordingly, IIAI was evidenced to enable the SMEs to tackle the grand challenges which SMEs had to confront with during the process of CEI by intensifying the effectiveness of RM. On the other end, RM was corroborated to be the prerequisite to assure CEI could be properly assessed pertaining to risk recognition, risk evaluation, risk control and risk mitigation. Effective risk management could, in fact, lead to standardization of risk treatment inside the business (Nichita & Vulpoi, 2016). Last but not least, the current manuscript also made great contribution to enriching FAS literature. Building on the perspectives of Kaur et al. (2022), while there is a sizable body of forensic accounting literature in the European and Western contexts, little attention has been given to developing nations. Besides, in light of the development of intelligent technologies, the integration of forensic accounting technologies with fraud detection has become a vital accounting and management issue (Yang & Lee, 2020) which thus led to the important role of FAS in leveraging the IIAI to manage risk for reaching the achievement in CEI.

Moving on to the practical aspect, the empirics of this study could generate valuable insights for practitioners in sensing and seizing the role of IIAI in CEI within SMEs. This was because of the fact that implementing CE principles in the setting of underdeveloped countries was challenging due to a number of possible obstacles (Moktadir et al., 2020). In addition, although the magnitude of RM has been acknowledged by numerous managers and organizational leaders, there have been several failures in addressing this in an appropriate manner. As a result, it has been required of the successful RM to lessen the risks that materialized in order to achieve organizational goals as effectively as possible. In this regard, the collected empirical data from the current study may offer insightful knowledge on RM to get the best CEI outcomes. Several financial scandals have occurred despite the fact that the governments and regulatory authorities of many nations have published guidelines for the ethical code of conduct of businesses (Kaur et al., 2022). In light of this, these findings also provided managers and organizational leaders in SMEs with profound insights into the critical role of FAS in the process of using IIAI in risk management and applying CE practices. Moreover, this work also supplied policymakers and governmental influencers a well-defined idea of how they could formulate and promulgate regulations and policies pertaining to advanced technologies application, while also providing insights for policymakers and governmental influencers to heart on enhancing CEI. Lastly, the outcomes of this work opened novel business opportunities which could be undertaken by hardware and software developers or retailers.

Apart from the introduction, the other sections in the present manuscript are structured in the following sequence. The theoretical understanding and foundation which concentrates on illuminating the theoretical foundation and determining the main concepts are discussed in Section 2. The hypothesized model and hypotheses substantiation are contemplated in Section 3. Afterwards, the methodology which focuses on the design and the deployment of empirical research is detailed in Section 4. This manuscript culminates with a result analysis which is demonstrated in the penultimate section. The theoretical and managerial ramifications are illustrated in the final section, along with some suggestions for the creation of new works.

2 THEORETICAL UNDERSTANDING AND FOUNDATION

2.1 Theoretical foundation

Contingency theory. Contingency theory was employed in this study as the theoretical lens to approach empirical context. One of the top theories providing a robust foundation for research on organizational design is contingency theory (Donaldson, 2007). Due to its deeply ingrained principles in the organizational literature, this form of theory, which was found by researchers from Ohio State University in 1950 (Nohria & Khurana, 2010), was employed in studies that were focused on organizational challenges (Sauser et al., 2009). The organizational design's main goal was to accommodate three critical contingencies: unpredictability, strategy, and size (Luo & Donaldson, 2013). In this context, Wang (2010) said that organizations needed to adapt to the challenging situations caused by the ambiguity of unit duties and act more vigorously. The organization's productivity strongly depended on the compatibility or alignment of the many sources of environmental instability, technology, organizational scale, organizational structure idiosyncrasies, and organizational information systems (Munyon et al., 2019). As technologies were considered to make a significant contribution to the differences in such organizational attributes (Woodward, 1958), formal structures were typically conglomerated or well-accorded with the use of various technologies according to the contingency aspect (Nohria & Khurana, 2010).

2.2 Conceptual respect

Intelligent internal audit informatization. Internal audit has long been recognized as a method of objective, independent validation and advice that improved organizational operations. It helped a company achieve its objectives by rigorously and systematically evaluating and improving the performance of governance and RM processes. Whereas the internal audit role in obtaining financial and accounting movement was to determine whether operations deserved efficacy in achieving objectives and complied with existing standards and expectations (Betti et al., 2021). Additionally, the internal audit attempted to reduce the likelihood that internal activities and transactions documented in the organizational accounting and main records would result in severe financial outcome distortions.

Under the background of big data epoch, the internal data of the organization could be revamped and balanced which thus helped the organization to reach the best development state (Zhu & Huang, 2019). The support of big data could minimize the subjective assessment and gain the objectivity of outcomes, foster the innovation of internal audit approaches and the fruitful integration of audit data (Zhu & Huang, 2019). Attributable to the age of profoundly informationalized big data, the generation and operation of organizations would create an enormous number of scattered and complicated accounting data. The advancement of big data technology would assist to put forward the constitution of audit informatization in an intelligent manner. This was because big data could provide a robust database information and a technical reinforcement to revamp audit efficiency (Zhu & Huang, 2019). On the other side, big data technology could be leveraged to audit to detect issues through comparative analysis of organizational financial data (Zhu & Huang, 2019). Additionally, the internal audit conducted investigations and evaluations of all financial accounting processes and components in order to provide management with useful information and objective advice on how to create an efficient and effective use of public assets (Munteanu et al., 2016).

Risk management. According to Staszkiewicz and Szelągowska (2019), risk is often defined as any uncertainties that could result in negative changes in profitability or losses. It could be understood as anything that causes obstacles to the achievement of a goal (Coetzee, 2016). According to Kerazan's (2017) definition, the risk is the likelihood that an incident may occur, usually with negative consequences. RM thought about the methods for reducing the risks and uncertainties that a business faced (Girangwa et al., 2020). It focused on lowering the risks associated with the chosen attempts to an extent that was acceptable to the company, building on the perspectives of Bromiley et al. (2015). Ascertaining hazards, analyzing or evaluating risks, and launching an appropriate response based on the analysis are the three main objectives of RM (Ahmeti & Vladi, 2017).

Circular economy implementation. There has not been agreement on the exact conceptualization for CE despite the advice of various academics (Lieder & Rashid, 2016; Schroeder et al., 2019). Planning, purchasing, and reprocessing activities were carried out as part of CE in order to maximize sustainability for both the ecosystem and human well-being (Murray et al., 2015). Another way to think of it is as an economic system designed for efficient resource use, waste reduction, and long-term value containment by keeping the resource in a closed loop for socioeconomic benefits and environmental protection (Morseletto, 2020). Briefly stated, the 6R principles – reuse, recycling, redesigning, reproduction, reducing, and recovering – formed the foundation of the definition of CE (Grdic et al., 2020). Through specific acts such as eco-design, refurbishing, reproduction (Nasr & Thurston 2006), product sharing, and industrial symbiosis (Chertow & Ehrenfeld, 2012; Lombardi & Laybourn, 2012), CEI mentioned on the change from the old linear manufacturing methods into a circular production practice (Dieckmann et al., 2020).

Forensic accounting skill. Based on the proposal of Enofe et al. (2015), forensic accounting was the application of expert understandings and specific knowledge to discover the evidence of economic transactions. In the meanwhile, Bhasin (2017) argued that the combination between accounting, auditing, and investigative procedures was employed in forensic accounting to perform fraud investigations in numerous situations. As such, it was considered as the three-pronged technique requesting merging accounting, auditing as well as inspecting abilities to spot or stop accounting fraud and other white-collar crimes (Kaur et al., 2022). Forensic accountants have been considered as special experts who could implement forensic accounting practices and obtain the complicated information about law, accounting, auditing as well as assessment (Renzhou, 2011). These accountants were recommended to obtain numerous skills in a wide range of scopes, namely skills in terms of accounting information system, digital forensics, auditing procedures (Crain et al., 2015), and cognitive skill (Okoye & Akamobi, 2009).

3 SUBSTANTIATION OF RESEARCH HYPOTHESES

A wide variety of operational and management practice alternations have been required as a result of the shift from a linear economy to a CE (Rizos et al., 2016; Ünal et al., 2018), this process of transition has thus raised an urgent claim on the valuable information for decision-making during operational and managerial implementation. In this regard, the IIAI would enable the SMEs to perform information integration, intelligent analysis as well as strategic implementation. More concretely, the information integration not only improved the objective criteria used to conduct internal audit investigations, but also provided favorable conditions for improving audit analysis. In order to meet the circumstances of audit system division, information integration could also deal with the barriers between organizational information and audit information. The phenomena of isolated audit information could be avoided by actively playing the role of mutual collaboration and sharing of audit-related information. The role of audit analysis may also be carried out intelligently with the aid of big data by continuously broadening the sources of information, which would capture more original, intuitive, and structured as well as unstructured data of audit-related information. In terms of strategic application, the IIAI would be both an extension of ideological methods to restructure organizational operations and a technique for internal audit work. The IIAI would offer quality control for business operations trend and prediction. In keeping with the same logic, the current study offered the following first hypothesis.

> **Hypothesis 1 (H1)**. *IIAI exerts a significant influence on CEI in a positive manner*

The implementation of CE has been difficult due to numerous risks that have presented obstacles, such as inadequate information management systems, a lack of appropriate technology, a lack of technical resources (Prieto-Sandoval et al., 2018; Rizos et al., 2016), and so on, even though CE has been well acknowledged to generate a number of environmental and economic benefits. The risks related with CEI could restrain its benefits and prevent it from functioning efficiently. They also resulted in several troubles in terms of quality and performance, and could even induce adverse influences on the environment unless risks were adequately handled or mitigated. As a result, when implementing CE, businesses risk missing their goals and alienating important stakeholders (Jakhar et al., 2019). The hazards associated with CEI should be carefully examined and addressed in order to reap the advantages to the best extent possible. Building on the perspectives of Oussii and Taktak (2018) and Sarens et al. (2012), the internal audit function comprises of evaluating the effectiveness of internal controls, providing oversight and assurance on governance processes, participating in risk management, and issuing warrants on whether an organization complied with the law.

Audit informatization could reduce the effort of auditors and filter the information needed by internal auditors from the vast data set, as informatization has proved beneficial for data exploitation (Opresnik & Taisch, 2015; Wu, 2021). By greatly improving and enhancing audit methodologies and procedures, the IIAI would alter the traditional control paradigm. Additionally, it reduced the likelihood of audit errors. The IIAI would give the SME the ability to efficiently and successfully carry out the auditing duty as well as timely and continuous data collecting throughout the project. This came as a result of IIAI's distinct advantages in data transfer, computing, and classification. With IIAI's assistance, supervision may be carried out from sampling to thorough inspection to reduce errors. The following hypotheses were taken into consideration for the current investigation based on the aforementioned analyses.

Hypothesis 2 (H2). *IIAI exerts a significant influence on RME in a positive manner*

Hypothesis 3 (H3). *IIAI exerts a significant influence on RI in a positive manner*

Hypothesis 4 (H4). *IIAI exerts a significant influence on RE in a positive manner*

Hypothesis 5 (H5). *IIAI exerts a significant influence on RR in a positive manner*

Hypothesis 6 (H6). *IIAI exerts a significant influence on RMR in positive manner*

Risk was a synonym for "hazard" and linked to the likelihood of unfavorable outcomes (Chapman & Ward, 2003). RM could be comprehended as the practices of minimizing uncertainties and risks which an entity had to confront with (Girangwa et al., 2020). The goal of RM was to enable for organizational managers to decline operating losses which might emerge in light of sudden and unexpected incidents. As proposed by Bromiley et al. (2015), RM targeted minimizing the risks pertaining to the chosen exerts at the level acceptable to the organization.

The implementation of CE has been difficult due to numerous risks that have presented obstacles, including a lack of financial support, an inadequate information management system, a lack of adequate technology, technical, and financial resources, a lack of customer interest in the environment, and a lack of support from public institutions, all of which collectively resulted in sluggish progress (Prieto-Sandoval et al., 2018; Rizos et al., 2016). The importance of risk management was that it aided companies in prioritizing risks based on their seriousness and choosing appropriate strategies and responses to those risks (Abdullah & Said, 2019). Organizations may be able to serve the interests of various stakeholders, boost confidence among stakeholders, and build stakeholder trust by implementing and adopting RM. In a similar vein, the hypotheses in the current investigation were posed as follows.

Hypothesis 7 (H7). *RME exerts a significant influence on CEI in a positive manner*
Hypothesis 8 (H8). *RI exerts a significant influence on CEI in a manner*
Hypothesis 9 (H9). *RE exerts a significant influence on CEI in a manner*
Hypothesis 10 (H10). *RR exerts a significant influence on CEI in a positive manner*
Hypothesis 11 (H11). *RMR exerts a significant influence on CEI in a positive manner*

With the rapid advancement of the digital economy and the adoption of emerging digital technologies, the analysis of the enormous volume of heterogeneous structured and unstructured data would result in a gloomy impact on the integrity and reliability of data. Although these immense outputs could be valuable for the end users, they, nonetheless, may result in an information overload for these target audiences (Brown-Liburd et al., 2015). Remarkably, the quality of internal audit abilities could no longer keep up with the evolving audit requirements, severely impeding the development of internal audit (Li, 2022). In this regard, internal auditors were required to constantly improve and update their professional skills in terms of forensic skill, analyzing data according to user requirements as well as offer a favorable environment for comprehensive, reliable, and accurate information (Li, 2022). On the other hand, RM has also been a challenging and complex action due to the integration of enormous amounts of gathered data and the uncertainty around vulnerabilities and disruptions (Papadopoulos et al., 2017) which impacted the efficacy and efficiency of operations (Comes et al., 2020). Forensic accounting involves a thorough investigation, and forensic accountants used expertise in authentication, auditing, and fraud investigation (Mishra et al., 2021; Okoye & Mbanugo, 2020). In this process, forensic accountants have been seen as playing a crucial role because of their knowledge, principles, and potential involvement in future audits that have made changes necessary (DiGabriele, 2009). Taken together, the SMEs have been required to be skillful at forensic accounting to be capable of gathering information from a wide range of sources, undertaking observations and connections, and turning analyses into more than the total of its elements. In a nutshell, the hypotheses of this research were postulated as follows (Figure 8.1).

Hypothesis 12A (H12A). *FAS moderates the positive effect of IIAI on CEI*
Hypothesis 12B (H12B). *FAS moderates the positive effect of IIAI on RMA*
Hypothesis 12C (H12C). *FAS moderates the positive effect of IIAI on RI*
Hypothesis 12D (H12D). *FAS moderates the positive effect of IIAI on RE*
Hypothesis 12E (H12E). *FAS moderates the positive effect of IIAI on RR*
Hypothesis 12F (H12F). *FAS moderates the positive effect of IIAI on RMR*

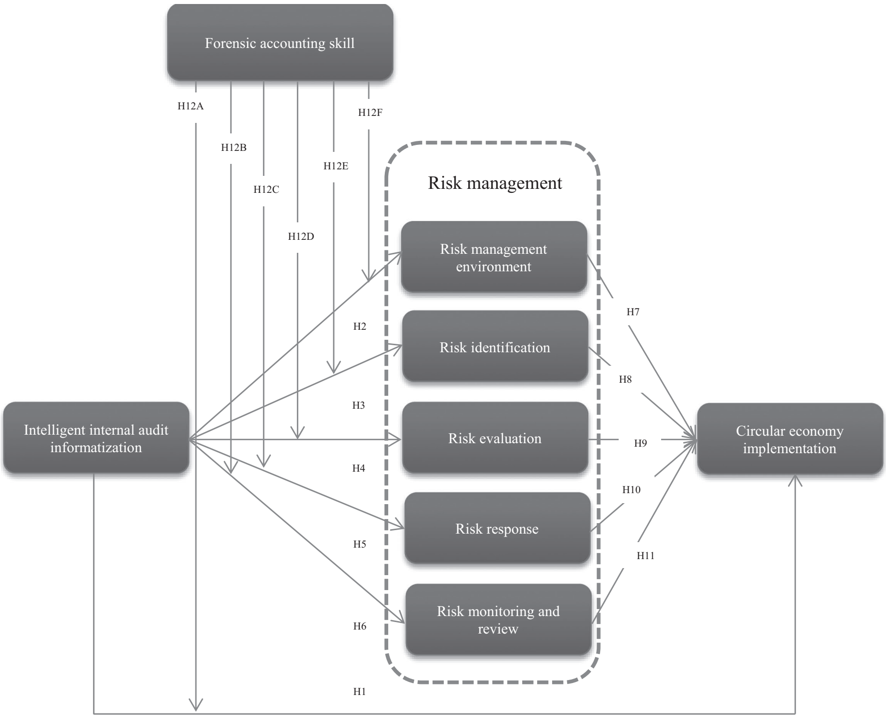


FIGURE 8.1 Conceptual model.

4 RESEARCH METHODOLOGY

In this empirical study, a cross-sectional survey approach was used. Against this backdrop, the statistical data were solely captured to assess the perspectives of the population on particular issues at a certain moment.

4.1 Research procedure

The questionnaire survey has been broadly leveraged in management studies to gather data. Additionally, the questionnaire has been determined to be an appropriate technique for participants to complete in order to collect data with a bigger sample size (Zhao et al., 2021). Extraordinarily, there has been a deficiency of readily available secondary databases in relation to digital transformation within SMEs, especially in developing economies like Vietnam.

Initially, a thorough literature survey was performed to determine the components in the conceptual model and it gave rise to a large pool of items for evaluation. In the current study, semi-structured interviews were used because they allowed for the collecting of open-ended data, the exploration of respondents' thoughts, the exploration of their sentiments and opinions towards a particular topic, as well as the analysis of deeply personal and delicate issues (DeJonckheere & Vaughn, 2019). To find the right sample of field respondents, purposive sampling was chosen since it allowed for the collection of a variety of viewpoints and fresh ideas for discussion (Kummer et al., 2021). Leaders of SMEs who could provide knowledgeable comments and in-depth information about the problem from the perspectives of their individual firms spoke up for the topic of the current study (Lincoln & Guba, 1985). The leaders of the SMEs were questioned in this regard. The selection criteria for the interviews included management seniority to ensure adequate decision-making authority and at least ten years of work experience in the organizations they represented. Based on Eisenhardt's (1989) recommendations, the number of instances should vary between 4 and 10, or until the acceptable degree of saturation was reached and the data became redundant. As a result, nine interviews with leaders of SMEs in the Southern regions of Vietnam were conducted.

The draft questionnaire was created based on the results of semi-structured interviews, and each item was scored on a five-point Likert scale that ranged from "vigorously disagree" to "vigorously agree." A five-point Likert scale was preferred to a seven-point scale because it would be simpler, have a higher response rate, and produce more reliable results (Berberoglu, 2018). Furthermore, the questionnaire was developed to safeguard respondent anonymity and reduce assessment apprehension (Campbell & Fiske, 1959; Han, 2015).

Since the questionnaire in this study was formulated in divergent settings – both culturally and environmentally, a pretest was undertaken to lessen unanticipated complexity (Alreck & Settle, 1995) as well as to warrant that the instrument quantified the constructs set out for evaluating the variables of the research, face validity of the questionnaire was corroborated by five university lectures who were suggested to measure the content of the questionnaire and provide feedback. Additionally, a pilot survey with 30 informants who were not incorporated in the primary sample of this work was carried out with the intent of investigating the viability of the questionnaire by testifying whether the questionnaire was proper and consistent plus whether the questions were obviously comprehended, well-defined, and presented constantly. In proportion to the feedback of the respondents, several adjustments were executed to the language as well as the wordings of statements. Even several statements were also reorganized and re-ordered.

4.2 Operationalization of the measured variables

Intelligent internal audit informatization. The measurement scale for IIAI consisted of five ingredients, namely Audit Efficiency, Audit Execution Ability, Audit Innovation Ability, Audit System Effectiveness, Audit Mechanism Effectiveness inherited from the findings of Gaosong and Leping (2021) and Li (2022).

Risk management. The measurement scale for RM comprised five components emanating from the integration of the contributions of Ariffin et al. (2009), Renault et al. (2018) and Moussa and Norhayate (2022), namely Risk management environment, Risk identification, Risk evaluation, Risk response, Risk monitoring, and review. More instrumentally, the measurement scales applied to evaluate Risk management environment were inherited from the contribution of Renault et al. (2018) and Moussa and Norhayate (2022). The measurement scales applied to evaluate Risk identification were inherited from the contribution of De Bakker et al. (2011); Kloss-Grote and Moss (2008), and Moussa and Norhayate (2022). The measurement scales applied to evaluate Risk evaluation were inherited from the contribution of Öztaş and Ökmen (2005); Nieto-Morote and Ruz-Vila (2011); Karimi et al. (2010), Moussa and Norhayate (2022). The measurement scales applied to evaluate Risk response were inherited from the works of Moussa and Norhayate (2022) and De Marco and Narbaev (2013). The measurement scales applied to evaluate Risk monitoring and review were inherited from the works of Hwang and Lim (2013); Kamau and Mohamed (2015) Spikin (2013); Renault et al. (2018), Moussa and Norhayate (2022).

Circular economy implementation. The CEI was operationalized by various CE practices stemming from the contribution of Zhu et al. (2010) and Khan et al. (2020).

Forensic accounting skill. The scales for FAS in the current research were determined as a first-order reflective construct stemming from the contribution of DiGabriele (2008).

4.3 Sampling procedure and data collection

In the current study, SMEs served as the analytic unit, while accountants were the respondents. Accountants were chosen because they were individuals in SMEs responsible for monitoring, disclosing, and ensuring all organizational information as well as decision-making procedures. Strikingly, due to the progressive spread of digital technology, accountants were considered to be a requirement for a favorable conclusion (Zybery & Rova, 2014).

As SEM has not provided an accurate sample size standard (Liao et al., 2022), Iacobucci (2010) suggested s a sample size of 200 or higher while Hair et al. (2010) recommended the sample-to-variable ratio preferred from 5:1 (minimal) to 10:1 (optimal) sample-to-item ratio was appropriate to identify the sample size. The sample of this work was set up on the basis of convenience and snowball sampling. Whilst convenience sampling was a form of nonprobability sampling in which individuals of the target population met criteria, namely easy accessibility and availability (Taufique & Vaithianathan, 2018), snowball sampling was a viable form of recruiting research participants difficult accessible or unknown to the researchers (Naderifar et al., 2017). As such, the survey happened between the beginning of March 2022 and the end of August 2022, targeting accountants in SMEs in the Southern areas of Vietnam, resulting in 783 samples for analysis after classifying the 900 samples circulated.

4.4 Statistical analyses and computations

Accordingly, all of the calculations for research model nomological validity in the current study were executed with the support of the SPSS v26 and AMOS v26. The analysis encompassed two stages, namely the convergent validity and discriminant validity, which were first performed and the structural model assessment was subsequently implemented. These approaches were recommended by Alzahrani et al. (2012) and Hair et al. (2011).

5 INTERPRETATION OF ANALYTICAL RESULTS AND DISCUSSION OBSERVATIONS

5.1 Sociodemographic characteristics

The samples' demographic data supported the conclusion that females made up the majority of participants, with a percentage of 82.76 versus 17.24% for males. In terms of age, the group "31–40" made up 79.57% of the entire sample, and the group "41–50" made up roughly 17.50%. Under the provided groupings, "under 30" and "over 50" made up a meager 2.30% and 0.64%, respectively, placing them last. Regarding academic qualifications and years of experience, the entire sample had at least an undergraduate degree and more than five years of experience working as accountants. Table 8.1 provides an illustration of the sociodemographic data of the participants.

5.2 Reliability and validity evaluation

The evaluation of the value of Cronbach's alpha and composite reliability served as the starting point for the reliability test of the constructs (CR). In order to demonstrate that the data was internally consistent, Cronbach's alpha and CR values were advised to be higher than the cutoff degree of 0.70 for each construct (Bagozzi & Yi, 2011). The value of factor loadings and average variance extracted (AVE) were

TABLE 8.1
Demographic information

Demographic profile	Usable responses	Weight (%)
	Gender	
Male	135	17.24
Female	648	82.76
	Age	
Under 30	18	2.30
31–40	623	79.57
41–50	137	17.50
Over 50	5	0.64
	Experience (years)	
Above 5–10	18	2.30
Above 10–15	228	29.12
Above 15–20	352	44.96
Above 20–25	180	22.99
Above 25–30	5	0.64
	Education	
Undergraduate	760	97.06
Postgraduate	23	2.94

used to evaluate the convergent validity of the applied scale (Elshaer et al., 2021). Each factor loading value that was recorded was judged to be good or higher than the cutoff of 0.5, according to the analysis (Hair et al., 2012). According to the literature, the result obtained for the AVE likewise exceeded the standard cut-off value of 0.50 (Hair et al., 2013). The reliability and convergent validity of the measurement model were attained in the current investigation based on the results in Table 8.2.

5.3 Discriminant validity evaluation

This study's evaluation of discriminant validity was based on advice from prior publications (Fornell & Larcker, 1981; Hair et al., 2014). The scores of the shared correlation of the other dimensions in both row and column should be lower than the values of the AVE square root for each individual dimension (Hair et al., 2014). Furthermore, the AVE scores must be higher than the highest possible shared value for each construct (Fornell & Larcker, 1981). Additionally, the variable's index value should be lower than 0.80 (Hair et al., 2012). The discriminant validity requirements in this investigation were met and warranted based on the results in Table 8.3.

TABLE 8.2
Results summary for the measurement model

Variables	Items (abbreviation)	Factor loadings ranges	Cronbach's alpha	AVE	Composite reliability	Discriminant validity
		Intelligent internal audit informatization				
Audit efficiency	AE	0.735–0.796	0.845	0.578	0.846	Yes
Audit execution ability	AEA	0.746–0.833	0.817	0.601	0.818	Yes
Audit innovation ability	AIA	0.762–0.859	0.843	0.644	0.844	Yes
Audit system effectiveness	ASE	0.730–0.808	0.802	0.579	0.805	Yes
Audit mechanism effectiveness	AME	0.796–0.886	0.876	0.705	0.877	Yes
		Risk management				
Risk management environment	RME	0.668–0.855	0.838	0.569	0.839	Yes
Risk identification	RI	0.528–0.899	0.819	0.561	0.831	Yes
Risk evaluation	RE	0.588–0.820	0.797	0.505	0.802	Yes
Risk response	RR	0.701–0.802	0.846	0.581	0.847	Yes
Risk monitoring and review	RMR	0.688–0.769	0.805	0.510	0.806	Yes
Circular economy implementation	CEI	0.672–0.814	0.893	0.543	0.893	Yes
Forensic accounting skill	FAS	0.728–0.812	0.880	0.600	0.882	Yes

5.4 Overall model fit evaluation

Since the χ^2 evaluation would become incongruous with empirical studies as it heavily rested on the theory of central χ^2 distribution (Byrne, 2016), a wide range of model fit indices were used for further model fitness assurance, namely Chi-square to degree of freedom (χ^2/df), goodness of fit index (GFI), comparative fit index (CFI), root mean square error of approximation (RMSEA), and Tucker–Lewis index (TLI) (Prasetyo et al., 2020). According to Table 8.4, all of the created indices satisfied the cut-off requirements specified by earlier researchers, proving that the measurement and structural models completely suited the received data.

5.5 Correlations among the constructs

Direct effect. As the SEM dropped within the admissible range, the statistical outcomes of the hypothesis tests were demonstrated in Table 8.5 as follows.

TABLE 8.3
Results of discriminant validity

	CEI	FAS	RR	AE	RME	RI	AME	RMR	RE	AIA	AEA	ASE
CEI	1											
FAS	0.087	1										
RR	0.235	0.099	1									
AE	0.269	0.021	0.228	1								
RME	0.122	0.231	0.068	0.036	1							
RI	0.310	0.087	0.024	0.037	−0.004	1						
AME	0.047	−0.055	0.100	0.168	0.061	0.053	1					
RMR	0.246	0.049	0.131	0.096	0.050	−0.028	0.113	1				
RE	0.300	−0.026	0.074	0.058	0.033	0.174	0.008	0.071	1			
AIA	0.153	−0.095	0.119	0.100	0.031	0.117	0.208	0.028	0.109	1		
AEA	0.051	−0.046	0.134	0.144	−0.010	0.060	0.185	0.090	0.083	0.184	1	
ASE	0.103	0.041	0.072	0.022	0.104	0.103	0.222	0.093	−0.011	0.354	−0.002	1

TABLE 8.4
Results of measurement and structural model analysis

Fit indices	Suggested value	Measurement model	Structural model	Global model fit	Reference
χ^2/df	≤2.5	1.490	1.632	Yes	Konecny and Thun (2011)
GFI	≥0.9	0.927	0.925	Yes	Hair et al. (2010)
CFI	≥0.9	0.969	0.962	Yes	Hair et al. (2010)
TLI	≥0.9	0.966	0.959	Yes	Hair et al. (2010)
RMSEA	≤0.07	0.025	0.028	Yes	Guo et al. (2019)

The first group of key hypotheses proposed that IIAI might significantly affect each RM dimension. While the effect of IIAI (**H3**: $\beta = 0.059$, $p < 0.01$) highlighted a strictly positive interconnection with RI, the effect of IIAI (**H2**: $\beta = 0.027$, $p < 0.05$) revealed a positive link with RME. IIAI's impact on RE (**H4**: $\beta = 0.056$, $p < 0.01$) received good support. Both the link between IIAI and RR (**H5**: $\beta = 0.097$, $p < 0.001$) and the relationship between IIAI and RMR (**H6**: $\beta = 0.067$, $p < 0.01$) were firmly established. Additionally, it was shown that IIAI had a favorable influence on CEI (**H1**: $\beta = 0.044$, $p < 0.05$). As such, **H1–H6** were buttressed.

According to the second set of key hypotheses, CEI could be significantly influenced by each RM dimension. To be more specific, the RME had a highly significant positive correlation with CEI (**H7**: $\beta = 0.135$, $p < 0.05$). Additionally, it was proven that the RI was significantly positively connected to CEI (**H8**: $\beta = 0.279$, $p < 0.001$). According to research on the association between RE and CEI (**H9**), the standardized path coefficient (β) was 0.225 at $p < 0.001$. As anticipated, the routes connecting RR and CEI were significantly positive (**H10**: $\beta = 0.163$, $p < 0.001$). The final set showed that RMR and CEI exhibited a significant, positive connection (**H11**: $\beta = 0.202$, $p < 0.001$). Thus, **H7–H11** were buttressed.

One could distinguish between partial and full mediation effects in the effects of mediation (Shankar & Jebarajakirthy, 2019). When both the direct and indirect effects were significant, partial mediation was assumed; however, full mediation was assumed when the direct effect was considerable but the indirect effect was minor (Cheung & Lau, 2008). The obtained outcomes depicted in Table 8.6 illustrated that only the RM partially mediated the interconnection between IIAI and CEI (direct effect: $\beta = 0.125$ and p-value < 0.05; indirect effect: $\beta = 0.177$ and p-value < 0.01).

Moderating effect. In order to determine the significance of the discrepancies between the parameters in the structural model between the groups proposed, MGA

TABLE 8.5
Structural coefficients (β) of the propounded model

Hypothesis No.	Causal interconnection			Estimate	S.E.	C.R.	Inference
H1	IIAI	→	CEI	0.044*	0.022	1.993	Buttressed
H2	IIAI	→	RME	0.027*	0.013	2.125	Buttressed
H3	IIAI	→	RI	0.059**	0.020	2.999	Buttressed
H4	IIAI	→	RE	0.056**	0.021	2.661	Buttressed
H5	IIAI	→	RR	0.097***	0.024	4.026	Buttressed
H6	IIAI	→	RMR	0.067**	0.022	3.077	Buttressed
H7	RME	→	CEI	0.135*	0.061	2.230	Buttressed
H8	RI	→	CEI	0.279***	0.043	6.545	Buttressed
H9	RE	→	CEI	0.225***	0.042	5.384	Buttressed
H10	RR	→	CEI	0.163***	0.045	3.589	Buttressed
H11	RMR	→	CEI	0.202***	0.043	4.737	Buttressed

Notes: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

TABLE 8.6
The summary of the mediation effects

Route of paths	Direct effect	Indirect effect	Mediation
IIAI → RM → CEI	0.125*	0.177**	Partial mediation

Notes: *p < 0.05; **p < 0.01.

TABLE 8.7
Research findings on the whole sample and the moderating role of degree of Forensic accounting skill

Causal relationship Estimate P	Low forensic accounting skill (n = 364)		High forensic accounting skill (n = 419)		Difference between parameters (high forensic accounting skill – low forensic accounting skill)		Hypothesis testing results
	Estimate	P	Estimate	P			
IIAI → CEI	0.061	0.241	0.067	0.006	0.006	−0.235	H12A was supported
IIAI → RME	0.025	0.419	0.027	0.003	0.002	−0.416	H12B was supported
IIAI → RI	0.102	0.079	0.030	0.025	−0.072	−0.054	H12C was supported
IIAI → RE	0.016	0.721	0.055	0.000	0.039	−0.721	H12D was supported
IIAI → RR	0.129	0.053	0.083	0.000	−0.046	−0.053	H12E was supported
IIAI → RMR	0.090	0.104	0.051	0.002	−0.039	−0.102	H12F was supported

Model fit	Chi-square	df	P	TLI	CFI	GFI	RMSEA
Constrained model	2297.100	1699	0.000	0.955	0.958	0.884	0.021
Unconstrained model	2272.373	1688	0.000	0.956	0.959	0.885	0.021
Differences	24.727	11	0	−0.001	−0.001	−0.001	0

could be used to investigate the existence of the moderating effects (Arbuckle, 2003; Byrne, 2001). In the present research, MGA was carried out with the support of AMOS 26.0 and the estimation for each analysis was conducted by maximum likelihood and the covariance matrix. In the MGA of invariance, invariance of the components was greatly critical. Unless it could be proved, the evaluation of the structural model had attained no value. In this case, path differences should be tested in order to seek the one which differentiated among the groups (Byrne, 2009). As such, this study endeavored to make the comparison on the χ^2 statistic between a constrained model and an unconstrained model. Prior to the measurement invariance test, this study divvied up the sample by FAS into two divergent groups including high FAS ($n = 419$) versus low FAS ($n = 364$) based on the median of the data. As the result was proved to be significant, at $p < 0.05$ ($\Delta\chi^2 = 24.727$, df = 11), it placed an emphasis on the conclusion that there were differences between the two groups. Table 8.7 depicted model fits of both unconstrained and constrained models and the values of the parameter coefficient of the path linking of IIAI with each dimension of RM and CEI between the two subgroups. As could be clearly seen from Table 8.7, the parameter coefficients for the group with high FAS were substantiated to be greater than those for the group with low FAS. This implied that high FAS could facilitate the process of IIAI application for the achievement of RM and CEI. Hence, H12A–H12F were accepted.

6 FINAL DELIBERATION AND FUTURE ENLARGEMENTS

6.1 Theoretical implications

This chapter also served as a groundbreaking investigation of the interactions between IIAI and CEI in SMEs in developing nations. In particular, the IIAI generated a more productive audit period when the internal auditing responsibilities were made as simple as possible, allowing the SMEs to conduct timely and continuous data collection throughout the project. The performance of the internal audit as well as the overall process would become more reliable with the help of IIAI, and the audit methods would also become clearer and simpler. By doing this, IIAI would guarantee the accuracy of business operations trends and prediction analysis. As a result, IIAI has been considered as both an extension of ideological tactics to energize organizational operations, particularly in the sectors of CEI and as a technique of internal audit techniques.

This study was groundbreaking because it was uniquely positioned to advance the most recent studies on the relationship between internal audit and risk management. In enterprises, there is a close connection between the internal audit function and risk management (RM), which has been supported by a number of academic studies (e.g., Coetzee & Lubbe, 2011; Kerazan, 2017; Ojo, 2019). Internal audit held a top place in firms, and its commitment to successful corporate governance has been sharply focused (Soh & Martinov-Bennie, 2011). Accordingly, IIAI was corroborated to demonstrate a marked impact on RM. Indeed, IIAI could formulate a risk-timely warning paradigm, the push and follow-up feedback system of risk-timely warning reports as well as a periodic reporting mechanism to the organizational managers. Additionally, with the support of IIAI, preliminary investigation, data procurement, and data processing could be carried out rapidly, simply, and systematically. The SMEs would simultaneously be able to identify hazards, weaknesses, and even automatic alerts that might lessen the burden on auditors' work.

SMEs have faced a variety of risks while implementing CE, including governmental risks (Moreno et al., 2017), economic risks (Boyer et al., 2021), organizational risks (Shahbazi et al., 2016), hazards related to the CE framework (Scheinberg et al., 2016), and market risks (Nikanorova et al., 2020). By advancing knowledge of the remarkable effects caused by RM on CEI, this manuscript enhanced the value of the existing literature. RM has been viewed as a process that includes early identification of potential risks, analysis of those risks, and implementation of preventative actions to totally reduce or control the risk. RM could provide organizations with benefits, namely gaining compliance with regulations, controlling expenses, established understandings and acceptance of risks within the organization (Berry-Stölzle & Xu, 2016; Choi et al., 2016). Therefore, a successful RM would reduce the risks that the SMEs had to deal with throughout the shift from the linear economy to the circular economy. An effective RM would consist of a risk management environment, risk identification, risk evaluation, risk response, as well as risk monitoring and review. Along this line, RM was also corroborated to act as a mediator on the interconnection between IIAI and CEI.

Moving to the moderating impact, the obtained findings in the current research offered interesting perspectives on the role of FAS. In this regard, FAS acted as a moderator on the interconnections between IIAI and CEI as well as the interconnection between IIAI and the components of RM. This was due to the notion that forensic accounting involved the application of science and cutting-edge technologies to identify dishonest accounting, financial, and business practices (Rezaee et al., 2016). Its functions centered on the analysis, interpretation, and presentation of difficult financial data with enough pertinent facts (Wijerathna & Perera, 2020). As a result, the FAS would allow SMEs to successfully utilize all of IIAI's advantages in RM and CEI.

6.2 Managerial and policy implications

Resting on the managerial outlooks, the acquired findings in this study offered numerous takeaway perspectives to put into practice. The statistical findings from the current study highlight the fact that IIAI would become the primary force behind RM enhancement and CEI success. Concerning IIAI, all the managers in SMEs should enhance their managerial cognitive capacities and place more concerns on this aspect. In order to warrant the smooth evolvement of IIAI, the leaders of SMEs should gain awareness of IIAI construction by improving and increasing the awareness of information-processing and providing a favorable atmosphere in the organization, so as to assure the stability of the internal audit environment. Simultaneously, tangible resources, namely infrastructure, digital platform as well as other essential resources to implement IIAI projects have been encouraged to be concentrated. Besides, all the managers in SMEs should concentrate on developing proficiency for the organizational workforce through proper training programs to keep them abreast of the state-of-the-art programming systems. Concretely, the quality of IIAI could be ameliorated by means of the following approaches namely intensifying the vocational training, training professional knowledge, and indispensable professional judgment skills. The internal auditors should also engage in the formulation and development of the information systems to seek for the troubles in the information system development process as well as generate suggestions in a timely manner so that the information systems could be effectively taken advantage of. In addition to the internal training programs, SMEs were urged to organize specialized outside training programs to take advantage of market conditions and learn cutting-edge information technology. The results of the current study also demonstrated the crucial role of FAS, which was shown to significantly and favorably increase the influence of IIAI on each component of RM and CEI. As a result, managers in SMEs should also focus on improving the forensic expertise of their team by sending them to participate in appropriate training programs.

The markedly significant interconnection between each component of RM on CEI should be taken into consideration. As such, all the managers in SMEs were advised to apply exceptional strategies in consistent with the organizational idiosyncrasies and meticulousness of risk control to succeed in operating. All the organizational strategies should be in compliance with governmental policies and risk control to prevent any potential crises as well as to gain more stability for organizational operations. Against to this backdrop, all the managers in SMEs were suggested to configure the internal processes and to decrease the potential causes of risk to achieve vigorous RM plans.

Due to the paramount part in revamping and developing the efficiency and effectiveness of advanced information technologies deployment amongst all organizations, policymakers should set up and promulgate policies as well as guidelines in relation to advanced information technologies adoption, and simultaneously, take specific actions in term of budget allocation, buttress and plan for implementing the advanced information technologies infrastructures. Moreover, the outcomes of this work were predominant for policymakers in configuring and promulgating policies and measures in relation to CEI.

Endeavors were also requested by hardware and software developers or retailers to comprehend much more explicitly the challenges and advantages of modern information technologies for more state-of-the-art system evolvement fitting appropriately with the SMEs' idiosyncrasies.

6.3 Research limitations

This analysis encountered several drawbacks, which bring out the trustworthiness of the obtained observations in this research as well as open novel avenues for upcoming academicians in the same field. Against this backdrop, one limitation of this analysis belonged to its geographical focus, as these analyses only contemplated the SMEs in the south of Vietnam. They could not be generalized to other SMEs across Vietnamese region without additional investigations. A future study should take into account other types of organizations, such as major firms and corporations, as SMEs were only one area of specialization. In order to demonstrate how the SMEs could profit from the CEI, follow-up researchers were also asked to do comparative studies and concentrate on emerging as opposed to developed economies. The third limitation was the sample comprised of accountants from SMEs in the southern areas of Vietnam, and its observations should be interpreted with caution. In this manner, it was recommended that future studies would enhance the scope of the research by pondering the staff from the other departments in the SMEs. Thus, it would assist to present the broader observations and substantial contributions to the prevailing literature. The convenience and snowball sampling methods created the next bottleneck, which may have affected the study's generalizability. Future research might benefit greatly from utilizing a quota sampling strategy to obtain sample data in order to produce representative and scientific results. The current research was conforming to other works that cross-sectional data was weak in illuminating the causal interconnections and thus, it was suggested that follow-up academicians should incorporate a longitudinal dataset for flawless causal interconnections. Finally, the conceptualization of IIAI depicted in the current research was predominantly rested on a semantic aggregation of substantiating clarifications procured from the flourishing body of literature. Nonetheless, it would incontrovertibly not be a drawback in the follow-up studies when the several scrupulous explorations would be carried out to reach a more vigorous and rigorous enlightenment for the above-mentioned concept.

ACKNOWLEDGEMENT

This chapter was funded by University of Economics Ho Chi Minh City (UEH).

REFERENCES

Abdullah, W. N., & Said, R. (2019). Audit and risk committee in financial crime prevention. *Journal of Financial Crime, 26*(1), 223–234. doi: 10.1108/JFC-11-2017-0116.

Ahmeti, R., & Vladi, B. (2017). Risk management in public sector: a literature review. *European Journal of Multidisciplinary Studies, 2*(5), 323–329.

Alhawari, O., Awan, U., Bhutta, M. K. S., & Ülkü, M. A. (2021). Insights from circular economy literature: a review of extant definitions and unravelling paths to future research. *Sustainability, 13*, 1–22. doi: 10.3390/su13020859.

Alreck, P. L., & Settle, R. B. (1995). The importance of word-of-mouth communications to service buyers. *Proceedings of American Marketing Association - Chicago, 1*(6), 188–193.

Alzahrani, A., Stahl, B. C., & Prior, M. (2012). Developing an instrument for E-public services' acceptance using confirmatory factor analysis. *Journal of Organizational and End User Computing, 24*(3), 18–44. doi: 10.4018/joeuc.2012070102.

Arbuckle, J. L. (2003). *AMOS 5.0: Update to the AMOS User's Guide*. Chicago, IL: Smallwaters Corporation.

Ariffin, N. M., Archer, S., & Karim, R. A. A. (2009). Risks in Islamic banks: evidence from empirical research. *Journal of Banking Regulation, 10*(2), 153–163. doi: 10.1057/jbr.2008.27.

Awan, U., Kanwal, N., & Bhutta, M. K. S. (2020). A literature analysis of definitions for a circular economy. In: Golinska-Dawson, P. (eds.), *Logistics Operations and Management for Recycling and Reuse* (pp. 19–34). doi: 10.1007/978-3-642-33857-1_2.

Bagozzi, R. P., & Yi, Y. (2011). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science, 40*(1), 8–34. doi: 10.1007/s11747-011-0278-x.

Baldacchino, P. J., Bugeja, S., & Grima, S. (2020). Social auditing and its applicability to maltese co-operatives*. *Contemporary Issues in Audit Management and Forensic Accounting*, 273–301. doi: 10.1108/s1569-375920200000102020.

Berberoglu, A. (2018). Impact of organizational climate on organizational commitment and perceived organizational performance: empirical evidence from public hospitals. *BMC Health Services Research, 18*(1). doi: 10.1186/s12913-018-3149-z.

Berry-Stölzle, T. R., & Xu, J. (2016). Enterprise risk management and the cost of capital. *Journal of Risk and Insurance, 85*(1), 159–201. doi: 10.1111/jori.12152.

Betti, N., Sarens, G., & Poncin, I. (2021). Effects of digitalisation of organisations on internal audit activities and practices. *Managerial Auditing Journal, 36*(6), 872–888. doi: 10.1108/MAJ-08-2020-2792.

Bhasin, M. L. (2017). Integrating corporate governance and forensic accounting: a study of an Asian country. *International Journal of Management Sciences and Business Research, 6*(1), 31–52.

Borg, G., Baldacchino, P. J., Buttigieg, S., Boztepe, E., & Grima, S. (2020). Challenging the adequacy of the conventional "three lines of defence" model: a case study on Maltese credit institutions*. *Contemporary Issues in Audit Management and Forensic Accounting*, 303–324. doi: 10.1108/s1569-375920200000102021.

Boyer, R. H. W., Hunka, A. D., Linder, M., Whalen, K., & Habibi, S. (2021). Product labels for the circular economy: are customers willing to pay for circular? *Sustainable Production and Consumption, 27*, 61–71. doi: 10.1016/j.spc.2020.10.010.

Bromiley, P., McShane, M., Nair, A., & Rustambekov, E. (2015). Enterprise risk management: review, critique, and research directions. *Long Range Planning, 48*(4), 265–276. doi: 10.1016/j.lrp.2014.07.005.

Brown-Liburd, H., Issa, H., & Lombardi, D. (2015). Behavioral implications of Big Data's impact on audit judgment and decision making and future research directions. *Accounting Horizons, 29*(2), 451–468.

Byrne, B. (2009). *Structural Equation Modelling with AMOS: Basic Concepts, Applications, and Programming* (2nd ed.). New York: Routledge/Taylor and Francis.

Byrne, B. M. (2001). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. Mahwah, NJ: Lawrence Erlbaum Associates.

Byrne, B. M. (2016). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming* (3rd ed.). New York: Routledge.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81–105. doi: 10.1037/h0046016.

Chapman, C., & Ward, S. (2003). *Project Risk Management: Processes, Techniques and Insights*. Chichester: John Wiley & Sons Ltd.

Chertow, M., & Ehrenfeld, J. (2012). Organizing self-organizing systems. *Journal of Industrial Ecology, 16*(1), 13–27. doi: 10.1111/j.1530-9290.2011.00450.x.

Cheung, G. W., & Lau, R. S. (2008). Testing mediation and suppression effects of latent variables: bootstrapping with structural equation models. *Organizational Research Methods, 11*(2), 296–325.

Choi, T.-M., Chiu, C.-H., & Chan, H.-K. (2016). Risk management of logistics systems. *Transportation Research Part E: Logistics and Transportation Review, 90*, 1–6. doi: 10.1016/j.tre.2016.03.007.

Coetzee, P. (2016). Contribution of internal auditing to risk management. *International Journal of Public Sector Management, 29*(4), 348–364. doi: 10.1108/ijpsm-12-2015-0215.

Coetzee, P., & Lubbe, D. (2011). Internal audit and risk management in South Africa: adherence to guidance. *Acta Academica, 43*(4), 29–60.

Comes, T., Van de Walle, B., & Van Wassenhove, L. (2020). The coordination-information bubble in humanitarian response: theoretical foundations and empirical investigations. *Production and Operations Management*, 1–42. doi: 10.1111/poms.13236.

Crain, M. A., Hopwood, W. S., Pacini, C., & Young, G. R. (2015). *Essentials of Forensic Accounting*. New York: American Institute of Certified Public Accountants, Inc.

De Bakker, K., Boonstra, A., & Wortmann, H. (2011). Risk management affecting IS/IT project success through communicative action. *Project Management Journal, 42*(3), 75–90. doi: 10.1002/pmj.20242.

De Marco, A., & Narbaev, T. (2013). Earned value-based performance monitoring of facility construction projects. *Journal of Facilities Management, 11*(1), 69–80. doi: 10.1108/14725961311301475.

Deen, S. J. (2020). An analysis of growth of MSMEs in India and their contribution in employment and GDP of the country. *Studies in Indian Place Names, 40*(65), 11–16.

DeJonckheere, M., & Vaughn, L. M. (2019). Semi-structured interviewing in primary care research: a balance of relationship and rigour. *Family Medicine and Community Health, 7*(2), 1–8. doi: 10.1136/fmch-2018-000057.

Dey, P. K., Malesios, C., De, D., Budhwar, P., Chowdhury, S., & Cheffi, W. (2020). Circular economy to enhance sustainability of small and medium-sized enterprises. *Business Strategy and the Environment*, 1–25. doi: 10.1002/bse.2492.

Dieckmann, E., Sheldrick, L., Tennant, M., Myers, R., & Cheeseman, C. (2020). Analysis of barriers to transitioning from a linear to a circular economy for end of life materials: a case study for waste feathers. *Sustainability, 12*(5), 1–21. doi: 10.3390/su12051725.

DiGabriele, J. A. (2008). An empirical investigation of the relevant skills of forensic accountants. *Journal of Education for Business, 83*(6), 331–338. doi: 10.3200/joeb.83.6.331-338.

DiGabriele, J. A. (2009). Implications of regulatory prescriptions and audit standards on the evolution of forensic accounting in the audit process. *Journal of Applied Accounting Research, 10*(2), 109–121. doi: 10.1108/09675420910984673.

Donaldson, L. (2007). Ethics problems and problems with ethics: toward a pro-management theory. *Journal of Business Ethics, 78*(3), 299–311. doi: 10.1007/s10551-006-9336-6.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review, 14*(4), 532–550.

Elshaer, I., Sobaih, A. E. E., Alyahya, M., & Abu Elnasr, A. (2021). The impact of religiosity and food consumption culture on food waste intention in Saudi Arabia. *Sustainability, 13*(11), 1–17. doi: 10.3390/su13116473.

Enofe, A. O., Agbonkpolour, O. R., & Edebiri, O. J (2015). Forensic accounting and financial fraud. *International Journal of Multidisciplinary Research and Development, 2*(10), 305–312.

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: algebra and statistics. *Journal of Marketing Research, 18*(3), 382–388. doi: 10.2307/3150980.

Gaosong, Q., & Leping, Y. (2021). Measurement of internal audit effectiveness: construction of index system and empirical analysis. *Microprocessors and Microsystems*, 1–26. doi: 10.1016/j.micpro.2021.104046.

Girangwa, K. G., Rono, L., & Mose, J. (2020). The influence of enterprise risk management practices on organizational performance: evidence from kenyan state corporations. *Journal of Accounting, Business and Finance Research, 8*(1), 11–20.

Grdic, S. Z., Nizic, M. K., & Rudan, E. (2020). Circular economy concept in the context of economic development in EU countries. *Sustainability, 12*(7), 1–13. doi: 10.3390/su12073060.

Guo, J., Marsh, H. W., Parker, P. D., Dicke, T., Lüdtke, O., & Diallo, T. M. O. (2019). A systematic evaluation and comparison between exploratory structural equation modeling and Bayesian structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 1–28. doi: 10.1080/10705511.2018.1554999.

Hair, J., Black, W., Babin, B., & Anderson, R. (2014). *Multivariate Data Analysis* (7th ed.). Harlow: Pearson.

Hair, J. F, Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis* (7th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Partial least squares structural equation modeling: rigorous applications, better results and higher acceptance. *Long Range Planning, 46*(1–2), 1–12. doi: 10.1016/j.lrp.2013.01.001.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2011). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science, 40*(3), 414–433. doi: 10.1007/s11747-011-0261-6.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science, 40*(3), 414–433. doi: 10.1007/s11747-011-0261-6.

Han, H. (2015). Travelers' pro-environmental behavior in a green lodging context: converging value-belief-norm theory and the theory of planned behavior. *Tourism Management, 47*, 164–177. doi: 10.1016/j.tourman.2014.09.014.

Hwang, B.-G., & Lim, E.-S. J. (2013). Critical success factors for key project players and objectives: case study of Singapore. *Journal of Construction Engineering and Management, 139*(2), 204–215. doi: 10.1061/(asce)co.1943-7862.0000597.

Iacobucci, D. (2010). Structural equations modeling: fit indices, sample size, and advanced topics. *Journal of Consumer Psychology, 20*(2010), 90–98. doi: 10.1016/j.jcps.2009.09.003.

Jakhar, S. K., Mangla, S. K., Luthra, S., & Kusi-Sarpong, S. (2019). When stakeholder pressure drives the circular economy. *Management Decision, 57*(4), 904–920. doi: 10.1108/MD-09-2018-0990.

Kamau, C. G., & Mohamed, H. G. (2015). The efficacy of monitoring and evaluation function in achieving project success in Kenya. *Science Journal of Business and Management, 3*, 82–94.

Karimi, A., Mousavi, N., Mousavi, S., & Hosseini, S. (2010). Risk assessment model selection in construction industry. *Expert Systems with Applications, 38*, 9105–9111.

Kaur, B., Sood, K., & Grima, S. (2022). A systematic review on forensic accounting and its contribution towards fraud detection and prevention. *Journal of Financial Regulation and Compliance*, 1–36. doi: 10.1108/JFRC-02-2022-0015.

Kerazan, F. H. (2017). The contribution of the internal audit function in risk management. *Journal of Management Research, 8*(3), 178–202. doi: 10.5296/jmr.v8i3.9096.

Khan, O., Daddi, T., & Iraldo, F. (2020). Sensing, seizing, and reconfiguring: key capabilities and organizational routines for circular economy implementation. *Journal of Cleaner Production*, 1–38. doi: 10.1016/j.jclepro.2020.125565.

Khanra, S., Kaur, P., Joseph, R. P., Malik, A., & Dhir, A. (2021). A resource-based view of green innovation as a strategic firm resource: present status and future directions. *Business Strategy and the Environment*, *3*(4), 1–19. DOI: 10.1002/bse.2961

Kloss-Grote, B., & Moss, M. A. (2008). How to measure the effectiveness of risk management in engineering design projects? Presentation of RMPASS: a new method for assessing risk management performance and the impact of knowledge management—including a few results. *Research in Engineering Design, 19*(2–3), 71–100. doi: 10.1007/s00163-008-0049-y.

Konecny, P. A., & Thun, J.-H. (2011). Do it separately or simultaneously—an empirical analysis of a conjoint implementation of TQM and TPM on plant performance. *International Journal of Production Economics, 133*(2), 496–507. doi: 10.1016/j.ijpe.2010.12.009.

Kumar, V., Sezersan, I., Garza-Reyes, J. A., Gonzalez, E. D. R. S., & AL-Shboul, M. A. (2019). Circular economy in the manufacturing sector: benefits, opportunities and barriers. *Management Decision*, 1–21. doi: 10.1108/md-09-2018-1070.

Kummer, T.-F., Pelzl, S., & Bick, M. (2021). A conceptualisation of privacy risk and its influence on the disclosure of check-in services information. *International Journal of Information Management, 57*, 1–11. doi: 10.1016/j.ijinfomgt.2020.102266.

Li, S. (2022). Discussion on the construction of enterprise internal audit informatization. *Journal of Advanced Transportation*, 1–11.

Liao, T., Tang, S., & Shim, Y. S. (2022). The development of a model to predict sports participation among college students in central China. *International Journal of Environmental Research and Public Health, 19*(3), 1–11, 1806. doi: 10.3390/ijerph19031806.

Lieder, M., & Rashid, A. (2016). Towards circular economy implementation: a comprehensive review in context of manufacturing industry. *Journal of Cleaner Production, 115*, 36–51. doi: 10.1016/j.jclepro.2015.12.042.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. Thousand Oaks, CA: Sage.

Lombardi, D. R., & Laybourn, P. (2012). Redefining industrial symbiosis. *Journal of Industrial Ecology, 16*(1), 28–37. doi: 10.1111/j.1530-9290.2011.00444.x.

Luo, B. N., & Donaldson, L. (2013). Misfits in organization design-information processing as a compensatory mechanism. *Journal of Organization Design, 2*(1), 2–10.

Mishra, K., Azam, M. K., & Junare, S. O. (2021). Role of forensic audit in controlling financial statement fraud: a case study of Satyam computers. *Psychology and Education Journal, 58*(2), 4016–4025.

Moktadir, M. A., Kumar, A., Ali, S. M., Paul, S. K., Sultana, R., & Rezaei, J. (2020). Critical success factors for a circular economy: implications for business strategy and the environment. *Business Strategy and the Environment*, 1–25. doi: 10.1002/bse.2600.

Moreno, M., Turner, C., Tiwari, A., Hutabarat, W., Charnley, F., Widjaja, D., & Mondini, L. (2017). Re-distributed manufacturing to achieve a circular economy: a case study utilizing IDEF0 modeling. *Procedia CIRP, 63*, 686–691. doi: 10.1016/j.procir.2017.03.322.

Morseletto, P. (2020). Targets for a circular economy. *Resources, Conservation and Recycling, 153*, 1–12. doi: 10.1016/j.resconrec.2019.104553.

Moussa, D. B., & Norhayate, W. D. W. (2022). Impact of cybersecurity risk management frameworks on Malaysian public universities business performance. *Global Journal of Management and Business Research: A Administration and Management, 22*, 27–34.

Munteanu, V., Copcinschi, L., Luschi, C., & Laceanu, A. (2016). Internal audit-determinant factor in preventing and detecting fraud related activity to public entities financial accounting. *Knowledge Horizons Economics, 8*(2), 14–21.

Munyon, T. P., Jenkins, M. T., Crook, T. R., Edwards, J., & Harvey, N. P. (2019). Consequential cognition: exploring how attribution theory sheds new light on the firm-level consequences of product recalls. *Journal of Organizational Behavior*, 1–16. doi: 10.1002/job.2350.

Murray, A., Skene, K., & Haynes, K. (2015). The circular economy: an interdisciplinary exploration of the concept and application in a global context. *Journal of Business Ethics, 140*(3), 369–380. doi: 10.1007/s10551-015-2693-2.

Naderifar, M., Goli, H., & Ghaljaie, F. (2017). Snowball sampling: a purposeful method of sampling in qualitative research. *Strides in Development of Medical Education, 14*(3), 1–7.

Nasr, N., & Thurston, M. (2006). Remanufacturing: a key enabler to sustainable product systems. *Rochester Institute of Technology, 23*, 15–18.

Nichita, M. E., & Vulpoi, M. (2016). Relationship between risk and transparency in the financial statements of professional services entities. *Audit Financiar, XIV*(137), 540–550.

Nieto-Morote, A., & Ruz-Vila, F. (2011). A fuzzy approach to construction project risk assessment. *International Journal of Project Management, 29*(2), 220–231. doi: 10.1016/j.ijproman.2010.02.002.

Nikanorova, M., Imoniana, J. O., & Stankeviciene, J. (2020). Analysis of social dimension and well-being in the context of circular economy. *International Journal of Global Warming, 21*(3), 299–316. doi: 10.1504/ijgw.2020.108678.

Nohria, N., & Khurana, R. (2010). *Handbook of Leadership Theory and Practice*. Boston: Harvard Business School Press.

Ojo, A. (2019). Internal audit and risk management in Nigeria's public sector. *International Journal of Business & Law Research, 7*(2), 1–15.

Okoye, E. I., & Akamobi, N. L. (2009). The role of forensic accounting in fraud investigation and litigation support. *The Nigerian Academic Forum, 17*(1), 39–44.

Okoye, K. R. E., & Mbanugo, C. I. (2020). Forensic accounting a tool for fraud detection and prevention in the public tertiary institutions in south-east Nigeria. *European Journal of Education Studies, 7*(6), 1–12.

Opresnik, D., & Taisch, M. (2015). The value of Big Data in servitization. *International Journal of Production Economics, 165*, 174–184. doi: 10.1016/j.ijpe.2014.12.036.

Oussii, A. A., & Boulila Taktak, N. (2018). The impact of internal audit function characteristics on internal control quality. *Managerial Auditing Journal, 33*(5), 450–469. doi: 10.1108/maj-06-2017-1579.

Öztaş, A., & Ökmen, Ö. (2005). Judgmental risk analysis process development in construction projects. *Building and Environment, 40*(9), 1244–1254. doi: 10.1016/j.buildenv.2004.10.013.

Papadopoulos, T., Gunasekaran, A., Dubey, R., Altay, N., Childe, S. J., & Fosso-Wamba, S. (2017). The role of Big Data in explaining disaster resilience in supply chains for sustainability. *Journal of Cleaner Production, 142*, 1108–1118. doi: 10.1016/j.jclepro.2016.03.059.

Prasetyo, Y. T., Castillo, A. M., Salonga, L. J., Sia, J. A., & Seneta, J. A. (2020). Factors affecting perceived effectiveness of COVID-19 prevention measures among Filipino during enhanced community quarantine in Luzon, Philippines: integrating protection motivation theory and extended theory of planned behavior. *International Journal of Infectious Diseases*, 1–12. doi: 10.1016/j.ijid.2020.07.074.

Prieto-Sandoval, V., Jaca, C., & Ormazabal, M. (2018). Towards a consensus on the circular economy. *Journal of Cleaner Production, 179*, 605–615. doi: 10.1016/j.jclepro.2017.12.224.

Rahman, T., Ali, S. M., Moktadir, M. A., & Kusi-Sarpong, S. (2019). Evaluating barriers to implementing green supply chain management: an example from an emerging economy. *Production Planning & Control, 31*(8), 673–698. doi: 10.1080/09537287.2019.1674939.

Reike, D., Vermeulen, W. J. V., & Witjes, S. (2018). The circular economy: new or refurbished as CE 3.0? — exploring controversies in the conceptualization of the circular economy through a focus on history and resource value retention options. *Resources, Conservation and Recycling, 135*, 246–264. doi: 10.1016/j.resconrec.2017.08.027.

Renault, B., Agumba, J., & Ansary, N. (2018). An exploratory factor analysis of risk management practices: a study among small and medium contractors in Gungormu. *Acta Structilia, 25*, 1–39.

Renzhou, D. (2011). Research on legal procedural functions of forensic accounting. *Energy Procedia, 5*, 2147–2151. doi: 10.1016/j.egypro.2011.03.371.

Rezaee, Z., Lo, D., Ha, M., & Suen, A. (2016). Forensic accounting education and practice: insights from China. *Journal of Forensic and Investigative Accounting, 8*(1), 106–119.

Rizos, V., Behrens, A., van der Gaast, W., Hofman, E., Ioannou, A., Kafyeke, T., Flamos, A., Rinaldi, R., Papadelis, S., Hirschnitz-Garbers, M., & Topi, C. (2016). Implementation of circular economy business models by small and medium-sized enterprises (SMEs): barriers and enablers. *Sustainability, 8*(11), 1–18. doi: 10.3390/su8111212.

Roy, S., Bhattacharya, S., & Das, P. (2019). Learning clusters, MOOCs, free videos and organization learning: a case study from Indian SMEs. *Development and Learning in Organizations: An International Journal, 34*(1), 16–20. doi: 10.1108/dlo-03-2019-0057.

Sarens, G., Abdolmohammadi, M. J., & Lenz, R. (2012). Factors associated with the internal audit function's role in corporate governance. *Journal of Applied Accounting Research, 13*(2), 191–204. doi: 10.1108/09675421211254876.

Sauser, B. J., Reilly, R. R., & Shenhar, A. J. (2009). Why projects fail? How contingency theory can provide new insights – a comparative analysis of NASA's Mars Climate Orbiter loss. *International Journal of Project Management, 27*(7), 665–679. doi: 10.1016/j.ijproman.2009.01.004.

Sawe, F. B., Kumar, A., Garza-Reyes, J. A., & Agrawal, R. (2021). Assessing people-driven factors for circular economy practices in small and medium-sized enterprise supply chains: business strategies and environmental perspectives. *Business Strategy and the Environment*, 1–15. doi: 10.1002/bse.2781.

Scheinberg, A., Nesić, J., Savain, R., Luppi, P., Sinnott, P., Petean, F., & Pop, F. (2016). From collision to collaboration – integrating informal recyclers and re-use operators in Europe: a review. *Waste Management & Research, 34*(9), 820–839. doi: 10.1177/0734242x16657608.

Schmidt, C. V. H., Kindermann, B., Behlau, C. F., & Flatten, T. C. (2021). Understanding the effect of market orientation on circular economy practices: the mediating role of closed-loop orientation in German SMEs. *Business Strategy and the Environment*, 1–17. doi: 10.1002/bse.2863.

Schroeder, P., Anggraeni, K., & Weber, U. (2019). The relevance of circular economy practices to the sustainable development goals. *Journal of Industrial Ecology, 23*(1), 77–95. doi: 10.1111/jiec.12732.

Shahbazi, S., Wiktorsson, M., Kurdve, M., Jönsson, C., & Bjelkemyr, M. (2016). Material efficiency in manufacturing: swedish evidence on potential, barriers and strategies. *Journal of Cleaner Production, 127*, 438–450. doi: 10.1016/j.jclepro.2016.03.143.

Shankar, A., & Jebarajakirthy, C. (2019). The influence of e-banking service quality on customer loyalty. *International Journal of Bank Marketing*, 1–25. doi: 10.1108/ijbm-03-2018-0063.

Siggelkow, B. F., Trockel, J., & Dieterle, O. (2018). An inspection game of internal audit and the influence of whistle-blowing. *Journal of Business Economics, 88*(7–8), 883–914. doi: 10.1007/s11573-018-0893-9.

Soh, D. S. B., & Martinov-Bennie, N. (2011). The internal audit function. *Managerial Auditing Journal, 26*(7), 605–622. doi: 10.1108/02686901111151332.

Sohal, A., Nand, A. A., Goyal, P., & Bhattacharya, A. (2022). Developing a circular economy: an examination of SME's role in India. *Journal of Business Research, 142*, 1–13. doi: 10.1016/j.jbusres.2021.12.072.

Spikin, I. C. (2013). Risk management theory: the integrated perspective and its application in the public sector, Estado, Gobierno. *Gestión Pública, 21*, 89–126.

Staszkiewicz, P., & Szelągowska, A. (2019). Ultimate owner and risk of company performance. *Economic Research-Ekonomska Istraživanja, 32*(1), 3795–3812. doi: 10.1080/1331677x.2019.1678499.

Taufique, K. M., & Vaithianathan, S. (2018). A fresh look at understanding green consumer behavior among young urban Indian consumers through the lens of Theory of Planned Behavior. *Journal of Cleaner Production, 183*, 46–55.

Ünal, E., Urbinati, A., & Chiaroni, D. (2018). Managerial practices for designing circular economy business models. *Journal of Manufacturing Technology Management, 30*(3), 561–589. doi: 10.1108/jmtm-02-2018-0061.

Wang, P. (2010). Restructuring to repair legitimacy - a contingency perspective. *Corporate Governance: An International Review, 18*(1), 64–82. doi: 10.1111/j.1467-8683.2009.00782.x.

Wijerathna, A. G. H. S. K., & Perera, H. A. P. L. (2020). A systematic literature review on forensic accounting. *Proceedings of the International Conference on Business & Information (ICBI) 2020*, University of Kelaniya, Sri Lanka, 105–119.

Woods, J. J., Seychell, S., Ozen, E., Spiteri, J., Suban, R., & Grima, S. (2020). Explaining heterogeneity in risk appetite and tolerance: the Turkish Case*,†. *Contemporary Issues in Audit Management and Forensic Accounting*, 111–130. doi: 10.1108/s1569-375920200000102011.

Woodward, J. (1958). *Management and Technology*. London: Her Majesty's Stationery Office.

Wu, Y. (2021). Database construction in computer audit under big data environment. *Journal of Physics: Conference Series, 1881*, 1–8.

Yang, C.-H., & Lee, K.-C. (2020). Developing a strategy map for forensic accounting with fraud risk management: an integrated balanced scorecard-based decision model. *Evaluation and Program Planning, 80*, 1–10. doi: 10.1016/j.evalprogplan.2020.101780.

Yi, S., Wu, N., Luo, P., Wang, Q., Shi, F., Zhang, Q., & Ma, J. (2008). Agricultural heritage in disintegration: trends of agropastoral transhumance on the southeast Tibetan Plateau. *International Journal of Sustainable Development & World Ecology, 15*(3), 273–283. doi: 10.3843/susdev.15.3:10.

Zhao, F., Kusi, M., Chen, Y., Hu, W., Ahmed, F., & Sukamani, D. (2021). Influencing mechanism of green human resource management and corporate social responsibility on organizational sustainable performance. *Sustainability, 13*, 1–15. doi: 10.3390/su13168875.

Zhu, Q., Geng, Y., & Lai, K. (2010). Circular economy practices among Chinese manufacturers varying in environmental-oriented supply chain cooperation and the performance implications. *Journal of Environmental Management, 91*(6), 1324–1331. doi: 10.1016/j.jenvman.2010.02.013.

Zhu, Y., & Huang, B. (2019). Summary of research on the application of big data in auditing. *2019 International Conference on Communications, Information System and Computer Engineering (CISCE)*, 1–4. doi: 10.1109/cisce.2019.00156.

Zybery, I., & Rova, L. (2014). The role of the accountants in the framework of the modern technological developments and digital accounting systems. *European Scientific Journal, 24*, 30–48. doi: 10.19044/esj.2011.v24n0p%25p.

9 Designing a framework for guest experience management in the hotel industry based on data analysis

Tam To Nguyen, Ha Thi Thu Nguyen, Linh Khanh Bui, and Xuan Trung Nguyen

1 INTRODUCTION

The customer experience is a critical area for a company to develop sustainably (De Keyser et al., 2015). Customer experience refers to customers' perception of a product or service's quality and features after they have used it. So, from a company's viewpoint, CX is a key strategic objective in the marketing of companies (Klaus & Maklan, 2013). And the customer experience is also considered the new competitive strategy for business (Klink et al., 2020). These issues have been mentioned many times in business strategies, marketing strategies, customer service, and business management. Therefore, many companies focus on customer experience management to understand the psychology and desires of customers and develop better products and services (Klink et al., 2020). Customer experience management is used in measuring customer satisfaction, customer loyalty, evaluating service quality, customer relationship management, etc. (Bhattacharya & Srivastava, 2020). In the age of technology and e-commerce, customer experience management is also expanded to many other goals, such as CX online, social network analysis, customer data analysis, business intelligence, hot trend prediction, content marketing, omi-channel engagement, and big data analysis (Blomberg-Nygard & Anderson, 2016; Gao et al., 2020; Holmlund et al., 2020).

The developments of e-commerce, information technology, and the popularity of social networks have made companies try to improve customers' experiences and their interaction with brands anytime, anywhere. The online experience of the customer is determined by every contact point such as Facebook, Twitter, website, and apps that the customer chooses to use to interact with the firm (Alamsyah et al., 2018; Jaiswal & Singh, 2020). In the fast-developing digital economy today, big data analytics is considered a potential tool that helps companies and organizations to understand the customer's feelings and customer journey and then make decisions to

DOI: 10.1201/9780367854690-9

improve customer experience (Kandampully et al., 2018; Alrawadieh & Law, 2019; Chen et al., 2019; Holmlund et al., 2020). Big data technology has opened a new door for analyzing large amounts of data automatically without too much human effort. Ignoring the complexity of simple calculations, computers can support maximum algorithms and task cycles in the shortest possible time (Alamsyah et al., 2021). Enterprise business data can also be based on these big data technologies for analysis. It can help to render new reports or future trends or competitor analyses. In an era where speed and information will dominate the market today, the use of new modern technologies applied in business is what companies are always looking for in order to maximize profits (Aluri et al., 2019; Alamsyah et al., 2021).

Tourism is a sensitive industry with technology, the trend of tourism is not out of the general trend of global e-commerce. Many previous studies highlight the importance of the Internet in the tourism sector (Antonio et al., 2018; Alrawadieh & Law, 2019; Kim et al., 2020). Growing international tourism and travel has increased competition among tourism destinations throughout the world (Kandampully et al., 2018). In the past, the opportunity to identify a destination was difficult, and mainly through travel agencies, today the choice of a destination is easier. Booking tours is also becoming easier through technology devices such as cell phones, social networks, and Internet-connected tools (Duan et al., 2016). The websites that support the promotion of online travel products and booking are booming, including popular websites such as booking.com, Agoda, and Tripadvisor. Online communities and social networks bring opportunities for customers: they can share their experiences, and their feelings, and interact with information (Duan et al., 2016). These online communities have transformed consumers' behaviors, and companies need to hold the chances to access information, enhanced communication, and improve social networking. In the online world, guests can find information about products or services, and reviews of experienced guests that support them in making decisions (Diéguez-Soto et al., 2017). These reviews are a great help for the next guests looking for information about hotels, tours, and other travel products (Jaiswal & Singh, 2020). Many hotels have begun to be interested in guest reviews because it can help them attract more tourists to know them and book their services. Managing guest experience which measures customer satisfaction has become a task for the hospitality industry in recent years (Kandampully et al., 2018). Mining online guest review data in the hospitality industry has attracted a lot of attention in recent years (Alrawadieh & Law, 2019). The power of this data implies part of the meaningful information to make business easier. The feature of this data is unstructured data or textual information. Hence, text mining was considered a perfect technology for extracting hidden information from these reviews. The text mining technique is considered a powerful tool when working with reviews in text form. It mines and discovers hidden information from the text (Tetzlaff et al., 2019). When combined with machine learning and statistic, text mining techniques can deal with a large amount of text as customer reviews on the Internet (Tetzlaff et al., 2019).

Currently, hotel managers think that collecting data on online booking rates, website hotel presence, online traveler presence, and social media use rate of visitors, analysis of shared content is still a pending topic, and it is only mentioned in academic studies (Tetzlaff et al., 2019). Therefore, despite the growing research in this field, but there are still many shortcomings in putting it into practice. A general model for exploiting customer experience in the hotel industry with text mining technology has not been established, and studies are still fragmented in different directions. In Klaus's study, a conceptual model for customer experience quality was proposed, which included research problems such as customer loyalty, customer satisfaction, and word-of-mouth (Klaus & Maklan, 2013). Several studies have expanded on this model for the hotel industry, but most of their conceptual models focus on measuring guest satisfaction (Diéguez-Soto et al., 2017; Alrawadieh & Law, 2019; Chen et al., 2019; Li et al., 2020). To enhance the online customer experience management in the hotel industry, further research is required to analyze different aspects, including hotel star ratings. This study's aim is to synthesize previous studies and extend the Klaus model to design a common framework for managing the Internet customer experience in the hospitality sector.

The next part of this chapter is structured as follows: Section 2 introduces related studies on online CX and shows the trend of the guest experience in the hotel industry, and Section 3 presents our designed framework for analyzing guest experiences from online review data. Section 4 shows texting techniques that can be used in analyzing guest experience, Section 5 presents the results and discussion and Section 6 provides the conclusion.

2 THEORETICAL BACKGROUND

2.1 Customer experience management

Customer experience is the response of the customer when they interact with the product. It is composed of the cognitive, emotional, sensorial, physical, and spiritual responses of the customer (De Keyser et al., 2015). In some studies, customer experience is a key to determining customer satisfaction (Gao et al., 2020; Jaiswal & Singh, 2020). In reality, this has positive implications when considering customer satisfaction in assessing the quality of products (Sánchez-Franco et al., 2019) and studying customer behavior. Customer loyalty will help companies make a larger customer network than the initial network. Hence, a link relationship is found between customer satisfaction and then how it affects customer loyalty. Not only that, a number of studies have found another effect of CX on buying behavior as word of mouth (Kim et al., 2020). Customer experience management has become an important task for companies today. Companies want to understand the customer experience in every interaction point to design touchpoints with customers, giving customers the most enjoyable experience possible. In many studies, customer experience management is the key to leading to customer satisfaction or customer loyalty (Aluri et al., 2019; Fida et al., 2020). The study by Klaus and Maklan (Klaus & Maklan, 2013) proposed a model (called CX). In this model, a scale measure of the customer experience model has four dimensions: moments of truth, outcome focus, product experience, and peace of mind. In which, product experience emphasizes customer behavior, the outcome focus on reducing customer's transactional costs, and moments of truth mention the importance of service recovery and flexibility. Peace of mind mentioned the positive feelings of the customer during all the interactions with the service provider: from before, during, and after the purchase process of the service. Klaus & Maklan also proves 03 directions to measure customer experience management quality includes: (1) Customer experience impacts positive customer satisfaction; (2) Customer experience has a positive effect on loyalty intentions; (3) Customer experience impacts word-of-mouth customer behavior. This model is used as a future research direction for customer experience management. A lot of research mentioned this model for exploitation, including three problems: customer satisfaction; customer loyalty, and word of mouth.

Customers' experience with an e-commerce site reflects their own feeling about the perceived service, perceived goods, and overall perception (Alamsyah et al., 2018). Arijit Bhattacharya (2020) proposed a framework for Online Customer Experience (OCE), their objective was to develop a model for Online Customer Experience comprising antecedent variables; component variables; and outcome variables. To evaluate the proposed theory, they performed the test of the mediating effect of online satisfaction and finally, tested the moderating effect of an online shopper's gender on OCE. However, what the model proposes leads to online repurchase intention through the effect of online satisfaction (Bhattacharya & Srivastava, 2020).

Big data is generated from a many difference sources on the Internet, a significant portion of which comes from reviews of customers on online shopping websites. Analyzing big data to find meaningful hidden knowledge by modern methods and tools has replaced traditional statistics methods (Holmlund et al., 2020). In comparison with traditional data, Big Data is different in the volume of data; variety; velocity; veracity, and it can create potential value. Mining big data in business helps accelerate companies to transform from business traditional models into business intelligence models. Big data generated by consumers has attracted much attention because of its value as public and community data. For instance, online consumer reviews have brought a positive meaning to improving the quality of services and products. Due to the development of unstructured natural language on the Internet, it is not easy to mine this data type. Opinion miningand sentiment analysis, collectively referred to as text mining, plays an important role in big data analytics.

A series of research on text mining including identification, entity extraction, data classification data clustering also becomes more meaningful when it is applied to customer experience data analysis (Liu, 2020). Maria Holmlund et al. (2020) developed a strategic framework for customer experience management through big data analytics. Big data can improve CX by providing analytics, insights, and actions. They are certain that text mining is the core technology in solving the problems in this model. More and more companies are realizing the important role positive of customer experience (CE) through electronic word of mouth (E-WOM) in all sectors. Customers using cutting-edge digital technology today expect a personalized experience at every point of interaction. Data analysis direction based on text mining techniques is developing and will support companies to understand more and more their experience of customers. From that, managers will have strategies or policies to improve the quality of products or services. In addition, they can use online customer reviews for their promotion, sales, and increase the company's revenue (Holmlund et al., 2020).

2.2 Guest experience management in the hospitality industry

The rapid growth of the tourism industry is direct to the growth of the hotel industry, so that hotels are trying to adapt to the different demands of consumers by developing different service models (Blomberg-Nygard & Anderson, 2016). Customers' experience and customer satisfaction are always interesting topics in the hotel industry. The satisfaction of customers is essential in securing customer loyalty. They also establish a good reputation and enhance the hotel's revenue (Kandampully et al., 2018). Therefore, knowing consumer psychology is key to understanding that hotels have a competitive advantage when compared to competitors. When understanding whether their consumers are satisfied or dissatisfied with the products and services provided by the hotel, the hotel manager will have plans for improving the quality of hotel services and products. It can raise the ratio of customer return. For the hotel industry, customer satisfaction is a complicated experience in all of the customer journey. Therefore, to measure consumers' satisfaction with hotels, there are many studies that use User – Generate – Content (UGC) such as consumer reviews to analyze opinions in this regard, to better understand consumers' experience (Kim et al., 2020). Recently, the explosion of online hotel booking websites has made the hotel industry completely change its appearance. Not only innovating in the way of doing business, the competition is also becoming fiercer. Customer satisfaction and customer loyalty are becoming more and more essential (Aluri et al., 2019; Li et al., 2020). Therefore, researchers also focus on mining reviews on booking pages to understand the customer experience with the provided services. Although there are a number of other research directions that deal with issues such as developing recommender systems or ratings (Ahmed & Ghabayen, 2022). However, that only focuses on developing the support tools on the system for hotel managers and customers. The work of Frederik Situmeang (2020) proposed an innovation framework for extracting some latent dimensions and analyzing sentiment. They used advanced techniques for mining information from customer online reviews. They analyzed 51,110 online reviews with 1,610 restaurants by using latent Dirichlet allocation (LDA), and they uncovered 30 latent dimensions for determining customer satisfaction (Situmeang et al., 2020).

Alrawadieh Zaid and Law Rob (Alrawadieh & Law, 2019) have confirmed that understanding the determinants of guest satisfaction has become an interesting area in the hospitality domain research. They found influential factors that determine the satisfaction of guests with hotels via examining online reviews and then they analyze 400 online travel reviews which contain 1,664 positive sentences and 236 negative sentences. They evaluated these factors as service quality, rooms, characteristics of the hotel (location, design, price…), and food.

A study by Hongxiu Li (Li et al., 2020) discovers the factors effects of customer satisfaction with hotel aspects by extracting 412,784 reviews created by consumers on TripAdvisor. These consumers are from different cities in China; the reviews have been used to analyze customer satisfaction based on the three-factor theory: basic, excitement factors, and performance factors. The study about relationship quality in hospitality services (Sánchez-Franco et al., 2019) analyzed the term occurrence to identify the topics that are mentioned many times in customer reviews. Then they investigated the hotel services that are associated with quality. This research also focused on topic modeling, sentiment analysis, text summarization, and latent

Dirichlet allocation. They collected data from 33 hotels in Las Vegas registered with Yelp and have 47,172 reviews.

The aim of Michela Fazzolari (Fazzolari & Petrocchi, 2018) is to extract important information from review data. The main purpose is for supporting hotel providers and potential customers. They proposed a Recommender System model. For customers, it will help to improve their decision processes. For providers, this system provides a Market Analysis Tool, which helps analyzes automatically the extracted information in all the phases, after that it will suggest a marketing plan to providers, such as tuning their offers according to the geographical origin of the visitors. They tested with the dataset containing more than 150,000 reviewers with more than seven million reviews.

3 PROPOSING A FRAMEWORK FOR ANALYZING ONLINE GUEST EXPERIENCE MANAGEMENT QUALITY IN THE HOSPITALITY INDUSTRY

Based on the studies that are summarized in (2.2.2) and the customer experience model in Section 2.2.1, the authors propose a framework for online guest experience management quality in the hospitality industry, including two main directions as guest satisfaction and hotel star rating in Figure 9.1. These tasks used text mining techniques to perform.

In Figure 9.1, there are two main directions for guest experience management quality: Guest satisfaction and Hotel star rating.

3.1 Guest satisfaction

The customer's psychological and behavioral states include the following:

- Satisfaction
- Loyalty
- Word of mouth

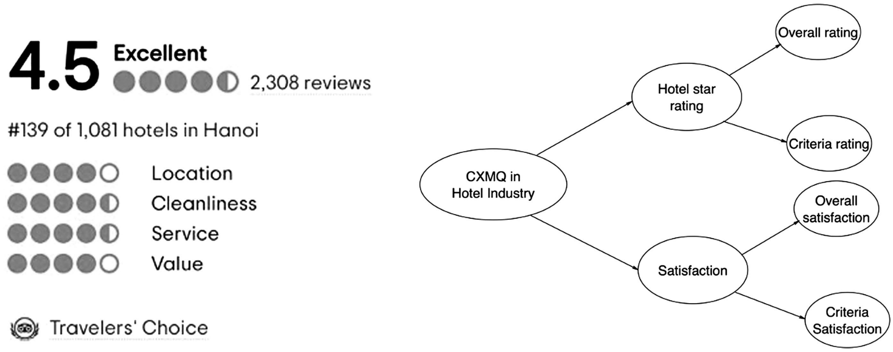


FIGURE 9.1 Framework for guest experience management quality in the hotel industry.

In psychological and behavioral states in which customer satisfaction is expressed as satisfaction with the service just used, the customers tend to perform the next level behavior such as sharing their feelings by word of mouth to relatives and friends. In the digital age, they will share their feelings on social networks (E-WOM). Customers also show loyalty by continuing to use the service again or referring relatives to use the service they have experienced. As a result, customer satisfaction is the starting point for customer follow-up activities, businesses now focus on improving customer experience across every touchpoint to enhance customer satisfaction.

A growing area of computer science includes data mining and opinion analysis, also called sentiment analysis. This is the task of automatically classifying reviews as negative or positive based on models and algorithms (Nguyen Thi Ngoc et al., 2019). The application of opinion analysis can assess customer satisfaction with the service they have experienced (Alrawadieh & Law, 2019). In the task of sentiment classification, techniques of natural language processing and text mining are used and combined with machine learning methods and statistical models to develop. The common point in many studies is to use a sentiment dictionary to measure the value of sentences, thereby classifying the positive and negative of sentences and inferring customer satisfaction. Customer satisfaction is clearly shown in positive sentences (Li et al., 2020). To create a sentiment lexicon, data will be collected from various sources, and an algorithm will be used to generate it.

The satisfaction of customers in the tourism sector is the most crucial part in customer experience management. A large number of studies are directly measuring customer satisfaction, this problem is divided into two tasks: evaluating overall customer satisfaction and analyzing customer satisfaction based on a set of criteria such as room, price, food, etc. This task uses sentiment analysis in text mining techniques:

+ Overall customer satisfaction indicates the percentage of customers who are satisfied with a hotel or a group of hotels in a certain region.
+ Satisfaction according to a set of criteria indicates the percentage of satisfied customers according to each criterion (room, staff, meal, location, etc.).

3.2 Hotel star rating

In the current hotel industry, the actual star ratings have many different standards. For instance, each country has its own standard, so when customers choose three stars, the hotel service provided is not as they think. Therefore, the star rating according to general standards based on customer reviews is also a problem that customer experience management in the tourism industry aims to solve:

+ Overall rating: Indicates star rating for the hotel, which is the average overall of the individual reviews
+ Rating for criterion: Star rating according to each criterion, such as room, location, value, etc.

4 METHODOLOGY OF CUSTOMER EXPERIENCE IN THE HOTEL INDUSTRY USING DATA ANALYSIS

4.1 Satisfaction

Based on related studies, customer satisfaction can be measured by specific values through the sentiment dictionary model. We use this vocabulary set to apply the measurement of customer satisfaction to Vietnam's tourism services. The data is collected using an automated tool for customer reviews on the TripAdvisor site and categorizes the positive and negative from the customer's free-text reviews (Figure 9.2).

The method of measuring satisfaction is carried out according to the following steps:

- Step 1. Vietnam hotel data is downloaded through an automatic data collection tool.
- Step 2. The evaluations are separated and the sentiment dictionary is used to calculate the value.
- Step 3. Categorize the reviews into two categories: Positive and Negative. The satisfaction rate of Hotel X:

$$\mathcal{S}(X) = \frac{\text{number of positive reviews}}{\text{total of collected reviews}} \times 100\%$$

- Step 4. Measure the overall satisfaction value of the hotel by the average value of the reviews according to the following formula:

$$\varphi(X) = \frac{\text{value of overall of reviews}}{\text{total of collected reviews}}$$

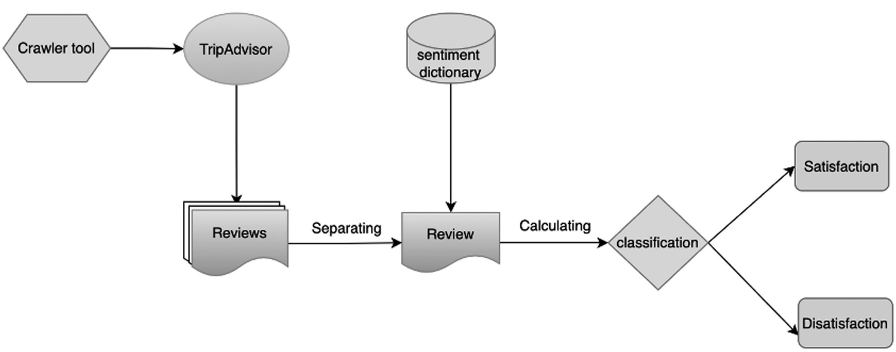


FIGURE 9.2 Satisfaction measuring model.

For example, Lotte Hanoi Hotel has a total of 358 reviews after collecting. After classification, the number of positive reviews is 276 reviews.

$$\text{Rate of satisfaction of guest } \mathcal{S}\left(\text{Lotte Hanoi}\right) = \frac{276}{358} \times 100\% = 77\%.$$

The total value of 358 reviews is +162.05. Hence the overall value $\varphi\left(\text{Lotte Hanoi}\right) = \frac{+162.05}{358} = +0.0464$.

Through the overall value, it can be seen that Lotte Hanoi Hotel has a high satisfaction rate. And the total mean for each review is +0.0464. Therefore, the hotel's overall satisfaction value is +0.0464. The higher this value, the higher the customer satisfaction with the hotel. This value can be used in hotel reviews and ratings.

4.2 Hotel star rating

Websites often let customers write their own reviews about hotels and rate themselves from 1 to 5 stars. The overall star rating for hotels on the TripAdvisor site is calculated as an average of the value of the users who have reviewed it. However, sometimes customer ratings and reviews content don't match. For example, the review content expresses satisfaction, but when rating, customers can rate 2 stars. Or some customers comment but don't rate stars. In addition, according to the customer's search, customers are interested in "location" or "room" or many different criteria. Therefore, we perform a star rating according to each criterion and finally aggregate it into an overall star rating using the following steps:

Step 1: Build a set of ranking criteria based on extracting aspects from customer reviews: room, location, check-in, service, price, surrounding, and staff.
Step 2: Identify the set of adjectives/adverbs with positive and negative meanings extracted from customer reviews.
Step 3: Assign value labels to this set of adjectives/adverbs on a scale of 1 → 5.
Step 4: Star rating according to criteria: Extract the pair of criteria + set of adjectives/adverbs to get the value of each aspect.
Step 5: Overall star rating of the hotel: Star rating according to each review

$$\text{Rank}\left(r_j\right) = \frac{\sum_{i=1}^{n}\left(\text{score}_i\right)}{n}$$

Where: r_j is the review jth;

score_i is the value of criteria ith in the review r_j
n is the total of criteria in review r_j.

Overall star rating on the set of criteria for each hotel

$$\text{Rank}(\text{hotel}_k) = \frac{\sum_{j=1}^{m}(\text{rank}(r_j))}{m}$$

And $\text{Rank}(\text{hotel}_k)$ is the ranking of hotel kth and m is the total of collected reviews.

5 RESULTS

5.1 Data

The study collected a set of data sources from TripAdvisor. These reviews are generated by customers who are experienced with Vietnamese hotels. TripAdvisor is the largest travel community, it has over 200 million monthly visits and over 500 reviews from these travelers. This website has the largest number of countries used, more than other similar websites. The TripAdvisor website is an address for travel agencies, travel companies, and guests to discover all reviews from customers – who have stayed at the hotel and have experiences with the hotels. First, we filter the Vietnamese hotels that are managed by the Vietnam National Administration of Tourism through the website http://vietnamhotel.org.vn to get a list of 3–5 star level hotels. In the next step, we use a crawler to collect data from these hotels. We store data including the name of the reviewer, the content of the review, and the date of stay. For Vietnamese hotels rated from 3 stars to 5 stars level, we take four cities including Hochiminh, Hanoi, Danang, and Nhatrang, the total number of hotels in 04 cities is 15. Each hotel has about 150–700 reviews. The total number of reviews of all hotels included 3,436 reviews.

5.2 Guest satisfaction

Tested on a dataset of 3,436 evaluations and followed the method described in Section 3. The sentiment dictionary was downloaded from Stanford's natural language processing database (https://nlp. stanford.edu/projects/socialsent/). Each word has two values, negative and positive. Table 9.1 shows some words in the Stanford's sentiment dictionary and their sentiment value.

The ratings are totaled based on the words present in them. For example, the sentence "It is smaller than I think" has a sentiment value = of −0.18. Therefore, this is a sentence expressing the negative equivalent of the customer's dissatisfaction. The performance results with the collected data set are as Table 9.2:

Figure 9.3 shows the rate of satisfaction and dissatisfaction of guests with 3 star-level hotel in Vietnam.

We perform measuring for all 15 hotels. This calculation is done simply by matching the sentiment dictionary with the set of words that appear in the data set of 3,436 reviews. Next, we only calculate the total value of those extracted words. After calculating the sentiment value of the reviews, the overall mean was +0.86614. Therefore, when measuring customer satisfaction for random hotels in Vietnam, the value is +0.86614.

TABLE 9.1
Some sentiment words and their value

Words	Negative score	Positive score
Infamous	−2.5	0.42
Stupid	−2.49	0.38
Guilty	−2.48	0.41
Foolish	−2.48	0.36
Heathen	−2.42	0.39
Pernicious	−2.41	0.71
Worse	−2.41	0.57
Silly	−2.39	0.39
Brutal	−2.37	0.47
Naked	−2.36	0.54
Despairing	−2.35	0.49
Mischievous	−2.34	0.56
Drunken	−2.29	0.46
Ragged	−2.23	0.52
Hungry	−2.2	0.53
Disgraceful	−2.19	0.51
Selfish	−2.16	0.15
Cruel	−2.06	0.15

TABLE 9.2
Results of classification in satisfaction and dissatisfaction

Evaluation	Number of reviews	Rate (%)
Satisfaction	2,848	82.89
Dissatisfaction	588	17.11
Total	3,436	100

5.3 Hotel star rating

We compare the results of hotel star ratings of hotels in Vietnam on the website http://vietnamhotel.org.vn to the customer reviews that show the difference between the hotel ratings with Vietnam standards, and online star ratings according to Vietnamese standards and according to the following formula:

$$\text{Deviation}\left(\text{hotel in VN}\right) = \frac{\sum_{k=1}^{c}\left(\text{standard\_rank}\left(\text{hotel}_k\right) - \text{rank}\left(\text{hotel}_k\right)\right)}{c} \times 100\%$$

Standard_rank (hotel_k) is the number of stars rated according to Vietnam's national standards for hotel k. And c is the total number of collected hotels. The average deviation with 23 hotels is approximately 56%. This result shows that online customers rate less than the number of stars that are rated by Vietnamese standards. The deviation is about 0.5 stars when compared to reality. Table 9.3 shows the different results between actual ratings and TripAdvisor site ratings and the article's analysis for a number of hotels in Vietnam.

The results in the table show that our rating method and the TripAdvisor rating method have less deviation than the deviation of the national standard rating and TripAdvisor site rating.

5.4 Discussion

This will lead to increased competitiveness of countries, so in the future, the hotel industry is well-positioned to take advantage of the boom in tourism. However, hotels are inherently under pressure from old business models and need to adapt and change

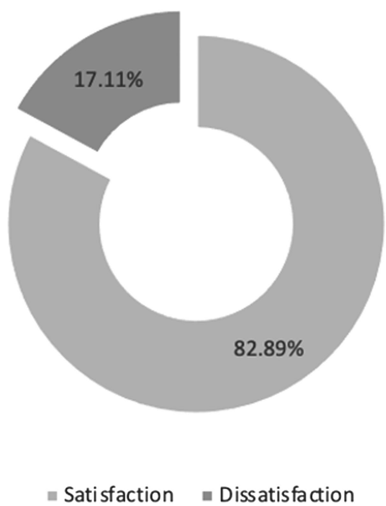


FIGURE 9.3 The rate of satisfaction and dissatisfaction of guests.

TABLE 9.3
The deviation between the national standard rating, the our rating method and the TripAdvisor site rating

Name of hotels	Hotel rating with Vietnam standard method	Hotel rating with our method	Hotel rating with TripAdvisor
Thang loi Hotel	4	4	3
Lotte Hanoi	5	4.7	5
Bao Son	4	3.8	3.5
Intercontinental Danang Hotel	5	4.5	5
Rex	5	4.3	4

more than many other types. In the age of the 4.0 Industry, booking platforms and social media all worked to bring benefits to the hotel industry. While Airbnb has disrupted and eliminated market share, at the same time Agoda, Booking, and other platforms allowed for greater reach. The impact of customer reviews is becoming more and more important in the age of social media, it brings benefits to both consumers and businesses. Travelers today rely on previous online reviews to make decisions on trip planning, while businesses use them to design marketing strategies more effectively. However, the large amount of data available makes it impractical to analyze each of the available online reviews. So, recently, a number of attempts have been made to propose methods of automatic analysis. Previous studies have applied a number of manual techniques to collect hotel data. The hotel can provide questionnaires or face-to-face communication. However, this work is limited by predesigned questions and it also lacks creativity. In addition, the data processing and analysis after collection becomes difficult and it is easy to be error-prone, while only a few hotels take advantage of these statistics for customer relationship management or service quality improvement (Gao et al., 2020).

Due to the constant evolution of information technology and automated analysis techniques, data collection has become simpler thanks to the built-in tools on the website. Hotel managers do not need to create questionnaires or face-to-face communication. These techniques come from data science and computer science fields that help in handling tasks: (1) Create a system that can support customer experience feedback (2) Automatically convert other languages into English (3) Collect automatic reviews data (4) Analyze data through text mining algorithms, big data. From here opens a new expectation for the hotel sector. The main objectives of the hotel management are promoting a brand, improving service quality, managing customer relationships, and maintaining customer loyalty (Aluri et al., 2019). These goals will be met by understanding customer psychology and customer satisfaction after experiencing the hotel services. The current technology is too perfect to support the maximum number of hotel managers. WOM on the Internet has significantly helped the goals of hotel managers set out in the era of hotel industry 4.0. Most recent research focuses on customer satisfaction measuring based on online reviews. Some other studies mentioned hotel ratings or the development of making decision systems for customers. These studies used machine learning and text mining techniques to process and analyze data (Alrawadieh & Law, 2019).

Due to the variety of issues to analyze around customer experience management, the main task of this chapter includes analyzing and summarizing research directions related to Internet customer experience management in the hospitality sector and expanding the basic Klaus's model. This model is complete and effective when techniques are described in detail for each of the Internet customer experience management tasks. It not only expands from Klaus' model, but also suggests techniques to process customer experienced data through online reviews.

6 CONCLUSION

Recent studies have focused on customer satisfaction evaluation, a trend in customer experience management, but this amount of data can also be explored for many other purposes and for improving the effectiveness of customer experience management, especially in the hotel industry.

This chapter presented a general framework for analyzing online customer experience in the hotel industry based on text mining technology. This framework covers the various tasks of customer experience management in the online hotel industry based on online reviews. The text mining techniques proposed to be able to quickly process the volume of millions and millions of records collected from the TripAdvisor page. The study applies text mining as a new suitable direction with the current e-commerce trend. These are also the general trends of many companies which want good customer experience management by understanding customers' psychology, and then companies will have specific strategies to promote products to increase sales, enhance customer relationships, and improve product quality. The article has introduced some effect text mining techniques in big data mining about customer experience management on the Internet. Text mining techniques also need to combine with some machine learning methods, and artificial intelligence to form an effective method with higher accuracy. The research directions are presented in Section 4; we will continue in more detail in the near future.

REFERENCES

Ahmed, B. H., & Ghabayen, A. S. (2022). Review rating prediction framework using deep learning. *Journal of Ambient Intelligence and Humanized Computing*, *13*(7), 3423–3432. https://doi.org/10.1007/s12652-020-01807-4.

Alamsyah, A., Laksmiani, N., & Rahimi, L. A. (2021). A core of E-commerce customer experience based on conversational data using network text methodology. In *arXiv [econ. GN]*. http://arxiv.org/abs/2102.09107.

Alrawadieh, Z., & Law, R. (2019). Determinants of hotel guests' satisfaction from the perspective of online hotel reviewers. *International Journal of Culture Tourism and Hospitality Research*, *13*(1), 84–97. https://doi.org/10.1108/ijcthr-08-2018-0104.

Aluri, A., Price, B. S., & McIntyre, N. H. (2019). Using machine learning to cocreate value through dynamic customer engagement in A brand loyalty program. *Journal of Hospitality & Tourism Research*, *43*(1), 78–100. https://doi.org/10.1177/1096348017753521.

Antonio, N., de Almeida, A. M., Nunes, L., Batista, F., & Ribeiro, R. (2018). Hotel online reviews: creating a multi-source aggregated index. *International Journal of Contemporary Hospitality Management*, *30*(12), 3574–3591. https://doi.org/10.1108/ijchm-05-2017-0302.

Bhattacharya, A., & Srivastava, M. (2020). A framework of online customer experience: an Indian perspective: an Indian perspective. *Global Business Review*, *21*(3), 800–817.

Blomberg-Nygard, A., & Anderson, C. K. (2016). United nations world tourism organization study on online guest reviews and hotel classification systems: an integrated approach. *Service Science*, *8*(2), 139–151. United Nations World Tourism Organization (UNWTO). https://doi.org/10.1287/serv.2016.0139.

Chen, M.-C., Hsiao, Y.-H., Chang, K.-C., & Lin, M.-K. (2019). Applying big data analytics to support Kansei engineering for hotel service development. *Data Technologies and Applications*, *53*(1), 33–57. https://doi.org/10.1108/dta-05-2018-0048.

De Keyser, A., Lemon, K. N., Klaus, P., & Keiningham, T. L. (2015). A framework for understanding and managing the customer experience. *Marketing Science Institute Working Paper Series*, *85*(1), 15–121.

Diéguez-Soto, J., Fernández-Gámez, M. A., & Sánchez-Marín, G. (2017). Family involvement and hotel online reputation. *BRQ Business Research Quarterly*, *20*(3), 151–163. https://doi.org/10.1016/j.brq.2017.05.001.

Duan, W., Yu, Y., Cao, Q., & Levy, S. (2016). Exploring the impact of social media on hotel service performance: a sentimental analysis approach. *Cornell Hospitality Quarterly*, *57*(3), 282–296.

Fazzolari, M., & Petrocchi, M. (2018). A study on online travel reviews through intelligent data analysis. *Information Technology & Tourism*, *20*(1), 37–58.

Fida, B. A., Ahmed, U., Al-Balushi, Y., & Singh, D. (2020). Impact of service quality on customer loyalty and customer satisfaction in Islamic banks in the Sultanate of Oman. *SAGE Open*, *10*(2), 215824402091951. https://doi.org/10.1177/2158244020919517.

Gao, L. (xuehui), Melero-Polo, I., & Sese, F. J. (2020). Customer equity drivers, customer experience quality, and customer profitability in banking services: the moderating role of social influence. *Journal of Service Research*, *23*(2), 174–193. https://doi.org/10.1177/1094670519856119.

Holmlund, M., Van Vaerenbergh, Y., Ciuchita, R., Ravald, A., Sarantopoulos, P., Ordenes, F. V., & Zaki, M. (2020). Customer experience management in the age of big data analytics: a strategic framework. *Journal of Business Research*, *116*, 356–365. https://doi.org/10.1016/j.jbusres.2020.01.022.

Jaiswal, S., & Singh, A. (2020). Influence of the determinants of online customer experience on online customer satisfaction. *Paradigm*, *24*(1), 41–55. https://doi.org/10.1177/0971890720914121.

Kandampully, J., Zhang, T., & Jaakkola, E. (2018). Customer experience management in hospitality: a literature synthesis, new understanding and research agenda. *International Journal of Contemporary Hospitality Management*, *30*(1), 21–56. https://doi.org/10.1108/ijchm-10-2015-0549.

Kim, B., Yoo, M. (myongjee), & Yang, W. (2020). Online engagement among restaurant customers: the importance of enhancing flow for social media users. *Journal of Hospitality & Tourism Research*, *44*(2), 252–277. https://doi.org/10.1177/1096348019887202.

Klaus, P. 'phil', & Maklan, S. (2013). Towards a better measure of customer experience. *International Journal of Market Research*, *55*(2), 227–246. https://doi.org/10.2501/ijmr-2013-021.

Klink, R. R., Zhang, J. Q., & Athaide, G. A. (2020). Designing a customer experience management course. *Journal of Marketing Education*, *42*(2), 157–169. https://doi.org/10.1177/0273475318818873.

Li, H., Liu, Y., Tan, C. W., & Hu, F. (2020). Comprehending customer satisfaction with hotels: Data analysis of consumer-generated reviews. *International Journal of Contemporary Hospitality Management*, *32*(5), 1713–1735.

Nguyen Thi Ngoc, T., Nguyen Thi Thu, H., & Nguyen, V. A. (2019). Mining aspects of customer's review on the social network. *Journal of Big Data*, *6*(1), 1–21.

Situmeang, F., de Boer, N., & Zhang, A. (2020). Looking beyond the stars: a description of text mining technique to extract latent dimensions from online product reviews. *International Journal of Market Research*, *62*(2), 195–215.

Sánchez-Franco, M. J., Navarro-García, A., & Rondán-Cataluña, F. J. (2019). A naive Bayes strategy for classifying customer satisfaction: a study based on online reviews of hospitality services. *Journal of Business Research*, *101*, 499–506.

Tetzlaff, S. J., Sperry, J. H., & DeGregorio, B. A. (2019). Effects of antipredator training, environmental enrichment, and soft release on wildlife translocations: a review and meta-analysis. *Biological Conservation*, *236*, 324–331.

10 Use of automated accounting information systems and operational risk in preparation of financial statements

An experimental study

Maciej Andrzej Tuszkiewicz and Ewa Wanda Maruszewska

1 INTRODUCTION

Risk has been a research topic for many years in many disciplines. In finance, up until the Global Financial Crisis (GFC),[1] risk research was mostly concentrated on identification, quantification, and measurement as well as risk management of credit and market risks (Chernobai et al., 2007; Mačerinskienė et al., 2014). The third risk – operational risk – was neglected (Chernobai et al., 2007). The research made a clear distinction between risk and uncertainty. Risk is not knowing what will happen in the future but having the ability to estimate the odds, while uncertainty is not even knowing the odds (Collier, Berry, 2002). With that, the focus of the research was only on the quantifiable risks. That is an unreasonable approach to take, as there is an unlimited situation in which assessment of the probabilities is impossible (Strzelczak, 2008). The unpredictable and uncountable events start from daily situations like a computer breakdown, through unreasonable behaviour from the employees to the energy crisis with a drastic increase of its prices. In all those situations, quantification is impossible; hence the classical risk management approach based on prior literature fails to provide a viable solution.

Since the GFC, risk research moved its focus to operational risk as many studies proved that one of the reasons for the GFC outbreak was the lack of proper operational risk management (ORM) (e.g., Andersen et al., 2012). It further resulted in the modification of regulations regarding risks issued by the Basel Committee on Banking Supervision (BCBS), Committee of Sponsoring Organizations of the Treadway Commission (COSO) and others to include or extend the regulations

 DOI: 10.1201/9780367854690-10

regarding operational risk. As indicated, the role of operational risk is expected to grow with time, as presented in Figure 10.1.

The operational risk definition specifies four main groups of operational risks: people risk, process risk, IT system risk and external risk. The above-presented classification encompasses various kinds of operational risk associated with day-to-day business activities regardless of the sector (financial vs. non-financial) or other entities' characteristics. It enables a multidimensional perspective of operational risks' attributes. It is also suitable for analysing risk linked to both the humans operating within the entities and to ERP as well as other systems helping managers make their decisions.

Further, Operational Risk Management (ORM) research proved that some types of operational risk could be assessed and quantified using top-down or bottom-up approaches (Marshall, 2001)[2] helping to handle it within business entities. The predominant opinion is that operational risk is largely a firm-specific risk (idiosyncratic) (Chernobai et al., 2007). Still, some operational risks cannot be quantified but they should be assessed. In such situations, especially considering idiosyncratic characteristics of operational risk, different methodologies developed in the literature can help define and handle various kinds of risks in a certain entity, including the operational risk arising from the use of information systems by human users.

The use of Information Systems (IS) in entities is an inherent element of business activities nowadays. Over the years, IS became more complex in terms of incorporating huge amounts of data describing complicated economic realities. Thus, IS can

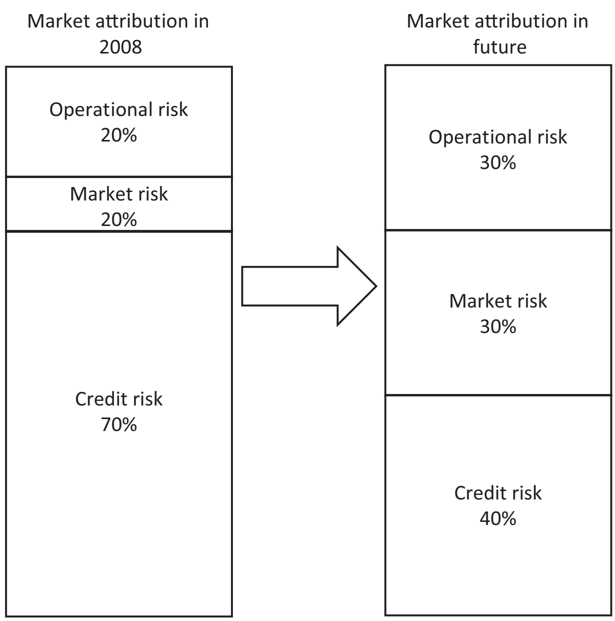


FIGURE 10.1 The growing role of operational risk.

Source: own elaboration based on (Strzelczak, 2008).

speed up and ease up many processes in entities. More advanced functions provide a deeper analysis of the entity's financial situation and become a source of business advantage if used well and purposefully. It is no different in the case of Accounting Information Systems (AIS), which allow to enter more transactions faster, create more complex analysis, provide financial reports at the hand of what is needed and synthesize information across entities within capital groups. However, despite AIS advantages, incorporating IS into decision-making processes in contemporary entities creates a risk of the wrong usage of these systems, which may lead to erroneous data, hence financial information, and in consequence, result in wrong decisions based on misleading data. Research around the world proved that beside the proper design of the system, and the implementation process, the crucial point is how its users use it. Studies have shown that even though the system may be well designed, it may not be used properly (Parasuraman, Riley, 1997) or willingly by the users (e.g., Nickerson, 1981; Swanson, 1974, 1982, 1988; Kocsis, 2019).

In this research, the authors focus on the operational risk associated with misuse of the automated AIS output by the users – an accountant in charge of financial statement preparation. Specifically, we combine people risk and IT system risk under the propensity to use erroneous data provided by automated AIS. The propensity to incorporate erroneous data into financial statements reflects uncertainties an entity faces in the accounting domain or a variety of unsystematic risks that are unique to the accounting domain present in every business entity. Therefore, we attempt to assess operational risk in our research. Moreover, as prior literature proved that there exists a risk associated with the misuse of automated AIS by its users (Klamut, 2018; Sunaryo et al., 2019), we build on our research on already confirmed relationships. By performing a 2 × 2 experiment, we try to assess the risk of incorporating erroneous data (driven from automated AIS) into the financial statement. We investigate the effect of two factors on accountant propensity to incorporate erroneous data provided by the automated AIS: the amount of workload expected to correct the erroneous data and the materiality of the data provided by the automated AIS. IT system risk in our research is represented by automated AIS output that presents erroneous data. In contrast, people risk is represented by the user, who is given the data, knows about the error in the data and decides whether to correct the erroneous data or incorporate them into the financial statement.

This chapter is organized as follows. First, to provide a theoretical background, we demonstrate previous research regarding AIS, AIS automation, risk classification, categories of operational risks, and operational risk management. Then, we present the development of our hypothesis and how we designed the research. Further, we describe the results and the respondents' characteristics. The final section discusses the results and presents the study's implications, limitations, and conclusions.

2 THEORETICAL BACKGROUND

2.1 Accounting information systems (AIS)

AIS are described as systems that involve helping decision makers get the information they want, when they want it, and how they want it (Borthwick, 1992). They are created and customized for accounting purposes (McCarthy, 1990). They provide accounting and financial information and other information derived from daily transactions (Astika, Dwirandra, 2020). AIS emerged from two disciplines: accounting and information systems (IS), as presented in Figure 10.2.

As shown in Figure 10.2, AIS touches on areas crucial for accounting and often enhances them with IS innovations. They affect work within four major areas of accounting, which are: financial accounting, taxation, auditing, and managerial accounting (Simkin et al., 2014), by using the new technology to be able to provide necessary information faster and more accurately.

Transactions entered in AIS are restricted by various regulations and accounting standards. ERP systems, which encompass AIS, represent the market's biggest and most complex AIS. They are created to unify all resources and tasks within the entity, from budgeting and planning to any ex-post reporting that may be needed. Thus, they often need to meet multiple opposing regulations to present unified financial data across many different countries to the management of the whole capital group. As a result, it is extremely difficult to design AIS that can present information with the same degree of accuracy despite input data source (Haight, 2007). That forces the constant attention of end users to overwatch system output and its exactness to look out for any possible errors.

Furthermore, AIS must be supervised by human users, who need to assess the output provided by the system and correct it if necessary. That creates operational risk as there can be mistakes made by AIS users. As a result, it is vital to understand the components affecting users of AIS and factors that lead to proper use or misuse of the output data provided by AIS, which is one of the streams of research in AIS, as presented in Figure 10.3.

Some research investigates the use of AIS. They show the research can head in myriad ways and consider the topic from many angles. A few possible routes already set in past studies are

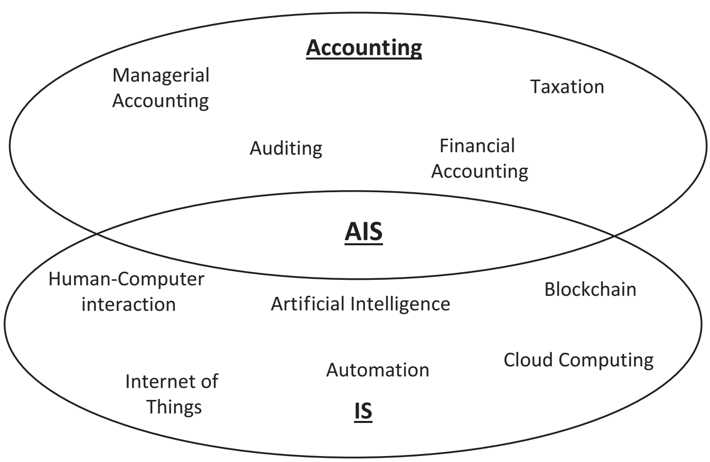


FIGURE 10.2 AIS as an emerge from accounting and IS.

Source: own elaboration.

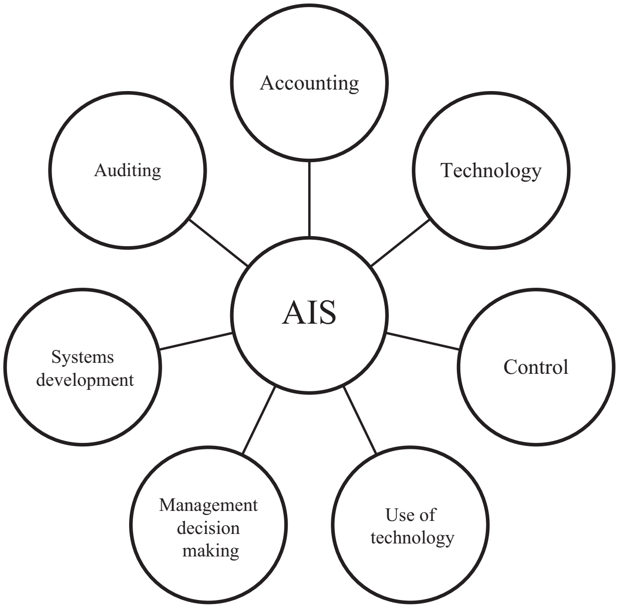


FIGURE 10.3 Directions of research regarding AIS.

Source: own elaboration based on (Gelinas et al., 2017).

- AIS users' satisfaction (Ilias, Razak, 2011; Well, 2014; Maruszewska, Tuszkiewicz, 2021),
- factors influencing ERP usage (Rajan, Baral, 2015), auditing of ERP systems (Grabski et al., 2011),
- critical success factors for ERP implementation from the user's perspective (Reitsma, Hilletofth, 2018),
- the role of accountants' trust towards the AIS on decisions (Huerta et al., 2012),
- the approach of the accounting user to automation and customization of the AIS (Tuszkiewicz et al., 2021),
- impact of AIS effectiveness on employee (user) performance (Awesejo et al., 2013; Astika, Dwirandra, 2020)

Accounting professionals' interactions with the AIS are human–machine interactions that are not always rational and are subject to many behavioural biases not limited to overconfidence in data derived from AIS, data anchoring, or loss aversion. No studies look at the matter from the point of view of risk, especially operational risk. In the authors' opinion, all the topics mentioned above investigated in the past can be a source of operational risk and, therefore, should be further investigated.

3 AUTOMATION IN AIS

Automation can be explained as an execution of a function or task by a machine, which earlier was performed by a human (Parasumaran, Riley, 1997). The range of functions covered by automation changes as time passes. Some tasks are considered automated until people change their way of thinking and consider it natural to be performed by a machine. When a task passes a point of being recognized as something obvious and natural to be automated, it becomes a machine's function. Consequently, studies regarding automation never end, as with time, the subject of the study changes, and so does its complexity and the questions that seek answers.

Many automations are created as a helping tool for AIS users to ease or speed up various accounting processes, but not to completely remove the user from the task. The reason is that the system designers cannot foresee all possible situations and possibilities of what can happen. Therefore, there must be some level of reliance on the human user (Parasuraman, Riley, 1997). The use of AIS automation can vary between AIS users, and the level of use and reliance on automation depends on many factors, for example, trust, workload, and others. Hoff and Bashir (2014) listed 127 research between 2002 and 2013 in the field of trust in automation lonely. The research within that field takes different approaches in investigating the role of trust using various methods. The effect of workload was tested and confirmed by Riley (1989), who also measured the risk associated with it, as well as Parasuraman and Riley (1997). In their findings, they noted a connection between the amount of workload and the likelihood of using automation. The operators with higher workloads were more eager to choose automation in opposition to operators with low/moderate workloads. The existing evidence regarding the effect of workload on users' decisions was so far researched within IS but not in AIS specifically.

AIS users may be given the authority to use the system output, override automation or disregard the automation. That may lead to possible errors, which were divided into three groups:

- automation disuse – a case where the user disregards the use of automation or uses only a part of it,
- automation misuse – over-reliance or under-reliance on automation that may result in a false outcome,
- automation abuse – a situation where the system is implemented to replace human operators without a detailed investigation of its implications (Parasuraman, Riley, 1997).

Within AIS, misuse can be referred to when, e.g., the accountant receives an output in the form of a complete balance sheet and decides to check every detail of it. On the opposite, misuse takes place as well when the user over-relies on AIS and never verifies its output. While it is normal to expect the system to produce correct output, the output should be periodically verified to check for any system failures. AIS disuse is when, e.g., a system has automatic schemes, but they get disabled by the accountant to operate manually. AIS abuse could be referred to as a situation when the implementation team prepares an automatic preparation of financial statement, which does not allow any adjustments and does not acknowledge the accounting department prior to it nor checks the correctness of it.

At the same, the use of new technology or change of the existing technology within the entity usually is connected with operational risk as competence and management systems need to be revised and supplemented to ensure proper use of new systems or tools (Andersen et al., 2012). Lack of proper training and no adjustments to internal processes, including control process, may result in misuse/disuse or abuse and therefore be a source of operational risk in the entity.

3.1 Risk classification

Risk can be defined as the effect of uncertainty on achieving objectives (ISO, 2009). While the word 'risk' has been used since ancient times and was always used in relation to uncertain situations (Hamberg, 2000; Picket, 2013), the interpretation of it changes with time. As per Hamberg (2000), the risk is associated with situations when the outcome probability can be measured, and when it cannot, it is defined as uncertainty. According to Kaplan and Garrick (1981), risk includes uncertainty, but it may also include loss or damage. Further, some authors point out that whatever is considered a risk at a given time may not be seen as one in the future (Cornia et al., 2014).

Research on risk was mostly focused on identifying risk, gauging it, and managing it (Kaplan, Garrick, 1981). In this stream of research, the important issue investigated is the ability to quantify the risk to calculate whether the benefits of managing (reducing it) are higher than the cost associated with it. In this approach, the risk is treated as a measure of uncertainty. It allows (1) the occurrence of the event that is positive, (2) views on the risk similar to investment – as long as calculated benefits outweigh the cost – the decision is considered optimal (Chernobai et al., 2007). However, while it looks well on paper, it led to bias in prediction and quantification. For example, research around GFC proved that many assessments were overly optimistic by either lowering the possible loss or increasing the potential profit, thus leading banks to way higher exposition on risks than the safe level (Anderson et al., 2012).

Risk assessment is part of risk management and includes identifying the risk, possible consequences, likelihood, and possible mitigations of consequences (Popov et al., 2016). Further, Popov et al. (2016) explain that consequences are the potential outcomes measured by severity while the likelihood is the estimation of chances. Risk assessment attempts to predict the worst possible situation that could take place and how likely it may occur. It is worth noting that Popov et al. (2016) argue that the estimations can be of qualitative, semiquantitative, or quantitative nature. Picket (2013) wrote: "Risks are assessed by the potential likelihood and financial impact they might have, representing a series of challenges to be met and assessed". By definition, the entity, after assessing risks, classifies them from the least to the most impactful. As a result, the entity holds a structured list of all risks from the most critical ones with high likelihood and high financial impact on the entity to the least ones allowing the entity management to make a decision accordingly to their strategy. Figure 10.4 indicates the described classification of risk determined by the potential likelihood and financial impact.

Figure 10.4 indicates that to assess the risk, one should have information about both dimensions: the measured likelihood and the description of financial impact. The above means that this approach to risk is possible only in case of risk that can be quantified in the form of likelihood, as demonstrated by prior literature on financial and operational risks. In case of human error arising from IS user interaction with the

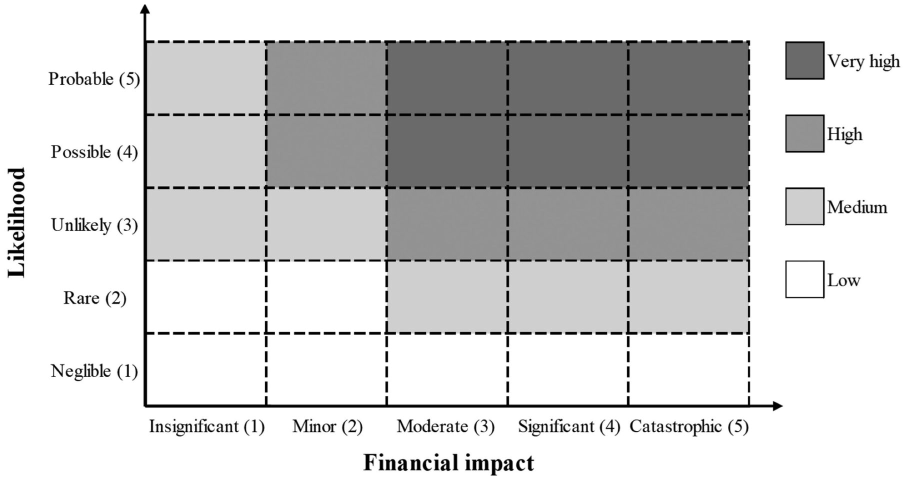


FIGURE 10.4 Risk financial impact – likelihood matrix.

Source: own elaboration based on (Samad-Khan, 2005; CIMA, 2008).

system, the likelihood of risk occurrence might be hard to precisely measure; thus, the description of likelihood as low or as a medium can serve this classification's purpose. So, discussing both AIS and users of AIS together with the likelihood of errors, it can be stated that the distinction between low and high likelihood of error can be diversified. For example, the likelihood of incorrect depreciation, foreign currency rate, or finished product calculations can be classified as low. They are based on the formulas that are basic equations in financial accounting, and little or no judgment is required during the accounting period. This results in a low likelihood of risk of erroneous data generated by the automated AIS. Apart from issues characterized by a low likelihood of errors in the automated AIS, provisions or impairment of long-term and current assets can be selected as accounting concerns with a high likelihood of errors in the automated AIS. Impairment or provisions can only partially be automated and, in many cases, require including a judgment of the AIS user, that in the case of financial statement preparation, is an accountant. Further, depending on the sector and resulting from the sector's characteristics, the financial impact of both low and high likelihood risks can have a very different financial impact. For example, depreciation amounts influence not only the net value of tangible and intangible assets but also the value of assets manufactured using certain tangible or intangible assets. Impairment of receivables, especially in the case of entities in the trading sector, can dramatically change the picture of total assets.

From the above, it can be concluded that financial data generated in the automated AIS that are the basis for financial statement preparation are subject to risk. The commencement of the risk is linked to humans, that is, users of automated AIS, and to automated AIS. The risk of erroneous financial accounting data is difficult to measure in the form of likelihood. Thus, we may use the suggested classification and describe the probability of the occurrence of a certain data error. In addition, the financial impact might be enormous, misleading the users of financial statements. Widening our knowledge of how users of automated AIS behave in response to the event of such risk seems important. It will enable taking appropriate, dedicated actions directed to risk mitigation.

4 CATEGORIES OF OPERATIONAL RISK

The financial sector developed one of the most commonly used definitions of operational risk (Strzelczak, 2008). Institutions: International Swaps and Derivatives Association (ISDA), British Bankers' Association (BBA), Risk Management Association (RMA), and Basel Committee on Banking Supervision (BCBS) define operational risk as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events (BIS, 2006). Strzelczak (2008) indicated that operational risk could be referred to as a range of possible failures in the entity's operation that is not related directly to market or credit risk. Operational risk is different from all other risks as it is related to all sorts of processes, people's decisions and behaviour, and internal and external events, therefore making it very difficult to measure (Ames et al., 2015; Muhtaseb, Eleyan, 2021). From all those definitions, the consensus is that operational risk is associated with potential management failures, faulty controls, IS errors, human mistakes, fraud, misuse of information, and further on.

Operational risk is often divided into four main categories distinguished in the definition (Chernobai et al., 2007):

- People risk – relates to all projections of loss due to errors, misuse of information and illegal actions undertaken by employees, improper organization of their work or their lack of qualification. It can involve human error, insufficient training or bad management of personnel, lack of duty segregation, and dishonesty.
- Process risk – it is the risk of loss associated with ineffective and/or inefficient processes. In other words, it is the risk you take on whenever you embark on completing a process. It may arise from products and service risk, imperfect control systems, and lack of security.
- System and technology risk (IT system risk) – it is a type of risk that should be considered in connection to other types of risks (Fheili, 2011). It may refer to the security of IS, its vulnerability to viruses and cyber-attacks, but also to poor quality and inadequacy of data provided by IS.
- External risk – it is the risk of loss due to the changes in the environment, such as changes in legislation, economics, politics, etc.

Examples of each of the risks are presented in Figure 10.5.

While the categorization presented above may exist separately, it is noticeable that the situations creating risk within a specific category may as well create another risk specified in a different type of risk. Therefore, one may say that some of the risks are closely associated with each other and specific events that may occur.

This is valid for financial accounting data derived from automated AIS that are the basis for financial statement preparation. The user of AIS can be the source of

Operational Risk			
People Risk	**IT System Risk**	**Process Risk**	**External Risk**
•Concealling losses •Misuse of information •Health & Safety •Frauds •Errors and mistakes	•Technology failure •Security •Inadequate testing •Inadequate setup •lack of maintance	•Inadequate supervision •Inappriopriate process setup •Failures in filling the reports	•External frauds •Natural disasters •Non-natural disasters

FIGURE 10.5 Categorization of operational risk.

Source: own elaboration based on (Moosa, 2007).

risk of erroneous data derived from automated AIS if the accounting system requires individual judgments to book for certain events or in the process of financial statement preparation. On the one hand, depreciation of tangibles and intangibles or manufactured goods measurement can be primarily classified as IT system risk as high automation of these calculations is achievable. On the other hand, foreign currency exchange risk can result from both: human error due to a mistake while entering the data from the invoice, as well as IT system risk arising from incorrect import of foreign currency input from the other IS that is external to the entity's AIS.

Other issues related to human risk are also valid for financial accounting conducted with the use of automated AIS. Intentional or unintentional misuse of information regarding reasons for impairment of assets, lack of IT qualifications of young or new accountants hired in the entity, or lack of segregation of duties in small and medium companies can be mentioned, among others. Regarding IT system risks, inappropriate processing methods can be mentioned in case of a lack of updating of AIS procedures, although the old procedures are no longer appropriate for the entity with new segments, products, business models, or changed manufacturing processes.

Analysis of operational risk indicates that risk linked to financial data generated in the automated AIS, that are the basis for financial statement preparation, can be classified as operational risk. The risk of erroneous financial accounting data can be categorized as human risk, mainly due to judgments required by accounting legislation. In addition, the risk of erroneous financial accounting data can be categorized as IT system risk due to the impossibility of automating certain accounting procedures for unlimited time with no regard to changing business environment. Thus, investigating how users of automated AIS change the operational risk with their decisions when working with the automated AIS seems crucial for the preparation of appropriate, dedicated actions directed to risk mitigation.

5 OPERATIONAL RISK MANAGEMENT

Operational risk management (ORM) plays a vital role not only in the finance and investment sectors but in each entity that fights to stay on the market or overtake its competitors. The task of ORM is to identify, assess, monitor, control, and mitigate operational risk in a comprehensive manner (Popov et al., 2016). Effective risk management improves stability and enhances financial performance (Hopkin, 2018; Sathyamoorthi et al., 2019), while inadequate or poorly designed risk management may lead to an economic slowdown (Fraser, Simkins, 2010; Bezzina, Grima, 2012). Unfortunately, operational risk linked to financial data generated in the automated AIS and its management did not gain much attention from researchers yet. This is surprising, especially when considering that operational risk management is viewed as a continuous responsibility of the board (Ingley, Walt, 2008) encompassing three main objectives: meeting/exceeding the organization's objectives, adhering to control-based objectives, and complying with regulatory requirements (COSO, 2013). Although the above-mentioned three objectives, at first sight, may not be considered as financial aspects of business, neglecting each may result in a financial loss at all levels of the organization (Jongh et al., 2013).

Several studies concluded that many of the economic breakdowns noted in the '90s and '00 resulted from a lack of ORM (Consiglio, Zenios, 2003; Holmes, 2003; Giraud, 2005; Chernobai et al., 2011; De Jongh et al., 2013). According to Anderson et al. (2012), the main causes of the GFC were severe violations regarding ORM, mostly due to the lack of attention towards it and its processes. Operational losses noted in 2008, at the peak of the crisis, exceeded those noted in 2007 by four times (Cagan, 2009). It is worth noting that the frequency of losses did not change significantly, but their severity (impact) increased during the crisis (Esterhuysen, 2010). This proves that research around operation risk and its management is necessary, and new tools for assessing risks as well as possible measurements of them are needed.

In the scope of accounting data and ORM, Enron, Parmalat, and WorldCom, among others, can be mentioned as examples of huge financial losses for investors and other stakeholders that resulted from neglecting the operational risk of financial data used for the preparation of financial statements. Most of the literature devoted to risk, risk management and accounting is concentrated on management accounting (Siti et al., 2011; Michalski, 2009; Soin, Collier, 2015; Hind, Derar, 2021). Operational risk linked to financial accounting with its product in the form of financial statements did not gain attention so far. It is all more surprising that financial statement analysis is considered the most related to operational risk management among product profitability analysis and total quality management. It seems that ORM within accounting data takes place only from the phase of financial statement analysis, taking for granted the data used for the preparation of financial statements. A few examples within literature examining risk management and financial accounting include the use of accounting for ORM and risk management in general (Klamut, 2018), the effect of the involvement of CFO in risk management implementation (Ojeka et al., 2019), the effect of risk management program on reduction of accrual estimation error (Johnston, Soileau, 2020), the effect of selected risk management tools on the quality of financial statement (Madu, Hassan, 2021), and the effect of auditing on financial reporting (Sunaryo et al., 2019; Mardessi, 2022). While they touch on problems within financial accounting, and some of them directly on the quality of financial reporting, they mostly discuss risk management in general, without a specific focus on ORM. In addition, although some of the research investigated a topic of people risk, none of them used methods that actually included people in the research, as their research was based on secondary data drawn from reports or financial statements.

6 HYPOTHESIS DEVELOPMENT AND RESEARCH DESIGN

The human-automation relationship has been a subject of research for many years. Examples of research are, e.g., job performance (Smith, Carayon, 1995), team effectiveness (Bowers et al., 1996), organizational behaviour (Gerwin, Leung, 1986), or the effect of automation design (Hancock, 1996). There is also a significant amount of research related to the decision-making process in IS (Swieringa, Weick, 1982; Snowball, 1986; Hannan et al., 2006; Obermaier, Müller, 2008), but to date, the research did not consider its implications to ORM. Misuse of information as a source of operational risk has been pointed out in the literature before (Moosa, 2007; Strzelczak, 2008). The lack of research around this topic may be due to the fact that research within ORM focuses on those risks that can be quantified and further insured or hedged. The crisis in the last 15 years has proven that there is a need for a wider variety of research methods in operational risk, as the to-date method does not meet the need as they were not able to measure all possible operational risks that take place during business activity.

Based on the literature review, the authors designed research to assess operational risk related to the human–machine relationship in the scope of automated AIS. We investigate the operational risk that can be categorized as people risks and IT system risks. In our experiment, IT system risk exists due to automated AIS that creates erroneous output data that was noticed by the user of AIS, the accountant. People risk is the possibility of misuse of information by the accountant, that receives erroneous data from automated AIS and makes the decision whether to correct it (right behaviour) or use the erroneous data in the financial statement. While this situation could be viewed as a process risk as well, the authors think it isn't crucial to this risk, as the process of preparation itself may be proper (and we do not examine whether there are mistakes in how the process of the financial statement is prepared). The focus of this study is on the decision of the accountant to correct or not the erroneous data received from automated AIS. As the erroneous data derived from the automated AIS and incorporated into financial statements can have a huge financial impact, we want to investigate the likelihood of the occurrence of a propensity to use erroneous data in the preparation of financial statements.

Theoretical background confirmed that past research investigated several factors influencing intention behaviour and, consequently – user decisions. Parasuraman and Riley (1997) and others mainly presented factors closely connected to the usage of the AIS. They can cause misuse, disuse, or abuse of IS – workload, understanding of the system, its appearance, trust, experience in using the system, and mistakes during the implementation. The factors that have come to light are also valid for the use of AIS. Thus, the authors decided to divide the possible factors into two categories:

- factors directly connected with AIS use,
- other factors (e.g., organizational).

Theauthors intended to include (in the research design) one factor from each category to investigate the main effect on the user's decision as well as to examine the joint effect. Out of many different factors, the authors decided to choose two.

The first factor is workload which is understood as an effort required to correct the erroneous output data derived from the automated AIS. A study investigating users who obligatorily have to work with IS was performed before, but it concerned simple classification automation (Parasuraman, Riley, 1997) or pilots (Harris et al., 1993). Therefore, we concentrate on AIS users who obligatorily have to work with partially or fully automated AIS, which are complex IS for performing their job tasks where they often need to multitask.

The second factor in our experiment is the materiality of data. In research regarding operational risk, the important piece is the assessment of the size of the risk and its possible implications of it. Using data that materially differs from the actual value that should be entered into the financial statement harms the quality of the financial statements and their reliability. The recipients of the financial statement, due to material differences, could make a different decision than they would normally make if they had a reliable and accurate financial statement available. While it is impossible to show an exact value of loss due to material differences in the financial statement, it is clear that accurate financial statements influence the value of an entity and its market capitalization, among other things that largely depend on financial information. Any published information about the unreliability of data in the financial statements results in a drop in the entity's market value, thus creating an actual loss of investors (a few examples from the past: Enron, General Electric Co., WorldCom, Lehman Brothers). Furthermore, the size of the error in the financial statements may influence the reaction size. Small errors – low material – will most likely result in a small drop in market value, if at all. In contrast, significant mistakes in the published financial statement may result in a large drop in market value, hence a large loss and impact on the entity's activity. Including the materiality of AIS output as one of the factors investigated in accountants' propensity to incorporate erroneous data into financial statements will show whether the operational risk associated with using erroneous AIS output data is affected by low/high materiality. Depending on the result, the implications for practice are different. Suppose the materiality of error is significant for the preparation of the financial statement. In that case, the mitigation of the risk should focus on a better control process regarding transactions of material value. In contrast, if the materiality of data is not significant for the decision-making process, then the processes for minimizing the risk may disregard this factor, as the cost of more restricted control will not result in risk reduction, which is counter cost-effective.

Based on the above theoretical review, we formulated three hypotheses:

H1: Accountants will be more likely to accept erroneous data incorporated into the financial statement when the workload to correct the automated AIS output is heavy (main effect of workload).

H2: Accountants will be more likely to accept erroneous data incorporated into the financial statement when the materiality of the automated AIS output is low (main effect of materiality).

H3: There is a joint effect of workload and materiality on the propensity to use erroneous AIS output. Specifically, we assume that when the workload is low, and materiality is high, the accountants will show the lowest propensity to use erroneous AIS output.

7 RESEARCH METHOD

In order to test the hypotheses, we developed a questionnaire. We conducted a computer-assisted web survey using Lime Survey software. The data gathered in the experiment were stored in MS Excel and uploaded into SPSS to perform the required tests. We used frequency, descriptive statistics, and univariate analysis of variance ANOVA for statistical analysis purposes. The significance threshold was set at 0.05.

The questionnaire consists of three sections: the experimental scenario, the manipulation check, and the demographics. The first section: experiment resembles other experiments aimed at decision-making by accountants and managers and is based on methodological clues described in Swieringa and Weick (1982), Snowball (1986), and Obermaier and Felix (2008). The aim of the experiment was to investigate whether the workload (low/heavy) and the materiality (low/high) influence the propensity to use erroneous AIS output in the financial statement preparation process. We conducted a 2×2 fully crossed factorial experimental design with the dependent variable defined as the propensity to use erroneous AIS output and two independent variables: low/heavy workload and low/high materiality of data used in the financial statement preparation process.

Participants of our laboratory experiment were provided with a scenario describing their role as an accountant working with automated software such as ERP or other software dedicated to financial accounting. Participants played the role of an accountant, in a manager position, who oversees financial statement preparation. The preparation of the financial statement is mainly based on the data provided by AIS that is automated due to algorithms developed in accordance with the assumptions adopted in the entity's accounting policy. During the actions aimed at financial statement preparation, the accountant noticed that the value of accounts receivables is overstated due to omitted write-offs of some customers. The identified erroneous data fall within the scope of the financial statement, and thus, the accountant should decide whether he/she will use the output data as generated by the automated AIS or will make an effort and will correct the data.

The first experimental group was presented with a sentence saying that the omitted write-offs' materiality is low, while the workload needed to correct the error is heavy and will result in neglecting other employee duties. Participants assigned to the second experimental group read the scenario with the information about low materiality and low workload required to correct the error, while in the third group, both the workload and the materiality were heavy/high. Finally, in the fourth experimental group, respondents were informed about the high materiality of error found and about the low materiality of the omitted write-offs. All scenarios were of descriptive character, with no numerical information anchoring the respondents on the value of receivables or write-offs to eliminate cognitive bias. As a result, the scenarios developed for this study enable us to investigate people risk (users of the automated AIS) in the form of accountants' decisions about the incorporation of erroneous data into financial statements.

Since the materiality of the error found was described as material with a description of low or high, the expected decision was to correct the value of receivables and book for the write-offs that were omitted. The decision to correct the erroneous automated AIS output is in accordance with the underlying assumptions of the financial accounting that are valid for all entities that are obliged to follow the accounting legislation. In contrast, the decision to use the erroneous automated AIS output indicates people risk resulting in the incorporation of erroneous data into financial statements. In order to make a decision, the respondents were asked a question:

> Indicate if you would choose to use the data provided by the system for the preliminary version of the financial statement prepared for the auditor OR you choose to manually correct the data provided by the accounting system. Use a scale from 1 to 10 to answer this question, where: 1–5 indicates "I would use the data provided by the accounting system for the preliminary version of the financial statement" and 6–10 means that "I would correct the data provided by the accounting system for the preliminary version of the financial statement".
>
> Marginal responses were:
>
> 1 – means that you would definitely use the data provided by the accounting system for the preliminary version of the financial statements
>
> 10 – means that you would definitely correct the data provided by the accounting system for the preliminary version of the financial statements.
>
> The degree of decisiveness (regarding the use of data or their improvement) decreases as you approach the middle of the scale.

After the experimental scenario, the respondents were presented with three questions used as verification of the cause-and-effect relationship between the scenario and the research outcome

1. Accounting information systems can generate incorrect reporting information.
2. The workload required to correct the information provided by the accounting information system may affect the decision of whether accountants will undertake to correct the information generated in the system.
3. The degree of materiality of accounting data influences the decision to manually correct their recognition in the financial and accounting information system.

The degree of compliance with the above-mentioned statements was marked by respondents using a 5-point Likert scale where "1" indicated strong disagreement and "5" indicated strong agreement. Participants who answered "1" or "2" to at least two questions were eliminated from the group of respondents.

Finally, respondents were asked about some demographic characteristics such as gender, experience with AIS, and professional undertakings in the scope of accounting.

Participation in the study was voluntary, and no remuneration was offered to participants. Participants were randomly assigned to four experimental groups.

7.1 Research results

We gathered 141 questionnaires collected from Polish accountants working with AIS. The final number of 138 answers was based on answers to three questions verifying the cause-and-effect relationship between the scenario and the research outcome

(Table 10.1) and verification of respondents' declaration in the scope of accounting education and professional experience with AIS (Table 10.2).

Three participants did not agree with two or three statements presented in Table 10.1 and were excluded from further analysis. The demographic data of 138 participants are presented in Table 10.2. No differences were found among respondents in the scope of demographic characteristics.

Among demographic data, respondents answered an open question indicating the name of the AIS software they are working with. Numerous respondents declared working with SAP, Comarch ERP, ENOVA, or Microsoft Dynamics. Others mentioned Insert, Symfonia, Rachmistrz, Optima, or Impuls. All the above-mentioned accounting software represents automated AIS.

Respondents who proclaimed an experience with automated AIS also in majority declared professional experience in finance or accounting departments. Most of our young respondents were women (79.7%) occupying lower-level (37.0%) and middle-level positions (35.5%), mainly in micro (26.1%) and small entities (25.4%).

We conducted the one-way analysis of variance (one-way ANOVA) to determine whether there were any statistically significant differences between the means of four experimental groups. The dependent variable was each respondent's propensity to incorporate erroneous data provided by the automated AIS for the preliminary version of the financial statement prepared for the auditor. Indicating "1" respondents showed the decision to withdraw from the correction of the erroneous data found and confirm the acceptance of erroneous data provided by the automated AIS. In other words, choosing "1," their decision was not compliant with the accounting underlying theoretical foundations and the accounting act, thus confirming the people's risk due to intentional or unintentional human error. Specifying "10" implied a propensity to correct the erroneous data provided by the automated AIS, that is, specifying professional decisions in accordance with both the accounting theory and legislature.

TABLE 10.1
Verification questions

	N	Min	Max	Mode	Mean	SD
Accounting information systems can generate incorrect reporting information	141	1	5	4	3.539	1.186
The workload required to correct the information provided by the accounting information system may affect the decision of whether accountants will undertake to correct the information generated in the system	141	1	5	4	3.829	1.068
The degree of materiality of accounting data influences the decision to manually correct their recognition in the financial and accounting information system	141	1	5	4	4.028	1.034

Where: 1 – definitely disagree, 2 – disagree, 3 – neither disagree nor agree, 4 – agree, 5 – definitely agree.

TABLE 10.2
Respondents' demographic data

	Frequency	Percentage		Frequency	Percentage
Professional experience			**Years of professional experience**		
I have already worked/I am currently working in the accounting department	101	73.2	Up to one year	51	37.0
I have no experience in the accounting department, but I have experience working in the financial/economic finance department	16	11.6	From one to four years	47	34.1
I have no experience in accounting, finance, or economics	4	2.9	Above five to nine years	20	14.5
Other (e.g., banking, insurance)	2	1.4	Above ten years	20	14.5
No experience	15	10.9			
Size of entity the respondent gained his/her experience			**Occupied position**		
0–10 employees (micro-enterprise)	36	26.1	Lower-level employee	51	37.0
>10–50 employees (small enterprise)	35	25.4	Middle-level employee	49	35.5
>50–250 employees (medium enterprise)	21	15.2	Middle-management employee	16	11.6
>250 employees (large enterprise)	31	22.5	Top management, owner	7	5.1
Gender of respondents			**Education**		
Female	110	79.7	Elementary	1	.7
Male	26	18.8	High school	57	41.3
Do not want to inform	2	1.4	Bachelor's degree	39	28.3
			Masters' degree	41	29.7

Descriptive statistics for all four groups show that the propensity to use erroneous AIS output is rather low among respondents, as most of them chose to manually correct the data derived from AIS. When the workload required to correct the data derived from AIS was heavy, the propensity to use erroneous data was 5.385, while when the workload was low, the mean was 7.51 ($F(1, 137) = 20.493$; $p < 0.001$). The

TABLE 10.3
Mean propensity to correct the error found for workload and materiality variables (including standard deviation and N)

	Workload		Totals
	Heavy (2)	Low (1)	Mean/SD/N
Materiality	Mean/SD/N	Mean/SD/N	
High (2)	(3) 5.37/2.921/35	(4) 7.55/2.488/34	6.42/3.011/69
Low (1)	(1) 5.40/2.982/35	(2) 7.47/2.699/34	6.44/2.913/69
Totals	5.385/2.930/70	7.51/2.577/68	6.43/2.952/138

Where: "1" – I would definitely use the data provided by the accounting system for the preliminary version of the financial statements, while "10" – I would definitely correct the data provided by the accounting system for the preliminary version of the financial statements.

TABLE 10.4
Two-way analysis of variance results for respondents' correction decision

Source	DF = degrees of freedom	SS = sum of square	MS = mean square	F = F-value	Pr > F
Model	3	156.489	52.163	6.738	<0.001
Error	138	1,037.424	7.742		
Corrected total	137				R-sq = 0.131

TABLE 10.5
Test of hypotheses 1–2 using the type III MS for the subject as an error term

Source	DF	Type III SS	F	p
Workload	1	156.342	20.194	<0.001
Materiality	1	0.031	0.004	0.950
Workload × materiality	1	0.118	0.015	0.902

materiality did not influence the propensity to use erroneous data, as the mean was 6.42 when materiality was high and 6.44 when materiality was described as low ($F(1, 137) = 0.003$; $p = 0.954$). Table 10.3 also indicates that respondents showed the highest propensity to correct output data when the workload was low and the materiality was high, as the assumption building **H3**.

The results presented in Tables 10.4 and 10.5 confirm the significance of the whole model and confirm **H1**, stating that accountants are more likely to accept erroneous data incorporated into financial statements when the workload to correct the automated AIS output is heavy. The results of the experiment do not confirm **H2** assuming that accountants will be more likely to accept erroneous data incorporated into financial statements when the materiality of the automated AIS output is low. Results also negate **H3** assuming the joint effect of workload and materiality.

In addition, we asked respondents about the most important responsibility that influences their decision about the use of erroneous data/about the manual correction of data derived from AIS. Respondents were given three possible answers: legal responsibility based on the accounting act, employee responsibility towards the supervisor, and ethical responsibility derived from the code of ethics. There was also the fourth possibility to write the respondent's own answer. The overwhelming majority indicated legal responsibility (n = 103; 74.6%). Similar numbers of respondents decided that employee responsibility (n = 15; 10.9) and ethical responsibility (n = 17; 12.3%) are the most important for the decision-making process in the described scenario. Three respondents chose to write that all three responsibilities are equally important (n = 3; 2.2).

7.2 Implications for practice and research

At present, the work of accountants is inevitably linked to the use of AIS. In times of numerous business operations, the use of automated systems helps not only to work faster and more efficiently but also seems the only possible way to embrace all data about the entity. Thus, deepening our knowledge about the relationship between the accountant and the AIS seems crucial to find out what the current accounting practice looks like and what risks are associated with it.

We contribute to the prior AIS and ORM literature by indicating that there exists an operational risk associated with the misuse of automated AIS output by users – the accountants in charge of financial statement preparation. Drawing on the advantages of the experimental method, which allows isolation of the effects of manipulated stimuli on user responses from other confounding variables and thus unveils relationships, we showed that the propensity to use erroneous AIS output is rather low among respondents gathered from accounting professionals working in Poland.

Although most of the respondents decided to correct the erroneous data derived from the automated AIS, we confirmed the main effect of workload, which suggest that automated AIS can make their users careless, generating an operational risk of data misuse. Respondents in our experiment showed a higher propensity to use erroneous data for the preparation of financial statements when the workload required to correct it was heavy compared to low-workload respondents. Further, we found no statistical significance when it comes to the materiality of erroneous data. This experimental finding is interesting, as the majority of respondents (n = 110; 78%) agreed with the statement, "The degree of materiality of accounting data influences the decision to manually correct their recognition in the financial and accounting information system". It suggests that although they acknowledge the importance of materiality as one of the qualitative characteristics of financial statements, they do not pay attention to it when it comes to decisions about the use of erroneous data, and as a result, workload prevails.

This finding is important for practice, although discouraging. Considering we haven't observed any differences among respondents based on their education or professional background, it suggests that the management of workload employees working in accounting departments is important for the quality of financial statements produced within the entity. It also suggests that despite the users' level of education or professional competence, accountants' propensity to incorporate erroneous data into financial statements is affected by the workload only. As follows, we call for a review of accounting personnel management, especially in the area of the amount of workload that the accounting personnel experience.

The assessed risk can be described mainly as people risk since it is derived from the misuse of automated AIS by its users. We did not investigate illegal actions taken by AIS users, although respondents were aware of the deterioration of the quality of financial statements in case of incorporating erroneous data into the report. Our results suggest that although accountants' propensity to incorporate erroneous AIS output into financial statements was low, the operational risk can be assessed as people, users of AIS, have an opportunity to misuse AIS. What is also important is that the misuse showed in our study is of discrete character which makes it hard to control within the entity.

Further, it can be described as an IT system risk as it is represented by automated AIS output presenting erroneous data. While we did not investigate in depth when the automated AIS may give output containing errors, it is important to note that when it does, there is a risk for the error to transfer to the financial statement. The approach to connect IT system risk with other types of risk was voiced before (Fheili, 2011). While ORM should focus on the management of users' workload, the risk can also be mitigated by increasing the quality of AIS and its output. When the chance of automated AIS giving erroneous output lowers automatically, the possibility to transfer it by AIS users lowers, as they face fewer situations in which they may intentionally or unintentionally do it.

We contributed to the literature on operational risk and ORM firstly by focusing on operational risk in financial accounting. Up to now, most of the research regarding risk in accounting has been focused on management accounting (Siti et al., 2011; Michalski, 2009; Soin, Collier, 2015; Hind, Derar, 2021). There is only a few research that focuses on operational risk in financial accounting (Klamut, 2018; Johnston, Soileau, 2020; Madu, Hassan, 2021; Mardessi, 2022), but they investigate the financial statement, which is already published, not the preparation of it. In addition, we contributed to the literature on ORM by introducing a different method of ORM – an experiment. The use of experiments in the research allows us to directly include people in the research, which is important when people's risk is investigated and even more important when we examine the risk regarding misuse or faults made by people. The majority of research on ORM in accounting investigates the issue using secondary data like financial statements or other reports (e.g., Sunaryo et al., 2019; Madu, Hassan, 2021; Mardessi, 2022). Only a small portion of them include people in the research, but with the use of questionnaires (e.g., Siti et al., 2011; Hind, Derar, 2021), which is a step forward but missing the ability to check the influence of specific factors (here, workload and materiality) on people's (here users of AIS) decisions.

7.3 Study limitations and future research

Our research is not free from limitations that should be considered when interpreting the above-described results. The first limitation refers to the small number of respondents encompassed in the study as well as their geographical homogeneity. Although our subjects were chosen from practitioners of accounting, the results should not be generalized to the whole population of accountants, although it offers insights into user misuse of accounting data derived from the automated AIS. The second limitation is linked to the method we used. As with all experimental laboratory studies, the results might only be valid in a particular setting, that is, in the accounting domain and should not be generalized to other types of technology. We used self-developed, simplified scenario of preparation of financial statements with the use of automated AIS (Maruszewska, Tuszkiewicz, 2023). Limitations of this study open up various avenues for future research.

First, for future research, this study proposes the use of experimental research in investigating the human–machine relationship in the scope of AIS. Most past research focused on the use of IS chooses exploratory factor analysis and/or confirmatory factor analysis. The result also suggests that multiple empirical approaches should be simultaneously considered in future research on human–machine relationships and their consequences for practice.

Second, further research can be directed to the role of accountants' trust towards output data derived from AIS in the case of automated systems helping in the preparation of financial statements. Also, studies investigating the subject of automation within AIS and its relationship with the decision to incorporate erroneous data into financial statements seem interesting. In our study, we used impairment of receivables, but there are many other financial statement positions that can result in different behaviour of users of AIS. Regarding AIS, we did not investigate the issue of when the automated AIS may give erroneous data. Therefore, one of the future research streams could focus on the improvement of AIS quality and its proper setup to mitigate the IT system risk and the possibility of an erroneous outcome. The focus in research may also be directed towards the use of specific AIS and whether there are differences in misuse of erroneous data between AIS systems. Future research could also focus on the approach of AIS users towards technology and possible risk assessment and mitigation in this area. The possibility of misuse of AIS output can be influenced by other factors like the approach to using technology (Venkatesh et al., 2003) or satisfaction from the use of AIS (Maruszewska, Tuszkiewicz, 2021). Finally, our current research design could be enhanced by introducing further questions that could suggest ways of mitigating the existing risks as well as modification of the experiment or using different experiment designs to better assess the risk associated with the use of automated AIS. This could be done by introducing different factors or performing the research in laboratory conditions by inviting the experiment participants and asking them to perform certain actions on specific AIS while their actions are monitored.

8 CONCLUSION

The study offers insights into AIS users' propensity to incorporate erroneous data into financial statements. In line with prior studies, we confirmed the importance of the human–machine relationship for the successful use of AIS and for high-quality financial statements. Further, we found that there can be assessed an operational risk associated with false usage of the automated AIS by the users that can directly influence the quality of financial statements produced in the entity. The combination of the widespread use of automated AIS in entities nowadays and the existing need to overwatch AIS output creates an operational risk that might be hard to measure its likelihood, but for sure, it can be described in terms of its financial impact on the entity. We hope that future research can build on the findings of this study and offer greater insights into the relationship between the user of AIS and the system itself, together with ORM suggestions dedicated to financial accounting and financial statement preparation.

NOTES

1 Refers to the period of extreme stress in global financial markets and banking systems between 2007 and 2009.
2 More information on specific methods can be found in (Finke et al., 2010).

REFERENCES

Ames, M., Shuermann, T., & Scott, H.S. (2015). Bank capital for operational risk: A tale of fragility and instability. *Journal of Risk Management in Financial Institutions, 8*(3), 227–243. https://doi.org/10.2139/ssrn.2396046.

Anderson, R.C., Duru, A., & Reeb, D.M. (2012). Investment policy in family controlled firms. *Journal of Banking and Finance, 36*(6), 1744–1758. https://doi.org/10.1016/j.jbankfin.2012.01.018

Andersen, L.B., Maberg, S., Häger, D., Næss M.B., & Tungland, M. (2012). The financial crisis in an operational risk management context – A review of causes and influencing factors. *Reliability Engineering & System Safety, 105*, 3–12.

Astika, I.B.P., & Dwirandra, A. (2020). Information technology duty, satisfaction and ais effectiveness moderate effect of benefits of informs technology on employee performance. *International Research Journal of Management, IT and Social Sciences, 7*(6), 9–20. https://doi.org/10.21744/irjmis.v7n6.996.

Awesejo, O.J., Kekwaletswe, R, M., Pretorius, P., & Zuva, T. (2013). The effect of accounting information systems in accounting. *International Journal of Advanced Computer Research, 3*(1), 142–150.

BIS, Basel Committee on Banking Supervision (BCBS) (2006). International convergence of capital measurement and capital standards. https://www.bis.orgS.

Bezzina, F.H., & Grima, S. (2012). Exploring factors affecting the proper use of derivatives. *Managerial Finance, 38*(4), 414–435. https://doi.org/10.1108/03074351211207554.

Borthwick, A.F. (1992). Helping users to get the information they want when they want it, in the form they want it: Integrating the choice and use of information. *Journal of Information Systems, 6*, v–ix.

Bowers, C.A., Randall, L.O., Eduardo, S., & Cannon-Bowers, J.A. (1996). Team performance in automated systems. In R. Parasuraman & M. Mouloua (eds.), *Automation and Human Performance: Theory and Applications*. Hillsdale, NJ: Erlbaum. https://doi.org/10.1201/9781315137957.

Cagan, P. (2009). Managing operational risk through the credit crisis. *The Journal of Compliance Risk and Opportunity, 3*(2), 19–26.

Chernobai, A., Jorion, P., & Yu, F. (2011). The determinants of operational risk in US financial institutions. *Journal of Financial and Quantitative Analysis, 46*(6), 1683–1725. https://doi.org/10.1017/S0022109011000500.

Chernobai, A., Rachev, S.T., & Fabozzi, F.J. (2007). *Operational Risk. A Guide to Basel II Capital Requirements, Models, and analysis*. Hoboken: John Wiley & Sons.

CIMA, The Chartered Institute of Management Accountants (2008). Operational Risk. Topic Gateway Series No. 51. https://www.cimaglobal.com/Documents/ImportedDocuments/51_Operational_Risk.pdf.

Collier, P.M., & Berry, A.J. (2002). Risk in the process of budgeting. *Management Accounting Research, 13*(3), 273–297. https://doi.org/10.1006/mare.2002.0190.

Consiglio, A., & Zenios, S.A. (2003). Model error in enterprise wide risk management: insurance policies with guarantee. In S. Jenkins (ed.), *Advances in Operational Risk: Firm-Wide Issues for Financial Institutions*, 2nd ed., 179–196. London: Risk Books.

Cornia, A., Dressel, K., & Pfeil, P. (2014). Risk cultures and dominant approaches towards disasters in seven European countries. *Journal of Risk Research, 19*(3), 1–17.

COSO, Committee of Sponsoring Organizations of the Treadway Commission (2013). https://www.coso.org/Shared%20Documents/COSO-McNally-Transition.pdf.

De Jongh, E., De Jongh, D., De Jongh, R., & Van Vuuren, G. (2013). A review of operational risk in banks and its role in the financial crisis. *South African Journal of Economic and Management Sciences, 16*(4), 364–382. https://doi.org/10.4102/sajems.v16i4.440.

Esterhuysen, J. (2010). The effect of stressed economic conditions on operational risk loss distributions. *SAJEMS NS, 13*(4), 476–492.

Fheili, M.I. (2011). Information technology at the forefront of operational risk: Banks are at a greater risk. *The Journal of Operating Risk, 6*(2), 47–67.

Finke, G.R., Singh, M., & Rachev, S.T. (2010). Operational risk quantification: A risk flow approach. *Journal of Operational Risk, 5*(4), 65–89. https://10.21314/JOP.2010.083.

Fraser, J., & Simkins, B. (2010). *Enterprise Risk Management: Today's Leading Research and Best Practices for Tomorrow's Executives*. Robert W. Kolb Series, Vol. 3. Hoboken: John Wiley & Sons.

Gelinas, U.J, Dull, R.B., Wheeler, P., & Hill, M.C. (2017). *Accounting Information Systems*, 11th ed. Stamford: Cengage Learning.

Gerwin, D., & Leung, T.K. (1986). The organizational impacts of flexible manufacturing systems. In T. Lupton (ed.), *Human Factors: Man, Machines and New Technology*. New York: Springer. http://doi.org/10.3233/HSM-1980-1306.

Giraud, J. (2005). Managing hedge funds' exposure to operational risks. In E.L. Davis (ed.), *Operational Risk: Practical Approaches to Implementation*. London: Risk Books, 275–283.

Grabski, S.V., Leech, S.A., & Schmidt, P.J. (2011). A review of ERP research: A future agenda for accounting information systems. *Journal of Information Systems, 25*(1), 37–78. https://doi.org/10.2308/jis.2011.25.1.37.

Haight, J. (2007). Automated control systems. Do they reduce human error and incidents? *Professional Safety, 52*(05), 20–27.

Hamberg, M. (2000). *Risk, Uncertainty & Profitability: An Accounting-Based Study of Industrial Firms Financial Performance*. PhD Thesis. Perth: Uppsala University.

Hancock, P.A. (1996). Teleology for technology. In R. Parasuraman & M. Mouloua (eds.), *Automation and Human Performance: Theory and Applications*. Hillsdale, NJ: Erlbaum, 461–497.

Hannan, R.L., Rankin, F.W., & Towry, K.L. (2006). The effect of information systems on honesty in managerial reporting: A behavioral perspective. *Contemporary Accounting Research, 23*(4), 885–918. https://doi.org/10.1506/8274-J871-2JTT-5210.

Harris, W., Hancock, P.A., & Arthur, E. (1993). The effect of task load projection on automation use, performance, and workload. In *Proceedings of the 7th International Symposium on Aviation Psychology*, 890A–890F. Ohio State University. Aviation Psychology Laboratory, Association of Aviation Psychologists. The Laboratory.

Hind, M., & Derar, E. (2021). The role of management accounting practices in operational risk management: The case of Palestinian commercial banks. *Journal of Operational Risk, 16*(4), 45–63.

Hoff, K.A., & Bashir, M. (2014). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors, 57*(3), 407–434. https://doi.org/10.1177/0018720814547570.

Holmes, M. (2003). Measuring operational risk: A reality check. *Risk, 16*(9), 84–87.

Hopkin, P. (2018). *Fundamentals of Risk Management: Understanding, Evaluating and Implementing Effective Risk Management*. London: Kogan Page.

Huerta, E., Glandon, T., & Petrides, Y. (2012). Framing, decision-aid systems, and culture: Exploring influences on fraud investigations. *International Journal of Accounting Information Systems, 13*(4), 316–333. https://doi.org/10.1016/j.accinf.2012.03.007.

Ilias, A., & Razak, M. (2011). End-user computing satisfaction (EUCS) towards computerized accounting system (CAS) in public sector: A validation of instrument. *Journal of Internet Banking and Commerce, 16*, 1–17.

Ingley, C., & Walt, N. (2008). Risk management and board effectiveness. *International Studies of Management & Organisation, 38*(3), 43–70.

International Organization for Standardization (ISO) (2009). ISO 31000:2009 *Risk Management – Principles and Guidelines*.

Jongh, E., Jongh, D.R., Jongh, R., & Vuuren, G. (2013). A review of operational risk in banks and its role in the financial crisis. *SAJEMS, 16*(4), 64–382.

Johnston, J., & Soileau, J. (2020). Enterprise risk management and accruals estimation error. *Journal of Contemporary Accounting and Economics, 16*. https://doi.org/10.1016/j.jcae.2020.100209.

Kaplan, S., & Garrick, B.J. (1981). On the quantitative definition of risk. *Risk Analysis, 1*, 11–27.

Klamut, E. (2018). Accounting and enterprise risk management. *International Journal of Contemporary Management, 17*, 233–252. https://doi.org/10.4467/24498939IJCM.18.044.10030.

Kocsis, D. (2019). A conceptual foundation of design and implementation research in accounting information systems. *International Journal of Accounting Information Systems, 34*. https://doi.org/10.1016/j.accinf.2019.06.003.

Mačerinskienė, I., Ivaškevičiūtė, L., & Railienė, G. (2014). The financial crisis impact on credit risk management in commercial banks. *KSI Transactions on Knowledge Society, VII*(1). http://www.tksi.org/JOURNAL-KSI/PAPER-PDF-2014/2014-1-01.pdf.

Madu, M., & Hassan, U. (2021). Enterprise risk management and financial reporting quality: Evidence from listed Nigerian non-financial firms. *Journal of Risk and Financial Studies, 2*(1), 43–70.

Mardessi, S. (2022). Audit committee and financial reporting quality: The moderating effect of audit quality. *Journal of Financial Crime, 29*(1). https://doi.org/10.1108/JFC-01-2021-0010.

Marshall, C.L. (2001). *Measuring and Managing Operational Risks in Financial Institutions: Tools, Techniques, and Other Resources*. Singapore: John Wiley & Sons.

Maruszewska, E.W., & Tuszkiewicz, M.A. (2021). Pomiar poziomu satysfakcji użytkowników końcowych systemów finansowo-księgowych – badanie ankietowe. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, 65*(2), 87–101.

Maruszewska, E.W., & Tuszkiewicz, M.A. (2023). The effect of supervisor's control and workload on AIS users' perceived usefulness and approach to misuse and automated system output. The moderating role of experience of AIS practitioners. In R.R. Daphne and J. Włodarczyk (eds.), *The Elgar Companion to Information Economics* (in print). Cheltenham: Edward Elgar Publishing.

McCarthy, W.E. (1990). The journal of information systems editorial guidelines—1990 (editorial). *Journal of Information Systems, 4*, iv–xi.

Michalski, G. (2009). Inventory management optimization as part of operational risk management. *Economic Computation and Economic Cybernetics Studies and Research*, 2 April 209, 213–222. Available at SSRN: https://ssrn.com/abstract=1562699.

Moosa, I.A. (2007). *The Taxonomy of Operational Risk.* Finance and Capital Markets Series. London: Palgrave Macmillan, 98–129. https://doi.org/10.1057/9780230591486_5

Muhtaseb, H., & Eleyan, D. (2021). The role of management accounting practices in operational risk management: The case of Palestinian commercial banks. *Journal of Operational Risk 16*(4), 45–63. https://doi.org/10.2131/JOP.2021.012.

Nickerson, R.S. (1981). Why interactive computer systems are sometimes not used by people who might benefit from them. *International Journal Man-Machine Studies, 15*, 469–483.

Obermaier, R., & Müller, F. (2008). Management accounting research in the lab – Method and applications. *Zeitschrift fur Planung und Unternehmenssteuerung, 19*, 325–351. https://doi.org/10.1007/s00187-008-0056-1.

Ojeka, S.A., Adegboye, A., Adegboye, K., Alabi, O., Afolabi, M., & Iyoha, F. (2019). Chief financial officer roles and enterprise risk management: An empirical based study. *Heliyon, 5*. https://doi.org/10.1016/j.heliyon.2019.e01934.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, misuse, disuse, abuse. *Human Factors, 39*(2), 230–253.

Picket, K.H.S. (2013). *The Essential Guide to Internal Auditing*. West Sussex: A John Wiley & Sons.

Popov, G., Lyon, B.K., & Hollocroft, B.D. (2016). *Risk Assessment: A Practical Guide to Assessing Operational Risks.* Hoboken: John Wiley and Sons.

Rajan, C.A., & Baral R. (2015). Adoption of ERP system: An empirical study of factors influencing the usage of ERP and its impact on end user. *IIMB Management Review, 27*(2), 105–117. https://doi.org/10.1016/j.iimb.2015.04.008.

Reitsma, E., & Hilletofth, P. (2018). Critical success factors for ERP system implementation: A user perspective. *European Business Review, 30*(3), 285–310. https://doi.org/10.1108/EBR-04-2017-0075.

Riley, V. (1989). A general model of mixed-initiative human machine systems. In *Proceedings of the Human Factors Society 33rd Annual Meeting*, 124–128. Santa Monica, CA: Human Factors and Ergonomics Society.

Samad-Khan, A. (2005). Why COSO is flawed. *Operational Risk, 6*(1), 1–6.

Sathyamoorthi, C., Mapharing, M., Mphoeng, M., & Dzimiri, M. (2019). Impact of financial risk management practices on financial performance: evidence from commercial banks in Botswana. *Applied Finance and Accounting, 6*(1), 25–39. https://doi.org/10.11114/afa.v6i1.4650.

Simkin, M.G., Rose, J.M., & Norman, C.S. (2014). *Core Concepts of Accounting Information Systems*, 13th ed. Hoboken: John Wiley & Sons.

Siti, Z.A.R., Abdul, R.A.R., & Wan, K.W.I. (2011). Management accounting and risk management in Malaysian financial institutions. *Managerial Auditing Journal, 26*(7), 566–585.

Smith, M.J., & Carayon P. (1995). New technology, automation, and work organization: stress problems and improved technology implementation strategies. *International Journal of Human Factors in Manufacturing, 5*(1), 99–116. https://doi.org/10.1002/HFM.4530050107.

Snowball, D. (1986). Accounting laboratory experiments on human judgement: Some characteristics and influences. *Accounting, Organizations and Society, 11*(1), 47–69.

Soin, K., & Collier, P. (2015). Risk and risk management in management accounting and control. *ORE Open Research Exter*, 14 April 2015, 1–13. Available at: http://hdl.handle.net/10871/16809.

Strzelczak, S. (2008). *Operational Risk Management*. Prace Naukowe: Organizacja i Zarządzanie Przemysłem, 21, 5–108.

Sunaryo, K., Astuti, S., & Zuhrohtun (2019). The role of risk management and good governance to detect fraud financial reporting. *Journal of Contemporary Accounting, 1*(1), 38–46. https://doi.org/10.20885/jca.vol1.iss1.art4.

Swanson, E.B. (1974). Management information system: Appreciation and involvement. *Management Sciences, 21*, 178–188. https://doi.org/10.1287/mnsc.21.2.178.

Swanson, E.B. (1982). Measuring user attitudes in MIS research: A review. *OMEGA, 10*, 157–165. https://doi.org/10.1016/0305-0483(82)90050-0.

Swanson, E.B. (1988). *Bridging the Gap between Design and Utilization*. Homewood, IL: Irwin.

Swieringa, R.J., & Weick, K.E. (1982). An assessment of laboratory experiments in accounting. *Journal of Accounting Research, 20*, 56–101.

Tuszkiewicz, M.A., Maruszewska, E.M., & Grzybek, O. (2021). An experimental study of the effect of AIS automation and customization on project recommendation decision. *Procedia Computer Science, 192*, 1629–1638.

Venkatesh, V., Morris, M.G., Davis, G.B., & Davis, F.D. (2003). User acceptance of information technology: Towards a unified view. *MIS Quarterly, 27*(3), 425–478.

Well, I. (2014). Manager satisfaction in using the enterprise resource planning (ERP) system and managerial performance. *Australian Journal of Information Systems, 18*, 119–135.

11 Machine learning in analytical procedures in audit

Anna Karmańska

1 INTRODUCTION

The widespread use of emerging technologies increases operational efficiencies, and the complexity and size of organizational information assets are ever-increasing (Sekar, 2022). This unavoidable process disturbs traditional audits and transforms techniques of financial analysis. The value of machine learning (ML) in financial analysis is becoming more apparent each day (Tatsat et al., 2020). The most basic of technological advancement, Robotic Process Automation, automates an audit process such as collecting data, whereas ML uses algorithms and builds models to analyze data and make correlations and predictions (with human oversight) (Hoggett et al., 2019).

One of the methods of audit, analytical procedures (APs) refer to a family of relatively inexpensive, expectation-based evidence-gathering tools available to an auditor to efficiently assure financial statements (Kwaku & Leiby, 2014).

The main objective of this study is to examine the usefulness of ML methods for analytical procedures in the audit of financial statements. For this purpose, an ML model was developed that predicts the value of costs of social securities and other benefits based on costs of salaries and wages.

As the model performs the analytical procedure using a simple algorithm, it can be applied by small audit firms without expert-level knowledge of ML. It is based on linear regression and the real dataset of firm-year observations of companies listed on the Warsaw Stock Exchange on Poland. It was implemented with the open-source programming language Python and its libraries.

The chapter focuses on addressing the following research question: Can ML models be useful in analytical procedures in an audit? According to the literature review, there is still an existing gap in the academic papers on this topic. The literature search, which was performed with the terms "machine learning" AND "audit" in bibliographic databases: Emerald Management (eight papers) and the ISI Web of Knowledge (60 papers), clearly confirms the research gap.

The chapter reads as follows: Section 1 is an introduction. The theoretical framework is presented in Section 2. Section 3 is devoted to the literature review. The methodology used in this research is described in Section 4, and the results in Section 5.

 DOI: 10.1201/9780367854690-11

Finally, the discussion is described in Section 6. Finally, Conclusions and policy implications/recommendations are drawn in Section 7.

2 THEORETICAL FRAMEWORK

A financial statement audit can be defined as the examination of an entity's financial statements by an independent auditor (Karmańska, 2022). It is an activity of collecting and evaluating evidence of financial and non-financial information to report the level of conformity with established criteria (Setiawan et al., 2021). The purpose is to enhance the degree of confidence of intended users in the financial statements by expressing an opinion to the auditor on whether the financial statements are prepared, in all material respects, by an applicable financial reporting framework (Sifa et al., 2019).

The opinion must be based on relevant and reliable evidence to reduce audit risk to an acceptable low level. The audit team performs procedures such as inspection, observation, external confirmation, recalculation, reperformance, inquiry, and analytical procedures to obtain audit evidence (ISA500.A14–25).

The International Standard on Auditing 520 (ISA 520) and the United States of America Auditing Standard 2305 (AS 2305) are pertinent to analytical procedures. The definitions of APs are presented in Figure 11.1.

Analytical procedures include among others the following techniques (Flood, 2020):

- comparisons of the current and comparable prior period(s),
- developing expectations among elements of financial information within the period,
- ratio analysis,

ISA 520	AS 2305
• APs are evaluations of financial information through analysis of plausible relationships among both financial and non-financial data. They also comprise of consideration of relationships, for example: • among elements of financial information that would be expected to conform to a predictable pattern based on the entity's experience, such as gross margin percentages, • between financial information and relevant non-financial information, such as payroll costs to a number of employees (ISA.520.A2).	• APs are an important part of the audit process and consists of evaluations of financial information made by a study of plausible relationships among both financial and nonfinancial data (Public Company Accounting Oversight Board (PCAOB), 2010). • APs involve: • comparisons of recorded amounts, or • ratios developed from recorded amounts, to expectations developed by the auditor.

FIGURE 11.1 Definitions of APs.

Source: The International Standard on Auditing 520 (ISA 520) and in the United States of America Auditing Standard 2305 (AS 2305).

- trend analysis,
- variance analysis,
- preparation of common-size financial statements, and
- regression analysis: linear and nonlinear.

The analytical procedures can reduce the cost, and increase the efficiency and effectiveness of the audit. The previous studies confirm a significant positive relationship between the implementation of analytical procedures and the quality of auditors' reports. For instance, it can be concluded from the results of studies conducted in Jordan (Al Qtaish et al., 2021) and Iraq (Matrood et al., 2019).

Analytical procedures help identify inconsistencies, unusual transactions or events, amounts, ratios, trends, and errors, such as duplicate expense claims, unauthorized expenditures, incorrect amounts, and suspicious suppliers or invoices that indicate matters that may have audit implications.

Unusual or unexpected relationships that are identified may assist the auditor in identifying risks of material misstatement, especially risks of material misstatement due to fraud (ISA315.A27).

Lachowski, a practitioner with rich experience in audit, noticed that the importance of analytical procedures has grown in the last 20 years (Lachowski, 2019). The author sees the reasons for this phenomenon in the increasing volume of business transactions carried out by enterprises, the development of advanced technologies for data processing, and the simultaneous pressure to limit the duration and cost of audits (Lachowski, 2019). Moreover, the author claims that the trend of usage APs will continue to intensify in the future (Lachowski, 2019).

Another practitioner from the audit firm BDO, Młyński classifies APs into three groups (Młyński, 2015):

- comparison to expectations – it consists of formulating the expected value for a given phenomenon, based on the knowledge of the entity's operations and the environment, the example can be a comparison of monthly costs in the current and previous year,
- examination of the main dependencies – the auditor must identify the main processes in the company and the interdependencies between them, the example can be a relationship between the costs of salaries and wages and the costs of social security,
- tests carried out with the use of advanced statistical tools – they require the auditor's advanced knowledge in econometrics and statistics as well as the use of advanced data processing programs.

According to the author, despite the fact that the usage of advanced statistical tools provides the best results, namely the highest certainty of detecting anomalies, the requirements concerning advanced knowledge make that, in practice, the first two groups of APs are rather used (Młyński, 2015).

APs are employed throughout the audit process by all audit firms, though auditors from Big 4 firms are found to use APs to a greater extent than auditors from non-Big 4 firms (Samaha & Hegazy, 2010).

The application of relevant knowledge and experience (professional judgment) is essential to the proper conduct of an audit (ISA200.A24). Using professional judgment, the auditor develops expectations by identifying plausible relationships between items of financial information that are expected to conform to a predictable pattern. For example, in a production company, the auditor can predict a higher share of fixed assets compared to a commercial company, in that current assets usually dominate. In the case of a company that finances its operations with a loan, the auditor expects interest in the profit and loss account. The margin should depend on the type of products sold, so the auditor awaits the relationship between revenues from sales and the cost of sold goods. The effective tax rate should be equal to corporate income tax of 19% or more in the case of tax non-deductible costs. Another example is seasonality, in periods of higher sales, inventories and receivables should be also higher.

Auditors use analytical procedures in any stage of the audit such as planning, execution, and conclusion. In planning, APs are adopted in preliminary analytical reviews. According to ISA 315, the auditor must perform a risk assessment to obtain an understanding of the entity's environment and the applicable financial reporting framework. For example, the auditor can examine an entity's financial performance relative to prior periods and relevant industries (ACCA, 2022). In an essential part of the audit, APs are used alongside tests of details of classes of transactions, account balances, and disclosures as substantive procedures to detect material misstatements at the assertion level. Finally, in the last stage of the audit, the auditor shall design and perform APs that assist the auditor when forming an overall conclusion as to whether the financial statements are consistent with the auditor's understanding of the entity (ISA.520.6). ACCA recommends four distinct steps inherent in the process to using the substantive APs (ACCA, 2022), that are presented on Figure 11.2:

The first step is the auditor's development of accurate and objective prediction of an amount or ratio, a percentage, a direction, or an approximation. In the next phase, the auditor should indicate the difference from the expectation that can be accepted without further investigation. In the third step, the auditor calculates the expected value and the difference between expected and recorded values. Finally, the auditor should investigate significant differences and draw conclusions based on the calculations (ACCA, 2022)

3 LITERATURE REVIEW

Adopting modern technologies is based on the technology acceptance model (TAM) that indicates two main factors: an individual's perceived usefulness is the degree to which a user believes that using the technology will improve the performance and perceived ease of use – the degree to which a user believes that using the technology will be free of effort (Davis, 1985).

ML is a key branch of Artificial Intelligence, it can be described as a broad discipline, that automatically learns correlations, patterns, and trends from historical data without being explicitly programmed by humans (Huang et al., 2022), while Supreme Audit Institutions define ML as a field of computer science dealing with methods to develop ('learn') rules from input data to achieve a given goal (Supreme Audit Institutions of Finland Germany the Netherlands Norway, and the UK, 2020).

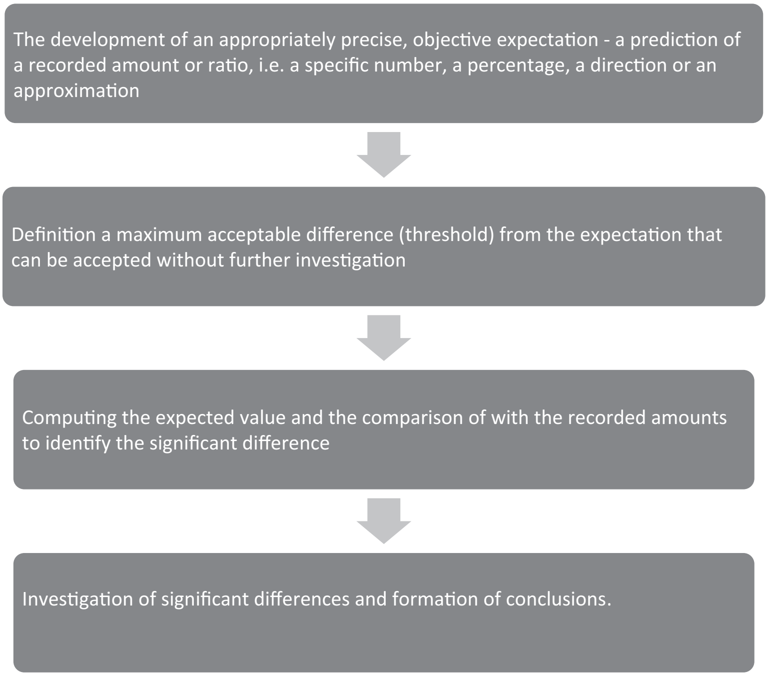


FIGURE 11.2 Four distinct steps inherent in the process of using the substantive APs (ACCA, 2022).

Source: ACCA (2022). Analytical procedures. https://www.accaglobal.com/my/en/student/exam-support-resources/professional-exams-study-resources/p7/technical-articles/analytical-procedures.html.

The technology of ML ensures powerful tools to isolate and analyze data from large databases (Damrongsakmethee & Neagoe, 2017). An audit data warehouse of historical data can be used to construct prediction models, providing suggested actions for audit procedures (Sun, 2019). These models can be used in analytical procedures to perform data analysis to understand patterns and make predictions.

Previous studies stipulate numerous benefits of implementing ML in the audit field. For instance, Ucoglu indicates the following advantages: consistency in decision-making, the ability to process huge amounts of structured and unstructured data, shorten the data processing cycle, error reduction and increased reliability of financial information, and boosted efficiency and effectiveness (Ucoglu, 2020). Another strength of algorithms compared to human work lies in reducing human error thus greater precision and accuracy are more probable (Tiron-Tudor & Deliu, 2021). Moreover, the automation of repetitive and redundant tasks significantly improves audit speed and quality (Dickey et al., 2019).

ML techniques transform the previously used methods of audit. Huang et al. claim that thanks to ML, it is possible to introduce the approach of full population testing (Huang et al., 2022). In this way, audit procedures can be executed on entire data sets rather than on samples of transactions taken from them. Full population testing eliminates the risk of sampling and improves audit quality (Huang et al., 2022).

The next significant issue of ML algorithms is enhanced communication with those charged with governance in an audit of financial statements (Elliot et al., 2020). Those charged with governance are the persons with responsibility for overseeing the strategic direction of the entity and obligations related to the accountability of the entity, for example, members of a management board, and owner-manager (ISA260.10). According to auditing standards, the auditor shall communicate with those persons significant findings from the audit (ISA260.16).

By creating complex ML-based models, auditors can also improve financial statement fraud detection. Financial statement fraud can be described as material omissions or misrepresentations resulting from an intentional failure to report financial information in accordance with generally accepted accounting principles (Hajek & Henriques, 2017). For another example, Hooda et al. analyzed the annual data of 777 firms from 14 different sectors and surveyed a wide range of ML methods such as logistic regression, Bayesian methods, decision trees, neural networks, and ensemble methods to establish a fraud early warning system (Hooda et al., 2018). Their results demonstrate an accuracy of 93% for suspicious firm classification.

There is also growing literature applying ML to predict economic indicators and trends. For example, Anand et al. employed a method from machine learning to generate out-of-sample predictions of directional changes (increases or decreases) in five profitability measures, return on equity (ROE), return on assets (ROA), return on net operating assets (RNOA), cash flow from operations (CFO), and free cash flow (FCF) (Anand et al., 2019). The authors use random forests with classification trees with a minimum set of independent variables. The classification accuracies are in the range of 57%–64%.

In another instance, Chen et al. use logistic regression and high-dimensional detailed financial data to predict the direction of one-year-ahead earnings changes. Their models show significant predictive power: the area under the receiver operating characteristics curve ranges from 67.52% to 68.66% and is significantly higher than the 50% of a random guess (Chen et al., 2022).

Meanwhile, Amel-Zadeh et al. compare a set of ML models in their ability to predict the sign and magnitude of abnormal stock returns around earnings announcements based on past financial statement data alone (Amel-Zadeh et al., 2020). The authors discovered that random forests dominated all other models and non-linear methods performed better for predictions of extreme market reactions, while the linear methods were better in predicting moderate market reactions (Amel-Zadeh et al., 2020).

ML technology is also useful in examining specific elements of financial statements such as notes. Neither standard within IFRS gives a uniform format for the notes. According to International Accounting Standard 1 (IAS 1.112), the notes should disclose any information required by IFRSs that is not presented elsewhere in the financial statements and provide additional information that is relevant to an understanding. The notes to the financial statements are usually extensive and long, moreover, results of studies show that the notes are significantly lengthier in the post-IFRS period (Cheung & Lau, 2016). Sifa et al. propose Automated List Inspection (ALI). It is a tool that utilizes methods from ML, and natural language processing, combined with domain expert knowledge to automate financial statement auditing (Sifa et al., 2019). According to the authors, the solution matches relevant text passages from the notes to the financial statement to specific law regulations, and ensures the accuracy and valuation of the reported numbers. In this way, ML techniques allow the automation of recurring and time-consuming tasks, which provides auditors more time to focus on more complex and value-added tasks. The further challenge for authors is checking the consistency of information both within the notes to the financial statement (Sifa et al., 2019).

According to their authors, all the models described in the above studies bring benefits in audit and are characterized by a high prediction ability. However, due to their complexity, in practice understanding, and usage of the models can be difficult for auditors. Though auditors must possess extensive knowledge of various fields such as accounting, finance, law, and management, in Poland the qualifying examinations for apprentices do not comprise numerical methods, statistics, or computer science. Continuously progressing technological development forces the change of curricula and hence the requirements for auditors. Many researchers indicate the need in the Digital Era 4.0, to pay attention to the knowledge, competencies, and learning of information technology related to auditing (Setiawan et al., 2021). Professional organizations, for example, Supreme Audit Institutions (SAIs), disclose the following areas of expertise for auditors:

- a good understanding of the high-level principles of ML algorithms and up-to-date knowledge of the rapid technical developments,
- ability to understand common coding languages and model implementations, and be able to use appropriate software tools,
- a basic understanding of cloud services (Supreme Audit Institutions of Finland Germany the Netherlands Norway and the UK, 2020).

The auditor is responsible for the engagement strategy and method selected (Tiron-Tudor & Deliu, 2021). Though the auditor may use the assistance of the expert, for example, a data scientist or a programmer, the sole responsibility for the audit opinion expressed rests with the auditor (ISA620.3). Therefore, the auditor should obtain a sufficient understanding of the field of expertise of the expert, determine the nature and aims of the work as well as evaluate the adequacy of that work's results for the auditor's purposes (ISA620.10). The above demands indicate that the auditor must know ML techniques if he or she wants to use them. It is noteworthy that the assistance of external experts is costly and may be beyond the reach of small businesses due to a lack of resources and competence. According to many authors, one of the most significant barriers to implementing algorithms is the high cost of their creation (Tiron-Tudor & Deliu, 2021). Additionally, the incorrect use of technology may create biases or a general risk of overreliance on the output of the audit procedure performed (IAASB, 2020).

Generally, ML techniques can be divided into supervised and unsupervised learning (Dickey et al., 2019). Supervised learning algorithms use labeled examples, which means that there are inputs with known outputs. Therefore, supervised learning is used in situations where historical data can be used to predict future outcomes

(Dickey et al., 2019), while unsupervised learning is adopted where there are no labels on the output variables (Dickey et al., 2019).

Hoogduin gives some examples of ML techniques in the audit (Hoogduin, 2019). As unsupervised learning, it can be the clustering of financial statement ratios or journal entries (Hoogduin, 2019). The most widely known supervised technique is regression analysis, which may be deployed among others in the analysis of depreciation charged against the historical cost of fixed assets, interest expense against the balance of long-term debt, or a margin analysis between revenue and cost of sales (Hoogduin, 2019).

Claudiu et al. examine the audit procedures and potential of ML algorithms and tools that can be used to enhance the quality of the financial statement audit (Claudiu et al., 2018). According to the authors, ML algorithms are beneficial in risk assessment, analytical procedures, substantive procedures, and tests of controls. For example, in analytical procedures, ML techniques of regression can be used in predicting the value of expenses and revenues as well as raw materials consumptions expectation to detect unusual amounts in financial statements (Claudiu et al., 2018).

Artificial intelligence systems based on ML models are currently under rapid development, with successful applications (Supreme Audit Institutions of Finland Germany the Netherlands Norway and the UK, 2020).

Although undoubtedly, ML is the future of audit, costs to develop and deploy the technology can be high. For instance, EY reports a US$10b investment plan for the years 2022–2024 to strengthen audit quality, through the implementation of technology-driven innovations in risk-based and other audit procedures, including the detection of fraud (Lloyd, 2021).

Especially audit firms from the Big 4 developed machine learning tools that are used for fully automated audits (in certain areas, such as cash audits), data analysis, risk assessment, and extracting information from unstructured documents such as contracts, pictures, and photos (Ucoglu, 2020).

As Handoko notices that audit firm size provides reinforcement in the affect in the context of technology and in the organizational context in the implementation of machine learning in auditing (Handoko, 2021). The author claims the big four audit firms are common with larger budgets than small firms, therefore greater investments can be used to provide support for technology adoption (Handoko, 2021).

Therefore, it is essential to develop ML models available for small audit firms with tight budgets. Such models should be designed for auditors with some knowledge of quantitative methods and based on open-source software.

4 METHODOLOGY

This research presents the fundamental knowledge on the design and implementation of a machine learning model based on selected algorithms tested on the real dataset of firm-year observations of companies listed on the Warsaw Stock Exchange.

The objective of this study is to develop an ML model to predict the value of the costs of social security and other benefits from the value of salaries and wages.

The model was implemented with the programming language Python. Thanks to its popularity, flexibility, and rate of adoption, Python is often the preferred language for ML development (Tatsat et al., 2020). In the study, many available open-source libraries and packages to support machine learning techniques were employed, among others: Pandas, NumPy, and Scikit-learn. Pandas is a library for data manipulation and analysis, NumPy allows many high-performance operations on single-dimensional and multidimensional arrays, while Scikit-learn offers a wide range of algorithms and utilities used in ML (Pedregosa Fabian et al., 2011; Tatsat et al., 2020).

The steps used in the research are the following:

- problem definition,
- load the data and packages,
- data preparation (data cleaning, data transformation),
- data analysis (calculation statistics, visualizations),
- splitting the sample into training and test sets,
- using training data to fit a model,
- making predictions based on the data,
- computing accuracy of the prediction (score),
- identification of anomalies.

The accounting records of costs are characterized by the occurrence of a high volume of similar transactions. When examining costs, the main aim of the auditor is to answer the following questions (Fedak, 1998):

1. Have the costs been presented completely, in the right amount, and in the appropriate period?
2. Have the costs been classified as a proper item in the profit and loss account?

To answer the first question the auditor usually carries out the substantive analytical procedures. When the data set is large, APs may be more effective and efficient than substantive tests of details (Appelbaum et al., 2018).

Undoubtedly, employee benefits are significant expenses in enterprises. Labor costs are the sum of gross wages and salaries (including advance personal income tax payments and contributions to compulsory retirement, disability, and sickness insurance paid by the insured employee) and non-wage expenses (including employers' retirement, disability, and accident insurance contributions as well as further training expenses, expenditures on occupational health and safety, social benefit funds) incurred by the employer to recruit, maintain, retrain and train further staff (Statistics Poland, 2021).

The primary form of employment in Poland is a fixed-term employment contract based on the Labour Code (consolidated text, *Journal of Laws of 2020*, item 1510). However, due to the principle of freedom of contract in Poland, civil law contracts (a contract of mandate and a specific-task contract) is a popular flexible form of employment, especially for young people. The flexible forms are attractive for employers because they do not provide employees with any protection or rights guaranteed by the Labour Code and the lower tax wedge associated with these contracts stems from various regulations on social security contributions (Lewandowski & Magda, 2017).

The Act of 13 October 1998 on the social insurance system (consolidated text, *Journal of Laws of 2022*, item 1009) regulates social insurance contributions in Poland. The employer (payer) is obliged to pay social insurance contributions based on the gross consideration using the following rates (The Social Insurance Institution, 2021):

- old-age pension insurance 9.76%,
- disability pension insurance 6.5%,
- and accident insurance 0.67%–3.33% (a differentiated rate for various groups of activities).

Additionally, depending on the circumstances such as age and occupation, the employer also pays contributions to Labour Fund 2.45%, Guaranteed Employee Benefits Fund 0.1%, Bridge Pension Fund 1.5%, and Employee Capital Plans 1.5%. To summarize in 2022, the contribution paid by employers for a fixed-term employment contract is within the range of 16.93% to 25.14% of the gross salary. For civil law contracts, the calculations are more complicated as some contracts are exempt from social security, for example, students up to the age of 26, copyright fees, and contracts for specific work.

According to Statistics Poland, in 2020 the main components of labor costs were personal wages and salaries at 81.1% and other costs at 18.9%, of which social insurance contributions accounted for 14.7% (Statistics Poland, 2021), i.e., on average the social insurance costs and other benefits constitute 22.3% costs of salaries and wages. Based on the above numbers, a thesis was formulated that there is a linear relationship between the costs of salaries and wages and costs of social securities and other benefits.

Simple (univariate) linear regression is a popular technique commonly used in widely studied and applied models across many fields and disciplines (Raschka, 2015). It is also used in this study to build the model that predicts the value of costs of social security and other benefits based on the value of costs of salaries and wages.

Regression is a process of predicting continuous values using previously labeled data. The relationship between the dependent variable and the independent variable or variables is defined in terms of a straight line. In determining the variables, the auditor will use knowledge of the client and previously audited historical data (Flood, 2020). Additionally, in developing regression models, the auditor also may use external independent variables, such as gross national product, disposable net income, unemployment rate, and so on (Flood, 2020).

The linear regression formula is as follows:

$$social\_security\_costs = \beta_0 + \beta_1\, salaries\_wages$$

In the ML algorithm, the relationship is trained on historical data. The regression analysis predicts the value of social security costs and other benefits and can help auditors identify unexpected outliers and anomalies.

The study employs a sample from the Notoria online database. It is a database containing an updated, standardized format of financial statements for all companies

TABLE 11.1
Description of variables

Variable	Definition
Ticker	Unique series of three letters assigned to a security for trading purposes
Market	GPW is the main WSE's regulated market
	NC – NewConnect, is the WSE's alternative trading platform for smaller and start-up issuers
Sector	A general field of business
Year	Reporting period
Framework	Financial reporting framework
	Local – the Act of Accounting
	IFRS
salaries_wages	Value of gross salaries and wages
social_security_costs	Social security and other benefits costs
share	*social_security_costs/salaries_wages*

Source: own elaboration based on analyzed data.

listed on the Warsaw Stock Exchange (WSE) (*Notoria Oferta, Dane Finansowe Dla Inwestorów*, 2022).

In the study, only the standalone financial statements were used. The data from consolidated financial statements were omitted in order not to duplicate the values. The second reason for the choice was that in the Notoria database, there were more statutory financial statements than consolidated ones. The bank and insurance sectors were not included in the developed model due to specific accounting principles and divergent forms of financial reporting. The financial statements were prepared both according to the local Polish rules (the Act of Accounting) and the International Financial Reporting Standards (IFRS). The vast majority of observations (99.5%) come from companies registered in Poland, in addition in the sample, there were observations from Germany, Cyprus, the Czech Republic, Estonia, and Luxembourg.

The data was downloaded from the database on 20.08.2022. The data of each company is in a separate Excel file, but the uniform format of spreadsheets allowed for automatic processing. The labor costs are presented in the Profit and Loss Account in the variant by function in two separate lines of operating expenses as salaries and social security and other benefits. Those lines were copied along with others such as ticker, market, sector, and year, to the file with observations.

The variables used in the study and their descriptions are presented in Table 11.1.

5 THE RESULTS

Initially, 27,360 firm-year observations were in the sample. However, in the next phase of cleaning data, observations with missing data were removed from the sample. Data is the basis upon which an ML model is designed and performed, therefore low-quality data can lead to incorrect decisions and unreliable analysis. In addition, some variables, such as *salaries_wages*, and *social_security_costs* were converted to numeric values and the variable *share* was calculated as the relationship between *social_security_costs* and *salaries_wages*. Finally, the sample consisted of 5,236-year observations from the period 2003 to 2022. In the sample, there is a total of 758 firms from 90 different sectors. To simplify the analysis, each observation was assigned to one of 11 sectors based on the Global Industry Classification Standard (GICS) (MSCI, 2022).

The next, very important, step is Exploratory Data Analysis (EDA). It always precedes machine learning algorithms (Raschka & Mirjalili, 2019). EDA involves graphics and numerical summaries and visualizations to explore data and identify potential relationships between variables.

Table 11.2 presents a description of the sample. A limited number of observations are made from the years 2003 to 2008, due to missing data the observations were removed from the sample. However, thanks to data homogeneity and a small population of observations from 2003 to 2008, quite large differences in the number of observations between years did not affect the research results. Additionally, the averages of the relationship between *social_security_costs* and *salaries_wages* (variable *share*), for market, years, and sectors were calculated.

The differences in the average *share* for sectors can be observed (Table 11.2). In Poland, in some industries for instance information technology or communication services, there are popular civil law contracts such as copyright fees, which are not subject to social security; therefore the average variable *share* is lower.

Table 11.3 presents the statistics of the numeric variables used in the study. A high variance indicates that the collected data has higher variability, i.e., the data is generally further from the mean.

Figure 11.3 shows the scatterplot matrix. It is a grid of several scatter plots of numeric variables that includes individual scatter plots for every combination of variables and shows the relationships between them.

Additionally, Figure 11.4 presents the fragment of the chart scatter plots for a combination of variables.

Figures 11.3 and 11.4 demonstrate the linear relationship between variables: *salaries_wages*, and *social_security_costs*. In further analysis, the heatmap was used to plot rectangular data as a color-encoded matrix (Figure 11.4). Additionally, to measure the linear correlation between *salaries_wages* and *social_security_costs*, the Pearson product-moment correlation coefficient was employed. The results presented in the heatmap (Figure 11.5) indicate that variables: *salaries_wages* and *social_security_costs* have a statistically significant strong linear relationship of 0.879, ($p < 0.01$). The direction of the relationship is positive (i.e., variables are positively correlated), meaning that these variables tend to increase together (i.e., greater *social_security_costs* are associated with greater *salaries_wages*).

The normality assumption was not tested, as training a linear regression model does not require that the explanatory or target variables are normally distributed (Raschka & Mirjalili, 2019).

The subsequent step is to divide the sample into two subsamples: the training sample serves to fit the model, and the test sample is used to evaluate the predictive ability of the model.

TABLE 11.2
The description of the sample

Variable	Number of firm-year observations	% of observations	Average of share
	Market		
GPW	1,847	35.28	16.68
NC	3,389	64.72	15.45
	Year		
2003	2	0.04	11.00
2004	11	0.21	17.34
2005	11	0.21	17.39
2006	8	0.15	13.96
2007	10	0.19	9.73
2008	42	0.80	13.01
2009	123	2.35	14.76
2010	252	4.81	15.17
2011	317	6.05	15.94
2012	350	6.69	15.03
2013	350	6.69	16.13
2014	447	8.54	15.91
2015	479	9.15	15.57
2016	486	9.28	16.43
2017	492	9.40	16.87
2018	483	9.23	16.05
2019	472	9.01	16.84
2020	452	8.63	15.49
2021	443	8.46	15.17
2022	6	0.11	19.89
	GISC sector		
Consumer discretionary	399	7.62	13.35
Consumer staples	359	6.86	19.71
Energy	170	3.25	16.49
Materials	406	7.75	19.96
Industrial	445	8.50	19.53
Healthcare	349	6.67	14.64
Financial	910	17.38	14.66
Information technology	930	17.75	14.14
Real estate	636	12.15	16.65
Communication services	514	9.82	13.27
Utilities	118	2.25	18.28
	Financial reporting standard		
Local	3,651	69.73	15.71
IFRS	1,585	30.27	16.27

Source: own elaboration based on analyzed data.

TABLE 11.3
The statistics of the numeric variables used in the study

	salaries_wages	social_security_costs	share
N	5,236	5,236	5,236
Mean	18,919.66	4,373.29	0.16
Variance	12,563,036,820.49	1,831,906,026.01	0.02
Std	112,084.95	42,800.77	0.12
Min	0.55	0.00	0.00
25%	379.72	37.10	0.11
50%	1,343.40	184.35	0.17
75%	4,564.35	768.45	0.20
Max	2,992,000.00	1,257,000.00	5.32

Source: own elaboration based on analyzed data.

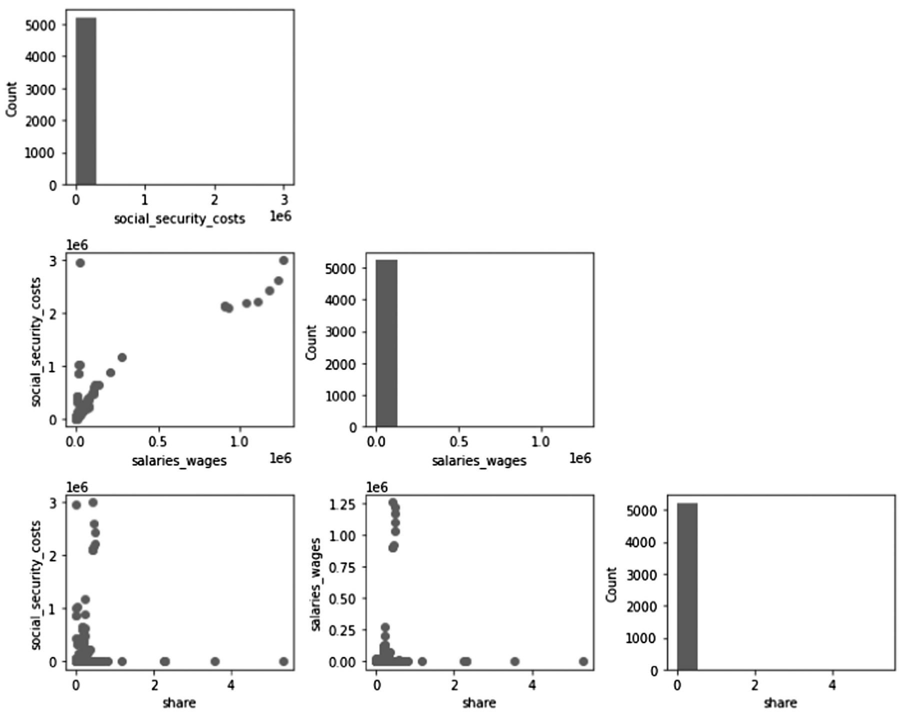


FIGURE 11.3 The scatterplot matrix.

Source: own elaboration based on analyzed data.

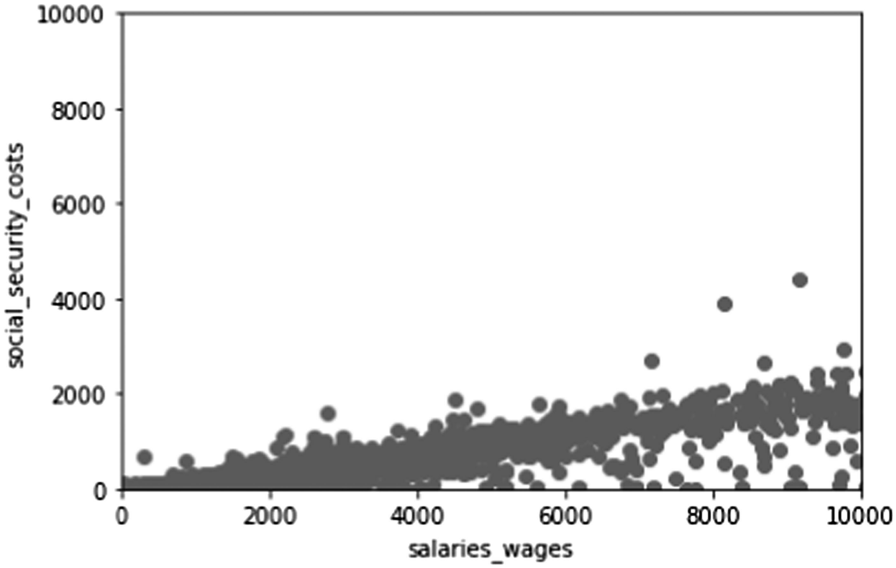


FIGURE 11.4 The fragment of the chart scatter plots for a combination of variables.

Source: own elaboration based on analyzed data.

FIGURE 11.5 The heatmap.

Source: own elaboration based on analyzed data.

One of the most important issues in developing supervised ML models is sufficient separation between training and testing (validation) data. When a part of the data is used both for shaping the model (during the training phase) and verifying the performance of the model (during the testing or validation phase), the performance metrics of the model will be inflated (Supreme Audit Institutions of Finland Germany the Netherlands Norway and the UK, 2020). Such a situation leads to 'overfitting' i.e., a loss of performance of the model on new data. As the testing dataset is not a part of the dataset used to train the model, it provides a more accurate evaluation. In practice, the most commonly used splits into training and testing sets are 60:40, 70:30, or 80:20, the split depends on the size of the initial dataset (Raschka, 2015). The model is trained on these training data (80% of the sample – 4,203 observations) and then tested on the independent new sample (20% of the sample – 1,033 observations) of testing data. To avoid overfitting the model, the split should guarantee randomness. A mask was created to select random rows of firm-year observations.

Afterward, the model is trained to predict the outputs using the package scikit-learn (Pedregosa Fabian et al., 2011). The model fits a linear with parameters: coefficient β_1 and intercept β_0 to minimize the residual sum of squares the observed targets in the dataset, and the targets predicted using the linear approximation.

The model is as follows:

$$social\_security\_costs = -1{,}280.81190291 + 0.31646039 * salaries\_wages$$

Figure 11.6 presents the linear regression as the red line for all samples, while Figure 11.7 shows the fragment of the chart.

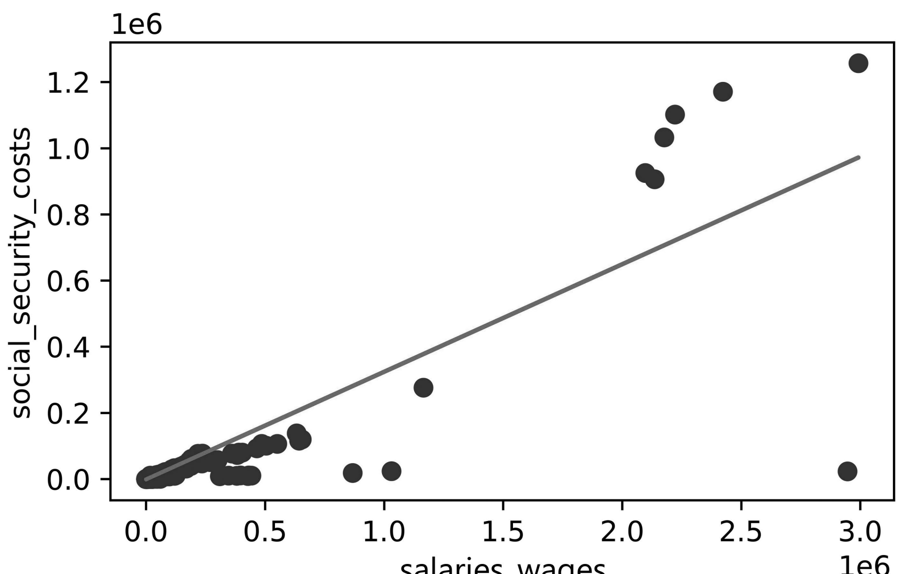


FIGURE 11.6 The linear regression.

Source: own elaboration based on analyzed data.

In the next step, the accuracy of the regression model was calculated. Evaluation metrics play a crucial role in the development of the model as they are a measure of how well a model performs and how well it approximates the relationship (Yashwanth, 2020). They also indicate areas to be improved. In ML regression techniques, different evaluation metrics are used to evaluate the accuracy of the model. However, the most common are: Mean Absolute Error (MAE), residual sum of squares (RSS), and R-squared that were calculated in the study. The results, presented in Table 11.4, confirm that the goodness of fit is high. It means that the values expected based on the model are close to the observed.

The mean absolute error (MAE) is the mean of the absolute values of the individual prediction errors over all instances in the test set. Each prediction error is the difference between the true value and the predicted value for the instance

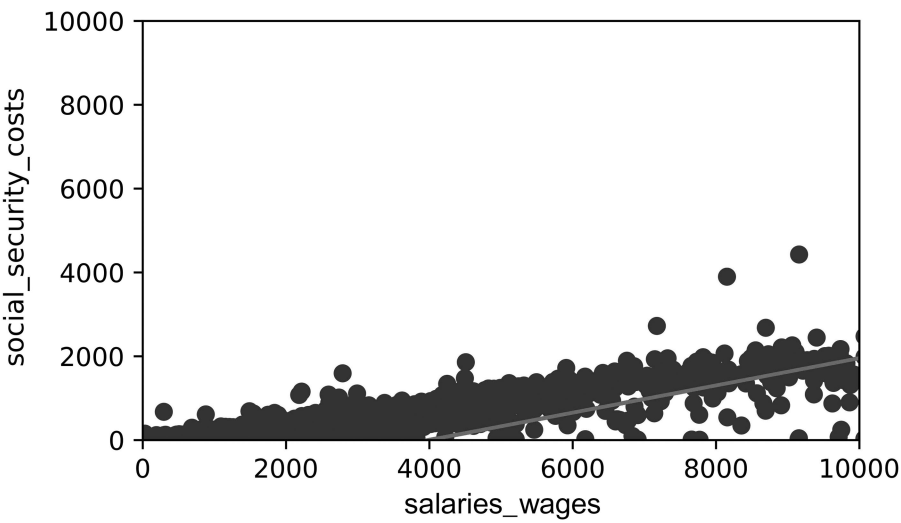


FIGURE 11.7 The linear regression.

Source: own elaboration based on analyzed data.

TABLE 11.4
The model evaluation metrics

Metrics	Value
The mean absolute error (MAE)	2,275.002
The residual sum of squares (RSS)	301,537
R-squared	0.916

Source: own elaboration based on analyzed data.

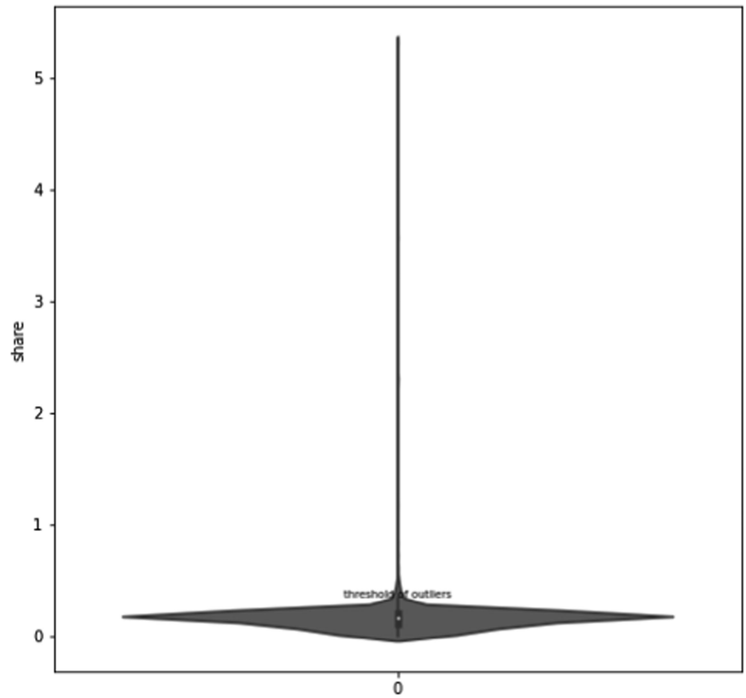


FIGURE 11.8 The violin plot for variable *share*.

Source: own elaboration based on analyzed data.

("Encyclopedia of Machine Learning," 2010). The larger the number the larger the error, for the model MAE is equal to 2,275.002. The residual is the difference between the actual and predicted value (Yashwanth, 2020), therefore the residual sum of squares measures the amount of variance in a data set that is not explained by a regression model. For the model, the residual sum of squares (RSS) is equal to 301,537. Finally, R-squared is a goodness-of-fit measure of the performance of the model. If R-squared is high, then the model represents the variance of the dependent variable (Yashwanth, 2020). The value of R-squared shows how close the data points are to the regression line. If they are very close to the fitted regression line, then the model accounts for a good amount of variance, thus resulting in a high R-squared value (Yashwanth, 2020). For the model, R-squared was equal to 0.916. It means that 91.6% of the variance in the dependent variable *social_security_costs* can be explained by the independent variable *salaries_wages*.

The last phase of the procedure is anomaly detection. For this purpose, the outliers, i.e., observations extremely high or low were identified for the variable *share*. In the case of continuous variables, one of the methods is to use the interquartile range (IQR), an outlier is any observation that is more than 1.5 IQR away from the first or third quartile (Walker, 2020). For the analyzed sample this distance is within the range −0.0286 ↔ 0.3392. As the variable *share* takes only a positive value, the outliers are observations with a variable share higher than 0.3392. In the sample, there are 66 such observations with a variable share higher than 0.3392, which represents 1.26% of the whole group. All outliers that were identified by the auditor should be reviewed and explained in detail. The outliers can be presented on the violin plot, which joins a box plot and a kernel density plot, and shows peaks in the data (Figure 11.8) that indicates outliers.

6 DISCUSSION

This study explores the use of ML algorithms for analytical procedures in audits. Both academia and as well as practitioners agree with the statement that ML algorithms and models are essential and beneficial in audits. However, often the models proposed by scientists are comprehensive and require advanced specialistic knowledge, which is a significant barrier for audit firms (Lachowski, 2019; Młyński, 2015), especially for small ones. Therefore, there is a need to create and promote simple, easily understandable, and low-cost ML models. In this way, it is possible to encourage auditors to implement emerging technologies that assist in a higher-quality audit process.

This study aimed to develop such an ML model, which can be used in analytical procedures performed by auditors to predict the value of the position of financial statements and detect outliers. For this purpose, the most widely known and least complicated machine-learning technique was chosen, namely regression analysis.

The advantages of the developed model include, among others:

- versatility, as the model was developed and tested on observations from the period 2003 to 2022, data from all markets and sectors (except banks and insurance companies) was used, and the financial statements in the sample were prepared under the local and international financial reporting framework,
- intuitiveness, simplicity of construction, and easy interpretation due to employed linear regression,
- no-cost, as an open-source programming language and its libraries were used to develop the model,
- high accuracy of the model – R-squared was equal to 91.6%, i.e., 91.6% of the variance in the dependent variable *social_security_costs* could be explained by the independent variable *salaries_wages*.

The model predicts relationships among two elements of financial information within the period. However, it can be expanded into more variables, including non-financial information. Further research lies in the area of developing, for particular sectors, more comprehensive ML models based on non-linear regression that can be used for auditors in conducting analytical procedures in different areas of financial statements.

7 CONCLUSIONS AND POLICY IMPLICATIONS/ RECOMMENDATION

Disrupting technologies such as ML brings significant changes for the audit profession. Thanks to the employment of algorithms, auditors can increase efficiency and reduce risk. The Big 4 firms: Deloitte, EY, KPMG, and, PwC are reporting usage of AI And ML applications in areas such as audit planning risk assessments, tests of transactions, analytics, and the preparation of audit work papers (Karmańska, 2022; Munoko et al., 2020). The chapter presents an example of using one of the ML algorithms for analytical procedures. Analytical procedures are used to obtain relevant and reliable audit evidence and can be more efficient than in reducing the risk of material misstatements at the assertion level to an acceptably low level (ACCA, 2022). One of the techniques of analytical procedures is regression analysis. It is a popular and uncomplicated algorithm used in machine learning. The aim of this study is to examine the usefulness of ML methods for analytical procedures in the audit of financial statements. The developed ML model based on linear regression predicts the value of costs of social securities and other benefits based on costs of salaries and wages and detects outliers with high accuracy (R-squared 91.6%).

The insights are likely to be of interest to different users of financial statements. First of all, the study can be useful for audit practitioners to understand the meaning of ML and its role to improve audit efficiency and effectiveness. To a greater extent, it can be interesting for auditors from smaller audit firms, because it shows the possibility and benefits of implementing simple and cost-effective ML models.

Second, the accountants and managers can use the study to create models to verify items of financial statements and accounting vouchers in the books.

Third, the government and regulatory bodies such as Social Insurance Institutions can use the findings of the study to develop ML models-targeted interventions aimed at firms that are more susceptible to avoidance of paying social contributions.

Fourth, investors can use machine learning models to predict future financial trends using the data from the financial statements.

Fifth, for academia, the study expands the science of machine learning in the audit. In addition, the study indicates a methodology that allows other researchers to explore the patterns in the big set of data.

However, the findings have to be interpreted in light of certain limitations. The main one concerns constraints of machine learning algorithms. While ML offers an innovative approach to problem-solving in various fields and domains that require processing a large amount of data, one should be aware of its potential limitations and pitfalls. The main problem of ML models is related to the lack of good-quality data. The ML techniques need a big amount of reliable data for efficient analysis, training, and performance. Lack of high-quality data means significant reduction in the model's accuracy and can lead to inadequate outcomes. Therefore, before using the data the audit team should assess its reliability by considering the source and the conditions under which it was gathered, as well as other knowledge that may have about the data (Public Company Accounting Oversight Board (PCAOB), 2010). The next concern of ML is related to the bias at any level of development of the models, the inherent error can come from various sources: data collection, retrieval, processing, etc. It can be also observed as a risk of overconfidence in the use of automated tools and techniques in audits (Karmańska, 2022). It should be underlined that adoption of ML does not eliminate the need to exercise the auditor's professional judgment and skepticism. The emerging technologies can increase audit efficiency but do not replace human-powered creativity, experience, and critical thinking. Finally, the algorithm is trained on a historical dataset as input to predict new output, but the output may not reflect past events.

Further limitations are related strictly to the research. The first one is the fact that despite the existence of many ML techniques and algorithms, the study focuses only on one of them. The other limitation concerns the fact that this study focuses mainly on companies listed in Poland.

After a thorough analysis, the following recommendations are hereby made:

1. The policymakers and professional bodies should verify legal provisions and standards in the field of accounting and audit to reflect the challenges of the increasing use of automated tools and ML techniques.
2. More emphasis on the education of accountants and auditors, changing universities' curricula, STEM (Science, Technology, Engineering, Mathematics) should be imperative.
3. Small audit firms should require specialized knowledge, invest in new technologies, and implement ML models; otherwise, they may lose their competitive advantage.

REFERENCES

ACCA. (2022). *Analytical Procedures*. https://www.accaglobal.com/my/en/student/exam-support-resources/professional-exams-study-resources/p7/technical-articles/analytical-procedures.html.

Al Qtaish, H., Hassan Makhlouf, M., & Joudeh, A. H. (2021). The Effect of Auditors' Use of Analytical Procedures in the Light of ISA 520 on Audit Quality: Evidence from Jordan. *Studies of Applied Economics, 40*(1). https://doi.org/10.25115/eea.v40i1.6637

Amel-Zadeh, A., Calliess, J.-P., Kaiser, D., & Roberts, S. (2020). Machine Learning-Based Financial Statement Analysis. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3520684.

Anand, V., Brunner, R., Ikegwu, K., & Sougiannis, T. (2019). Predicting Profitability Using Machine Learning. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3466478.

Appelbaum, D. A., Kogan, A., & Vasarhelyi, M. A. (2018). Analytical Procedures in External Auditing: A Comprehensive Literature Survey and Framework for External Audit Analytics. *Journal of Accounting Literature*, *40*, 83–101. https://doi.org/10.1016/j.acclit.2018.01.001.

Chen, X., Cho, Y. H., Dou, Y., & Lev, B. (2022). Predicting Future Earnings Changes Using Machine Learning and Detailed Financial Data. *Journal of Accounting Research*, *60*(2), 467–515. https://doi.org/10.1111/1475-679X.12429

Cheung, E., & Lau, J. (2016). Readability of Notes to the Financial Statements and the Adoption of IFRS. *Australian Accounting Review*, *26*(2), 162–176. https://doi.org/10.1111/auar.12087.

Claudiu, B., Muntean, M., & Didraga, O. (2018). Intelligent Decision Support in Auditing: Big Data and Machine Learning Approach. *Proceedings of the IE 2018 International Conference*, *November*, www.conferenceie.ase.ro, 425–430. https://www.researchgate.net/publication/328858614_INTELLIGENT_DECISION_SUPPORT_IN_AUDITING_BIG_DATA_AND_MACHINE_LEARNING_APPROACH

Damrongsakmethee, T., & Neagoe, V.-E. (2017). Data Mining and Machine Learning for Financial Analysis. *Indian Journal of Science and Technology*, *10*(39), 1–7. https://doi.org/10.17485/ijst/2017/v10i39/119861

Davis, F. D. (1985). A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results. *Management*, *Ph.D.*, 291. https://doi.org/oclc/56932490.

Dickey, G., Blanke, S., & Seaton, L. (2019). Machine Learning in Auditing Current and Future Applications. *The CPA Journal*, *89*(6), 16–21. https://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=136901845&lang=fr&site=ehost-live.

Elliot, V. H., Paananen, M., & Staron, M. (2020). Artificial Intelligence for Decision-Makers. *Journal of Emerging Technologies in Accounting*, *17*(1), 51–55. https://doi.org/10.2308/jeta-52666.

Encyclopedia of Machine Learning. (2010). In C. Sammut & G. I. Webb (Eds.), *Encyclopedia of Machine Learning*. https://doi.org/10.1007/978-0-387-30164-8.

Fedak, Z. (1998). *Metody i technika rewizji sprawozdań finansowych*. Stowarzyszenie Księgowych w Polsce, Warszawa.

Flood, J. (2020). Analytical Procedures. *Wiley Practitioner's Guide to GAAS 2020*, 223–236. https://doi.org/10.1002/9781119596042.ch19.

Hajek, P., & Henriques, R. (2017). Mining Corporate Annual Reports for Intelligent Detection of Financial Statement Fraud – A Comparative Study of Machine Learning Methods. *Knowledge-Based Systems*, *128*, 139–152. https://doi.org/10.1016/j.knosys.2017.05.001.

Handoko, B. L. (2021). How Audit Firm Size Moderate Effect of TOE Context toward Auditor Adoption of Machine Learning. *Journal of Theoretical and Applied Information Technology*, *99*(24), 5972–5980.

Hoggett, E., Dubois, S., Connor, S. O., & Jamieson, R. (2019). Machine Learning and the Audit : Rise of the Machines? *KPMG*, *April*. https://home.kpmg/au/en/home/insights/2019/04/audit-technology-machine-learning.html.

Hooda, N., Bawa, S., & Rana, P. S. (2018). Fraudulent Firm Classification: A Case Study of an External Audit. *Applied Artificial Intelligence*, *32*(1), 48–64. https://doi.org/10.1080/08839514.2018.1451032.

Hoogduin, L. A. (2019). Using Machine Learning in a Financial Statement Audit. *Compact*, *4*, 4–8.

Huang, F., No, W. G., Vasarhelyi, M. A., & Yan, Z. (2022). Audit Data Analytics, Machine Learning, and Full Population Testing. *The Journal of Finance and Data Science*, *8*, 138–144. https://doi.org/10.1016/j.jfds.2022.05.002.

IAASB. (2020). *Non-Authoritative Support Material Related to Technology: Frequently Asked Questions (FAQ)——The Use of Automated Tools and Techniques When Identifying and Assessing Risks of Material Misstatement in Accordance with ISA 315 (Revised 2019)*. https://www.ifac.org/system/files/publications/files/IAASB-Technology-FAQ-Automated-Tools-Techniques.pdf.

Karmańska, A. (2022). Artificial Intelligence in Audit. *Prace Naukowe Uniwersytetu Ekonomicznego We Wrocławiu Research Papers of Wroclaw University of Economics and Business*, *66*(4), 87–99, https://doi.org/10.15611/pn.2022.4.06

Kwaku, S. A., & Leiby, J. (2014). Analytical procedures. In J.C. Bedard & L. Graham (eds.), *The Routledge Companion to Auditing*. London: Routledge. https://doi.org/10.4324/9780203094921-29

Lachowski, W. K. (2019). Stosowanie procedur analitycznych do badania sprawozdań finansowych. *Rachunkowość-Audytor*. https://www.pibr.org.pl/static/items/publishing/20190116-ce-rachunkowosc.pdf.

Lewandowski, P., & Magda, I. (2017). Temporary Employment, Unemployment and Employment Protection Legislation in Poland. In A. Piasna & M. Myant (Eds.), *Myths of Employment Deregulation: How It Neither Creates Jobs Nor Reduces Labour Market Segmentation*. Brussels: ETUI, The European Trade Union Institute, 143–163.

Lloyd, R. (2021). *EY reports global revenues of US$40b in 2021 and outlines record US$10b investment plan over next three years*. https://www.ey.com/en_gl/news/2021/09/ey-reports-global-revenues-of-us-40b-in-2021-and-outlines-record-us-10b-investment-plan-over-next-three-years

Matrood, A. K., Abd alrazaq, D. N., & Khilkhal, N. S. (2019). The impact of applying analytical procedures by external auditor in accordance with ISA 520 on audit performance improvement: An exploratory study in the Iraqi audit firms and companies. *Academy of Accounting and Financial Studies Journal*, *23*(1).

Młyński, J. (2015). Testy analityczne – efektywna metoda badania sprawozdania finansowego. https://www.bdo.pl/pl-pl/publikacje/biuletyn-podatki-i-rachunkowosc/2015/testy-analityczne-—-efektywna-metoda-badania-sprawozdania-finansowego.

MSCI. (2022). The Global Industry Classification Standard (GICS®). *MSCI*. https://www.msci.com/our-solutions/indexes/gics.

Munoko, I., Brown-Liburd, H. L., & Vasarhelyi, M. (2020). The Ethical Implications of Using Artificial Intelligence in Auditing. *Journal of Business Ethics*, *167*, 209–234. https://doi.org/10.1007/s10551-019-04407-1.

Notoria Oferta, dane finansowe dla inwestorów. (2022). https://ir.notoria.pl/oferta.html.

Pedregosa, F., Michel, V., Varoquaux, G., Thirion, B., Dubourg, V., Passos, A., Pedregosa, V., Perrot, M., Grisel Oliviergrisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Vanderplas, J., Cournapeau, D., Pedregosa, F., Varoquaux, G., Gramfort, A., Thirion, B., Grisel, O., … Brucher, M. (2011). Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. http://scikit-learn.sourceforge.net.

Public Company Accounting Oversight Board (PCAOB). (2010). *AS 2305: Substantive Analytical Procedures. Auditing Standard (AS)*. https://pcaobus.org/oversight/standards/auditing-standards/details/AS2305.

Raschka, S. (2015). *Python Machine Learning*. Birmingham: Packt Publishing.

Raschka, S., & Mirjalili, V. (2019). *Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-Learn, and Tensorflow 2*. Birmingham: Packt Publishing.

Samaha, K., & Hegazy, M. (2010). An empirical investigation of the use of ISA 520 "analytical procedures" among Big 4 versus non-Big 4 audit firms in Egypt. *Managerial Auditing Journal, 25*(9), 882–911. https://doi.org/10.1108/02686901011080053

Sekar, M. (2022). *Machine Learning for Auditors: Automating Fraud Investigations Through Artificial Intelligence*. New York: Apress. https://doi.org/10.1007/978-1-4842-8051-5.

Setiawan, S., Hanny, Lestari Dewi, N., Handayani, R., Purnama Sari Eddy, E., & Purba, M. (2021). Supporting Auditor Competence in the Digital Era 4.0. through the Audit Learning System in Universities. *ACM International Conference Proceeding Series*, 51–58. https://doi.org/10.1145/3494583.3494604.

Sifa, R., Ladi, A., Pielka, M., Ramamurthy, R., Hillebrand, L., Kirsch, B., Biesner, D., Stenzel, R., Bell, T., Lübbering, M., Nütten, U., Bauckhage, C., Warning, U., Fürst, B., Khameneh, T. D., Thom, D., Huseynov, I., Kahlert, R., Schlums, J., … Loitz, R. (2019). Towards Automated Auditing with Machine Learning. *Proceedings of the ACM Symposium on Document Engineering, DocEng 2019*. https://doi.org/10.1145/3342558.3345421.

Statistics Poland. (2021). *Labour Costs in 2020*. https://stat.gov.pl/obszary-tematyczne/rynek-pracy/pracujacy-zatrudnieni-wynagrodzenia-koszty-pracy/koszty-pracy-w-2020-r-,8,6.html.

Supreme Audit Institutions of Finland Germany the Netherlands Norway and the UK. (2020). *Auditing Machine Learning Algorithms*. https://www.auditingalgorithms.net/

Sun, T. S. (2019). Applying deep learning to audit procedures: An illustrative framework. *Accounting Horizons*, *33*(3), 89–109. https://doi.org/10.2308/acch-52455

Tatsat, H., Puri, S., & Lookabaugh, B. (2020). *Machine Learning and Data Science Blueprints for Finance*. Sebastopol: O'Reilly Media. https://www.amazon.com/Machine-Learning-Science-Blueprints-Finance/dp/1492073059

The Social Insurance Institution. (2021). *Social Security in Poland*. https://www.zus.pl/documents/10182/167615/Social+Security+in+Poland/71ffe1b1-c142-48fa-a67b-0c7e1cec6eb6.

Tiron-Tudor, A., & Deliu, D. (2021). Reflections on the Human-Algorithm Complex Duality Perspectives in the Auditing Process. *Qualitative Research in Accounting and Management*. https://doi.org/10.1108/QRAM-04-2021-0059.

Ucoglu, D. (2020). Current Machine Learning Applications in Accounting and Auditing. *Pressacademia*, *12*(1), 1–7. https://doi.org/10.17261/pressacademia.2020.1337.

Walker, M. (2020). *Python Data Cleaning Cookbook Modern Techniques and Python Tools to Detect and Remove Dirty Data and Extract Key Insights*. Birmingham: Packt Publishing.

Yashwanth, N. (2020). Evaluation Metrics & Model Selection in Linear Regression. *Towards Data Science*. https://towardsdatascience.com/evaluation-metrics-model-selection-in-linear-regression-73c7573208be.

12 Application of advanced tools to bolster the business performance of companies in the new normal

Subhash Chander Arora and Vinod Kumar Singh

1 INTRODUCTION

In this digitalized world, the use of automated and advanced tools like Artificial Intelligence (AI), Machine Learning (ML) and blockchain technology is getting increased acceptance so as to better serve the market demands (Enholm et al., 2021). In layman's language, the concept rotates around coming up with machines and systems that operate like humans (Makarius et al., 2020). The term was first coined by John McCarthy in the year 1956 to define scientific and technical knowledge of developing smart computer programs so as to carry out activities of visualization, decision-making, language translation, answering queries, navigating journey, speech recognition thereby reducing human workload (Mikalef et al., 2021).

Today, these terms have common parlance (Figure 12.1), *social experts* are studying the legal and ethical implications of these tools to run the business enterprise; *engineers* and *computer scientists* are working tirelessly to develop new algorithms whereas people from *management* background are more concerned about the effect of such technologies on customers, competitors, firms, logistics and market response (Boukherouaa et al., 2021). In order to increase the market competitiveness, companies are reengineering their existing products and services, reconsidering their business strategies and have been identifying attractive business prospects that may result in significant advantage through the deployment of AI (Campbell et al. 2020).

AI is being viewed as an innovative technology with potential to change the company's prevalent business model and organizational system (Lee et al., 2019). It promises numerous advantages for organizations in terms of added business value (Mishra & Pani, 2020). However, to understand the notion of AI, it is necessary to first understand the two terms individually. *Intelligence* can be defined as one's ability to learn, reason, understand and comprehend things (Lichtenthaler, 2019). *Artificial*, on the

 DOI: 10.1201/9780367854690-12

other hand, refers to anything created by humans, rather than occurring naturally in the external environment (Mikalef & Gupta, 2021).

By bringing these two terms together, Artificial Intelligence (AI) can be described as building machines especially computer programmes capable of simulating or performing actions or tasks without human intervention (Wamba-Taguimdje et al., 2020). Thus, AI is a technology of making machines learn from actions over a period of time and take decisions on your behalf. The machine is able to do the job of a human without any manual intervention (Zheng et al., 2017). AI includes programming computers for certain qualities such as information processing, giving logic, solving problems, studying behavioural patterns, self-learning and strategizing, etc. (Collins et al., 2021).

An example of this is the chatbots that helps clients, customers at banks and other corporate offices through voice and text display interface (Castillo et al., 2020). This includes conversational intelligent agents such as *Cortana*, *Apple's Siri and Amazon's Alexa* which provide the ease of making phone calls, controlling other smart home devices like light bulbs and playing music or providing necessary update/news/information on voice commands (Prentice et al., 2020). But AI is not just a chatbot installed to improve customer affairs, it is much more than that and influences not just the business model but the entire operating paradigm of an organization (Riikkinen et al., 2018).

Another word that is commonly used and frequently mixed up with AI technology is cognitive technology (Zheng et al., 2017). According to Bytniewski et al. (2020), cognitive technologies mimic or duplicate human thought and behaviour, giving computers the capacity to think and behave much like people. Moreover, prior knowledge of the popular belief that such technologies should function as augmentation agents for tough and time-consuming tasks is necessary (Mikalef & Gupta, 2021).

1.1 Background of the problem

AI is being used in manufacturing to provide services to banking and fintech industries (Duan et al., 2019) and in varied disciplines such as tourism, hospitality, marketing and psychology to name a few (Alsheibani et al. 2018). The tool has the power to help companies redefine the ways in which they used to function, create groundbreaking products or services for their customers and change the overall experience of various stakeholders operating in the market (Enholm et al., 2021). Today, companies are incorporating AI systems into their operations and managerial activities due to increased digitalization in the form of IT-enabled infrastructure which can ultimately lead to enhanced market value (Afiouni, 2019). Tough times during the past two years necessitated companies to embrace, imbibe and weave AI ecosystem into their business functioning and strategies (Boukherouaa et al., 2021).

Although past studies identified and studied different facets of AI implementation in different segments ranging from banking industry to fintech companies, limited literature was extracted in terms of its application in the context of companies operating in other diverse fields such as manufacturing and other operations-based organizations. Technology giants worldwide and concerned leaders are still understanding how AI works and how it can be used in their own domain. Moreover, much work has been carried out in technology-rich and other advanced economies like those of Europe, Japan, USA and South Korea. Therefore, what is the present state of AI adoption in a developing country such as India and how it can act as a boon in the functioning of a company came to the attention of this venture. By investigating the potential of these tools for Indian business houses, the study can fulfil the existing gap in AI–Business relationship in the Indian market. Moreover, actual benefits from these technologies can only be reaped if one has desired knowledge of their application and consequences.

2 REVIEW OF LITERATURE AND PROPOSED WORK

Organizations are incorporating new methods, processes and technologies so as to ensure their long-term survival (Soni et al., 2020). Implementation and upgradation of innovative tools like AI and other informatics science into the company is being catalyzed by many factors working simultaneously (Sestino & De Mauro, 2022). This brings us to the question as to what are the things that make companies go for AI adoption in their internal and external processes.

One thing is clear and predominant, that the nucleus for AI is data and to make sense of this data, companies are increasingly resorting to business analytics (BA) which not only helps in generating actionable insights but also makes organizations ready for any discrepancy based on market requirements (Fosso Wamba et al., 2015). This BA involves the use of advanced capabilities and technologies to collect, transform and analyze data to support better decision-making (Ashrafi et al., 2019). The concept is not new but has recently re-emerged as a new and important research direction for the development of capabilities to handle big data. Past researchers have demonstrated a positive relationship between BA and organizational performance (Ramanathan et al., 2017). Studies by GE and Accenture argued that almost 90% of global firms believe in AI-driven systems and models to guard against volatile tendencies of the market and to stay competitive (Delen & Zolbanin, 2018).

Furthermore, the term *Big Data* is used for data sets that are enormous or large in size (Mariani & Wamba, 2020). These data sets are information assets with characteristics like high variety, large volume and great velocity and require distinct and varied processing so as to support better decision-making, generate new insights and optimize business processes (Fosso Wamba et al., 2015). Big data has gradually penetrated all walks of life and has become an attractive avenue for scholarly research. AI can process these large amounts of information at a high speed that goes beyond humans' cognitive capabilities (Jarrahi, 2018). The purpose here is to bring AI and human expertise together for better efficiency and synergy with little intent to replace humans with AI (Schmidt et al., 2020).

Besides data that is beyond manual handling, there are other aspects like infrastructure and external environment which lead to the deployment of AI into businesses (Pumplun et al., 2019). To handle this data in an apt manner, the organizations primarily need two things: state-of-the-art processing infrastructure and advanced algorithms (Alsheibani et al., 2018; Wamba-Taguimdje et al., 2020). But, it is possible that many of the companies do not have timely access to all these required resources, therefore, large companies such as Google and Microsoft have started to provide this infrastructure facility (Google Cloud AI) for machine learning over the cloud (Borges et al., 2020). This mechanism allows other organizations to have online access to the necessary infrastructure for deploying AI within their own premises.

Further, intense competition or serving customers better than others also catalyze the process of AI adoption in the macro environment (Demlehner & Laumer, 2020; Pumplun et al., 2019). In order to preserve the intended competitive advantage, a firm must respond to changes in the external environment (Roberts & Grover, 2012). In order to improve customer experience over time, firms are driven to acquire cutting-edge IT tools and/or innovations in order to maintain the desired competitive edge (Ashrafi et al., 2019). With the availability of resources, every firm is trying to grab the attention of potential customers in the market by putting their own products as superior to others (Chatterjee et al., 2021). Thus, organizations must continuously respond to such dynamics so as to preserve their position in the minds of the customers (Dedrick et al., 2013).

2.1 AI technologies

Automation, data analytics and Natural Language Processing (NLP) are among the top applications of AI (Enholm et al., 2021). People are no longer required to undertake repetitive activities as a result of automation (Davenport et al., 2020). It frees up employees' time to focus on higher-value work by completing monotonous or error-prone tasks. Data analytics allows organizations to gain insights that were previously non-existent or inaccessible due to limited capacity, by discovering new patterns and correlations in data (Ramanathan et al., 2017).

A crucial part of most AI systems, that is, the learning process takes the form of ML, which is based on statistics, algebra and decision theory (Canhoto & Clear, 2020). A branch or application of AI techniques that is frequently employed by commercial businesses is machine learning (Baier et al., 2019). It gives the systems the capacity to autonomously learn from experience and get better without explicit programming. The majority of the most recent technological successes including self-driving cars, digital assistants, robotics and other smart home gadgets are the result of advances in ML, particularly deep learning algorithms (Boukherouaa et al., 2021).

The business world is changing very fast and people are optimistic about the importance of AI to their enterprises over the next few years (Ransbotham et al. 2017). A study by MIT in 2018 revealed that 58% of corporations believed AI would significantly change their business models by 2023 (Gupta, 2022). AI/ML systems, being argued as a force of disruption (Davenport & Ronanki, 2018) are reshaping client experiences, interpersonal communication, investment decisions and borrowing and lending interface through automated mortgage underwriting and identification of clients (Khandani et al., 2010). The technological revolution is transforming the functioning of business enterprises in terms of employee efficiency, cost savings through automation routes, better product offerings, fraud detection and risk management, operational efficiency and other regulatory compliances (Alsheibani et al. 2018).

The COVID-19 pandemic has further fuelled the intake of AI technologies in the commercial sector (Sahay et al., 2020). Banks are gaining by adopting AI tools in their processing of loan applications, underwriting services, risk management,

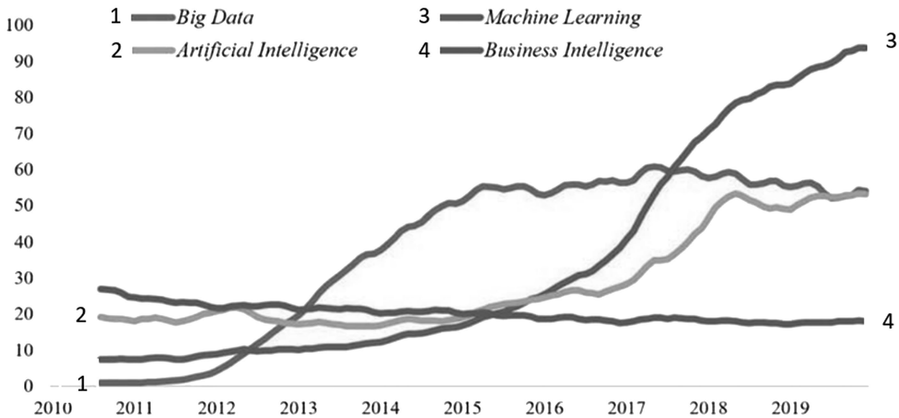


FIGURE 12.1 Popularity of related terms among web users between 2010 and 2019 (Sestino & De Mauro, 2022).

strengthening oversight procedures and detection of fraudulent activities (Riikkinen et al., 2018).

Businesses today are getting heavily reliant on AI and ML tools on account of the enormous data that is at their disposal (Enholm et al., 2021). AI has the potential to carefully and successfully analyze this large amount of data, minimizing human error in interpreting this data and providing useful insights to understand future developments (Makarius et al., 2020). Besides, it has the potential to upscale operational and employee efficiency, trim down safety risks at the premises, enable better decision-making and come up with tailor-made solutions for individual clients (Lui & Lamb, 2018). Business houses, big or small, are incorporating the AI-enabled technologies in their day-to-day operations and back-office related activities almost uncompromisingly (Roh et al., 2019). Major applications include sales increase and prediction, profit maximization, inventory management, security, fraud detection and portfolio management (Eriksson et al., 2020).

Machine learning can be either shallow or deep (Wang et al., 2019). Shallow-structured learning architectures are the most traditional where the system learns from the data described by predefined features (LeCun et al., 2015). In contrast, deep machine learning or deep learning, can derive structure from data in a multi-layered manner (Afiouni, 2019). What differentiates deep learning from more traditional machine learning is the use of an artificial neural network architecture (Wamba-Taguimdje et al., 2020). Neural network solutions refer to the human brain's functionality by imitating human neurons (Schmidt et al., 2020).

Some other key AI technologies that have been studied in empirical studies include *Natural language processing* where AI systems can communicate by understanding and generating human language (Jarrahi, 2018), it is beneficial because it empowers search engines to be smarter, chatbots to be more helpful and boosts accessibility for those with disabilities such as hearing impairments; *Image recognition* where facial and signature recognition is being used to carry out the due diligence process, identify and verify the customers, strengthening systems security (Enholm et al, 2021); *Speech Synthesis System* which includes text-to-speech and speech-to-text solutions (Lichtenthaler, 2019).

Today, most of these technologies are used in combination with machine learning or deep learning to provide industry-based solutions. For instance, in the case of chatbots, language models are being built using both semantic analysis and machine learning (Baby et al., 2017). The functionality enabled through NLP allows chatbots to understand and communicate using the human language. On the other hand, machine learning algorithms facilitate chatbots to learn and evolve as they get access to more data (Castillo et al., 2020).

2.2 Application of AI tools and capabilities

As already discussed, the discipline of AI could be defined as a set of studies and techniques, dealing with computer science and mathematical aspects of statistical modelling, carrying significant economic and social implications, aimed to create technological systems capable of solving problems and carrying out tasks and duties, normally attributable to the human mind (Wamba et al., 2017). The capability of acquiring large sets of data from the environment and processing it with said tools is changing the corporate sector landscape (Dhanabalan & Sathish, 2018). AI is a crucial growth factor for multiple sectors across all industries and enables companies to know and address weaknesses, build upon their strengths and improve their teams' skill sets (Sestino & De Mauro, 2022).

2.2.1 Risk and compliance management

Recent developments in AI/ML are altering the scope and functioning of technology in regulatory compliance (Financial Stability Board, 2020). In reaction to regulatory tightening and increased compliance costs in the aftermath of the Harshad Mehta scam of 1992, Satyam scandal in 2009 and Global Financial Crisis of 2008, regulatory technology has grown in prominence. For the most part, compliance and reporting processes have been digitized using technology (Zetzsche et al., 2017). But recent developments in AI/ML are redefining risk and compliance management by utilizing large data sets, frequently in real time and automating compliance judgments (Boukherouaa et al., 2021). As a result, there has been a major improvement in compliance quality and cost reduction. Due to the significant expenses associated with human selection, AI and ML are replacing human analysts in executing corporate tasks (Ransbotham et al., 2017). AI is built on ML which continuously learns new things and help in improvising the previously time-consuming tasks (Boukherouaa et al., 2021).

2.2.2 Forecasting

Novice algorithms' data sets are being used for forecasting macro-economic externalities, meeting customer demands, foreseeing payment capacity, monitoring business conditions and other financial variables and their inter-relationships (Vidgen et al., 2017). Compared to conventional statistical or econometric models, they typically perform better (Khandani et al., 2010). This is utilized in the financial industry in areas including credit risk assessment, determining risk appetite and tracking the movement of economic and financial factors (Bolhuis & Rayner, 2020). Expert sellers and sales companies are rethinking the balance between humans and machines in sales. Automation by AI is already affecting sales and will continue to do so (Kumar et al., 2019). According to a study by *Harvard Business Review*, companies using AI for sales can increase their leads by more than 50%, reduce call time by 60%–70% and have cost reductions of 40%–60%. Artificial intelligence enables the creation of automated and accurate sales projections based on all client interactions and historical sales results (Kleinings, 2022). Moreover, AI aids in lead prioritization. These tools help sales professionals prioritize customers based on their probability to convert (Malthouse et al., 2013). With AI, the algorithm can rank the opportunities or leads in the pipeline based on their chances of closing successfully by compiling historical information about a client and social media postings and the salesperson's customer interaction history (Chatterjee et al., 2021).

Such AI/ML technologies could assist in more precise monetary and macroprudential policies, enhance internal operations and help banks better comprehend economic and financial events (Lui & Lamb, 2018). Additionally, they might be able to foresee the development of systemic problems, strengthen the framework for risk management and hasten crisis response (Boukherouaa et al., 2021). Despite these potential advantages, using AI/ML processes for policymaking should still be done with caution. In order to make better-informed policy decisions, banks and other financial institutions have concentrated their experiments and research on enhancing their ability to foresee short-term events and monitor market sentiment (Khandani et al., 2010).

However, there are challenges especially in case of non-traditional data with social media content, browsing history and others (Malthouse et al., 2013). A number of issues are raised by the use of non-traditional data in financial forecasting including the governing legal and regulatory framework, ethical and privacy implications and data quality in terms of cleanliness, accuracy, relevance and potential biases (Rana et al., 2022).

2.2.3 Transferability of funds

The application of these tools is also influencing the investment environment of the business enterprises. Managing finance is a challenging task in the materialistic world for people and business houses. But this problem won't exist any longer due to emergence and utilization of AI and its enabling technologies in the fund management field (Ashta & Herrmann, 2021).

Conventional technology is already helping the companies and institutions in managing their big data and high trading activities but modern tools are reshaping the interaction interface between the respective parties leading to product customization, tailor-made portfolio management, risk quantification, enhanced efficiency, better decision-making in terms of analytics and cost reduction due to enhanced automation (Mikalef et al., 2021). This is especially true in the case of evolution of BHIM UPI which makes the transferability of funds easier and convenient and other digital platforms which provide the ease of monitoring various lucrative investment opportunities (Riikkinen et al., 2018). The application of AI has proved to be successful in many other areas such as providing automated customer support, claims management and financial advisory services (Dhanabalan & Sathish, 2018).

2.2.4 Anomaly and/or fraud detection

Algorithms can be applied to detect rare items, outliers and/or missing/incomplete data. In the financial sector, issues related to insider trading, fraudulent use of credit card and insurance policy manipulation are some of the identified applications (Chandola et al., 2009). Giant organizations have excess funds to manage on a regular basis (Ashta & Herrmann, 2021). There are continuous inflows and outflows and thus companies resort to AI-enabled technologies so as to avoid financial pollution (Rajamohan & Dhanabalan, 2013). AI is the topmost technology that comes with security and fraud detection techniques (Yigitcanlar et al., 2020). It can be used to assess the earlier spending pattern and behaviour of an individual towards various transactions and from which it would identify the odd behaviour such as whether there has been any attempt to take out the money unusually from an account which is in a state of query (Ashta & Herrmann, 2021).

2.2.5 Prudential supervision

In their risk-based supervision process, many supervisory agencies and corporate houses are actively investigating the usage of AI/ML systems (Rana et al., 2022). There is a place for AI/ML, especially in data collecting and data analyses, even if decisions will ultimately depend on the supervisors' judgment (Financial Stability Board, 2020). With AI, managers can get more information from any sort of data and use that information to make more informed, data-driven decisions. AI can notice patterns that humans miss, improving the effectiveness of supervision. By alerting managers to irregularities in real time, it can help improve supervisory agility (Ashrafi et al., 2019). Furthermore, AI-powered solutions for supervisory technology can deliver predictive analyses that have the potential to raise the standard of supervision (Roberts & Grover, 2012). Despite all of its potential, AI/ML is not a panacea and human judgement and an organization's risk culture will always play a major role in how well supervision works (Boukherouaa et al., 2021).

AI/ML in micro-prudential supervision could enhance the evaluation of risks in business and financial institutions such as credit and liquidity hazards, governance and risk culture allowing supervisors to concentrate on risk analysis and other futuristic assessments (Yudkowsky, 2008). Risks and trends could be assessed at the sectoral level improving the quality of macro-prudential supervision (Roberts & Grover, 2012). For market surveillance objectives, AI/ML systems are also used to look for collusive behaviour and price manipulation in the securities market (Mizuta, 2022). Moreover, the resultant impact of Covid-19 in the form of distant working also drives authorities to regulate remote supervision so as to sustain corporate efficiency.

2.2.6 Effective promotional campaigns

To improve operational efficiency while increasing customer experience, many organizations and marketing teams are rapidly adopting intelligent technology solutions such as AI (Vidgen et al., 2017). This permits marketers to attain beneficial customer insights, automatize everyday assignments and enhance workflows (Davenport et al., 2020). By evaluating piles of data about the consumers, business can leverage AI to forge more personalized experiences (Kumar et al., 2019). The compilation and analysis of past purchase behaviour patterns and previous on-site activities and interactions can help deduce a defined customer persona (Hair, 2007). AI instruments help produce email content and personalized documents using algorithms that look at a subscriber's website experience and browsing data history (Mikalef et al., 2021). This saves a lot of time for the business by recognizing what areas the marketing team should focus upon (Eriksson et al., 2020). ML algorithms also help identify the products or the content for your business displays to match the buyer's interests, diminishing the product return rate (Kumar et al., 2019).

With AI, marketers can acquire a deeper insight into their target consumers (Chatterjee et al., 2021). The data obtained can then be utilized to increase conversions while also reducing marketing teams' workload. While Artificial Intelligence is still far from being able to construct whole new websites, it can assist in improving the visitor experience on a website through intelligent personalization (Kleinings, 2022). In search engine optimization, the term "search volume" informs us of how many people are searching for specific terms and phrases while looking for items or services. ML algorithms are now being used to gain a better grasp of the intent behind search term usage as well as the substance of searches (Malthouse et al., 2013).

Additionally, AI may be used to analyze attitudes, views and feelings about a specific good or service (Jelonek et al., 2019), which is becoming increasingly important for businesses as it allows them to gain in-depth knowledge about how their clients view their products (Bytniewski et al., 2020). AI can be used in public relations to track social media and forecast media trends (Galloway & Swiatek, 2018). AI can be used in marketing to categorize clients based on their tastes and lifestyles (Mishra & Pani, 2020). AI is utilized in the fashion industry to predict future trends, anticipate consumer behaviour and improve recommendation systems (Wamba-Taguimdje et al., 2020). Better market and consumer targeting enables businesses to give one-to-one marketing by tailoring the mutual experience (Afiouni, 2019; Mishra & Pani, 2020). As a result, AI improves marketing accuracy and effectiveness by focusing on the proper target market. Companies might offer products that stop customers from leaving, forward specialized services or tailored deals which satisfies needs and leads to customer satisfaction (Schmidt et al., 2020).

2.2.7 Customer being priority

Businesses work hard to ensure that consumers appreciate their experience with an organization or service and want to tell others about it (Chatterjee et al., 2021). Promptly resolving issues is one of the most efficient ways to keep customers happy (Hair, 2007). But, for scaling businesses, this may be challenging as more and more requests come in. As a result, businesses are introducing AI into customer service teams to enhance the customer service experience (Kumar et al., 2019). Chatbots, powered by AI, as front-line customer service agents can help organizations handle jobs, fix issues and even communicate with potential consumers (Kleinings, 2022). AI-powered chatbots can assist a vast number of consumers 24/7 without a human employee. Using chatbots as a first encounter can assist in identifying consumer needs and guarantee that they get connected to the right people (Galloway & Swiatek, 2018). At the same time, it can cut down on wait times and the amount of time that customer service agents spend answering the same questions (Wang et al., 2022).

On the other hand, voice-bots reduce the number of calls and are capable of listening to the caller, interpreting their mood and detecting the seriousness of the situation (Baby et al., 2017). An AI voice-bot, similar to a chatbot, may be trained to offer responses to common questions on its own, much like a chatbot but in a speech format. Because they are capable of understanding queries, they may also be used to improve call categorization and make sure that callers get connected to the most appropriate department (Wang et al., 2022).

2.2.8 Usage in operations

The use of AI in Business Operations (AIOs) helps many organizations tread the path toward a successful digital transformation (Enholm et al., 2021). What's more, the need for corporate agility made operations complex, making it challenging for humans to keep up (Ashrafi et al., 2019). AIOs help with this by automating certain processes and freeing up staff time for higher priority projects. Businesses can save a lot of money with an AI application that employs machine learning (Gupta, 2022). AI can automate cybersecurity and software maintenance tasks. It can also detect possible threats faster than humans, potentially saving companies from cyber-attacks (Lee et al., 2019). AI applications help IT staff with maintaining the organization's systems and keeping things running smoothly. Globally, businesses are actively adopting advanced technology. Business processes are being automated by first digitizing them and then allowing users to use apps based on new technology, whether it is in banking, travel, healthcare or e-commerce (Miller, 2018). The use of robotics and blockchain technology can assist with managing information, process efficiency and reducing IT operational friction (Ransbotham et al., 2017).

2.3 Risks underway

AI has been acknowledged as revolutionary with its ability to change business world dynamics (Rana et al., 2022). However, the concept is not free from drawbacks and the integration of AI into organizational activities heralds the emergence of new obstacles and difficulties (Duan et al., 2019). Businesses that rely on cutting-edge methods should exercise caution when implementing AI/ML procedures into their daily operations (Vidgen et al., 2017). This is due to the fact that the application of AI/ML depends on the availability of substantial amounts of timely, high-quality data (Lichtenthaler, 2019). Additionally, even though businesses devote time, energy and resources to the adoption process, the benefits of AI may not materialize as anticipated (Makarius et al., 2020). The employment of AI, therefore, has a number of negative effects beyond only skewed results including black-box algorithms, security issues, lack of transparency and accountability, harm to society and the environment and concerns about systems' security (Yudkowsky, 2008). Several current instances demonstrate that with a lack of proper governing AI policies, serious and unwanted repercussions are bound to occur (Keding, 2020).

2.3.1 Privacy of data

Consider the example of digital watches these days that capture the human body's blood pressure level, oxygen level, heartbeat rate and pulse rate besides providing time. The beauty of data science is that the company manufacturing these watches can record the individual data for each of these watches and can trade this data with some third party like insurance companies. These insurance providers, with the help of data analytics, can then decide the premium to be charged from different customers on predetermined health parameters.

Here, businesses must exercise caution and protect the privacy of customer and employee data (Lee et al., 2019). As a result, companies' ability to fulfil their social obligations will face numerous new obstacles as a result of the deployment of AI (Kumar et al., 2019). Other instances include bigotry and discrimination-related issues. When designing and training AI/ML systems, human bias has the potential to introduce bias into the algorithms. There have been numerous news stories on biased AI results involving racial and gender prejudice (Vigdor, 2019; Zuiderveen Borgesius, 2020). For instance, *Amazon* found that its internal hiring tool was excluding female candidates because it was programmed to favour men over women in previous hiring choices. Such outcomes may negatively impact the companies' reputations and can result in customer defection and huge financial losses including those imposed by the government in the form of compensations, fines and penalties (Engler, 2021).

2.3.2 Clustered data

The landscape of processing and analyzing data has been completely changed since the dawn of this century. The calibre of the data itself is an important part of the data utilized to train the AI (Baier et al., 2019; Lee et al., 2019). Having high-quality data is essential for making accurate forecasts (Alsheibani et al., 2018). AI is data-based, thus if the underlying data is prejudiced or discriminatory, the outputs may also be. AI systems do not comprehend the inputs they receive or the outputs they produce (Keding, 2020). According to Lee et al. (2019), the guiding concept of AI is "garbage-in, garbage-out", which states that if loaded data is of poor quality, so will be the insights produced by the AI, making them useless in the context of organizational goals. Locating and cleaning up these many data sources, then combining them with already existing AI systems needs special attention (Davenport & Ronanki, 2018).

2.3.3 Cyber security

Adoption of these advanced technologies widens the risk of potential attacks and introduces new, distinctive security dangers (Rana et al., 2022). AI/ML systems are susceptible to novel vulnerabilities in addition to the usual cyber threats from human or software errors. The uses of AI have allowed researchers to accomplish a lot of good. However, AI systems can be employed maliciously or even dangerously if, in the wrong hands (Boukherouaa et al., 2021). These attacks concentrate on changing data at a certain point in the AI/ML lifecycle to take advantage of built-in limitations in AI/ML algorithms (Comiter, 2019). Such manipulation enables attackers to avoid detection and causes AI/ML to make poor judgment calls or gather/produce irrelevant data (Mizuta, 2022). The intricacy of ML models and the potential effects they could have on financial sector institutions necessitate ongoing oversight to ensure that such malware attacks are noticed in due time and prevented accordingly (Collins et al., 2021).

Along with general cybersecurity worries, the other unique cyber dangers to AI/ML malfunctioning can be roughly categorized into: *Data poisoning* – By including unique samples in a machine learning algorithm's training data set, data poisoning attacks seek to alter the algorithm's behaviour during the phase (Enholm et al., 2021). The AI will improperly learn to categorize or recognize information as a result of these attacks. Data poisoning attacks necessitate privileged access to the model but once carried out successfully and provided the malicious behaviour does not interfere with routine diagnostic testing, compromised models may be invisible to human oversight; *Input attacks* allow attackers to manipulate data inputs using manipulated algorithms to trick AI systems while they're operating; *Model extraction* or model inversion allows attackers with read-only access to the model to recover pre-existing inputs or data using model inversion assaults (Rana et al., 2022). The General Data Protection Regulations passed by the EU fails to address these kinds of concerns (Boukherouaa et al., 2021).

Moreover, regulators in the banking sector are growing more concerned about cybersecurity related to AI/ML. Cyber dangers from AI and ML could jeopardize the financial sector's integrity and trust (Fosso Wamba et al., 2015). The ability of the financial sector to effectively assess, price and manage risks could be compromised by corrupt systems, which could result in the accumulation of hidden systemic problems (Miller, 2018).

2.3.4 Rigid culture

Job automation is generally viewed as the most immediate concern, especially for the employees. It's no longer a matter of if AI will replace certain types of jobs, but to what extent (Dedrick et al., 2013). In many industries particularly but not exclusively where workers perform predictable and repetitive tasks, disruption is well underway (Demlehner & Laumer, 2020). According to a Brookings Institution study, 36 million people work in jobs with "high exposure" to automation, meaning that before long, at least 70% of their tasks ranging from retail sales and market analysis to hospitality and warehouse labour will be performed using AI (Gupta, 2022).

Machines are being used instead of humans to take more efficient decisions (Mishra & Pani, 2020). The implementation of AI may modify the roles of the organization's employees because in many of these situations, people were in charge of doing these routine activities. The adoption of AI is a difficult process that involves numerous organizational and technological challenges (Enholm et al., 2021). The abilities of the organization's human resources are crucial to the application of AI (Nam et al., 2019).

Negative impressions can cause digital transformation to be stiff and to become lifeless (Lichtenthaler, 2019). Prior thoughts in the form of the intricate nature of AI can invite new risks (Lee et al., 2019). According to a study, a company's culture might have a significant impact on the decision to implement AI (Mikalef & Gupta, 2021; Pumplun et al., 2019). Healthy organizational cultures are more likely to adopt AI technologies because they are passionate about and eager to capitalize on novel, opportunistic ideas (Mikalef & Gupta, 2021). However, forceful implementation of such tools, the absence of skilled workforce or apprehension of existing employees to technology adoption can have serious negative outcomes for the organizational productivity (Enholm et al., 2021).

2.3.5 Trust issues between human and technology

It has been demonstrated that AI systems can automate manual operations or carry out actions that mimic human intellect (Zheng et al., 2017). Employees would have to collaborate or make choices using AI technology. They must therefore have faith in the AI system and comprehend how it functions in order to draw conclusions (Makarius et al., 2020). Building trust between people and robots can be challenging since human–machine interaction is a complex process. Thus, employee faith in AI implementation can be a barrier to its implementation with employees showing strong resistance to change for the sake of technology (Enholm et al., 2021). Increased tensions, disagreements and sentiments of mistrust towards the technology itself as well as towards the top management units that support its deployment are anticipated to result from the introduction of AI and the replacement or relocation of a number of conventional employment roles (Huang et al., 2019).

Further, if left unchecked, it's possible for AI's imperfections to cause physical harm. Let's look at self-driving cars, an AI application that is beginning to take hold in today's automobile market. If a self-driving car malfunctions and goes off-course, that poses an immediate risk to the passenger, other drivers on the road and pedestrians. This is going to put the functioning of automobile makers who are manufacturing AI-enabled vehicles in deep peril. Therefore, the twin challenges of managing the people within the organization in dealing with AI and ensuring successful corroboration between users of technology and the technology itself deserves careful attention (Yigitcanlar et al., 2020). The security issues here could be enhanced in light of the culmination of AI/ML opacity (Rana et al., 2022).

2.3.6 Increased regulations

The use of AI/ML by the supervisory team is important to carefully assess the risks and challenges that come with authority (Lichtenthaler, 2019). Data standardization, quality and completeness will be necessary for AI/ML-driven supervision to be effective, which could be difficult for management and regulated institutions to achieve, especially when using non-traditional sources of data like social media (Financial Stability Board, 2020). Resource and talent shortages could make it difficult for supervisory agencies to use AI/ML in an efficient and secure manner (Ashrafi et al., 2019).

Data scientists and experts in AI/ML could be added to the workforce to address this. Finally, implementing AI/ML systems exposes managers to dangers like those related to embedded bias, privacy, cybersecurity and outcome explainability (Boukherouaa et al., 2021). Stakeholders may regard bias as a possible source of operational and reputational issues given that it may develop as an unexpected result of AI/ML systems (Mayson, 2019). Therefore, appropriate plans need to be chalked out to properly mitigate and detect biasedness, ensure robustness of algorithms and risk management in terms of operational framework for the organization (Miller, 2018).

Computer programs that systematically and unfairly discriminate against some people or groups of people in favour of others are known as embedded bias (Friedman & Nissenbaum, 1996). Processes utilized by AI/ML for customer categorization may be biased due to differences in pricing or service quality (Silberg & Manyika, 2019). AI/ML models can learn to be prejudiced by using biased methods and data sets which frequently leads to bias in choices made by these systems (Rana et al., 2022). Data biases, such as incomplete and erroneous information may increase financial exclusion and foster mistrust of technology, especially among the poorest (Sahay et al., 2020).

Explainability of AI/ML systems outcomes is an important but complex and multifaceted issue, particularly when used in the financial sector (Arrieta et al., 2020). These applications are being described as black boxes because they are not easy to comprehend, they are not directly explainable by the user, their inputs might be unknown and they are an ensemble of models rather than a single independent model (Guidotti et al., 2019). This characteristic could make detection of the appropriateness of ML decisions difficult and could expose organizations to vulnerabilities such as unsuitable modelling techniques or incorrect decision-making with potential to undermine the trust factor in a dynamic business environment (Silberg & Manyika, 2019).

Stronger explainability, however, could make it easier for outsiders to decipher and manipulate the algorithms, posing concerns for the entire financial system (Huang et al., 2019). As a result, there is a trade-off between model explainability which refers to the model's ability to approximate various functions and model flexibility, which is directly tied to the model's parameter count (Guidotti et al., 2019). Compared to linear models, which generate results that are straightforward to interpret but are less accurate, ML models are more adaptable and accurate but less explicable (Boukherouaa et al., 2021).

In May 2018, the General Data Protection Regulations (GDPRs) were enforced by the European Union to regulate activities such as processing of personal data of citizens living in Europe and Scandinavian countries. As a result of this new rule, businesses using AI solutions struggle to upload or disclose the personal data gathered for use in training their intelligent robots (Pumplun et al., 2019). To comply with these new regulatory restrictions, many data sets must be anonymized, which makes the employment of sophisticated, self-learning algorithms more challenging or perhaps impossible (Pumplun et al., 2019). Explaining the reasoning behind AI/ML-based financial decisions is becoming an increasingly critical problem (Ashta & Herrmann, 2021). GDPR makes AI deployment more complicated (Baier et al., 2019), which may prevent AI from being gradually adopted. The intellectual property involved in AI algorithms and the data are further legal considerations that may prove to be barriers to the adoption of AI (Baier et al., 2019; Demlehner & Laumer, 2020).

2.3.7 Financial instability

Businesses that lack in-house skills or are unfamiliar with AI often have to outsource, which is where challenges of cost and maintenance come in (Duan et al., 2019). Due to their complex nature, smart technologies can be expensive and businesses can incur further costs for repair and ongoing maintenance (Dhanabalan & Sathish, 2018). The computational cost for training data models can also be an additional expense (Soni et al., 2020). Moreover, financial stability concerns regarding the AI/ML algorithms' resilience in the face of structural changes and increased interconnection due to a disproportionate reliance on a small number of AI/ML service providers may potentially surface (Rana et al., 2022).

In the event of significant and abrupt changes in the input data leading to the breakdown of established correlations (especially in response to a crisis), these sophisticated models and the underlying algorithms may not perform well, potentially leading to inaccurate decisions with unfavourable effects for business and financial institutions or their clients (Boukherouaa et al., 2021). To protect the financial system's integrity and safety, strong governmental responses will be needed in a time bound manner (Sahay et al., 2020).

In addition, ML algorithms are still vulnerable to abrupt structural changes in the data caused by unforeseeable events and the inaccuracies that arise could reduce the effectiveness of an institution, such as a bank's capacity for crisis monitoring and reaction (Afiouni, 2019). The acquisition of accurate and representative data as well as security issues may offer difficulties for various sectors (Bolhuis & Rayner, 2020). If the concerned companies lacked the tools and expertise necessary to operate AI/ML properly or to reduce the dangers involved, these worries would only grow. Other difficulties can include the costs in installing the necessary infrastructure, hiring right and capable employees and relying on external partners in providing the required expertise and assistance (Lui & Lamb, 2018).

In a recent paper, the European Commission outlines seven crucial factors that businesses should take into account when using AI applications (European Commission, 2019). These include characteristics like openness of AI applications, accountability, safety and security, societal and environmental well-being, design for universal access, as well as human agency and oversight which go beyond factors like data sensitivity. The rationale behind incorporating such things is to maximize public welfare and minimize the potential risks being faced by organizations in terms of data revolution (Arrieta et al., 2020).

3 THE DISGUISED SIDE OF AI SYSTEMS AND REMEDIES AHEAD

The tech community has long debated the threats posed by artificial intelligence. Automation of jobs, the spread of fake news and a dangerous arms race of AI-powered weaponry have been proposed as a few of the biggest dangers posed by AI (Borges et al., 2020). Moreover, the high success of these processing applications and the resultant benefits can further bridge the digital distance between the developed and developing countries such as USA and India respectively (Yigitcanlar et al., 2020). These economies lack the requisite investment and are lagging behind in terms of access to research and human resources (Dedrick et al., 2013). To close this gap, policymakers will need to create a framework built on four major policy pillars: infrastructural investment, conducive business environment policy, skill investment and risk management framework (International Monetary Fund, 2020).

Collaboration between nations and between the public and private sectors may be able to reduce the threat of a growing digital divide (Dedrick et al., 2013). Global initiatives have so far called for cooperation on investing in digital infrastructure, such as the creation of norms, to avoid ethical hazards linked with AI (UNESCO, 2021). Multilateral organizations could be crucial in information transfer, capital-raising, capacity-building and facilitation of a peer-learning strategy that can help strengthen digital policy efforts in emerging economies of Asia and Africa (Jelonek et al., 2019). In a similar vein, less developed and underdeveloped economies might be added to the membership of globally recognized AI groups like the Global Partnership on Artificial Intelligence and the OECD Network of Experts on AI (OECD, 2019).

Despite the possibility of bias in AI/ML systems, they might be able to lessen current prejudices (Duan et al., 2019). As per the study conducted by Mayson (2019), AI/ML systems might lessen decision makers' irrational biases since ML algorithms can minimize biases that result from the subjective interpretation of facts. Miller (2018) pointed out that despite being prone to embedded bias, AI systems can nonetheless enhance decision-making by reducing human prejudice. Finally, even though many AI/ML systems may be seen as being opaque to humans and being beyond their comprehension, their prediction and decision-making processes can be examined more thoroughly than those by other individuals, allowing for the identification and correction of any biases that may already exist (Jarrahi, 2018). By creating and implementing a larger framework for the governance and moral application of AI/ML, policy responses to embedded bias issues in AI/ML applications could be enhanced (Boukherouaa et al., 2021).

The financial sector's regulatory perimeter for *cybersecurity* standards should be widened to include cyber threats specifically related to AI/ML (Duan et al., 2019). It should be mandatory for developers and users of AI/ML applications in the financial sector to implement mitigating measures as part of their overall cybersecurity strategy (Keding, 2020). These can include methods for detection and reporting, strong safeguards for training data streams and measures to protect the privacy of models and data (LeCun et al., 2015).

AI/ML raises new, distinct *privacy* concerns. There are well-known privacy risks with big data, and technologies have been created to help safeguard data anonymity and data subjects' privacy (Fosso Wamba et al., 2015). To address these issues, legal data policy frameworks are being implemented all around the world (Yudkowsky, 2008). However, the ability of AI/ML models to effectively stop data leaking from the training data set presents fresh privacy issues (Roh et al., 2019). Similarly, after using the data, AI/ML may remember details about the training set's participants, or the model's conclusions may infer or explicitly reveal private information (Wamba et al., 2017). More work is needed to update the legal and regulatory framework that mandates AI/ML systems and related data sources to adhere to enhanced privacy standards as well as pertinent anti-money laundering/combating the financing of terrorism requirements (Boukherouaa et al., 2021).

Strong AI/ML algorithms will protect *financial stability* and the integrity of the financial system while also fostering public confidence in an AI-driven financial system (Lui & Lamb, 2018). Given the relatively low concentration of AI/ML service providers and the insufficient oversight capacity in many countries to effectively engage financial institutions utilizing AI/ML systems, this is particularly critical (Khandani et al., 2010). In order to guarantee the robustness of ML models with respect to cyber threats and privacy protection, as stated above, as well as with respect to their performance, a joint effort by industry and regulators would be needed (Baier et al., 2019). The latter deals with concerns linked to having an adequate governance system in place for the AI/ML development process and avoiding misleading signals produced by ML models during times of structural upheaval (Borges et al., 2020).

In a data environment that is largely stable and generates reliable signals, AI/ML systems have done well in the financial sector, but this might swiftly alter during times of fast structural depletion (Sahay et al., 2020). In an environment that is sufficiently stable, AI/ML models can reasonably take into account changing data trends without suffering a large loss in forecast accuracy (Chatterjee et al., 2021). When the structure of their data environment changes or when a formerly trustworthy signal loses its reliability or when behavioural correlations significantly change, they are faced with a more difficult task (Davenport & Ronanki, 2018). A good example is the newly noted misalignment of AI/ML-generated risk evaluations during the COVID-19 epidemic. ML models' performance was significantly impacted by the crisis because they weren't initially trained for it (Enholm et al., 2021).

For the development of AI/ML systems, new governance frameworks are required to increase prudential monitoring and prevent unforeseen outcomes (Gupta, 2022). Given how much the process of developing AI/ML systems resembles that of developing software, effective quality control and process agility are needed (Wang et al., 2019). The procedure might cover every stage of development, testing and deployment while concentrating on the pertinent risks and controls. The best practices of software development such as the separation of roles and ongoing monitoring should be followed while conducting tests for inherent bias, data tainting, other security threats and performance (Zetzsche et al., 2017).

Though their impact on financial stability has not yet been completely determined, the broad adoption of AI/ML technologies in the financial sector has the potential to be transformative (Ransbotham et al., 2017). As mentioned above, AI/ML systems may bring increased efficiencies, better risk assessment and management, appropriate pricing, improved regulatory compliance and new tools for prudential surveillance and enforcement, all of which will positively affect financial stability (Canhoto & Clear, 2020). However, due to the opaqueness of AI decisions, these systems pose new and distinct hazards like manipulation of information, vulnerability issues, privacy concerns and lack of robustness (Collins et al., 2021). These risks undermine public confidence in the honesty and security of a financial system powered by AI and ML (Galloway & Swiatek, 2018).

Additionally, they have the potential to create new systemic risk sources and channels (European Commission, 2019). More specifically, because of the high degree of specialization in AI/ML systems and the network effects that could make the financial system more susceptible to single points of failure, AI/ML service providers could become systemically significant players in the infrastructure of the financial market which might give birth to unexpected systemic risks throughout economies (Financial Stability Board, 2020). Therefore, democratization of AI is needed so as to build a pool of experts and technology-savvy engineers to narrow the skill gap and to take care of any future contingencies.

Additionally, the extensive application of AI/ML may make financial conditions more procyclical (Comiter, 2019). The procyclicality may be automated, accelerated and maybe hidden using AI/ML. Additionally, a flawed risk assessment and response by ML algorithms in the instance of a tail risk event could hasten the shock's propagation throughout the financial system, complicate or even impair the effectiveness of the policy response (Canhoto & Clear, 2020). These issues call for extensive regulatory action as well as cooperative initiatives.

A sufficient governmental response necessitates the creation of unambiguous minimum standards and norms for the industry as well as a larger emphasis on gaining the required technical capabilities (Mizuta, 2022). Collaboration between financial institutions, central banks, financial supervisors and other stakeholders is crucial so as to avoid duplication of effort and to help mitigate potential risks (Lui & Lamb, 2018). Many prominent countries in the AI/ML space have relied on clearly stated national AI strategies to encourage the development of AI/ML while preventing regulatory lapses (Kleinings, 2022).

4 CONCLUSION, LIMITATIONS AND FUTURE RESEARCH AGENDA

Today, AI has become the de-facto technology used by all large and small, start-up and fintech companies to build their platforms. It is estimated that within the next few years, AI-powered financial services will be the only medium of interaction for users, making financial products and lending available to the masses even in the remote towns of the country, thereby making financial inclusion a close reality (Riikkinen et al., 2018). It could be understood from the above-furnished details that AI would be the future for the business segments of the country in serving their customers and in achieving greater financial progression (Anon, 2020). But, careful thought needs to be given to find out where the parity lies in investing in AI resources and the expected return that might occur along the overall supply chain (Ashrafi et al., 2019).

Businesses are still having trouble integrating and using AI in their day-to-day operations. A comprehensive understanding is consequently required so as to pick out how AI technologies create business value and what kind of business value is anticipated (Canhoto & Clear, 2020). Although it is believed that organizations may get significant business value from bringing in AI, only a very limited number of businesses have, yet, embraced and deployed AI applications beyond pilot projects (Anon, 2020). Due to recent improvements in computer technology, network speeds, the large amount of available data and processing methods, AI has attracted a lot of attention in recent years (Baby et al., 2017). However, there is much ambiguity on what AI actually is (Alsheibani et al., 2018).

Bots are the kind of innovation in the AI-powered sector which are reinstating the role of humans and their usage has been growing gradually (Prentice et al., 2020). Hence huge investments are being made on this by different industries those who are envisaging this technology as a prolonged cost-cutting investment (Wang et al., 2022). It helps industries in reducing human-induced losses, achieving better customer satisfaction and saving money from recruiting humans and also keeps away from human errors in this process (Chatterjee et al., 2021).

Historically, analytics were used almost exclusively to support internal decisions. That's still useful, of course, but now companies are also using data analytics to create new products and services (Campbell et al., 2020). And it's not just the digital players you would expect like Google and LinkedIn but also players in FMCG, automobiles, fashion and clothing (Mariani & Wamba, 2020). This is a new option for organizations that managers need to understand and explore. And in today's business world, not knowing about analytics can be dangerous to you and your company's prosperity (Chandola et al., 2009). Furthermore, AI-BA integration will feed no tangible convenience or become opaque in light of malformed data or poor governance (Rana et al., 2022).

The deployment of AI/ML systems in businesses providing financial and other services will continue to accelerate. Rapid advances in modelling and use-case adaptations as well as major advancements in processing power, data storage capacity and big data are what are driving this development (Castillo et al., 2020). The COVID-19 epidemic is quickening the transition to a contactless world and increasing use of digital financial services will increase the appeal of AI/ML systems to those who provide these services and those who avail it (Kleinings, 2022).

Using AI/ML will have many advantages. For instance, companies and financial institutions can benefit from AI/ML systems' potential in terms of significant cost reductions and efficiency improvements, new markets, products, improved risk management, new customer experiences and lower prices as well as strong tools for regulatory compliance and prudential oversight. However, it will also keep the window of considerable financial policy issues and uncertainties wide open (Chatterjee et al., 2021).

These technologies raise moral concerns and new, distinct threats to the general integrity and safety of the financial system, the full scope of which has not yet been determined (Castillo et al., 2020). The fact that these innovations are still developing and changing as new technologies are put into use, makes the task that commercial firms are facing even more difficult (Lee et al., 2019). These developments necessitate enhancing oversight monitoring mechanisms and actively engaging stakeholders in order to pinpoint potential hazards and adopt corrective regulatory measures (Lui & Lamb, 2018).

The development of AI/ML in consumer markets should be universally welcomed by regulators and they should make the necessary preparations to take advantage of its potential benefits and reduce its hazards (Mayson, 2019). This entails timely institutional capacity upgrading, the acquisition of pertinent skills, the accumulation of knowledge, the enhancement of external communication with stakeholders and the expansion of consumer education. The financial sector has found that deploying AI/ML systems works best when there is a national AI strategy in place that incorporates all pertinent public and commercial authorities (Soni et al., 2020).

At the regional and international levels, cooperation and knowledge sharing are becoming more and more crucial (Ashta & Herrmann, 2021). As a result, operations may be coordinated and information and experience could be shared to assist the safe deployment of AI/ML systems. To make sure that less developed economies have access to knowledge about techniques and processes, use cases and regulatory and supervisory approaches, cooperation will be especially crucial (Miller, 2018).

Finally, because AI/ML technology is still emerging and is being used by a variety of partners, neither users, technology providers, developers nor regulators fully comprehend the technology's merits and shortcomings at this time (OECD, 2019). Due to the possibility of numerous unanticipated hazards, countries will need to improve their monitoring and prudential oversight (Riikkinen et al., 2018). Making machines smarter is the wisest way to make effective use of current human resources towards innovating new strategies for the betterment of the organization and the economy (Sestino & De Mauro, 2022).

AI assists in human decision-making. This is because algorithms are evolving and computing power is getting smarter (Mikalef et al., 2021). True, AI cannot complete all common-sense jobs yet it processes and interprets data faster than the human brain. It then provides us with many composite courses to figure out the possible consequences of actions undertaken and streamline business decision-making (Mariani & Wamba, 2020). When this coalesces with human interaction, it helps to meet consumers' demands. We live in a world where virtual interactions/shopping is increasing day by day. Irrespective of B2B or B2C businesses, the virtual work is buzzing 24×7 where customers can interact with your business at every possible hour (Soni et al., 2020). Hence, it is high time that we look closely and understand what the customer needs really are. Only when one has this clarity, can one begin looking for the correct data and its AI applications.

AI and Machine Learning have revolutionized and will continue to revolutionize businesses for many years to come. From IT operations to sales, implementing AI into business environments cuts down on time spent on repetitive tasks, improves employee productivity and enhances the overall customer experience (Ramanathan et al., 2017). It also helps avoid mistakes and detect potential crises at a level unattainable to humans. No wonder organizations are leveraging it to improve a number of business areas, from logistics all the way through to recruiting and employment. Before incorporating AI into internal operations, it is essential to have a clear understanding of what you want AI to do to your data. It's our conviction that companies at the forefront of AI will reap the financial advantages and dominate the competition in the future.

4.1 Limitations

The present work ought to investigate the increased acceptance of modern sophisticated tools such as AI/ML in the day-to-day working of the domestic companies and how it can better the performance of companies in times to come. It has been revealed that earlier adoption of these techniques can significantly contribute to business value and can help companies to remain competitive in the longer run. Although the study provides a comprehensive view regarding the present state of AI implementation, the study is not free from limitations. *To start with*, the study puts forward a general view of the AI environment in the country which may not present the true picture of AI acceptance. Therefore, future studies can investigate AI implementation by carrying out empirical research so as to find out the actual factors that lead to the adoption of AI or which factors contribute to AI integration. *In addition*, future researchers could identify select companies or sectors which are more prone to AI adoption than others. *Lastly*, a comparative study can be undertaken to investigate the performance of companies working in the same or different (manufacturing vs. services) category so as to find out how AI really contributes to the firm's success in terms of added business value.

REFERENCES

Afiouni R (2019) Organizational learning in the rise of machine learning. In *International Conference on Information Systems*, Munich, Germany.

Alsheibani S, Cheung Y, Messom C (2018) Artificial intelligence adoption: AI-readiness at firm-level. *Artificial Intelligence* 6: 26.

Anon (2020) *AI Is Essential to India's Future of Work*. Press wire.

Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, Benjamins R (2020) Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* 58: 82–115. https://doi.org/10.1016/j.inffus.2019.12.012.

Ashrafi A, Ravasan AZ, Trkman P, Afshari S (2019) The role of business analytics capabilities in bolstering firms' agility and performance. *International Journal of Information Management* 47: 1–15. https://doi.org/10.1016/j.ijinfomgt.2018.12.005.

Ashta A, Herrmann H (2021) Artificial intelligence and fintech: An overview of opportunities and risks for banking, investments, and microfinance. *Strategic Change* 30(3): 211–222. https://doi.org/10.1002/jsc.2404.

Baby CJ, Khan FA, Swathi J (2017) Home automation using IoT and a chatbot using natural language processing. 2017 Innovations in Power and Advanced Computing Technologies (iPACT).

Baier L, Jöhren F, Seebacher S (2019) Challenges in the deployment and operation of machine learning in practice. In *Proceedings of the 27th European Conference on Information Systems (ECIS)*, Stockholm, Sweden.

Bolhuis MA, Rayner B (2020) *The More the Merrier? A Machine Learning Algorithm for Optimal Pooling of Panel Data*. International Monetary Fund.

Boukherouaa EB, AlAjmi K, Deodoro J, Farias A, Ravikumar R (2021) Powering the digital economy: opportunities and risks of artificial intelligence in finance. *Departmental Papers* 24. https://doi.org/10.5089/9781589063952.087.

Borges AF, Laurindo FJ, Spínola MM, Gonçalves RF, Mattos CA (2020) The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. *International Journal of Information Management* 57: 102225.

Bytniewski A, Matouk K, Chojnacka-Komorowska A, Hernes M, Zawadzki A, Kozina A (2020) The functionalities of cognitive technology in management control system. In *Asian Conference on Intelligent Information and Database Systems*, Phuket, Thailand.

Canhoto AI, Clear F (2020) Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential. *Business Horizons* 63(2): 183–193. https://doi.org/10.1016/j.bushor.2019.11.003.

Campbell C, Sands S, Ferraro C, Tsao HYJ, Mavrommatis A (2020) From data to action: How marketers can leverage AI. *Business Horizons* 63(2): 227–243. https://doi.org/10.1016/j.bushor.2019.12.002.

Castillo D, Canhoto AI, Said E (2020) The dark side of AI powered service interactions: Exploring the process of co-destruction from the customer perspective. *The Service Industries Journal* 1–26. https://doi.org/10.1080/02642069.2020.1787993.

Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: A survey. *ACM Computing Surveys* 41(3): 1–58. https://doi.org/10.1145/1541880.1541882.

Chatterjee S, Chaudhuri R, Vrontis D, Thrassou A, Ghosh, SK (2021) Adoption of artificial intelligence-integrated CRM systems in agile organizations in India. *Technological Forecasting and Social Change* 168: 120783. https://doi.org/10.1016/j.techfore.2021.120783.

Collins C, Dennehy D, Conboy K, Mikalef P (2021) Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management* 60: 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383.

Comiter M (2019) Attacking artificial intelligence: AI's security vulnerability and what policymakers can do about it. Belfer Center for Science and International Affairs, Harvard Kennedy School. https://www.belfercenter.org/publication/AttackingAI.

Davenport T, Guha A, Grewal D, Bressgott T (2020) How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science* 48(1): 24–42. https://doi.org/10.1007/s11747-019-00696-0.

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review* 96(1): 108–116.

Dedrick J, Kraemer KL, Shih E (2013) Information technology and productivity in developed and developing countries. *Journal of Management Information Systems* 30(1): 97–122. https://doi.org/10.2753/MIS0742-1222300103.

Delen D, Zolbanin HM (2018) The analytics paradigm in business research. *Journal of Business Research* 90: 186–195. https://doi.org/10.1016/j.jbusres.2018.05. 013.

Demlehner Q, Laumer S (2020) Shall we use it or not? Explaining the adoption of artificial intelligence for car manufacturing purposes. In *Proceedings of the 28th European Conference on Information Systems (ECIS)*, Marrakesh, Morocco. Online.

Dhanabalan T, Sathish A (2018) Transforming Indian industries through artificial intelligence and robotics in Industry 4.0. *International Journal of Mechanical Engineering and Technology* 9(10): 835–845.

Duan Y, Edwards JS, Dwivedi YK (2019) Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda. *International Journal of Information Management* 48: 63–71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021.

Engler, A. (2021). Auditing employment algorithms for discrimination. Retrieved 25/09 from https://www.brookings.edu/research/auditingemployment-algorithms-for-discrimination/.

Enholm IM, Papagiannidis E, Mikalef P, Krogstie J (2021) Artificial intelligence and business value: A literature review. *Information Systems Frontiers* 1–26. https://doi.org/10.1007/s10796-021-10186-w.

Eriksson T, Bigi A, Bonera M (2020) Think with me, or think for me? On the future role of artificial intelligence in marketing strategy formulation. *TQM Journal* 32(4): 795–814. https://doi.org/10.1108/TQM-12-2019-0303.

European Commission (2019) Proposal for a Regulation laying down harmonised rules on artificial intelligence. Retrieved 15/08 from https://digital-strategy.ec.europa.eu/en/library/proposal-regulationlaying-down-harmonised-rules-artificial-intelligent.

Fosso Wamba S, Akter S, Edwards A, Chopin G, Gnanzou D (2015) How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics* 165: 234–246. https://doi.org/10.1016/j.ijpe.2014.12.031.

Friedman, B, Nissenbaum, H (1996). Bias in Computer Systems. *ACM Transactions on Information Systems* 14(3): 330–347. https://doi.org/10.1145/230538.230561.

Financial Stability Board, FSB (2020) The use of supervisory and regulatory technology by authorities and regulated institutions: Market developments and financial stability implications. Financial Stability Board, Basel, Switzerland. https://www.fsb.org/wp-content/uploads/P091020.pdf.

Galloway C, Swiatek L (2018) Public relations and artificial intelligence: It's not (just) about robots. *Public Relations Review* 44(5): 734–740. https://doi.org/10.1016/j.pubrev.2018.10.008.

Gupta (2022) https://techgraph.co/opinions/how-ai-is-crucial-for-business-strategy-management/.

Guidotti RA, Monreale S, Ruggieri F, Turini D, Pedreschi F, Giannotti (2019) A survey of methods for explaining black box models. *ACM Computing Surveys* 51(5): 1–42. https://doi.org/10.1145/3236009.

Hair JF (2007) Knowledge creation in marketing: The role of predictive analytics. *European Business Review* 19(4): 303–315. https://doi.org/10.1108/09555340710760134.

Huang MH, Rust R, Maksimovic V (2019) The feeling economy: Managing in the next generation of artificial intelligence (AI). *California Management Review* 61(4): 43–65. https://doi.org/10.1177/0008125619863436.

International Monetary Fund, IMF (2020) *Regional Economic Outlook for Sub-Saharan Africa*, April. Washington, DC.

Jarrahi MH (2018) Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons* 61(4): 577–586. https://doi.org/10.1016/j.bushor.2018.03.007.

Jelonek D, Mesjasz-Lech A, Stępniak C, Turek T, Ziora L (2019) The artificial intelligence application in the management of contemporary organization: Theoretical assumptions, current practices and research review. In *Future of Information and Communication Conference*, San Francisco, CA.

Keding C (2020) Understanding the interplay of artificial intelligence and strategic management: Four decades of research in review. *Management Review Quarterly* 71(1): 91–134. https://doi.org/10.1007/s11301-020-00181-x.

Khandani A, Adlar K, Lo A (2010) Consumer credit-risk models via machine learning algorithms. *Journal of Banking & Finance* 34(11): 2767–787. https://doi.org/10.1016/j.jbankfin.2010.06.001.

Kleinings H (2022) https://levity.ai/blog/8-uses-ai-business.

Kumar V, Rajan B, Venkatesan R, Lecinski J (2019) Understanding the role of artificial intelligence in personalized engagement marketing. *California Management Review* 61(4): 135–155. https://doi.org/10.1177/0008125619859317.

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Mature* 521(7553): 436–444. https://doi.org/10.1038/nature14539.

Lee J, Suh T, Roy D, Baucus M (2019) Emerging technology and business model innovation: The case of artificial intelligence. *Journal of Open Innovation, Technology, Market and Complexity* 5(3): 44. https://doi.org/10.3390/joitmc5030044.

Lichtenthaler U (2019) An intelligence-based view of firm performance: profiting from artificial intelligence. *Journal of Innovation Management* 7(1): 7–20. https://doi.org/10.24840/2183-0606_007.001_0002.

Lichtenthaler U (2019) Extremes of acceptance: employee attitudes toward artificial intelligence. *Journal of Business Strategy*. https://doi.org/10.1108/JBS-12-2018-0204.

Lui A, Lamb GW (2018) Artificial intelligence and augmented intelligence collaboration: Regaining trust and confidence in the financial sector. *Information & Communications Technology Law* 27(3): 267–283. https://doi.org/10.1080/13600834.2018.1488659.

Makarius EE, Mukherjee D, Fox JD, Fox AK (2020) Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. *Journal of Business Research* 120: 262–273. https://doi.org/10.1016/j.jbusres.2020.07.045.

Malthouse EC, Haenlein M, Skiera B, Wege E, Zhang M (2013) Managing customer relationships in the social media era: Introducing the social CRM house. *Journal of Interactive Marketing* 27(4): 270–280. https://doi.org/10.1016/j.intmar.2013.09.008.

Mariani MM, Wamba SF (2020) Exploring how consumer goods companies innovate in the digital age: The role of big data analytics companies. *Journal of Business Research* 121: 338–352. https://doi.org/10.1016/j.jbusres.2020.09.012.

Mayson SG (2019) Bias in, bias out. *The Yale Law Journal* 128(8): 2218–300. https://ssrn.com/abstract=3257004.

Mikalef P, Conboy K, Krogstie J (2021) Artificial intelligence as an enabler of B2B marketing: A dynamic capabilities micro-foundations approach. *Industrial Marketing Management* 98: 80–92. https://doi.org/10.1016/j.indmarman.2021.08.003.

Mikalef P, Gupta M (2021) Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Information & Management*. https://doi.org/10.1016/j.im. 2021.103434.

Miller AP (2018) Want less-biased decisions? Use algorithms. *Harvard Business Review*, July 26. https://hbr.org/2018/07/want-less-biased-decisions-use-algorithms.

Mishra AN, Pani AK (2020) Business value appropriation roadmap for artificial intelligence. *VINE Journal of Information and Knowledge Management Systems* 51(3): 353–368. https://doi.org/10.1108/VJIKMS-07-2019-0107.

Mizuta T (2022). Artificial intelligence for financial markets: A good AI for designing better financial markets and a bad AI for manipulating markets. *Digital Designs for Money, Markets, and Social Dilemmas* 305–329. https://doi.org/10.1007/978-981-19-0937-5_13.

Nam D, Lee J, Lee H (2019) Business analytics use in CRM: A nomological net from IT competence to CRM performance. *International Journal of Information Management* 45(4): 233–245. https://doi.org/10.1016/j. ijinfomgt.2018.01.005.

Organisation for Economic Co-operation and Development, OECD (2019). Recommendation of the Council on Artificial Intelligence. OECD Legal Instruments, OECD, Paris. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

Prentice C, Lopes D, Wang XS (2020) The impact of artificial intelligence and employee service quality on customer satisfaction and loyalty. *Journal of Hospitality Marketing & Management* 29(7): 739–756. https://doi.org/10.1080/19368623.2020.1722304.

Pumplun L, Tauchert C, Heidt M (2019). A new organizational chassis for artificial intelligence-exploring organizational readiness factors. In *Proceedings of the 27th European Conference on Information Systems (ECIS)*, Stockholm & Uppsala, Sweden.

Rajamohan S, Dhanabalan T (2013) Rural development through ICT: Opportunities and challenges. *International Journal of Retailing & Rural Business Perspectives* 2(2): 354.

Ramanathan R, Philpott E, Duan Y, Cao G (2017) Adoption of business analytics and impact on performance: A qualitative study in retail. *Production Planning & Control* 28(11–12): 985–998. https://doi.org/10.1080/09537287.2017.1336800.

Rana NP, Chatterjee S, Dwivedi YK, Akter S (2022). Understanding dark side of artificial intelligence (AI) integrated business analytics: Assessing firm's operational inefficiency and competitiveness. *European Journal of Information Systems* 31(3): 364–387. https://doi.org/10.1080/0960085X.2021.1955628.

Riikkinen M, Saarijärvi H, Sarlin P, Lähteenmäki I (2018) Using artificial intelligence to create value in insurance. *International Journal of Bank Marketing* 36(6): 1145–1168. https://doi.org/10.1108/IJBM-01-2017-0015.

Roberts N, Grover V (2012) Investigating firm's customer agility and firm performance: The importance of aligning sense and respond capabilities. *Journal of Business Research* 65(5): 579–585. https://doi.org/10.1016/j.jbusres.2011.02.009.

Roh Y, He G, Whang SE (2019) A survey on data collection for machine learning: A big data-ai integration perspective. *IEEE Transactions on Knowledge and Data Engineering* 33(4): 1328–1347. https://doi.org/10.1109/TKDE.2019.2946162.

Ransbotham S, Kiron D, Gerbert P, Reeves M (2017) Reshaping business with artificial intelligence: Closing the gap between ambition and action. *MIT Sloan Management Review* 59(1): 145–155.

Sahay R, Von Allmen UE, Lahreche A, Khera P, Ogawa S, Bazarbash M, Beaton K (2020) The promise of Fintech: Financial inclusion in the post-COVID-19 Era. Departmental Paper 20/09, International Monetary Fund, Washington, DC.

Schmidt R, Zimmermann A, Moehring M, Keller B (2020) Value creation in connectionist artificial intelligence–A research agenda. AMCIS proceedings – Advances in information systems research, Online.

Sestino A, De Mauro A (2022) Leveraging artificial intelligence in business: Implications, applications and methods. *Technology Analysis & Strategic Management* 34(1): 16–29. https://doi.org/10.1080/09537325.2021.1883583.

Silberg J, Manyika J (2019) *Notes from the AI Frontier: Tackling Bias in AI (and in Humans)*. McKinsey Global Institute.

Soni N, Sharma EK, Singh N, Kapoor A (2020) Artificial intelligence in business: From research and innovation to market deployment. *Procedia Computer Science* 167: 2200–2210. https://doi.org/10.1016/j.procs.2020.03.272.

United Nations Educational, Scientific and Cultural Organization, UNESCO (2021) Intergovernmental meeting of experts (category II) related to a draft recommendation on the ethics of artificial intelligence. https://unesdoc.unesco.org/ark:/48223/pf0000373434.

Vidgen R, Shaw S, Grant DB (2017) Management challenges in creating value from business analytics. *European Journal of Operational Research* 261(2): 626–639. https://doi.org/10.1016/j.ejor.2017.02.023.

Vigdor N (2019) Apple card investigated after gender discrimination complaints. *The New York Times*. Retrieved 15/09 from https://www.nytimes.com/2019/11/10/business/Apple-credit-cardinvestigation.html.

Wamba, SF, Gunasekaran A, Akter S, Ren SJF, Dubey R, Childe SJ (2017) Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research* 70: 356–365. https://doi.org/10.1016/j.jbusres.2016.08.009.

Wamba-Taguimdje SL, Wamba SF, Kamdjoug JRK, Wanko CET (2020) Influence of artificial intelligence (AI) on firm performance: The business value of AI-based transformation projects. *Business Process Management Journal* 26(7): 1893–1924. https://doi.org/10.1108/BPMJ-10-2019-0411.

Wang H, Huang J, Zhang Z (2019) The impact of deep learning on organizational agility. In *Proceedings of the 40th International Conference on Information Systems (ICIS)*, Munich, Germany.

Wang X, Lin X, Shao B (2022) How does artificial intelligence create business agility? Evidence from chatbots. *International Journal of Information Management* 66: 102535. https://doi.org/10.1007/978-981-19-0937-5_13.

Yigitcanlar T, Desouza KC, Butler L, Roozkhosh F (2020) Contributions and risks of artificial intelligence (AI) in building smarter cities: Insights from a systematic review of the literature. *Energies* 13(6): 1473. https://doi.org/10.3390/en13061473.

Yudkowsky E (2008) Artificial intelligence as a positive and negative factor in global risk. *Global Catastrophic Risks* 1(303): 184.

Zetzsche DA, Buckley RP, Barberis JN, Arner DW (2017) Regulating a revolution: From regulatory sandboxes to smart regulation. *Fordham Journal of Corporate & Financial Law* 23: 31.

Zheng N, Liu Z, Ren P, Ma Y, Chen S, Yu S, Wang F (2017) Hybrid-augmented intelligence: Collaboration and cognition. *Frontiers of Information Technology & Electronic Engineering* 18(2): 153–179. https://doi.org/10.1631/FITEE.1700053.

Zuiderveen Borgesius FJ (2020) Strengthening legal protection against discrimination by algorithms and artificial intelligence. *International Journal of Human Rights* 24(10): 1572–1593. https://doi.org/10.1080/13642987.2020.1743976.

13 Examine manipulation of financial statements of commercial banks – Evidence from an emerging country

Tam To Nguyen, Ha Thi Thu Nguyen and Nga Thi Hong Nguyen

1 INTRODUCTION

Commercial banks are the lifeblood of the economy; hence, many related parties are interested in commercial banks' information. Financial statements provide specific information to evaluate the performance and financial status of commercial banks. The financial information helps to direct, supervise the bank's capital, and assess the bank's capital mobilization capacity of the operation. Based on their information, several indicators are calculated to evaluate comprehensively in the bank. In addition, it also helps investors, bank managers, and other related parties to make strategic decisions. In general, a bank's financial statements are an important basis for analysis, development management, and a basis for building a business strategy to help business operations more efficiently and maximize profits.

There is a requirement of a whole economy and society that the financial statements of commercial banks must be true and fair. Fraudulent financial statements of commercial banks have a serious impact and have always been of interest to researchers. In the current context, the process of using published financial statements has an increasingly important role in management and investment. With a risk-limiting mentality, investors want to know clearly whether the cash flow they invest in will be used effectively.

In Vietnam, recently, the difference in the results between public entities' financial statements before and after auditing has created doubts (Tam, 2014; VNBusiness, 2022). This factor has a negative impact on the decision-making of investors. It is also a big challenge for managers as well as for auditors in detecting material misstatements in financial statements. Therefore, fraud has always been a topic of interest in all fields of the economy.

DOI: 10.1201/9780367854690-13

Many previous researchers have conducted studies to detect and identify fraud in financial statements, such as the "Fraud Triangle" model, F-Score, M-score, etc. Fraud- and errors-detecting in financial statements aims to improve the transparency and reliability of the information, facilitating decision-making of investors.

This study uses the M-score model to identify fraudulent financial statements of 35 commercial banks in Vietnam to give recommendations for bank managers, and investors, and other parties in using financial information to make business decisions. This chapter has three main objectives: (1) a literature review of the manipulation probability of financial statements in banks and other industries; (2) a theoretical basis of material misstatement in financial statements, and models of identification of manipulation probability in financial statements; (3) application of the M-score model in Vietnamese commercial bank, and giving several recommendations of banks' financial statements.

2 LITERATURE REVIEW

2.1 Model to detect fraud or manipulation probability in non-financial companies

Beneish (1999) built the M-score model to detect fraud in companies' financial statements by using a unit probability model based on weighted exogenous sample maximum likelihood probit as well as an unweighted probit. This research results show that the threshold value of the model is −1.78, any company having M-score greater than −1.78 shows a sign of profit manipulation.

The Beneish model (M-score) has been used frequently in studies of the prediction of financial statements' manipulation of Vietnamese listed companies (Nguyen Cong Phuong, 2014; Tran, 2014; Nguyen Huu Anh, 2016; Hariri & Widjajanti, 2017). Nguyen Cong Phuong (2014) showed that the M-score model is one of the effective tools in detecting the earnings management of companies and can be used to improve the quality of financial statements and protect investors. The Beneish model has been used in a sample of 30 listed companies in Vietnam that have discovered material misstatements by auditors. The Beneish model gave the probability of correct prediction to the rate of 53.33% of the companies in the sample. Tran (2014) uses the Beneish model to predict the probability of material misstatement of financial statements of listed companies on the Vietnamese stock market.

Hariri and Widjajanti (2017) used the qualitative descriptive study related Beniesh M-Score model to detect material misstatement in the financial statements of two companies. The results showed that the management has carried out manipulating numbers on the financial statements. The chapter recommended regulators should inspect the company's financial status that has been booked in the financial statements and the professional accountants who have been engaged to audit the financial statements to be more independent. These independent examinations detect material misstatement of financial statements rather than using the ratio analysis method.

The logit regression method has been applied to indicate the impact of the fraud triangle under the guidance of auditing standards in detecting fraud in companies that were listed on the Ho Chi Minh Stock Exchange. The model used a sample of

78 companies listed on HOSE in 2012 and showed that the possibility of fraud is statistically significant with the samples in the fraud triangle, and the model is capable to predict 83.33% of the companies in the sample (Tran Thi Giang Tan et al., 2014).

Nguyen Anh Hien (2015) has researched and applied three research models of earning management, namely, Jones's model, Dechow, and the model of Kothari to apply Vietnamese listed companies. The research sample was conducted on 280 companies in Hanoi (HNX) and Ho Chi Minh (HOSE). The research results show that the models of Dechow and Kothari are meaningful in identifying earnings management.

Several studies use the derivative model of the original Beneish model, such as the mathematical model of American researcher Messod Beneish (Natalia V. Feruleva, 2017). The research focuses on identifying companies that overestimate net assets and financial results. As a result, it is possible to improve a fraud detection probit model and a linear model (integrated M-Score) that allows stakeholders to uncover fraud with an accuracy of 83, respectively, and 60%.

Patricia M. Dechow et al. (2011) studied a fraud prediction model at all three levels and named it the F-score. The research results show that, with a threshold value of 1, when the F-score is greater than 1, there is a probability that the project is correct in 65.9% of the companies' financial statements. This figure is 65.78% and 66.38% for the F-score 2 and F-score 3 models. Rasa Kanapickiene (2015) investigated the influence of financial ratios by the logistic regression method. Research results show that when the probability of predicting fraud is greater than 50%, 84.8% of the predicted results are correct.

Patric Bimpong et al. (2020) showed that Altman (2013), Taffler (1983), and Beneish (1999) models should be useful in predicting fraudulent financial statements in the banking and mining industries in Ghana. This study helps to take into account mergers and acquisitions.

2.2 Model to examine fraud or manipulation probability in financial institutions

There are little numbers of studies on manipulation probability in the banking field than manipulation probability in non-financial entities. In financial institutions, M-score is used to detect financial statement manipulations (Hassnain Raghib Talab et al., 2017; Samuel Nyakarimi, 2020; Van, 2021).

Samuel Nyakarimi (2020) used M-score and Probit regression models to assess financial statement manipulations of banks in Kenya. This study concluded that several banks were involved in manipulation of financial statements. Moreover, this study indicated that both internal and control auditors should calculate several indicators to identify whether manipulations were incurred in the process of financial statement preparation. Hassnain Raghib Talab et al. (2017) used M-Score model to uncover earning management practices of banks that have been listed in the Iraqi stock exchange. The sample of 23 listed commercial banks applied the M-score model and indicated all of them were existent financial statement manipulations. The M-score model is confirmed as the best model to examine financial statement manipulations.

Using the M-score model of Beneish (1999) and Van (2021) has added some indicators to examine the fraudulent financial statements of listed commercial banks in Vietnam. The accuracy of the model in predicting financial statement manipulation of listed commercial banks in Vietnam at 20% was 63, 16% of that was based on the results of independent auditors.

The appropriateness of the M-score in the banking industry is assessed and not included in added elements in the model. A comparision of the financial situation of banks after calculating M-score is needed to examine the manipulation probability.

3 THEORETICAL BACKGROUND

3.1 Financial statement fraud

Fraud is a serious problem in the world. It causes most costly and affects entities in every industry and place (ACFE, 2022). Previous research indicates three types of fraud, namely, asset misappropriation, corruption, and financial statement fraud. Asset misappropriation involves an employee stealing or misusing the organization's resources. Corruption, such as bribery and conflicts of interest, relates the power and control in organizations. Financial statement fraud is material misstatement or omission in the organization's financial statements. Banking service is the industry that was affected by a great amount of fraud cases (351 cases) in the year 2022 (ACFE, 2022). In which, according to the ACFE 2022 report, there are 2,110 cases from 133 countries, causing total losses of more than $3.6 billion. Financial statements of fraud are reported for 675 cases: (36%) in the United States and Canada, 429 cases (23%) in Sub-Saharan Africa, 194 cases (10%) in the Asia Pacific, 145 cases (8%) in Western Europe, 138 cases (7%) in the Middle East and North Africa, 138 cases (7%) in Southern Asia, 95 cases (5%) in Latin America and the Caribbean, and 78 cases (4%) in Eastern Europe and Western/Central Asia; 11% of 351 cases are financial statement fraud in the banking industry.

Financial statement manipulation is a sample of accounting fraud or fraudulent accounting and mentions the practice of using "creative accounting tricks" to make an entity's financial report look better than its financial statement and performance. There is truth in the saying that fraudulent financial statement is not actually committed by the entity that published the financial statements. Fraudulent financial statement is "committed" by top management such as the president, CEO, or CFO (Wm. Dennis Huber, 2021).

All stakeholders including shareholders, managers, employees, investors, suppliers, customers, and financial institutions are concerned about organizational financial health. The ability to forecast fraudulent financial statements is significant to take the necessary preventive methods in organizations. In addition, corporate governance and ethics have provided the basis to prevent fraud; however, a previous prediction is necessary for investors to safeguard their investments (Mahama, 2015). The prediction of business failure can enable companies to reduce manipulation costs, prevent fraud, and help improve their financial statements. The financial health of any firm can be measured by its financial indicators and the ability to predict financial statement manipulation is significant for society and investors (Patric Bimpong et al., 2020).

In businesses, financial goals pressure managers to achieve financial targets and force financial statement manipulation (Supri et al., 2018, Sunardi & Amin, 2018, Andrew et al., 2022). Previous research has indicated that financial statement manipulations are serious problems and are threats to a country and society. The stakeholders in the world admit that their fraudulent financial statements have a tendency to be realized too late. The lateness in identifying fraudulent financial statements has harmful impacts and adverse effects on the society, economic, and legal in a country or the world as a whole (Abdullahi & Mansor, 2018). Therefore, it can be implemented by detecting its financial statements to handle and predict the status before corporate organizations commit fraudulent financial statements. Examining financial statement manipulations can be realized by previous predicting models.

3.2 M-score model

Beneish (1999) used eight (8) financial indicators to make M-score model to predict the probability of a company manipulating its earning management. Eight financial indicators, namely Days sales in receivables; Gross margin; Asset quality; Sales growth; Depreciation; Sales general and administrative expenses; Leverage; and Total accruals to total assets.

The "Days Sales in Receivables" (DSRI) is calculated as flows:

$$DSRI = (\text{Net Receivables}_t/\text{Sales}_t)/(\text{Net Receivables}_{t-1}/\text{Sales}_{t-1})$$

The "Gross Margin" (GMI) is calculated as flows:

$$GMI = [(\text{Sales}_{t-1} - \text{COGS}_{t-1})/\text{Sales}_{t-1}]/[(\text{Sales}_t - \text{COGS}_t)/\text{Sales}_t]$$

The "Asset Quality" (AQI) is calculated as flows:

$$AQI = [1 - (\text{Current Assets}_t + \text{PP\&E}_t + \text{Securities}_t)/\text{Total Assets}_t]/$$
$$[1 - ((\text{Current Assets}_t\text{-1} + \text{PP\&Et}_{-1} + \text{Securities}_{t-1})/\text{Total Assets}_{t-1})]$$

The "Sales Growth" (SGI)) is calculated as flows:

$$SGI = \text{Sales}_t/\text{Sales}_{t-1}$$

The "Depreciation" (DEPI)) is calculated as flows:

$$DEPI = (\text{Depreciation}_{t-1}/(\text{PP\&E}_{t-1} + \text{Depreciation}_{t-1}))/$$
$$(\text{Depreciation}_t/(\text{PP\&E}_t + \text{Depreciation}_t))$$

The "Sales General and Administrative Expenses" (SGAI)) is calculated as flows:

$$SGAI = (\text{SG\&A Expense}_t/\text{Sales}_t)/(\text{SG\&A Expense}_{t-1}/\text{Sales}_{t-1})$$

The “Leverage” (LVGI)) is calculated as flows:

$$LVGI = [(\text{Current Liabilities}_t + \text{Total Long Term Debt}_t)/\text{Total Assetst}_{]}/ \\ [(\text{Current Liabilities}_{t-1} + \text{Total Long Term Debt}_{t-1})/\text{Total Assets}_{t-1}]$$

The “Total Accruals to Total Assets” (TATA)) is calculated as flows:

$$TATA = (\text{Income from Continuing Operations}_t - \\ \text{Cash Flows from Operations}_t)/\text{Total Assets}$$

The M-score is formulated:

$$\text{M-score} = -4.84 + 0.92 \times DSRI + 0.528 \times GMI + 0.404 \times AQI + 0.892 \times SGI + \\ 0.115 \times DEPI - 0.172 \times SGAI + 4.679 \times TATA - 0.327 \times LVGI$$

The threshold value of M-score model is −1.78. If M-score is less than −1.78, the company is unlikely to be a manipulator. For example, an M-score value of −2.0 suggests a low probability of manipulation. If M-score is greater than −1.78, the company is likely to be a manipulator. For example, an M-score value of −1.3 suggests a high probability of manipulation.

Eight indicators of the M-score model (Beneish, 1999) are divided into two groups: One identifies fraud and another reflects fraud motivation. The variables indicate fraud likelihood, namely DSRI, AQI, DEPI, and TATA. The remaining variables identify fraud motivation, including GMI, SGI, SGAI, and LVGI. The meaning of variables is detailed as follows:

- AQI ≥ 1.254: Capitalized costs in financial statements.
- DSRI ≥ 1.465: The recorded assets increased incorrectly compared to the reality, the value of the receivables was overstated in financial statements.
- DEPI ≥ 1.077: Overstates the use value of assets and income is not increased compared to reality.
- GMI ≥ 1.193: Economy is difficult
- LVGI ≥ 1.111: There is a sign of profit manipulation.
- SGAI ≤ 1.041: There is a sign of profit manipulation.
- SGI ≥ 1.607: There is a sign of error in revenue recognition.
- TATA ≥ 0.031: Revenue is incorrectly reflected compared to reality.

The M-score model is widely applied to identify fraudulent financial statements of companies and industries. However, the accuracy of fraud detection using M-Score is confirmed at 50% and not absolutely.

4 RESEARCH METHODOLOGY AND RESULTS

4.1 RESEARCH METHODOLOGY

Usage and calculation indicators of the M-score model of commercial banks sample are conducted in the chapter. The financial information of the commercial banks

TABLE 13.1
List of Vietnamese commercial banks

No	Bank name	Code	Capital (million USD)
1	Asia Commercial Joint Stock Bank	ACB	453
2	Tien Phong Bank	TPB	270
3	DongA Bank	DAB	241
4	South East Asia Bank	SeABank	220
5	An Binh Bank	ABB	214
6	Bac A Bank	BacABank	220
7	Viet Capital Bank	VietCapitalBank	120
8	Vietnam Maritime Joint Stock Commercial Bank	MSB	473
9	Vietnam Technological and Commercial Joint Stock Bank	TCB	1,407
10	Kien Long Commercial Joint Stock Bank	KienLongBank	120
11	Nam A Bank	NAB	121
12	National Citizen Bank	NCB	121
13	Vietnam Prosperity Bank	VPB	632
14	Ho Chi Minh City Housing Development Bank	HDB	395
15	Orient Commercial Joint Stock Bank	OCB	265
16	Military Commercial Joint Stock Bank	MB	730
17	Vietnam Public Joint Stock Commercial Bank	PVcombank	362
18	Vietnam International and Commercial Joint Stock Bank	VIB	227
19	Sai Gon Commercial Bank	SCB	575
20	Saigon Bank for Industry and Trade	SGB	123
21	Saigon – Hanoi Commercial Joint Stock Bank	SHB	484
22	Sai Gon Thuong Tin Commercial Joint Stock Bank	STB	758
23	Viet A Bank	VAB	141
24	Bao Viet Bank	BVB	141
25	Vietnam Thuong Tin Commercial Joint Stock Bank	VietBank	130
26	Petrolimex Group Commercial Joint Stock Bank	PG Bank	121
27	Vietnam Export Import Commercial Joint Stock Bank	EIB	497
28	Lien Viet Post Joint Stock Commercial Bank	LPB	260
29	Bank for Foreign Trade of Vietnam	VCB	1,448
30	Vietnam Joint Stock Commercial Bank for Industry and Trade	CTG	1,498
31	Joint Stock Commercial Bank for Investment and Development of Vietnam	BID	1,376
32	Construction Bank	CB	326
33	Ocean Bank	Oceanbank	232
34	Global Petro Sole Member Limited Commercial	GPBank	121
35	Vietnam Bank for Agriculture and Rural Development	Agribank	1,287

Collected by the authors.

TABLE 13.2
Audited and unaudited profits of commercial banks in the year 2020 (units: USD)

Bank	Code	Audited profit	Unaudited profit	Difference	% difference
A Chau Bank	ACB	386,194	386,194	0	0.00
Vietcombank	VCB	927,712	927,551	161	0.02
Vietinbank	CTG	687,623	687,060	563	0.08
Techcombank	TCB	635,943	635,943	–	0.00
VPBank	VPB	524,009	524,009	–	0.00
Agribank		531,374	523,244	8,130	1.55
MBBank	MBB	430,187	430,187	–	0.00
BIDV	BID	363,292	370,819	–7,527	–2.03
HDBank	HDB	234,172	234,172	–	0.00
VIB	VIB	233,568	233,447	121	0.05
Nam a Bank	NAB	32,186,144	32,186,144	–	0.00

Calculated and collected by the authors.

TABLE 13.3
M-score of banks

Bank	Period	DSRI	GMI	AQI	SGI	DEPI	SGAI	TATA	LVGI	M-score
A Chau Bank	2018–2019	0.728	1.009	1.004	1.150	0.760	0.013	0.004	0.991	–2.424
	2019–2020	0.676	0.934	1.003	1.130	0.509	0.003	–0.003	0.992	–2.591
Tienphong Bank	2018–2019	0.878	0.982	1.009	1.505	0.608	0.771	–0.030	0.998	–2.292
	2019–2020	0.904	0.925	1.003	1.224	1.273	1.038	0.085	0.998	–1.982
Maritime Bank	2018–2019	0.969	1.085	1.024	1.000	0.656	0.856	–0.010	1.006	–2.520
	2019–2020	0.756	0.740	1.016	1.523	1.056	0.941	0.050	0.999	–2.120
Sacombank	2018–2019	0.780	0.874	0.895	1.253	0.635	0.943	–0.022	1.002	–2.700
	2019–2020	0.761	0.992	0.921	1.180	0.452	0.996	0.001	1.000	–2.635
Nam A Bank	2018–2019	3.237	1.060	0.988	1.373	1.029	0.980	–0.017	1.004	–0.139
	2019–2020	1.868	1.070	0.985	1.295	0.965	0.810	0.001	1.003	–1.357
Kienlong Bank	2018–2019	0.748	1.189	1.020	1.155	0.719	0.971	–0.113	1.016	–3.029
	2019–2020	1.498	1.068	1.007	0.855	0.518	1.218	–0.042	1.006	–2.404
National Citizen Bank	2018–2019	1.072	0.943	1.002	1.002	1.002	0.848	–0.068	0.991	–2.729
	2019–2020	2.668	0.910	0.903	1.333	1.333	0.698	0.014	1.006	–0.583
Eximbank	2018–2019	0.986	1.133	1.007	1.008	1.166	0.924	–0.030	1.004	–2.522
	2019–2020	1.280	0.930	0.997	0.991	1.162	0.911	–0.018	0.988	–2.314
Saigon Hanoi Bank	2018–2019	1.405	0.868	0.980	1.393	1.095	0.881	–0.036	1.133	–2.013
	2019–2020	1.009	0.458	1.017	1.300	0.983	0.837	0.023	1.299	–2.449
Agribank	2018–2019	1.519	0.975	0.898	1.116	0.524	0.916	–0.033	0.998	–2.146
	2019–2020	1.083	1.084	0.956	0.979	1.124	1.084	–0.021	1.001	–2.493

Calculated by the authors.

that have the manipulation sign is collected after the analyzing period. We examine the relevant application of the M-score in the banking industry for Vietnamese commercial banks.

4.2 Research results

Vietnam has thirty-five (35) commercial banks, including four (04) limited liability banks with a single member and thirty-one (31) that are joint stock banks. Ten commercial banks in 35 banks are chosen and scores are calculated (see Table 13.1). This research was carried out by analyzing ten bank financial statements in the period from 2018 to 2020. Other banks had the same profit prior to and after auditing, whereas some had a difference in audited profit and previous auditing. These used banks are targeted at state commercial banks and joint-stock commercial banks.

Using the M-score model that is assessed and is relevant to the banking industry, we compute M-scores. Audited financial statements are used to calculate indicators of the model, and thereby show the M-score.

To assess material misstatements in the financial statements of commercial banks, audited results are used to test the difference. Table 13.2 shows audited and unaudited profits after examining audited and unaudited financial statements.

In terms of absolute value, after the audit, Agribank's profit after tax go up to USD 531,374, which is an increase of USD 8,130 compared to unaudited profit. The main reason is that the adjustment increased revenue and other profits (Table 13.2).

Using audited financial statements of ten (10) commercial banks, M-score indicators are calculated and compared with the M-score, the threshold value is −1.78. A financial statement with an M-score greater than 1.78 is probably fraudulent. This is the case for Nam A Bank and National Citizen Bank

In reality, National Citizen Bank has some problems with bad debt and loss in financial performance in the third quarter of the year 2022 (Dantri, 2022; HNX, 2022a). In the first quarter of 2021, Nam A Bank did not make provision for risks. Therefore, Nam A Bank reported more than three times the profit in the same period. Bad debt increased by 19% compared to the beginning of the year. Meanwhile, Nam A Bank's net cash flow from operating activities was negative to VND 2,337 billion (HNX, 2022b; VNFinance, 2021). The above information indicates the financial problem in two banks that has M-score greater than −1.78.

5 CONCLUSIONS

The analysis, calculation, and collection of data from banks show that not all banks have completely accurate figures, and some banks have used fraudulent techniques to manipulate financial statements for their own benefit. Applying the M-score model to evaluate, we find that the M-score model has a higher capability to predict and detect material errors in Vietnamese commercial banks' financial statements. The application M-score is similar the result in previous (Hassnain Raghib Talab, 2017; Van, 2021). As a result of its relevance to predicting financial statement manipulation in the banking industry, the M-score model can help investors and related parties evaluate the quality of financial information in the banking field.

REFERENCES

Abdullahi, R., & Mansor, N. (2018). Fraud prevention initiatives in the Nigerian public sector: understanding the relationship of fraud incidences and the elements of fraud triangle theory. *Journal of Financial Crime*, 1–18. https://doi.org/10.1108/JFC-02-2015-0008.

ACFE. (2022). Occupational Fraud 2022 - a report to the nations. [Press release].

Altman, E. I. (2013). *Predicting financial distress of companies: revisiting the Z-Score and ZETA® models*, in: A. R. Bell, C. Brooks, & M. Prokopczuk (ed.), *Handbook of Research Methods and Applications in Empirical Finance*, chapter 17, pp. 428–456. Edward Elgar Publishing.

Andrew, C., Robin. (2022). Detecting fraudulent of financial statements using fraud score model and financial distress. *International Journal of Economics, Business and Accounting Research (IJEBAR), 6*(1), 211–222.

Beneish, M. D. (1999). The detection of earnings manipulation. *Financial Analysts Journal, 55*(5), 24–36. https://doi.org/10.2469/faj.v55.n5.2296

Dantri. (2022). NCB has loss 180 billion VND after 9 months in the year 2022 and bad debt incresedover 14%.

Hariri, A. W. S. P., & Widjajanti, S. L. (2017). Predicting financial statements corporate fraud: Bênish M-Score model. *Jurnal Ilmiah Bidang Akuntansi dan Manajemen (JEMA), 14*(2), 93–100.

Hassnain Raghib Talab, H. H. F., Sallama Ibrahim Ali. (2017). Role of Beneish M-score Model in Detecting of Earnings Management Practices: Empirical Study in Listed Banks of Iraqi Stock Exchange. *International Journal of Applied Business and Economic Research, 15*(23), 287–302.

HNX. (2022a). https://www.hnx.vn/cophieu-etfs/chi-tiet-chung-khoan-nyNVB.html?_des_tab=3.

HNX. (2022b). https://owa.hnx.vn/ftp///cims/2022/4_W5/000000011509229_2022Q1BCTCRL.pdf.

Mahama. M (2015). Detecting corporate fraud and financial distress using the Altman and Beneish models. *International Journal of Economics, Commerce and Management, 3*(1), 1–18.

Natalia V. Feruleva, M. A. S. (2017). Detecting nancial statements fraud: the evidence from Russia. *Journal of Corporate Finance Research, 12*(2), 32–45.

Nguyen Anh Hien, P. T. T. (2015). Examining and identifying research models on earning management of Vietnamese listed companies *Journal of Science and Technology Development, 18*(Q3), 7–13.

Nguyen Cong Phuong, N. T. N. T. (2014). Predict material misstatements in the financial statements by the Beneish model. *Journal of Economic and Development, 206*(8), 54–60.

Nguyen Huu Anh, N. H. L. (2016). Using the M-score model in detecting earnings management: evidence from non-financial Vietnamese listed companies. *VNU Journal of Science: Economics and Business, 32*(2), 14–23.

Patric Bimpong, I. A., Khelanan, T. H., Danso, E., Opoku, P., Benedict, A., & Tettey, G. (2020). Assessing predictive power and earnings manipulations. Applied study on listed consumer goods and service companies in Ghana using 3 Z-Score models. *Expert Journal of Finance, 8*(1), 1–26.

Patricia M. Dechow, W. G., Chad R.Larson, Richard G.Sloan (2011). Predicting Material Accounting Misstatements. *Contemporary Accounting Research, 28*(1), 17-82. doi:https://doi.org/10.1111/j.1911-3846.2010.01041.x

Rasa Kanapickiene, Z. G. (2015). The model of fraud detection in financial statements by means of financial ratios. *Procedia - Social and Behavioral Sciences, 213*, 321–327. https://doi.org/10.1016/j.sbspro.2015.11.545.

Samuel Nyakarimi, S. K., & Kariuki, P. (2020). Financial statements manipulations using beneish model and probit regression model. a case of banking sector in Kenya. *European Online Journal of Natural and Social Sciences, 9*(1), 253–264.

Sunardi, S., & Amin, M. N. (2018). Fraud detection of financial statement by using fraud diamond perspective. *International Journal of Development and Sustainability, 7*(3), 878–891.

Supri, Z., Rura, Y., & Pontoh, G. T. (2018). Detection of fraudulent financial statements with fraud diamond. *Journal of Research in Business and Management, 6*(5), 2347–3002. www.questjournals.org.

Taffler, R. J. (1983). The assessment of company solvency and performance using a statistical model. *Accounting and Business Research, 13*(52), 295–308. https://doi.org/10.1080/00014788.1983.9729767.

Tam, N. T. (2014). Improving organizing controls to enhance the financial accounting information quality of listed companies in the Vietnam stock exchange (Hoàn thiện tổ chức kiểm soát nhằm tăng cường chất lượng thông tin kế toán tài chính của các công ty niêm yết trên thị trường chứng khoán Việt Nam). (PhD). National Economics University.

Tran, N. T. N. (2014). Research on material misstatment in financial statements of listed companies on Vietnamese stock market. (Master). Da Nang University.

Tran Thi Giang Tan, N. T. T., Tu, Đ. N., Hiep, H. T., & Nguyen, Đ. H. U. (2014). Fraud assessment in financial statement of Vietnamese listed companies. *Journal of Economic and Development, 26*(1), 74–94.

Van, H. T. H. (2021). The accuracy of the model in predicting fraud in financial statements of commercial banks in Viet Nam. *Banking Science & Training Review, 231*, 49–60.

VNFinance. (2021). https://vnfinance.vn/khong-trich-lap-du-phong-rui-ro-loi-nhuan-va-no-xau-cua-nam-a-bank-tang-manh-16308.html.

VNBusiness. (2022). https://vnbusiness.vn/giao-dich/phia-sau-nhung-con-so-nhay-mua-tren-bao-cao-tai-chinh-1087596.html.

Wm. Dennis Huber, J. A. D. (2021). Financial statement fraud and the failure of corporate financial statement fraud prediction. *Journal of Accounting and Finance, 21*(4), 30–41.

14 Investments & alternate investment options in India

Puja Roshani, Divya Bansal, Shivani Agarwal and Abhay Bhardwaj

1 INTRODUCTION

Investments are assets purchased or sums of money invested with the intent of generating revenue in the future. Investments are also made in order to profit from a future increase in an asset's value.

An investment is a future-focused purchase of products with the intention of generating income or building wealth in the future. A person might potentially try to profit by selling the asset later for a bigger sum. Investments can also include funds allocated for starting a new business, growing an existing one, buying stock or shares in a company, or investing an asset in a company. Making your money work for you or allowing your money to grow is the goal of investing.

A certain degree of uncertainty and risks always exists with investments. Risk is the possibility of receiving a return on your investment. The risk is minimal when investing in government securities. The risk is more when investing in stocks, starting new businesses, growing existing businesses, and similar activities.

Investments fall across the following general categories:

Bonds and debentures that pay a fixed rate of return, such as interest, are examples of investments with fixed income.

Equities and real estate are examples of variable income investments that don't offer a guaranteed annual return. Each fiscal year, the dividends or rental payments change. Additionally, their worth grows over time.

Points to be considered before making investments:

- It is undertaken to achieve short as well as long-term financial objectives.
- The motive behind investments is to raise funds for the future.
- A target return on investment is considered when making investments.
- Bonds, stocks, real estate, and other assets can all be purchased as investments.
- Businesses invest in things like labour, real estate, research and development, and equipment and machinery. In a similar vein, creating a factory, plant, or other asset would constitute an investment.

 DOI: 10.1201/9780367854690-14

- Investment in future skills and knowledge is yet another motive to spend funds on higher education.

Investments made in assets other than cash, stocks, and bonds are referred to as alternative investments. Alternative investments can be made in movable goods like fine wine or precious metals. They could also include financial assets like hedge funds, distressed securities, and private equity investments.

Alternative Investment Fund (AIF) can be defined as an investment mechanism that gathers money or creates a pool of investments from wealthy individuals. The investment can be domestic or foreign in nature. The investments are made for the benefit of the shareholders in accordance with a predetermined investment programme.

There are majorly three categories and subcategories within those categories, applicants may apply for registration as an AIF which are as follows:

1. Category 1
2. Category 2
3. Category 3

1.1 Why alternate investments

As a prudent investor, it is never a good idea to put all the investments in one place as advised by any financial planner or asset manager. The most probable investors to receive better value and safety for their investments are those that have varied portfolios that contain investments besides FDs, mutual funds, equities, gold, real estate, etc. Alternative investments refer to those assets or investments that are not typically available.

Worldwide venture capital, structured financing, leasing, and more lately, cryptocurrency investment and non-fungible tokens are included in the list of investments.

Some alternative investments have been popular in India for many years, but they are only available to the very wealthy. These are often offered at very high entrance prices to the top-tier clients of wealth management companies.

Alternative investments with fixed income act as a buffer against inflation. This is so because there is no direct relationship between them and the market. Other perks include interest payments, which can be a great post-retirement benefit, and the development of passive income. Alternative investments are a wonderful method to diversify a portfolio because they reduce risks, give income stability and predictability, and produce positive earnings.

1.2 Current scenario of alternate investments in India

In response to a shift in lifestyles and an increase in additional money that can be prudently invested, small-scale individual investors now have more options that can help them diversify their portfolios. Return on investment is an important area of concentration and the goal for any investor. However, the significance of safety, liquidity, and the presence of a trustworthy system to monitor the investment throughout its life cycle is now getting increasingly apparent to individual investors.

The assets of AIFs surged by over 32% year over year to reach Rs. 5.35 lakh crore at the end of September 2021, according to the Securities and Exchange Board of India's (SEBI) most recent data, reflecting the market's rapid growth in previous years. As of September 2020, there were Rs. 4.05 lakh crore in total assets owned by the AIF. Although the majority of these are linked to private equity and growing capital funds, it also demonstrates the increasing interest in such investment vehicles and the intention to make them widely accessible.

In addition to the elevated prices which only ultra- and high-net-worth investors (HNIs) could endure, the uncertainties, an absence of regulation, and a manque of liquidity had historically prohibited the general public from participating in these asset classes. The constraints are being surmounted, though, with the emergence of technologically sophisticated platforms that can expedite and make the approach more accessible.

Even while technologies are very much accountable for efficiency, the world of investing requires meticulous protocols, thorough due research, and vigilant post-investment monitoring. The creation of financial instruments and investor portfolios in India is a great opportunity for startups.

1.3 The future of AIF in India

Prospects in this area right now comprise lending money to corporations in the form of a corporate loan, invoice discounting, leasing, etc. The Indian market has a lot of potential, nevertheless, to gradually open to more exotic goods such as purchasing farmland, agricultural goods, artwork, footwear, and collectibles. Utilizing blockchain technology and smart contracts, which aid in ownership identification, creative items have also been made possible by technological advancement. Future improvements in this area's innovation and transparency may greatly benefit from this. People will have the chance to invest in and earn money from the goods and services they use daily in the future, in addition to being able to consume them.

Everyone's portfolio absolutely must include fixed-income alternative assets given the recent market turbulence. Regulatory authorities are assisting in safeguarding investors and strengthening financial markets by making the process more transparent. After the early 1990s, a considerable portion of the population started building wealth via having access to public stock markets. India is set to witness a similar opportunity, wherein involvement in alternative ventures would spark the country's economic succeeding wealth-creating revolution.

2 WHAT ARE ALTERNATIVE INVESTMENTS?

The classic investments that come to mind when most people think of investing are cash, stocks, and bonds. These conventional assets, like the index fund in your 401(k) or the cash in your savings account, are typical for most individual investors. But that only gives a partial view. Beyond traditional investments, there is another type of investing known as alternative investments.

The term "Alternative Investment Funds" (AIF) refers to the resources established or developed in India as a surreptitiously collective investment vehicle to raise money from participants in accordance with a predetermined investment policy.

The development of an alternative investment fund with a Rs. 25,000 core limits have also received approval from the Union Cabinet.

AIFs are made up of investment funds that have been combined and utilized to invest in private equity, hedge funds, and other types of investments.

2.1 Types of alternative investment funds/different categories of AIFs

AIFS are classified into three groups by the SEBI. The list is as follows:

2.1.1 Category 1

The funds in this category are invested in start-ups, small trades, and medium-sized enterprises that have the potential to grow financially. The government promotes investments in these businesses because of their high output and capacity for job generation, which benefit the economy. The investments made in new companies, non-profit organizations, SMEs, infrastructure and other industries or regions regulated by the government are believed to be socially or economically suitable. Venture capital funds, SME funds, social venture funds, infrastructure funds, and angel funds are all included in Category 1 of AIF.

Examples of this category are as follows:

- Infrastructure Funds
- Angel Funds
- Venture Capital Funds
- Social Venture Funds

2.1.2 Category 2

The investments made in both debt and equity instruments fall under this category. Included are any monies that do not already fall within Categories 1 and 3, respectively. These AIFs only borrow or employ leverage as necessary for their daily operations. These include debt funds, real estate funds, distressed asset funds, funds of funds, and private equity funds. The government does not grant any discounts for investments for Category 2 AIFS.

Examples of this category are as follows:

- Fund of Funds
- Debt Funds
- Private Equity Funds

2.1.3 Category 3

AIFs under these categories are investments that offer returns in a short-lived amount of time. To achieve their objectives, these funds employ a wide range of intricate trading tactics. These are the traders who use a variety of complex trading strategies and may use leverage, such as buying public or unregistered derivatives. There is no information regarding any special government incentives or concessions for these funds.

Examples of this category are as follows:

- Hedge Funds
- Private Investment in Public Equity Funds

2.2 AIFs becoming popular

AIFs are providing originality in terms of product differentiation since, in the opinion of experts, "Conventional investment pools are progressively getting fragmented and encountering iss ues in product innovations, uniqueness, and alpha production owing to size and other considerations."

Investors are nowadays confronted with a variety of traditional investing options, such as mutual fund schemes, unit-linked insurance plans (Ulips), equity finance, and services that manage portfolios. But the main purpose of these options is to trade equities. Investor demand for alternative investments is therefore still unmet.

As a result, AIFs are increasingly used to introduce themed and curated merchandise because it is challenging to convey innovative concepts on traditional platforms, according to the report.

According to Sunil Singhania, founder of Abakkus Asset Manager LLP, "As the number of wealthy individuals and affluent people increases, investors have grown increasingly demanding and need creative services, which have differentiated and a boutique style of managing risk and reward."

2.3 Who can invest in alternative funds?

AIF investments are open to everyone who is an Indian, including NRIs, PIOs, and OCIs. However, they must meet the prerequisites, which includes a minimum capital of Rs. 10 crores for angel funds and Rs. 20 crores for each program. Each investor must put up a minimum of Rs. 1 crore ($25 lakh).

If you are a risk-taking investor who desires to diversify financial risks, you can invest in AIFs that have been registered with SEBI and enjoy its benefits.

2.4 Reason to invest view in AIF

HNIs have the prospect of just using AIFs to diversify their investment portfolio because it offers a significantly higher return along with a quite high risk. AIFs enable investors to put their money in instruments apart from equities, bonds, collective investment schemes, and other mainstream investments. AIFs access higher-yielding assets and help in the diversification of the investors' existing holdings.

2.5 Investors' limit regarding AIF

All types of AIFs (apart from angel funds) are permitted to have up to 1,000 investors. An angel fund may contain up to 49 angel investors. Additionally, AIF can only raise money through private memorandum issuances and cannot publicly solicit investors to subscribe to its units.

According to the reports of outlookindia.com, AIFs are to grow by 25% by the year 2025.

India and Asia are prepared for the next significant expansion of alternative investment products. AIFs in India have the same amount of assets under management (AUM) as mutual funds did in 2009. The industry and the market are currently poised towards an alternative investment funds paradigm shift.

According to a forecast by Anand Rathi, the total amount of investments made through AIFs would increase at a 25% CAGR between 2022 and 2025, driven by wealth managers that provide products from AIF as substitutes to family offices, insurance firms, and HNIs.

A report claims that India and other growing economies in Asia are particularly well-positioned for the upcoming wave of alternative growth.

As of May 2022, more than 900 AIFs have already been registered with SEBI, with capital contributions accumulating at a 63% CAGR between 2012 and 2022. In the study, it was predicted that the worldwide alternative investment will increase from $4.1 trillion in 2010 to $10.7 trillion in 2020 and $17.2 trillion in 2025.

The research also claims that 1,625 investment transactions totaling $38 billion were made by Indian start-ups in the fiscal years 2021 and 2022. By the first week of May 2022, India had added 15 unicorns, bringing its total to 100. The research also states that investments in AIFs stand to benefit from the speedily increasing percentage of the populace that is becoming "wealthy."

The report also emphasizes that essential components exacerbating worldwide inflation expectations due to the Russia–Ukraine War, increased interest rates, and upcoming elections may prevent the country from achieving its $400 billion in planned capital commitments by 2030.

AIFs are categorized as commercially organized financial products that gather assets to invest in accordance with an established policy for the benefit of its investors. The Securities and exchange board (Alternative Investment Fund) Regulations, 2012 regulate AIFs.

3 CONCLUSION

Numerous traders are facing unprecedented problems nowadays because of the steadily declining bond rates. Many traders nowadays are venturing outside the conventional asset classes in quest of yield, which has boosted movements towards alternative investments. The rapid expansion of alternative investments is also evident from the rise in hedge funds. Numerous investigations on the effectiveness of commodity markets and hedge fund managers have already been produced in the old days, which concluded that they typically perform well. Investors who integrate resources and investment companies in their portfolios benefit through diversification in both instances.

The presence of hedge funds on the financial markets and their use of dynamic trading techniques, such as various timing methods, are anticipated to have a significant impact on the dynamics of the financial markets. The marketization of the commodities market and its repercussions are also projected to intensify as a result of the significant flows into those markets. Overall, alternative investments can offer realistic alternatives to investors seeking income, but further study is needed to show the long-term effects of directing capital to these types of investments in an environment with low interest rates.

4 LIMITATIONS

The chapter focuses only on alternate investments and types of alternate investments. The inclusions of various options in alternate investments would have added more weightage to the chapter. The prospects of alternate investments in India and the comparison of the same with other countries would have provided more insights into the chapter.

BIBLIOGRAPHY

Abdullahi, D., & Ahmed, A.D., 2016. Integration of financial markets, financial development and growth: Is Africa different? *Journal of International Financial Markets, Institutions and Money*, 42, 43–59.

Bao, X., Johan, S., & Kutsuna, K., 2016. Do political connections matter in accessing capital markets? Evidence from China. *Emerging Markets Review*, this issue.

Blazy, R., & Letaief, A., 2016. When secured and unsecured creditors recover the same: The emblematic case of the Tunisian corporate bankruptcies. *Emerging Markets Review*, 30, 19–41.

Cumming, D., & Zhang, Y., 2016. Alternative investments in emerging markets: A review and new trends. *Emerging Markets Review*, 29, 1–23.

Chordia, T., Subrahmanyam, A., & Tong, Q., 2014. Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? *Journal of Accounting and Economics*, 58(1), 41–58.

Cole, R., Cumming, D.J., & Li, D., 2016. Do banks or VCs spur small firm growth? *Journal of International Financial Markets, Institutions and Money*, 41, 60–72.

Dai, N., Jo, H., & Kassicieh, S.K., 2012. Cross-border venture capital investments in Asia: Selection and performance. *Journal of Business Venturing*, 27, 666–684.

Groh, A., & Liechtenstein, H.v., 2011. The first step of the capital flow from institutions to entrepreneurs: The criteria for sorting venture capital funds. *European Financial Management*, 17, 532–559.

Henderson, B., Pearson, N., & Wang, L., 2015. New evidence on the financialization of commodity markets. *The Review of Financial Studies*, 28, 1285–1311.

Iriyama, I., Li, Y., & Madhavan, R., 2010. Spiky globalization of venture capital investments: The influence of prior human networks, *Strategic Entrepreneurship Journal*, 4, 128–145.

Jurek, J.W., & Stafford, E., 2015. The cost of capital for alternative investments. *The Journal of Finance*, 70(5), 2185–2226.

https://aifpms.com/alternative-investment-funds-aif/.

https://byjus.com/free-ias-prep/alternative-investment-funds-aif/.

https://www.iaaif.com/wp-content/uploads/2017/04/Structuring-an-AIF.pdf.

https://www.indusind.com/iblogs/categories/investment/understanding-the-difference-between-alternative-investment-fund-aif-and-portfolio-management-service-pms-indusind-pioneer/.

https://www.outlookindia.com/business/all-you-need-to-know-about-sebi-s-alternative-investment-fund-news-198378.

https://www.tatacapital.com/blog/wealth-management/lets-understand-the-difference-between-pms-and-aif/.

Kräussl, R., Lehnert, T., & Rinne, K., 2017. The search for yield: implications to alternative investments. *Journal of Empirical Finance*, 44.

Li, C., Shi, Y., Wu, C., Wu, Z., & Zheng, L. Policies of promoting entrepreneurship and angel investment: Evidence from China. *Emerging Markets Review*, 29, 1–23.

Liang, B., 2005. Alternative investments: CTAs, hedge funds, and funds-of-funds. In *The World of Hedge Funds: Characteristics and Analysis* (pp. 109–127).

Nahata, R., Hazaruka, S., & Tandon, K., 2014. Success in global venture capital investing: Do institutional and cultural differences matter? *Journal of Financial and Quantitative Analysis*, 49(4), 1039–1070.

Otchere, I., & Pou Iu Vong, A., 2016. Venture capitalist participation and the performance of Chinese IPOs. *Emerging Markets Review*, this issue.

Peltomäki, J., 2008. Emerging market hedge funds and the yen carry trade. *Emerging Markets Review*, 9(3), 220–229.

Robb, A.M., & Robinson, D.T., 2014. The capital structure decisions of new firms. *Review of Financial Studies*, 27(1), 153–179.

Wright, M., Pruthi, S., & Lockett, A., 2005. International venture capital research: From cross-country comparisons to crossing borders. *International Journal of Management Reviews*, 7, 135–165.

Yang, T., & Hou, W., 2016. Pay-performance sensitivity and risk-taking: Evidence from closed-end funds. *Emerging Markets Review*, this issue.

Zhou, Z., & Zhou, J., 2010. Chinese IPO activity, pricing, and market cycles. *Review of Quantitative Finance and Accounting*, 34(4), 483–503.

15 Risk and return dynamics in portfolio theory

Vikas Gupta and Sripal Srivastava

1 INTRODUCTION

1.1 Portfolio risk & return

We have clarified many aspects of investment and trading behavior. However, bearing in mind the recent adage that he can't put all his eggs in one basket, I'd like to find a technique to diversify my investments to limit the opportunities. This is because each investment style and its theme have vastly different levels of risk and reward. Therefore, it is always desirable to have a wide range of investment strategies to mitigate these risks.

Let us help you build a portfolio that supports your risk appetite and investment goals. A portfolio is the overall combination of all assets managed by an individual, together with financial assets such as bank deposits, stocks, bonds, and non-financial assets such as land and residential real estate. Some of these assets are transferable and some are not inheritable. Because assets are transferable, individuals have the opportunity to work out their asset mix through the performance of their current assets and the commercialization of their current possessions. Portfolio management essentially means planning the careful mix of these assets in line with your personal goals.

1.2 Beta (market risk)

Each stock's contribution to the overall risk of a well-diversified portfolio depends on the equity diversification of the different stocks within the portfolio. Of course, this depends on how each stock reacts to fluctuations within the broader financial markets. Beta can be the volatility or systematic risk of a security or portfolio relative to the overall market. Beta is used within the Capital Quality Rating Model (CAPM). CAPM is a model that uses its beta and expected market return to calculate the expected comeback of quality. Beta is calculated through multivariate damage analysis and can reflect beta as securities returns tend to respond to fluctuations within the market. A beta of 1 indicates that the security's value can fluctuate with the market. A beta of just 1 means the protection is less volatile than the market. A beta greater than 1 indicates that the security's value is more volatile than the market.

 DOI: 10.1201/9780367854690-15

1.3 Phases of portfolio management

There are five phases which will be known during this process:

1. Security Analysis
2. Portfolio Analysis
3. Portfolio Choice
4. Portfolio Revision
5. Portfolio Analysis

1.3.1 Security analysis

Security analysis allows well-trained capitalists to derive a rough price for a company from all cash transactions, buy when the market inevitably falls below the price of many transactions, and recover satisfactorily. Represents the assumption that, and is not actually at risk of permanent loss. This was already explained in module 2.

1.3.2 Portfolio analysis

A portfolio is a group of assets managed together with an investment. For risk aversion, investors invest their money in portfolios of securities rather than single securities. By building a portfolio, an investor seeks to unlock risk rather than putting all her eggs in one basket. The analytical part of portfolio management consists of distinguishing potential portfolio diversities deeply rooted in particular security groups, complicating their resurgence and further analysis risks. What determines your portfolio?

1. Age of the capitalist
2. Risk-bearing capability
3. Perspective toward risk
4. Family responsibilities
5. Education background
6. Liquidity desires
7. Tax saving desires
8. Time span for investment

1.3.3 Portfolio choice

The ultimate goal of portfolio construction is to obtain the portfolio that produces the highest return for a given risk level. A portfolio with this characteristic is considered an economic portfolio. Inputs from portfolio analysis typically determine a set of economic portfolios.

From this set of economic portfolios, we need to select the best portfolio for investment. Harry Markowitz's Portfolio Theory provides anyone with an abstract framework and analytical tools for determining optimal portfolios in a disciplined and objective manner. Choose the option that best suits your purpose according to your priorities. When making an investment, remember that he has three criteria necessary to evaluate investment opportunities. This is often called LSR.

1. Liquidity
2. Safety
3. Return

1.3.4 Portfolio revision

Over time, the certainty that was attractive may not be so. New high-comeback, low-risk securities may emerge. Portfolio review is just as necessary as portfolio analysis and selection.

1.3.5 Portfolio analysis

This is a method of evaluating a portfolio's performance over a specified period of time in terms of risk and return. The actual realized yield, i.e., the risk borne by the portfolio, is measured quantitatively in relation to the amount invested. This method provides a feedback mechanism to improve overall portfolio management practices.

The portfolio management method is the progression method. It begins with security analysis, goes through portfolio construction, and progresses to portfolio revision and analysis. Analysis provides essential feedback for planning a higher portfolio next time.

1.4 Portfolio optimization

Optimal portfolios are selected based on risk-return swaps between different sectors. An economy's performance can be manipulated by various industrial sectors whose rates of return change over time. This changing correlation pattern b/w sector is essential for portfolio optimization purposes. A common goal of financial investors is to achieve the optimal risk/return combination. As a result, an analysis of all the risks and returns of all sectors in the form of portfolios was performed in order to evaluate the optimal portfolio selection.

Expected Risk and Returns – Mean, Standard Deviation, and Portfolio Choice:

The risk and return of the portfolios have been analyzed with the help of arithmetic mean or average mean, standard deviation, and the ratio between these two. The ratio was computed with the help of the formula:

$$\mathbf{R = (\mu - d)/\sigma}$$

Here

'μ'denotes the arithmetic mean of the daily data;
'σ' refers to standard deviation;
'd' refers to the disaster level, which is also called by the name lowest level of return.

These three results show the relationship b/w the expected return and the risk of the different portfolios.

2 BACKGROUND AND OVERVIEW OF CONTEMPORARY PORTFOLIO THEORY

2.1 Modern portfolio theory

It was developed by Harry Markowitz, who states that it is insufficient to indicate the expected risk and return of an explicit stock. By investing in multiple stocks, MPT quantifies the benefits of diversification, also known as non-Golf. Each stock has its own deviation from the average, which MPT calls risk. Markowitz starts by saying that all investors want to avoid as much risk as possible. He defines risk as the typical deviation from the expected return. The only risk in a portfolio of many individual stocks is the risk of owning a single stock among the individual stocks (assuming the risks of different stocks are not directly related).

But Markowitz assumes that all investments in a portfolio are likely to remain valuable under the same conditions. This is called correlation and it measures how different you think the values of different securities or positive categories are compared to each alternative. For example, high fuel prices make sense for oil companies, but it's unhealthy for airlines that have to buy fuel. As a result, the stock prices of companies in these industries might typically be expected to move in opposite directions. These industries have negative (or low) correlations. If you own airlines and companies instead of oil majors, you have good diversification in your portfolio. Putting all this together, it is perfectly possible to create a portfolio that contains a much higher average return than the risks involved. It can be managed and reverted in other words, investing is not about choosing stocks, but about choosing the right mix of stocks to allocate that precious money. It is when the alternative asset or portfolio of assets does not provide a higher expected return for the same (or less) risk or a lower risk for the same (or more) risk). Expect to come back. This theory states that the risk to an individual stock's return consists of two parts: structured risk and unstructured risk (also called specific risk). Unstructured risks are specific to individual stocks and can become widespread as you increase the number of stocks in your portfolio. This represents the portion of stock returns that is unrelated to general market movements. In a well-diversified portfolio, each security's risk (or mean deviation from the mean) contributes little to portfolio risk. Instead, the difference or covariance between the risk levels of each stock determines the overall risk of the portfolio. As a result, investors prefer to hold diversified portfolios rather than individual stocks.

3 EFFICIENT FRONTIER

How to determine the most effective level of diversification? The solution lies in entering the efficiency frontier. It represents the set of portfolios that give the highest return for a given level of risk, or the set of portfolios that give the lowest risk for any level of return. Each return level has a portfolio that offers the lowest possible risk, and each risk level has a portfolio that offers the highest return. These combinations are intentionally plotted on the chart, and the next line is the economic limit. It shows the economic frontier for two stocks: high-risk, high-reward technology stocks (Infosys) and high-risk, low-reward customer product stocks (ITC Ltd). This works for different plus categories. The best portfolio is the economic frontier portfolio that has the best utility for a particular capitalist. As a rational capitalist, it's easiest to keep your portfolio aligned with valid limits. The highest risk a capitalist can take determines the position of the portfolio on the street.

3.1 Capital asset pricing model (CAPM)

The CAPM model states that the expected rate of return an investor can demand is equal to the safe security's velocity and risk premium.

If the expected return is insufficient or higher than the expected return, the investor may refuse to take the position and no investment will be made. CAPM categorizes portfolio risks into systematic and specific risks. Systematic risk is the risk of holding a market portfolio. All the pluses are affected to a greater or lesser extent as soon as the market moves. To the extent that pluses participate in such general market movements, they carry systemic risks. Specific Risks are risks that are unique to Private Plus. According to CAPM, the market compensates investors for taking scientific risks, but not for taking chosen risks. This can be heterogeneous for selected risks. When an associate degree capitalist holds a market portfolio, every quality of that portfolio carries certain risks. However, through diversification, an investor's exposure to the internet is only a systemic risk to the market portfolio.

3.2 CAPM formula

Expected security come = risk-free come + Beta X (expected market risk premium)

or:

$$R = RF + \text{Beta } X(RM - RF)$$

Where:

R: is that the expected come rate on a security.
RF: is that the rate of a 'risk-free' investment, i.e. cash.
RM: is that the market rate of the acceptable quality category.

Beta is the overall risk of investing in large markets such as bovine spongiform encephalitis and National Stock Exchange.

Each company includes a beta version collectively. A company's beta is the company's risk relative to the general market's beta (risk). If the company holds his 3.0 beta, it's said to be three times riskier than the broader market. Beta indicates the security's volatility relative to its quality category.

3.3 Asset allocation

Asset allocation is an important tool in constructing a diversified portfolio. In other words, it is often important to find the proportion of a portfolio that is blessed with a number of quality classes such as stocks, bonds, or money. For example, if stocks traditionally rise 10% annually and bonds rise 5% annually, a 50% stocks and 50%

bonds combination would be expected to rise 7.5% annually. Associate degree quality categories include securities with similar characteristics, attributes, and risk-return relationships. Broader quality categories such as bonds fall into smaller quality categories such as government bonds, corporate bonds, and high yield bonds. Reputation for quality assignments is not an isolated selection of relevant grades. Rather, it is part of a portfolio management technique. Many of the associate's quality attribution strategies rely on the investor's policy statement, which includes investor objectives, limitations, and investment tips. However, your personal budget should be related to your age, financial situation, future plans, risk-averse traits, and desires.

Before starting an investment program for an associate's degree, you want to make sure that your various needs are met. Serious investment deals should not be made until would-be capitalists have earned enough financial returns to cover their living expenses and have secure internet access in case the unexpected happens. When it comes to quality allocation strategies, remember that the younger you are, the more risk you can tolerate. As you get older and nearer retirement, you may be less interested in expanding your portfolio and more interested in wealth and asset preservation to prevent the value of your portfolio from declining. Maintaining a portfolio when the required retirement age is reached requires additional investment because excessive declines in the value of holdings may affect how you retire or prevent you from retiring as planned and becomes a necessity. Asset categories have completely different market dynamics and different interaction effects. Therefore, the allocation of cash to quality categories can have a significant impact on fund performance. As mentioned earlier, a portfolio consisting of completely different stocks can also be part of an associate's investment. Then, in this case, diversify the class or sector backed by equities.

3.4 Proposed work: constructing smart portfolio

3.4.1 Portfolio construction

Many investors do not have a basic understanding of what an honest portfolio is. In fact, there is no correct result when it comes to portfolio modifications. Many times I bought and sold securities without observing the portfolio. The biggest result of such a ruthless approach is that few companies can deliver extraordinary returns, but portfolios are forgotten. If you're looking to build a portfolio, the checklist below will help you make the right calls.

3.4.2 Primary and secondary

Ideally, any portfolio should be divided into two main components,a primary portfolio and a secondary portfolio. The former should have a long-term view, and the latter should include a short to medium-term script to play the market, perhaps throughout momentum.

Don't rush to sell your first portfolio. The majority of your investment should be in your first portfolio.

3.4.3 Skewed sector bias

In terms of weights, you should not be overly exposed to any sector. The maximum weight of the selected sector should not exceed 20%. This strategy provides a volatility grip on your portfolio.

3.4.4 Keep future leaders

Unfortunately, no one shows up for the long term because everyone has to ride the current leader. Increase your bottom line by benefiting from security delivered at a lower cost.

3.4.5 Stocks in momentum

Many momentum stocks can even see their market cap increase many times over as the trend progresses. However, the most effective way to deal with this problem is to defer 100% of the initial investment once Momentum stock has risen at least 50% from its price. To modify this for the Associate's example, let's say he bought his 1,000 shares in First Principle Ltd for 27 rupees. When shopping security comes to Rs. 41. This is often the case when you want to sell enough shares to recover the value of Rs.27,000. In other words, sell 650 shares for 41 rupees to hide your value. Your credit amount will allow you to race at the top of your associate's degree. If the market goes down, don't be too upset because it has already recovered its value.

In fact, momentum stocks should never be part of a major portfolio.

3.4.6 Dividend concerns

In a volatile sell, it's best to own a company that pays dividends. Note that dividend payments actually show that money is flowing to shareholders. Companies that systematically pay dividends to investors consistently outperform companies that do not pay dividends within the same business. Such companies bring additional stability to your portfolio.

3.4.7 Take profit in time

At the end of the day, what matters is the money you make from your portfolio. People are often not smart enough to book a profit. Also, try out the core portfolio. Remember, you have to maintain an honest balance between your arena and your company's portfolio. If one certificate suddenly jumps up, it is a weight adjustment. Here, we account for profits of such magnitude that the weighting is very high. For example, if you buy a certificate with a weight of 10%, some favorable terms increase the weight of the certificate to 20% now. In this scenario, we need to take more of the profit so that the weight goes back to 10%.

3.4.8 No churning

Many investors are in the habit of buying and selling stocks without good reason. Portfolio churning is just a cost associated with the portfolio and usually results in a commercially viable script at lower levels.

Don't change your portfolio just for that. Examine changes in company fundamentals. If there is no change, do not change the portfolio. If you're thinking about it, hold scrip and wait until it's wise to sell them.

3.4.9 Right breadth of portfolio

Investors often have multiple companies in their portfolios. This doesn't make it harder to watch, but it also results in below-average returns. At the same time, investors typically have only a handful of companies in their portfolios, making them riskier. In general, the maximum weight of the certificate should not exceed 15%.

3.5 Return concerns of portfolio

3.5.1 Portfolio returns

You need to know how much you expect and whether it is consistent with your risk tolerance. When evaluating expected arrivals, consider percentage arrivals rather than absolute arrivals. For example, if you speculate Rs.20,000, you have two options: buy TISCO at Rs.320 per share on the secondary market or speculate on one of the initial offerings offered to the market at Rs.10 per share from the primary market. There are two options. They expect the first offering to be listed two months later when each share is listed at an expected value of Rs. 15. At this point, TISCO's share price is expected to rise to INR 370. In terms of absolute value per share, TISCO is sure to offer a much better profit of Rs.50 (Rs.370–320) compared to Rs.5 (Rs.15–10) for the first issuing unit. To do this, however, here is the source of the misunderstanding.

In fact, the first issue will give you a much better return in terms of total shelf life as you will get 150% of the first issue versus TISCO's 15.63%. Due to the limited range of funds, the percentage values are important here as we focus on equities. For example, if you invest Rs. 20,000 in TISCO shares, you will get 62 shares in eight months offering 370.62 Rs. 22,940. Instead, if you invest Rs. 20,000 in the initial issue at Rs. 10 each, you will buy a pair of 1,000 oversubscribed units at Rs. 15, 1,000 Rs. 30,000 after eight months. Therefore, TISCO's earnings will be Rs. 22,940 Rs. 20,000 Rs. 2,940 from the investment, but if invested within the first issue unit, it will yield Rs. 30,000 20,000 Rs. 10,000. Note that even small price movements when prices are low consistently provide better stocks and vice versa. Live returns are also found in the form of guaranteed financial benefits. The youngest graduate widow living off the financial gains from her husband's investments will be interested in a guaranteed return on the investment. Therefore, the most effective investments for them are stocks with financial returns, blue chip stocks that often offer high dividends, and high-yielding bonds. But if you're a brave young man with a college degree and can't wait to make more money, your main target should be growth, not quality stock and decent dividend payouts will become stock.

4 CONCLUSION

Portfolio management is still a science and does not provide clear answers on portfolio construction. Building a portfolio that yields excess return on investment seems more of an art than a science. The results show that investing in the same sector cannot fully diversify risk, but there is evidence that it is possible to look for increased risk-return trade-offs by creating different portfolios. Investing in portfolios that combine manufacturers and suppliers provided the lowest asset-to-risk ratio.

The global economic recession caused by the financial crisis had little impact on the global automotive industry. Further research may reveal whether the decline in yields during and after the crisis was also driven by regional factors or was only affected by the global economic crisis. Financial markets are showing the first signs of an ailing economy. Financial globalization, reflected in highly correlated stock markets, has reduced the range of international portfolio diversification. Same industry returns may not be a profitable target for investors who tend to find negatively correlated securities. Systematic shocks that can occur affect most assets in a portfolio in the same way. The findings show that the Indian stock market does not follow efficient market theory, as an efficient market includes all kinds of information such as historical market data and public–private information. However, the Indian market is responding quickly to domestic and global cues. Expected returns remain inadequate compared to the systematic increase in risk. Therefore, portfolio selection should be made in such a way that the correlation value of beta values beats the market. This shows the financial market volatility in the Indian economy. This requires compliance with financial regulatory mechanisms.

BIBLIOGRAPHY

Alexander, C., &Dimitriu, A. (2005). Indexing and statistical arbitrage. *The Journal of Portfolio Management*, 31(2), 50–63.

Behr, P., Guettler, A., &Miebs, F. (2013). On portfolio optimization: Imposing the right constraints. *Journal of Banking & Finance*, 37(4), 1232–1242.

Brands, S., & Gallagher, D. R. (2005). Portfolio selection, diversification and fund-of-funds: A note. *Accounting & Finance*, 45(2), 185–197.

Brealey, R., Myers, S. C., & Marcus, A. J. (2007). *Fundamentals of Corporate Finance*. Mc Graw Hill, New York.

Butler, K. C., & Joaquin, D. C. (2002). Are the gains from international portfolio diversification exaggerated? The influence of downside risk in bear markets. *Journal of International Money and Finance*, 21(7), 981–1011.

Chen, M. H. (2003). Risk and return: CAPM and CCAPM. *The Quarterly Review of Economics and Finance*, 43(2), 369–393.

DeMiguel, V., Martin-Utrera, A., & Nogales, F. J. (2013). Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking & Finance*, 37(8), 3018–3034.

Drake, P.P.,& Frank, J. F.(2010). *The Basics of Finance*. Hoboken, NJ: John Wiley & Sons, Inc.

Dwyer, G.,& Hafer, R. (1988). Are national stock markets linked? *Federal Reserve Bank of St. Louis Review*, 70, 3–14.

Fernandez, P. (2015). CAPM: An absurd model. *Business Valuation Review*, 34(1), 4–23.

Jeon, B. N., & Vonfurstenberg, G. M. (1990). Growing international co-movement in stock-price indexes. *Quarterly Review of Economics and Business*, 30(3), 15–30.

Khan, T. A. (2011). Co-integration of international stock markets: An investigation of diversification opportunities. *Undergraduate Economic Review*, 8(1), 7.

Kozelský, T., & Novák, R. (2015). *Automotive industry: Future trends*. EU Office ČS.

Louton, D. A., & Racine, M. D. (2007). Diversification in portfolios of individual stocks: 100 stocks are not enough. *Financial Review*, 42(4), 557–570.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.

Maxton, G. P., & Wormald, J. (2004). *Time for a Model Change: Re-engineering the Global Automotive Industry*. Cambridge University Press.

Medo, M., Yeung, C. H., & Zhang, Y. C. (2009). How to quantify the influence of correlations on investment diversification. *International Review of Financial Analysis*, 18(1), 34–39.

Mohamad, S., Hassan, T., & Muhamad S. Z. (2006). Diversification across economic sectors and implication on portfolio investments in Malaysia. *International Journal of Economics and Management*, 1(1), 155–172.

Nieuwenhuis, P., & Wells, P. (eds.) (2015). *The Global Automotive Industry*. Wiley.

Observations of Economic Complexity (OEC). (2016). Retrieved from http://atlas.media.mit.edu/en/ profile/country/cze/.

Olibe, K. O., Michello, F. A., & Thorne, J. (2008). Systematic risk and international diversification: An empirical perspective. *International Review of Financial Analysis*, 17(4), 681–698.

Ray, C. (2010). *Extreme Risk Management:Revolutionary Approaches to Evaluating and Measuring Risk*. McGraw-Hill Education.

Sentana, E. (2004). Factor representing portfolios in large asset markets. *Journal of Econometrics*, 119(2), 257–289.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), 425–442.

Shukla, R. (2004). The value of active portfolio management. *Journal of Economics and Business*, 56(4), 331–346.

Soros, G. (2003). *The Alchemy of Finance*. John Wiley & Sons.

Tang, G. Y. N. (2004). How efficient is naive portfolio diversification? An educational note. *Omega*, 32(2), 155–160.

Tofallis, C. (2008). Investment volatility: A critique of standard beta estimation and a simple way forward. *European Journal of Operational Research*, 187(3), 1358–1367.

Tola, V., Fabrizio, L., Mauro, G.,& Rosario, N.M. (2008). Cluster analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, 32(1), 235–258.

Van Biesebroeck, J., & Sturgeon, T. J. (2010). Effects of the 2008–09 crisis on the automotive industry in developing countries: A global value chain perspective. In *Global Value Chains in a PostCrisis World*. Policy Research Working Paper 5330. The World Bank, 209–244.

16 Use of machine learning for software project cost estimation based on ISO/IEC standards

Beata Czarnacka-Chrobot

1 INTRODUCTION

The scarcity of benchmarking (historical) data is considered to be one of the fundamental, if not essential, problems of proper cost estimation of software projects. Incorrect cost estimation of such projects means that many of them are not completed or are completed at significantly greater costs than planned, which in turn not only reduces the success rate of software projects, but above all causes the waste of vast amounts of money invested in IT projects. Large losses in this respect are borne not only by companies, but most of all by state budgets, and therefore by us – citizens. Wasting resources is an act that is contrary to sustainability as a horizontal principle, because it inhibits economic growth and investment in other factors of sustainable development.

The main purpose of this chapter is to demonstrate that machine learning (ML)-based software project cost estimation models can help to minimize this problem, provided that the software Functional Size Measurement (FSM) methods approved by ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) are used as their basis. On the one hand, in order to make sense, the use of ML requires substantial amounts of historical data of appropriate quality, which somehow forces their collection and cleaning, and on the other – on this basis, they give estimation results more accurate than other approaches to cost estimation if they based on ISO/IEC standards.

That is why the following issues will be presented in the chapter: (1) problems of estimating the costs of software projects and their practical effects; (2) software project cost estimation based on functional size measurement methods approved by ISO/IEC; (3) the importance of generalized benchmarking data in cost estimation of software projects; (4) analysis of the use of machine learning algorithms for software project cost estimation along with its results and discussion; and (5) conclusions and future work.

 DOI: 10.1201/9780367854690-16

2 PROBLEMS WITH PROPER COST ESTIMATION OF SOFTWARE PROJECTS – THEORETICAL FRAMEWORK AND EFFECTS IN PRACTICE

The latest research conducted by the Standish Group (Standish Group, 2022), a US research institution which has specialized, for many years now, in analyzing the effectiveness of software system project execution (currently based on more than 50 thousand projects concerning such systems), shows that only 31% of them are successful, i.e., end as planned in terms of costs, time, and scope, whereas 69% of them are either a complete failure, i.e., were abandoned and so the product was not delivered (19%), or a partial failure (50%), i.e., ultimately, the product was delivered, but the costs were higher and/or the delivery time longer than planned and/or it did not satisfy the target requirements (Johnson, 2021). This means that the projects ended with a partial failure constitute over 60% of the undertakings completed with product delivery (31% + 50% = 81% of the projects closed with product delivery). The estimated costs are exceeded by approx. 65% of the projects ended with a partial failure, and the average exceedance is around 60%, which often translates into multi-million losses in a single project. The planned duration of a project is exceeded also by approx. 65% of the projects ended with a partial failure and its average exceedance is approx. 75%, which also leads to major losses, in particular in the case of projects with strict deadlines (e.g., a settlement system that needs to be put into operation specifically from the beginning of a year or a system, which supports a sports event scheduled for specific days, etc.). Around 50% of the projects ended with a partial failure are ones that have not met the set requirements, with the average delivery of the required functions and features being approx. 70% (Standish Group, 2018).

Estimation is an anticipatory quantitative assessment of unknown attributes of an undertaking, here: software system development project, conducted based on the general knowledge of the requirements for a given product and its development considerations, and so made while having incomplete information. Hence, it is a probabilistic value obtained with a certain confidence level, relying on specific assumptions, which are not seldom modified at a later stage of the project cycle. And yet, the contracting entity requires that the contracted software system be priced, which means estimating the system cost either before the development cycle starts or at the very beginning of the project life cycle in order to be able to make a reasonable investment decision. Moreover, the estimation should be reliable and objective, justifiable and documentable. This is not impracticable: if suitable software project cost estimation methods and tools are employed correctly, it is possible to arrive at estimates with up to ±10% of accuracy already at this stage.

If a project to develop a new or enhance an existing software system covers all components (layers), then – in general terms – its costs include: IT infrastructure costs, commercial software costs, and effort (personnel) costs, that is the remuneration of the people involved in the development/enhancement of a software system (Flasinski, 2007). As a rule, effort costs are the decisive components of the costs of such an undertaking, since they involve changing a software system, but at the same time they are the most difficult to estimate correctly and can be easily manipulated. While infrastructure and commercial software costs can be determined based on market prices, the proper effort cost estimation involves a proper estimation of the effort required to develop/enhance a software system beforehand. The basis for the proper effort cost estimation, in turn, is the correct estimation of the size of the software system under development/enhancement (the required product of a project), but – to this end – it is also crucial to take into account the suitable productivity (see Figure 16.1) or the suitable model conditioning effort on software system size. Such an effort cost estimation is the appropriate basis for software system pricing.

However, the following principal problems hampering the proper estimation of effort costs and hence the proper pricing of such systems are frequent in the practice of software project:

1. Application of improper cost estimation methods, i.e., ones disregarding the size of the required software system while, generally, relying on the available resources rather than product requirements, which is reverse logic of the actual dependence. As a result, defective contracts, i.e., contracts based on fixed price as well as time and material contracts, are in use (Parthasarathy, 2007; State Government of Victoria, 2000). In the former case, the amount due for the product of a project is established based on the assumed fixed expenditure determining the fixed price. In the latter type of contracts, the calculation of the amount due for the product is based on the rate set for the contractors' person-hours. This means that the unit effort cost is measured in relation to a working time unit rather than a product size unit and therefore it is the time rather than the required/delivered product size that determines the total effort costs. Nonetheless, given that it is popular to implement fewer functions and features in software systems than specified in the user/client requirements and that software projects are relatively rarely completed within the expected timeframe, the most commonly employed approaches encourage exceedance of the planned expenditure on

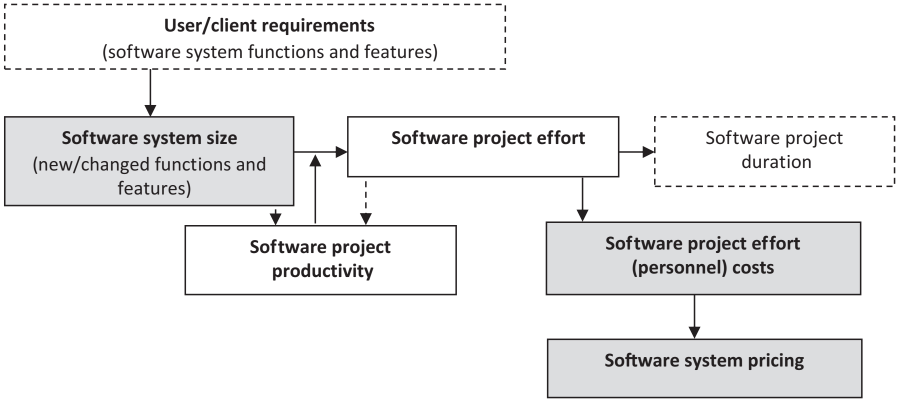


FIGURE 16.1 Model of dependence between the key parameters/attributes of a software project and software system size.

Source: Author's own elaboration.

the delivery of a product with the assumed functions and features. Where the contracting entity and the contractor conclude a contract based on an hourly rate, the contractor might prolong the duration of the project. And there is no guarantee that, as an effect, even if the contractor extends its duration excessively, thus leading to uncontrolled rise in costs, the delivered product will have the required functions and features. As regards fixed price contracts, in turn, apart from a likely situation where the actually developed product might be delivered with a smaller range of functions and features than required, yet another problem emerges: contractors are highly reluctant to any requirement changes, which are in fact typical of software projects. This is why these types of contracts could prevent cost exceedance on the one hand, but they do not guarantee that the product delivered at the agreed costs will have the required functions and features and, additionally, make it difficult to introduce changes to the original user/client requirements on the other hand. Moreover, the contracting entity could pay for the functions and features that have not been actually delivered.

2. Even if methods based on product size are employed when estimating effort costs, the common difficulty involves its proper *ex ante* determination. This is because in order to estimate it, an appropriate measure of software product size and a method based on it are needed. Only then it is possible to apply effort and, hence, cost estimation models based on the correctly estimated system size. Nevertheless, the problem is that no unequivocal software size measure to be commonly recognized by IT branch has been devised yet, and hence also the *ex ante* pricing of a software system is often random, intuitive, and manipulative rather than based on reasonable, measurable, and objective foundations, and the price calculation is difficult to justify and document. In practice, such a pricing is often prepared using an expert assessment, and such an assessment depends to a considerable extent on the competences, skills, and informal connections between the experts engaged in such an assessment.
3. Difficulties with adapting an effort estimation method to specific project conditions. M. Jørgensen (2007) suggests that three effort adaptation (calibration) levels should be distinguished:
   a) low adaptation level, where organizational historical (benchmarking) data are not analyzed and effort is determined using either productivity – defined as the ratio between the size of the required system and the effort involved in its development/enhancement, and hence it is an inverse of the unit effort measured with respect to the product size unit – or estimation models built on generalized benchmarking data;
   b) medium adaptation level, where productivity values specific to a software development organization are used to estimate effort;
   c) high adaptation level, where effort estimation models (still usually regression) based on organization-specific historical data are employed.

Therefore, it would be ideal if every software development organization built effort estimation models adapted to the specific conditions of its operation, which in fact entails the necessity to derive its own organizational dependence, and it is not that simple to do, as it requires a sufficient quantity of its own benchmarking (historical) data on the projects executed in the past, which would, additionally, reflect the dependence of effort (a response variable) on the software product size (the necessary explanatory variable) measured in appropriate units. This is why such a situation is rarely encountered in reality. Therefore, a certain intermediate form of adapting an effort estimation method to specific project conditions, which is significantly simpler in practice, is the inclusion of the organization-specific productivity (the medium adaptation level), which does not require as much suitable historical data as when constructing models. Due to the fact that this attribute is strongly dependent on the specific nature of the project team's work and on various characteristics of the project, including in particular the software system type, its complexity level, and field of its application, it should be derived relying on the software development organization's own experiences gained from similar past projects. The estimated effort can be then determined as the quotient of the estimated system size (in appropriate units) and productivity, calculated *ex post* based on the past. This adaptation method is not popular either, since collecting the sufficient quantity of own relevant benchmarking data by software system development organizations still is not common. Hence, the reality is usually a low adaptation level of the effort estimation models, that is the employment of generalized historical data that are not organization-specific. What counts is that they should be of high quality and, first of all, illustrate the dependence of effort/productivity on the software system size measured in appropriate units. This is when two out of the three aforementioned difficulties are removed, thus making the accurate estimation of effort costs highly probable.

Therefore, the base of both *ex ante* and *ex post* pricing of a software system should be its size: the required (estimated) one for *ex ante* pricing and the actually delivered (measured) one for *ex post* pricing. As a consequence, firstly, the unit effort costs should be taken into account; they obviously condition the total effort costs, which should not be assumed in advance (as the case is with fixed price contracts) without referring to the estimated and then measured product (software) size. Secondly, the unit effort costs should be referred to the product size unit rather than the working time unit (as the case is with time and material contracts). Such an approach to pricing requires a proper software size measure (unit). But it is only then that the software system pricing is objective and reliable. This is because the contracting entity can plan the project costs depending on the required functions and features of the system (its size) and, as a result, will pay for the actually delivered product size rather than for its requirements that have not been satisfied by the contractor or for the contractor's additional working time. As demonstrated by the research conducted by the Standish Group, the products delivered currently as part of over 60% of the software projects ended with product delivery lack over 30% of requirements on average, while their planned duration is extended by approx. 75% on average. This approach should facilitate the reduction of both the number of unsuccessful projects and the average cost exceedance, which is approx. 60% of the estimated costs for the said over 60% of the projects with product delivery.

Such an approach, although still applied too seldom, works well in practice. The study ordered by the State Government of Victoria in Australia (Sage Technology,

2007) revealed that the software system pricing based on a software size unit (in this case the so-called functional size) results in a reduction of the average exceedance of the estimated costs to less than 10% only. The reasonability of this approach was confirmed also by the analyses conducted by the International Software Benchmarking Standards Group (2012). They permitted the finding that the projects where a software system was priced using a unit of the (functional) size of a product were characterized by accurate estimates: for 90% of the cases, the cost estimates showed a deviation of up to 20% of their actual value, with the exceedance being up to 10% of the actual costs for 70% of the projects.

Moreover, the approach to the pricing of a software system based on its size expressed in appropriate units motivates the prospective contractor to propose a reasonable effort cost of the product size unit, as the risk of its potential underestimation and the related consequences are transferred to the contractor. Hence, it will not have the motivation to underestimate the offered price of the system to be developed. This requires the awareness of its own productivity or the relevant dependence models, ability to manage it properly, including to take measures for its enhancement and hence for the improvement of the software development processes.

In summary, it needs to be stated that the following three factors are of principal significance for the accuracy of effort cost estimation in software projects, that is, the pricing of their products:

- application of an estimation method based on software system size,
- expression of software system size in appropriate units, and
- availability of relevant historical (benchmarking) data and the proper way of using them (e.g., data mining techniques, including ML).

3 SOFTWARE PROJECT COST ESTIMATION BASED ON THE FUNCTIONAL SIZE MEASUREMENT METHODS APPROVED BY ISO/IEC – LITERATURE AND STANDARDS REVIEW

3.1 Software system functional size measurement

The methods serving the purpose of estimating software system size in appropriate units should be: (1) reliable, i.e., ensure that the actual assessments will not deviate significantly from the estimated assessments made at early stages of a project; (2) objective, i.e., produce identical results regardless of factors such as system development technology or the person making the estimations; (3) based on the system attribute that is of crucial significance to the software users and at the same time understandable to them, that is on system functionality: here, the most important is "what" the user can do with the software rather than "how" it is developed and how many lines of code it will take, which is irrelevant to the user. Additionally, the proper methods should: (4) make it possible to capture the modifications made in the user/client requirements, which are a natural phenomenon during software project; (5) be relatively easy to apply; (6) be universal, i.e., be applicable to any type of software; (7) serve as the basis for establishing the effort of all stages constituting the project life cycle, and (8) be compliant with the economic definition of productivity, which means that the use of a more efficient programming language should result in reduction of not only total project effort and costs but also its unit cost, which means an increase in project productivity. Moreover, (9) the desired method is one enabling a comparison of the actually delivered product size with the size required by the contracting entity, which permits the determination of the degree to which the contractor has fulfilled its obligations as well as the accurate *ex ante* and *ex post* pricing of the product.

"The most critical step in estimating anything is to understand what it is" (Peter Korda; Galorath, 2004). In general terms, software functionality is defined as the software product's capability to provide functions that are consistent with the needs of its users/clients. "If software fails to provide a set of required functions, the whole discussion about other software attributes is pointless" (Kobylinski, 2005). According to the author, if a software product fails to provide the required functions, its purchase makes no sense. Hence, it needs to be functional in the first place even if it is difficult to operate, fails a lot, works inefficiently, and its maintenance requires much effort.

The fundamental software functionality unit, which served as the basis for developing all of the similar units, is function point (FP). The author of this concept, and at the same time the concept of software functional size measurement, was Allan Albrecht, an IBM employee, who defined function points as "a dimensionless number (…) which we have found to be an effective relative measure of function value delivered to our customer" (Albrecht, 1979). He argued that software product size should be measured in units that are important to software users rather than to its developers, the most important one being its functionality (e.g., program length is completely irrelevant to users). Additionally, he assumed that the units should be independent of the development technology, including the programming language used. As a result, Albrecht devised a method known as the Function Point Analysis (FPA), based on estimation and measurement of the functionality required by and delivered to the users/clients. Hence, the group of methods resulting from the development of that approach is called software Functional Size Measurement Methods (FSMM).

Function points are then the measure of software functionality: both the one to be delivered to the contracting entity/users to meet its/their requirements (the estimated one) and the one actually delivered (the measured one). The units are derived based on certain measurable characteristics, which determine software product size: "They are a quantitative representation of (…) functions (…) and data working together as a computer application, (…) relying on program components corresponding to requirements" (Parthasarathy, 2007). Hence, the described units allow for the perspective of the user/client, which is reflected in particular in:

- software size measurement from the viewpoint of an attribute that is of primary significance for the user/client, i.e., functionality, which is treated as the principal benefit of the project;
- capability to estimate software size based on the specification of requirements relatively early in the project life cycle and, on this basis, to estimate its costs in a sufficiently objective and reliable manner at the same point in

time, that is to price the product, which enables a comparison and appraisal of competitive proposals in terms of the proposed functionality, delivery costs, and duration;
- capability to control the progress of the project by the client relatively easily – based on the specified required functionality, even a client/user with no knowledge or experience in the area of software system development, the applied project methodology, or the used technology can monitor the activities related to system development, since the product size expressed in this manner is independent of them and allows for all standard stages of the project life cycle, which enables the determination of the effort and costs required for the execution of its individual stages and a relatively easy identification of the stage at which the project should be;
- capability to measure the functionality actually delivered by the contractor and to compare it with the functionality required by the client in order to establish the degree to which the contractor has fulfilled its obligations and to determine the actual amounts due for the delivered system;
- compliance with the economic definition of productivity, and hence the accuracy of the indicators obtained on their basis from the economic perspective – this means that if the total costs of a project decrease as a result of using a more efficient programming language, the application of functionality units indicates a decrease in unit cost and so an increase in productivity.

Hence, functionality units serve the purpose of estimating and measuring software size based on criteria that are not only understandable to contracting entities, with no need to familiarize them with technical and implementation details, but also perceived by them as the most important ones. Therefore, the units facilitate the contracting entity's active participation in the project, which raises the probability of a precise specification of requirements and acceptance of a product as compliant with the requirements, which in turn contributes to the success of the entire project. However, methods based on the discussed units are characterized by high usability also from the viewpoint of software developers, since they permit an analysis, evaluation, and comparison of the attributes of the completed projects and their products, even when using various development technologies/programming languages, which in turn enables, among others, determination of the productivity that is specific to a given software development organization and constitutes the basis for estimating the effort and development/enhancement costs, and thus pricing the products of future projects (data obtained in this manner should be collected in the organizational historical data repository), and determination of the software size change arising from the modification of the requirements by the contracting entity during the project life cycle and pricing of the change.

Research on the objectivity and reliability of software system size estimation in functionality units early in the project life cycle was conducted by ISBSG (2005). It concerned both the estimation of the product size expressed in this manner based on the data model and its methodical calculation based on a specification of requirements. In both cases, it was found that the product size expressed in functionality units is estimated in a sufficiently objective and accurate manner compared to the calculations made based on the final product, with the estimates derived from the specification of requirements being characterized by higher reliability than those obtained from the data model: in the former case, 70% of the estimates were not lower than the actual product size, while in the latter case, the ratio was 62% of the estimates. In both instances, the acceptable error of estimate was assumed at ±10%. As regards the objectivity of estimates for product size using functionality units, it is ascertained that two specialists get results differing by merely ±10%, yet on condition that the requirements are specified properly. The findings of ISBSG's research are confirmed also by e.g., M. Parthasarathy (2007), who stated the objectivity and a relatively high level of reliability of software size estimation based on function points. Therefore, functionality units are increasingly commonly used for determining the costs of a project, that is for pricing its product in system development contracts.

However, functionality units are not flawless. The source method (Albrecht's FPA) was criticized primarily for its lack of universality – there were doubts whether it was possible to measure the size of all software categories properly by means of such units. While adequate measurement of business software system (data-driven) functionality was never controversial, in the case of software products characterized by a low number of inputs and outputs and at the same time a high internal processing complexity (e.g., real-time systems, embedded systems, telecommunications software, process control systems, scientific software with complex mathematical algorithms, expert systems, CAD), size measurement based on functionality units was regarded as incomplete. Therefore, variants of the described units were proposed in order to expand the capability to measure the functionality of a software product to cover, in the first place, real-time systems (time-driven), for example full function points developed by COSMIC (Common Software Measurement International Consortium). It displayed usability for both data- and time-driven software systems.

Another critical argument raised against software size estimation and measurement based on functionality units is quite a high degree of its complexity. However, if considering the fact that there are numerous tools to support such measurement, the argument loses some of its validity. Some, because the level of complexity of methods based on such units makes the tools incapable of supporting all operations needed for their correct employment. This is why a substantial portion of the process is still a non-automated operation requiring high skills in the discussed area although an international standard has even been introduced recently to standardize the automation process (ISO/IEC 19515:2019). Nevertheless, software systems are inherently complex and hence it is unrealistic to expect that their size measurement method will be effective and simple at the same time, and the costs of the process should be treated as an investment in the improvement of the maturity of software processes in an organization.

In light of the acceptance of several measurement methods based on functionality units by international standardization organizations, the currently voiced criticism concerned with subjectivity of such measurement also begins to lose its validity – objectivization was one of the fundamental conditions for them being recognized by the standardization organizations.

3.2 Measurement standardization in software engineering

According to the definition of software engineering adopted by the Institute of Electrical and Electronics Engineers (IEEE), it "is the application of a systematic, disciplined, quantifiable approach to the design, development, operation, and maintenance of software" (IEEE Std 610.12–1990). The quantifiable approach means that software process and product measurement should be an immanent characteristic of this discipline of knowledge and life. However, software engineering cannot boast an elevated level of maturity in terms of measurement of its subject matter, including in particular various software attributes, with the main focus on its size. Hence, the lack of an unambiguous measure of software product size is the fundamental problem with estimating the main attributes of software system projects. This is the source of the multiannual endeavors to create objective and reliable software system size measurement methods as an attribute required for the proper estimation of the project effort and hence its costs, i.e., product pricing.

Moreover, the employment of standard approaches to measurement in the practice of software engineering is important to shareholders of IT projects. This is because it helps build contracting entities' confidence in prospective contractors, thus facilitating the selection of the product developer based on objective and reliable criteria. It can also be a formal requirement arising from the adopted laws or other regulations (e.g., the condition of participation in a tendering procedure, a contractual clause) or an informal requirement originating from the fact that customers prefer businesses certified in a given area. This is why, in order to gain a competitive advantage or, at times, stay in business, software development organizations should be motivated to implement the so-called good practices provided for in such standards and subject themselves to certification processes, which bring benefits not only to their customers but also to themselves. This is because they regulate the activities of such organizations, rationalize the planning based on reliably collected historical data, as a result of which the difference between the estimated values and the actually obtained ones diminishes, permit the ongoing control of the key attributes, which allows the shortening of the response time to undesirable situations and the enhancement of the capability to develop software systems of a suitable quality, and the identification of the areas to be improved, thus reducing the uncertainty of the organization's operation, improving its effectiveness and market image (Kobylinski, 2005). Although the implementation of formal approaches often involves high expenditure, might initially impose constraints on operation flexibility, requires relatively frequent adaptations to subsequent improved versions, and certificates do not guarantee that the client's requirements will be satisfied and are not the necessary condition to do so, it definitely facilitates that.

The best-known and most renowned international standardization organizations, in particular in the area of software engineering, include ISO (International Organization for Standardization) and IEC (International Electrotechnical Commission) as well as IEEE, which joined some standards developed earlier by ISO/IEC. ISO, IEC, and IEEE are independent, world-leading, non-governmental international organizations that bring together experts to share knowledge as well as develop consensus- and market-based standards supporting innovation and providing solutions to global challenges (IEC, 2023; IEEE, 2023; ISO, 2023). ISO does this with membership in 167 national standards bodies (ISO, 2023). IEC focuses on all electrical, electronic, and related technologies (IEC, 2023). IEEE is a professional technical organization dedicated to technological innovation, development, and excellence for the benefit of humanity (IEEE, 2023).

The ISO/IEC standards on measurement in software engineering can be divided into the following categories with the increasing level of specificity of the standards (the basic published standards are enumerated here, with the exclusion of accompanying standards and ones under development):

1. Standards covering measurement, including ones regarding:
   - measurement process: ISO/IEC/IEEE 12207:2017, ISO/IEC/IEEE 15288:2015, ISO/IEC 14598:1998–2001;
   - software process measurement: ISO/IEC/IEEE 90003:2018, ISO/IEC 15504:2003–2013;
   - software product measurement – ISO/IEC SQuaRE (Systems and Software Quality Requirements and Evaluation) series standards, i.e., some with numbers 250xx, including in particular: ISO/IEC 25000:2014, ISO/IEC 25001:2014, ISO/IEC 25010:2011, ISO/IEC 25030:2019;
2. Standards dedicated to measurement, regarding:
   - measurement process: ISO/IEC/IEEE 15939:2017;
   - software product measurement – some ISO/IEC SQuaRE series standards, i.e., ISO/IEC 25020:2019, ISO/IEC 25021:2012, ISO/IEC 25023:2016, and ISO/IEC 14143:2002–2012;
3. Standards regarding individual software product functional measurement methods: ISO/IEC 19761:2011, ISO/IEC 20926:2009, ISO/IEC 20968:2002, ISO/IEC 24570:2018, ISO/IEC 29881:2010.

Standards from categories two and three focus explicitly on measurement, whereas the ones from category one perceive the process as one of the recommended categories of activities. Furthermore, some of them are more general, i.e., they standardize the very process of measurement, while others concentrate on measuring software products (software systems) or software processes. In the case of software products, in turn, standards regarding product quality measurement and standards regarding product functional size measurement – and only such size – can be distinguished.

The process of selecting and applying software process and product measurement standards should include the following steps:

- identifying the business needs and goals of an organization;
- identifying the measurement needs and goals, e.g., product delivery cost and/or time reduction, productivity increase, high-quality product delivery, accurate estimation of project attributes;
- selecting the standard covering/dedicated to the measurement process in the context of measurement needs and goals;
- selecting the relevant standard covering software process quality and/or software product quality measurement and/or dedicated to software functional size measurement – depending on measurement goals;

- if a software quality measurement standard has been selected, selecting the relevant SQuaRE standard, and if a software functional size measurement standard has been selected, selecting a standard regarding the measurement method that is adequate to a given functional domain; and
- performing the measurement for the purpose of obtaining the information required for making reasonable management decisions.

The integration of the activities covered by measurement is intended to support, among others, objective planning, including estimation of attributes of software system projects, monitoring of the actual progress of a project in terms of its compliance with plans and goals, and identification and resolution of the emerging problems related to the measurement process. Therefore, in organizations characterized by high maturity of software product and process measurement (Czarnacka-Chrobot, 2012):

- project execution costs and time are estimated more accurately due to the proper collection of reliable historical (benchmarking) data: organizations at the highest maturity level virtually never exceed the estimates, whereas organizations at the first level exceed time by 150% and costs by approx. 200% on average;
- the quality of software products increases, since the number of errors decreases as a result of their control starting from the earliest phases of the project life cycle, which leads to lower costs of their maintenance: the average cost of remedying errors in organizations at a high level of maturity is approx. 4% of the total expenditure on software development, while in organizations at a low level – over 50% of such expenditure;
- due to the reduced costs of improving inadequate quality software products, software development costs decrease: the average cost of creating one function is over 3 times lower in organizations at the highest level relative to the cost in organizations at the lowest level;
- as an effect of the reduction of the aforementioned unit cost, productivity grows;
- due to the smaller number of errors in software products and the resulting reduction of the time dedicated to their repair, the total delivery time shortens, too.

As shown above, the application of standards regarding software process and product measurement contributes to the improvement of such a process and product through increased productivity of project activities and quality of the product, which permits a faster and cheaper delivery of a better functionality. Therefore, the implementation of the discussed standards in line with an organization's business needs and goals, which encompasses understandability to and approval by their users/clients, might be perceived as an investment in software process and product improvement.

In summary, it is worth mentioning that a major progress in the development of new standards for measurement in software engineering, in particular as regards software products, has been recorded in recent years. Additionally, standardization institutions attach significant importance to the up-to-datedness of the standards (they are updated on an ongoing basis if necessary).

3.3 ISO/IEC 14143 standard for software system functional size measurement

Software systems, like any other product, in particular one of an engineering nature, are characterized by certain attributes, which should be subject to measurement. The basic attribute of each engineering product is its size. As was already mentioned, software engineering cannot boast such a high level of maturity in terms of units for measuring software size as other engineering disciplines (e.g., construction engineering). This is the basic reason for difficulties with objective and reliable cost estimation of software projects, that is a proper software product pricing.

> Measurement of software size (…) is as important to a software professional as measurement of a building (…) is to a building contractor. All other derived data, including effort to deliver a software project, delivery schedule, and cost of the project, are based on one of its major input elements: software size.
>
> *(Parthasarathy, 2007)*

From among the several software system size measures applied in practice, i.e., (1) programming units (e.g., source lines of code); (2) construction complexity units (e.g., object points); (3) functionality units, it is functionality units in the form of the so-called function points (FPs) that currently enjoy the greatest worldwide recognition. This is reflected in them being accepted by ISO and IEC in ISO/IEC 14143:2002–2012 norm, which standardizes the concept of the so-called software functional size measurement (FSM), as the only proper ones, i.e., sufficiently reliable, and objective software system size units.

In accordance with the ISO/IEC 14143 standard, software functional size is "size of the software derived by quantifying the Functional User Requirements" (ISO/IEC 14143:2002–2012). While Functional User Requirements (FUR) are defined as "the sub-set of the User Requirements describing what the software does, in terms of tasks and services" (ISO/IEC 14143:2002–2012). Hence, FUR is an ISO expression that denotes a sub-set of user requirements, based on the user's point of view. The standard specifies the definition of FUR by identifying their fundamental components and the elements that are not normally included in such requirements. The former group includes primarily the requirements concerning: transferring data (e.g., entering customer data, transmitting a control signal), processing data (e.g., calculating a bank's profit, determining the average temperature), collecting data (e.g., customers' orders, ambient temperature over time), and searching data (as regards e.g., the current employees, aircraft position). Functional User Requirements do not include, however, above all the following: qualitative constraints (e.g., usability, reliability, efficiency, maintainability, portability), organizational constraints (e.g., operation placement, target equipment, compliance with standards), environmental constraints (e.g., interoperability, security, privacy), and implementation constraints (e.g., programming language used, testing tools applied, delivery schedule). Hence, the non-functional requirements constrain the software project, but do not describe the services (functions) to be supplied by the software (COSMIC, 2008). Therefore, based on the older version of the discussed standard, FUR can be defined as representing "the user practices and procedures, that the software must perform to fulfil the user's needs. FUR exclude Quality Requirements and any Technical Requirements" (ISO/IEC 14143:2002–2012), in order to ensure measurement objectivity and comparability of its results. The software functional size measurement method, in turn, is defined in the discussed standard as "a specific implementation of functional size measurement defined by a set of rules, which conforms to the mandatory features of such measurement" (ISO/IEC 14143:2002–2012). The software functional size measurement method, in turn, is defined in the discussed standard as "a specific implementation of functional size measurement defined by a set of rules, which conforms to the mandatory features of such measurement" (ISO/IEC 14143:2002–2012).

Hence, the document standardizes the concept of software system FSM, among others:

- providing a guide on how to use FSM to support the management of software systems projects;
- defining the FSM method;
- identifying the common characteristics of FSMM and determining the set of obligatory requirements they must satisfy;
- facilitating the selection of the FSM method that is best suited to user's needs;
- defining and describing the so-called functional domains (software classes) for which FSM methods can be used;
- facilitating the process of defining the functional domains based on the preset FUR;
- providing the rules for selecting the FSM method that will be best suited to a given functional domain; and
- recommending the steps in the FSM method application process.

According to the ISO/IEC 14143 standard, the fundamental principles, that is obligatory features of software functional size measurement methods, include (ISO/IEC 14143:2002–2012):

- basing the method on the representation of the user's functional requirements, defined from the users' perspective;
- the possibility to apply it as soon as FUR are defined and made available;
- the outcome of the method is the functional size obtained from the evaluation of the Base Functional Components (BFC), which are the elementary unit of FUR;
- the FSMM should be as independent of software development methods and technologies as possible.

The FSM method application process, in accordance with ISO/IEC 14143 standard, should be composed of the following steps:

1. determination of the scope of functional size measurement, that is the set of FUR taken into account in a given case of FSM determined by the software measurement goal;

2. identification of the FUR falling within the scope of FSM;
3. identification of the BFC falling within the scope of FUR;
4. classification of the BFC in terms of their size;
5. assignment of the correct value to each BFC;
6. calculation of the functional size of a software product.

What follows from the above is that when measuring a given set of FUR, determined by the purpose of the process (the so-called scope of measurement), in particular based on a preset functional model (hence, the degree of specificity is defined), the results obtained by independent specialists measuring software size by means of FSMM should be very similar, if not identical. This is the primary sense of using the methods: the results concerning software size are supposed to be reliable and objective and, additionally, comparable. Moreover, as regards the result of FSM methods (software system size) the unique technology applied by contractors or their individual know-how are irrelevant, since the result, as mentioned above, must be independent of these attributes, as it reflects FUR, which are independent of them. The only thing that matters is the application of specific measurement procedures, which – as an effect of taking the requirements into consideration – must lead to standard results in the form of software system size.

According to the ISO/IEC 14143:2002–2012 standard, the following are the areas supported by FSM methods:

- project management, in the area of both planning and control, by enabling:
  early estimation of resources, including effort costs, which are required for project execution,
  monitoring of the progress of project execution, including expenses,
  management of changes in the required project product size – depending, among others, on the estimated costs of such changes,
  determination of the degree of FUR satisfaction by the purchased commercial software product or the product developed as contracted by the contracting entity,
  post-completion analysis of the project and comparison of its attributes, including effort, costs, productivity, to other projects,
- software development management by:
  managing parameters (attributes) of software development/enhancement and maintenance projects, including e.g., productivity, costs, speed of delivery, in particular in terms of the influence of various factors on the parameters,
  managing software quality, in particular reliability,
  managing the maturity of an organization and capabilities of software processes,
  enabling the determination of the organizational value of the whole or a part of the software in order to estimate the cost of its potential replacement, the so-called reengineering and outsourcing,
  enabling the estimation of the budget required for software maintenance,
  managing software delivery contracts.

3.4 ISO/IEC standards for software system functional size measurement methods

As an effect of a multiannual verification of the reliability and objectivity of individual FSMM, ISO and IEC considered five (out of approx. 25) of them as compliant with the rules provided for in ISO/IEC 14143 standard. They include:

A. First generation methods:
   1. International Function Point Users Group (IFPUG) method – standardized in ISO/IEC 20926:2009; the measure unit in the method is 1 IFPUG FP;
   2. Netherlands Software Metrics Association (NESMA) method – standardized in ISO/IEC 24570:2018; the measure unit in the method is 1 NESMA FP, which is currently considered equivalent to 1 IFPUG FP;
   3. Mark II (MkII) method developed by the United Kingdom Software Metrics Association (UKSMA) – standardized in ISO/IEC 20968:2002; the measure unit in the method is 1 UKSMA FP.

B. Second generation methods:
   4. Common Software Measurement International Consortium (COSMIC) method – standardized in ISO/IEC 19761:2011; the measure unit in the method is 1 COSMIC FP (COSMIC Function Point – CFP);
   5. FSM method developed by the Finnish Software Measurement Association (FiSMA) – standardized in ISO/IEC 29881:2010; the measure unit in the method is 1 FiSMA FP.

The first three of the abovementioned FSMM have not been approved by ISO/IEC in their full versions compared with the versions proposed by the organizations developing the methods. This is why they are called first generation methods. However, it must be emphasized that the parts that are the most important, not only from the perspective of the aspects discussed in this chapter, i.e., the parts concerning software FSM, have been approved. The methods developed by COSMIC and FiSMA, in turn, have been approved by international standardization organizations in whole as proposed by both organizations developing them – hence, they are at times named second generation methods.

As mentioned before, FSM methods differ from each other in terms of the functional domains to which they can be applied. And so, the ISO/IEC 14143 standard states that the IFPUG, NESMA, and FiSMA methods are suitable to all software system classes, the Mark II method is more appropriate for business software, i.e., data-driven one, for which it has been specifically devised, whereas the COSMIC method is equally relevant to data-driven systems as it is to (real) time-driven and hybrid ones (ISO/IEC 14143:2002–2012).

The said standards, like ISO/IEC 14143, are compliant with ISO/IEC/IEEE 15939:2017, which determines the principal rules of the software measurement process. One of the key steps in the software measurement process, as defined in ISO/IEC/IEEE 15939 norm, is the procedure of selecting the method to be used for measuring its size. According to the procedure, the selection of the FSMM that is best suited to the needs of the software recipient should be composed of the following actions:

1. description of the organizational units of the software recipient from the point of view of the measurement process;
2. identification of their information needs with respect to the measurement process;
3. selection of the relevant FSM method based on the identification of applicable methods.

The requirements to be met by the relevant FSM method differ depending on the nature of an organization. For instance, financial institutions normally choose a method that correctly measures business software, while chemical enterprises, due to their core activities, require a measurement method that is more relevant to real-time systems. Hence, the selection of the proper method needs to begin with dividing the organizational software into functional domains. The selection of a method that is suited to needs depends also on the planned method of employing its result. If an organization intends to use the measurement results also for the purpose of comparing its own productivity with industry data, it is recommendable to choose a method that is relatively popular in a given industry, one for which such data exist. When it needs only a rough, tentative estimation of the functional size, the requirements for the relevant measurement method will be reduced.

As practice shows, the most popular FSM method is still IFPUG in the version accepted by ISO/IEC (ISBSG, 2016). Consequently, the resources of the benchmarking data based on which software process attributes, such as effort, productivity, and costs, are estimated are the greatest for this method. However, a younger method, i.e., COSMIC, has been dynamically gaining popularity worldwide for a dozen or so years now although the historical data resources existing for the method are still considerably smaller than the ones for the IFPUG method. As regards the NESMA method, it is assumed to produce virtually identical results as the IFPUG method – it is, in fact, its simplification. The UKSMA and FiSMA methods are essentially used only locally: the former in the United Kingdom and the latter in Finland.

What follows from the above is that there are two FSM methods with the broadest global reach: IFPUG and COSMIC. They are described in sufficient detail not only in the aforementioned ISO/IEC standards, but also in the relevant literature, where the reader will also find a comprehensive description of the differences between these two methods. Hence, their description will not be included in this publication.

However, it is worth mentioning here that the ISO/IEC/IEEE 32430:2021 Software engineering – Trial use standard for software non-functional sizing measurements, which attempts to measure Non-Functional User Requirements as a supplementation of the ISO/IEC 20926:2009 standard (the IFPUG FP method), has just been published. This is because the authors of the standard make the assumption that a combination of the functional size and the non-functional size should constitute the entire system size. Nevertheless, serious objections have already been raised against this standard, the most significant of which include: (1) in reality, it is an upgraded version of influence factors from the IFPUG method (the part of the method that was not accepted by ISO/IEC in standard 20926 on grounds of its subjectivity); (2) the measurement process uses different scales and different measurement units for different parts of the evaluation process and reduces them to one scale, which is a certain contradiction of mathematical rules; contradiction of mathematical rules; (3) the considered process does not provide any added value with respect to the one described in the ISO/IEC 29881 standard (the FiSMA method). Additionally, it will be necessary to wait relatively long for the appropriate resources of relevant benchmarking data to be collected, taking into account also Non-FUR, in order to estimate the attributes of software system development projects, including costs and pricing.

4 IMPORTANCE OF GENERALIZED BENCHMARKING DATA IN SOFTWARE PROJECT COST ESTIMATION – BASIS OF RESEARCH

As noted above, a proper cost estimation and pricing of software systems, in particular dedicated ones, requires not only the application of a reliable and objective software size unit, such as functional size unit, and an appropriate method intended for this purpose, that is one of the standardized FSMM, but also the knowledge of the software development/enhancement productivity or models pointing to the dependence of effort on the functional size of such a system. Both methods require access to appropriate resources of historical (benchmarking) data. This because they enable, based on the result obtained with the selected FSM method, estimation of the following attributes of the projects: total effort and total effort costs, unit effort and unit effort cost, duration, speed of delivery, and productivity of project activities. Their knowledge, in turn, enables reasonable decisions regarding: selection of the appropriate contractor – on the part of the contracting entity, participation in the tendering procedure and possible launch of the project – on the part of the contractor, and correct contracting of the subject matter of the contract and the possibility to compare the delivered values with the planned ones – on the part of both parties.

The appropriate resources of their own historical data, which would permit the correct determination of their own, specific to a given software development organization, attributes of the projects, are held by a relatively small number of software development organizations, since the conditions for holding them include (Czarnacka-Chrobot, 2009):

1. the effective implementation of measurement programs and their application to as many software products and processes as possible such that at least the median (the medium value is in this case a more reliable one than the arithmetic mean, since it permits the avoidance of the influence of several untypical, i.e., outlying, projects) and, ideally, the model suitable for the project attributes that is representative of a given organization, can be determined on this basis, which is not often;
2. collection of historical data for a relatively high number of projects executed in the past that were similar in terms of application (including the nature and size of the software product) and technology such that at least the median and, ideally, the model suitable for the project attributes that is representative for a given organization when taking into account the essential

factors affecting the attributes (mainly: project type – system development/enhancement; product type and size – a business software or another application, of what size; generation of the programming language used) can be determined;
3. reference of the historical data to the appropriate software size units.

In such a situation, also when comparing the organizational value of the attributes with the standards applicable as the so-called good practices, the usefulness of repositories with generalized and generally available historical data becomes apparent. The usefulness of such data is additionally reinforced by their quantity as well as objectivity and reliability since they are subjected to the relevant qualitative treatment and cleaning.

The largest commonly acknowledged and available repositories of generalized benchmarking data for software projects the products of which are measured with FSM methods, including in particular the still most popular IFPUG method, are currently held by the International Software Benchmarking Standards Group (ISBSG, 2020a; ISBSG, 2020b). The repositories are considered to be the most reliable sources of data on the attributes of a high number of latest software projects from numerous industries.

ISBSG is a non-profit organization founded in the mid-1990s with a view to improve IT resource management processes in both economic entities and public administration institutions. It performs the task by maintaining, developing, and using repositories of historical data provided by the leading enterprises in the IT industry and ones dealing with measurement. One of them, concerned with software development projects (new development and enhancement projects) and at the same time the largest one, includes (in the version of 2020) data from approx. 9,600 projects executed in more than 30 countries in various industries and business areas. The data are related, among others, to: software product size, effort, project schedule, architecture, development technology, details of the hardware platform – overall, there are a few hundred variables describing a single project. The organization holds also a repository collecting data for software system maintenance projects (nearly 1,200 projects in the repository version of 2020). Both repositories are standardized in line with ISO/IEC/IEEE 15939 norm, representative of the current technology, and – obviously – verified, since ISBSG takes particular care of data quality.

The data collected in the repository for software development projects are classified by ISBSG by, among others, the following criteria that are significant for each project attribute (ISBSG, 2020b):

- country where the project was executed (i.e., effort was made – this is not the same as the place from which the data were made available), approx. 80% of which are data from the so-called developed countries;
- context of the project, including: economy sector (the broadly defined financial sector prevails – approx. 36%) and business area (the insurance area prevails – nearly 19%);
- project type, including: type of project activities (system enhancement – approx. 70%, new development – approx. 30%), intended use of the

product of the project (products realized for the needs of the organization transmitting data to the repository – approx. 92%), and size of the project team (products with teams of up to nine members – approx. 60%);
- product type, including: application type (including business applications, i.e., data-driven ones, which constitute over 91%) and product type – expressed mainly in IFPUG FPs (approx. 72%, and if also adding the one measured in NESMA FPs – over 80%), but the products measured in COSMIC FPs are also represented (approx. 740 projects, i.e., almost 8%);
- project execution environment, including programming language (more than 120 programming languages are represented, 68% of which are 3GLs, and 30% – 4GLs), and a hardware platform (approx. 32% are mainframe projects);
- the development methodologies applied: almost 73% of the projects used the waterfall methodology (characteristic of large projects), and nearly 15% – agile/RUP methodologies.

When using the data collected by the organization, it should be remembered that they are representative of extraordinary projects, which arises from the following:

- the criteria for collecting data in the ISBSG repository allow for only those organizations which employ FSM methods and do it properly, and such organizations are considered more mature than others, since they execute software measurement programs;
- contractors choose the projects whose data are provided by them to the ISBSG repository themselves – these can be both projects that are typical of them and those with the best parameters, including the highest productivity and/or speed of delivery (for obvious reasons, data on projects characterized by poor parameters are not provided), but it needs to be remembered that most data come from clients;
- the ISBSG repository does not contain vast quantities of data on large projects (such projects are rarely completed, and they succeed even more rarely, i.e., their parameters usually fail to fit into the planned boundaries and exceed them markedly in an undesired direction).

It needs to be stressed that the ISBSG data are subjected to a process of strict qualitative verification. This is why they are appreciated in the IT industry. The comprehensiveness, reliability, quality, and up-to-datedness of the data sets are the main advantage of ISBSG, which offers repositories that are, for these reasons, unique relative to other available sets of data on completed projects of the relevant type.

To recapitulate on the above discussion, software system pricing can be said to depend on the estimated costs of its execution, which – in the case of the analyzed types of projects – are determined by the estimated effort costs. They, in turn, are conditioned on the estimated effort of project activities, which is determined by the estimated size of the software system to be developed/enhanced as a result of the execution of the project. The size is estimated by means of standardized FSM methods based on the contracting entity's requirements. The dependence of effort on size, in turn, can be derived from the known or estimated productivity, or, instead, using models which show such a dependence. The former option, which takes into account productivity of project activities, is simpler and hence more frequently applied. However, when productivity is unknown yet and therefore it still needs to be estimated, the basis for that is benchmarking data, preferably organizational ones (for the abovementioned reasons), but in practice they are usually generalized. Such generalized data are available in the ISBSG repository, so no complicated calculations are needed. Nevertheless, there are attempts at building models to estimate the effort of software system project depending on its size without using productivity. The majority of such models are still regression models, but there are increasingly more frequent attempts at constructing models based on data mining, especially ML techniques, using the ISBSG data repository as the largest and the most objective and reliable one in terms of quality.

5 USE OF MACHINE LEARNING ALGORITHMS FOR SOFTWARE PROJECT COST ESTIMATION – PROPOSED WORK, RESULTS, DISCUSSION

5.1 Data mining

> Data mining, known also as knowledge discovery in databases is the nontrivial attempt at extracting implicit, previously unknown, and potentially useful information from data. This encompasses a range of techniques (…) in order to identify information or decision-making knowledge in a database and extract it such that it can be used in processes such as decision-making support, prediction, and estimation. The data are often extensive, and it is impossible to derive information from them directly.
>
> *(Pujari, 2013)*

Data mining techniques are divided, primarily, by goals of the conducted analyses, which could be pattern recognition and predictions. In the case of predictions, it is estimation and generalization of the values of unknown attributes based on the patterns found by means of e.g., artificial neural networks (non-resistant to missing data) and decision trees (resistant to missing data). In the case of pattern recognition, the recognized patterns are used for describing data and finding general characteristics, thus obtaining interesting and significant information, taking into account e.g., clustering methods (Fronczak & Michalcewicz, 2010).

There are numerous data mining techniques originating from established fields of science, including ML. Data mining uses the state-of-the-art ML algorithms for extracting useful knowledge from databases/repositories through the automated learning process based on the supplied data (Ben-David & Shalev-Shwartz, 2014).

Broad research on the application of data mining techniques to the estimation of software system project attributes (Sehra et al., 2017), and in particular machine learning algorithms (Wen et al., 2012), has been conducted for just over two decades. ML is a form of artificial intelligence, which can be used for decision-making process automation and prediction, i.e., ML enables computers to improve efficiency through data: exemplary or experimental. "Various parameters define a model (that can be predictive, descriptive, or both), and the learning is optimizing those parameters by a program that uses the training data or experience data" (Mezouar & El Afia, 2022). The techniques are supposed to eradicate shortcomings of various simpler, traditional estimation techniques based on a smaller quantity of data, which do not take into account the multitude of factors influencing productivity, effort, and costs of project activities, and – as a result – increase of the chances for success of such projects and their better adaptation to modern approaches to project management and software development. However, in the software project management practice, ML is still rarely employed despite the fact that the existing research demonstrates its usefulness (Mezouar & El Afia, 2022), e.g., in the following areas: effort/cost and project duration estimation (e.g., López-Martín, 2015; Wen et al., 2012), project monitoring (Azzeh et al., 2010), software maintenance cost determination (e.g., Shukla et al., 2012), and quality and risk management (e.g., Moeyersoms et al., 2015).

The main advantage of ML algorithms is their adaptability to the changing environment. It is a particularly important factor in estimating attributes of software system projects, where a dynamic progress in the area of technologies, new programming tools and languages, methodologies, and project team skills is noticeable. This can be achieved through ML algorithms' ability to learn and improve based on the generated estimations that best match the observed data and the entire previous knowledge (Mitchell, 1997). Hence, ML algorithms are suitable to modeling complex problems that are difficult to program and imitate, up to a point, the human learning process, which can be unsupervised (descriptive tasks; Linoff & Berry, 2011) or supervised – in which case it needs explicitly defined inputs that correspond to the known outputs by the specified target variable (Cios et al., 2007), which is used for classification or prediction. For both of them, there are many algorithms that can be used depending on the desired outcome. In the case of the issue described in this chapter, the point is to predict the target variable of effort/cost and, potentially, duration.

5.2 Use of machine learning algorithms – Related work

According to the latest Systematic Literature Review (SLR) carried out from the perspective of the application of ML algorithms in software engineering (Mezouar & El Afia, 2022), only several studies (4) are dedicated to the use of ML algorithms in estimating software effort/cost, and some of them are not interesting from the perspective discussed in this study, as they concern either software size estimation methods that are not recognized by international standardization organizations or project team efficiency.

To the best knowledge of the author, the most comprehensive overview of the ML algorithms applied to estimating effort/cost for software projects was conducted by J. Wen and his co-authors (2012). For this purpose, they analyzed 84 studies from slightly over two decades. In accordance with the findings, the authors focused primarily on adapting individual algorithms to their best efficiency, and these were mainly the following algorithms: Artificial Neural Network (ANN), Case-Based Reasoning (CBR), and decision trees. The accuracy level of the ML models was acceptable and better than the statistical one. The researchers also indicated that, depending on the set of data applied to the creation of models and the approach to data preprocessing, ML can generate diverse results: due to the outliers, missing values, and the possibility of overfitting. The focus in López-Martín & Abran (2015) is, in turn, on the precision of effort/cost estimation of software projects by means of diverse types of neural networks – the ISBSG data set was also used to that end.

In general terms, ML algorithms are considered highly effective in dealing with uncertainty and risk, which are characteristic of estimation in this area, and the obtained results show their pretty high capabilities with respect to estimation of the effort/cost and duration of the discussed projects at initial stages of their life cycles (e.g., Lopez-Martin et al., 2012). Additionally, through the automated prediction process based on benchmarking data, they can minimize the effect on the process of estimating undesirable factors in the form of human attitude, psychological factors, or ones arising from an organization's policy. However, still not many, if any, models relying on such algorithms are used in practice (Pospieszny et al., 2018). This state of affairs might be caused by narrowing down the research concentrating on finding the most accurate data mining algorithm possible, which was frequently performed on small and obsolete benchmarking data sets of completed projects (e.g., Pai et al., 2013). In addition, for the preparation of the data that play a key role in devising effective models, diverse, often conflicting, methods were employed (e.g., García et al., 2016). And so, the results concerning the accuracy of individual algorithms are ambiguous, even if they were applied to the same data set.

As mentioned above, the ambiguous results were to a considerable extent caused also by the application of small, obsolete data sets for ML algorithms. Therefore, it needs to be stressed that it is the data set offered by ISBSG that has been broadly used by researchers for estimating the attributes of the discussed projects by means of data mining algorithms for a more then decade now (González-Ladrón-de-Guevara et al., 2016). Nonetheless, due to the use of a variety of ISBSG data set versions and diverse purposes of devising such models as well as the fact that ISBSG repositories contain quite a high number of missing values, no standard approach to their application has been developed (López-Martín & Abran, 2015; Mittas et al., 2015). Yet, despite the diversity of the approaches adopted to the construction of ML models based on the ISBSG database, it is possible to distinguish recommendations supporting their practical implementation in the case of estimating the effort/cost and duration of a software project at its early stages:

- Given ML's sensitivity to disruptions in data sets, the models should not rely on individual algorithms; instead, they should be based on algorithms used in combination, which additionally increases the accuracy of predictions (Minku & Yao, 2013; Pospieszny et al., 2018). Their construction should involve a set of varied but numerically limited algorithms, and a simple approach, such as averaging the obtained estimates to reduce the cost of the estimation (e.g., Ho, 1998).
- When training ML models, emphasis should be placed on data preprocessing, in particular in case of outliers or missing data since they have a substantial influence on the accuracy of ML algorithms. Apart from various available techniques (deletion, imputation), their use depends on a data set (Huang et al., 2015).

- For outliers, it is recommended to apply the common rule of three standard deviations from the mean (Ruan et al., 2005).
- Missing values should, if possible, be rejected to eliminate the errors that might change the accuracy of a ML prediction (Strike et al., 2001).
- Some studies (e.g., Berlin et al., 2009; López-Martín & Abran, 2015) show that the log transformation of effort/cost and duration generates more accurate estimates.
- Normally, in order to evaluate effort/cost and duration estimation models by means of ML algorithms, most studies use the following (Wen et al., 2012): MMRE (Mean Magnitude of Relative Error) and PRED (Percentage Relative Error Deviation), sometimes MdMRE (Median Magnitude of Relative Error), as they enable comparison of results and are still commonly and broadly applied by researchers. However, they can be criticized at times, especially MMRE as an asymmetrical measure (Myrtveit & Stensrud, 2012) that is sensitive to outliers (Port & Korte, 2008).
- In order to avoid overfitting of ML models, the k-fold cross validation is conducted for validation purposes (Idri et al., 2016).

Moreover, the models combining algorithms confirm the earlier deliberations, as they demonstrate that effort/cost and duration of a software project depend to the greatest extent on the size of the software. Their application is most beneficial at early stages of the project life cycle, when uncertainty is the highest, and the obtained values are the most desirable from the viewpoint of both the contracting entity and the contractor. The models can be used with both the waterfall approach to development and agile methodologies since early estimation of the budget and timeframes of a project is often required for its approval by decision-makers and enables its reasonable justification. Additionally, an accurate estimation of effort/cost and duration at this stage is often the most difficult activity for each development practitioner, and this is what determines, to a considerable extent, not only the successful completion of a project but also the correct pricing of a software system and selection of a specific contractor.

It would be ideal if models relying on ML algorithms were organization-specific and based on the organizational repository of data on the projects completed in the organization. Only then is it possible to accurately estimate the attributes of the projects to be commenced. The repository size can differ depending on the maturity of an organization and the stage of implementing software product and process measurement programs there. Nevertheless, for the purpose of an accurate estimation using ML algorithms, such a base should contain data on tens or, preferably, hundreds of projects. There is a view that data on 40–60 completed projects could suffice (Pospieszny et al., 2018). Another substantial factor is obviously quality – data should be complete and cleaned. To this end, the contractor can use data from the Enterprise Project Management (EPM) system, which is relatively popular in software development organizations, which apply it for monitoring project portfolios. On this basis, it is possible to employ standardized FSM methods for estimating software size, as they normally ensure the most accurate approximation, followed by ML algorithms for estimating effort/cost and duration of a project (Pospieszny et al., 2018). However, such an approach can be applied only in large and medium-sized mature organizations where a software measurement program has been implemented and which have data on a substantial number of completed projects at their disposal.

5.3 Example of using machine learning algorithms

Apart from the mentioned above study by J. Wen et al. (2012), from among the few studies concerning software projects effort/cost estimation by means of ML algorithms found in the SLR in (Mezouar & El Afia, 2022), the (Pospieszny et al., 2018) was found noteworthy, since it provides examples of applying several ML algorithms to the purpose discussed in this study.

The ambiguity of the results and unclear approaches to building data mining models encouraged the authors of the said study to continue their explorations. Their effects were presented with intention to demonstrate an approach that would allow reduction of the gap between the current outcome of the research and implementations in organizations, while being effective and practical. The set of ISBSG data, which were subjected to proper pretreatment, and three ML algorithms were used for that purpose. The obtained results in the form of project effort/cost estimation models might serve as a tool supporting decision-making by all parties engaged in any way in the software project. At the same time, emphasis was put on the said software project attributes, because their accurate estimation at an early stage of the project is the most difficult due to uncertainty and limited knowledge. Any major deviations of the constraints during the project life cycle could have a serious effect on the actually delivered software functionality, quality, and the ultimate successful completion of the project.

Undoubtedly, it is worth mentioning that one of the principal causes of unreliable results and inconsistent of the predictions of ML models is poor quality of data, in particular significant quantities of missing values and outliers. Hence, the proper data preparation is a critical task in the ML model building process – data should be pretreated through selection, cleaning, reduction, transformation, and feature selection (García et al., 2016). As noted above, the best source of generalized benchmarking data for the software projects where software size is estimated/measured in function points is the data collected by ISBSG and this is why that data set was used for the purpose of verifying ML models. The version employed for the research comprises data from more than 6,000 software projects executed over the past two and a half decades by numerous enterprises from almost 30 countries – but only data concerning projects from the last decade were extracted from them. Other data sets available for software projects (e.g., PROMISE Software Engineering Repository, COCOMO, SourceForge) are obsolete, or small, or unreliable. Nonetheless, the ISBSG benchmarking data also have flaws, including in particular heterogeneity (the data come from various organizations) and quite a high number of missing values, which – along with heterogeneity – poses a challenge for the proper preparation of data and construction of ML models (Fernández-Diego & González-Ladrón-De-Guevara, 2014). However, despite the flaws and the fact that there are various versions of ISBSG repositories, the ISBSG data set has been extensively applied by researchers for estimating software projects by means of data mining algorithms (González-Ladrón-de-Guevara et al., 2016).

In the case of the issue described in this chapter, the point is to predict the target variable of effort/cost of software project. To that end, the following three ML prediction algorithms were employed:

1. Support Vector Machines (SVM) – it is capable of modeling complex linear and non-linear problems and generating extremely accurate results, even on noisy data (Han et al., 2006). However, their training can be time-consuming in the case of a large quantity of data and resources of benchmarking data on the completed software system projects might contain data on thousands of such projects.
2. Multi-Layer Perceptron Artificial Neural Network (MLP-ANN) – it is a parametric algorithm and is very robust in relation to noisy data (Larose & Larose, 2015). Yet, it is prone to overfitting in the event of a prolonged training process, and hence the application of another algorithm, namely:
3. Generalized Linear Models (GLM) – an algorithm being a generalization of simple linear regression, which – depending on several properties – serves the purpose of representing relations between attributes (Clarke et al., 2009). Due to its flexibility, it is efficient in handling non-linear variables and ensures a broad range of inference tools.

Afterwards, ensemble averaging, and cross validation were conducted. The model construction process took place in accordance with the cross industry standard process for data mining (CRISP-DM) methodology. The IBM SPSS Modeler was employed. Details of the procedure and specific model verification results can be found in (Pospieszny et al., 2018). This study presents only generalized findings:

1. All algorithms employed separately generated substantially accurate estimates – for both effort/cost (and also duration) of software project. When applied separately, the SVM algorithm performed best, and error measures were almost identical for MLP and GLM and insignificantly higher than they were for SVM. The cause might be the SVM algorithm's extraordinary capability of handling complex dependences in heterogeneous data, but it could also result from overfitting of the algorithm. The differences between the accuracy of the algorithms indicate that the factor that has the greatest influence on precision might be the approach applied for processing, cleaning, and transforming ISBSG data.
2. Although ensemble models were slightly less accurate than the SVM algorithm, due to their resistance to outliers, noises in the data and their adaptability, it is more recommendable to apply the combined approach since it is a more stable and practical procedure. This arises from the fact that if single algorithms are applied to another data set, e.g., one specific to a given software development organization, other algorithms might prove more accurate when estimating effort/cost (and also duration) of the project. The ensemble approach averages the predictions of three robust ML algorithms, stabilizes the model, reduces the effect of noise in data and the effect of incorrect behavior of algorithms.

3. It can be inferred from the findings of the described research that ensemble models generate accurate estimations due to: (a) appropriate pretreatment of the heterogeneous, diversified, noisy data set with missing values, (b) three effective ML algorithms, and (c) cross-validation method for prevention of overfitting.
4. Compared to other studies on ML algorithms for estimating effort/cost of software projects, a noticeable benefit from the proposed approach was found. What can be inferred from the most extensive (to the best knowledge of the authors of the discussed approach) comparison of 143 publications on effort presented in (Wen et al., 2012) is SVM's superiority over any other ML algorithm – the accuracy of other algorithms, such as ANN, was lower. However, it needs to be taken into consideration that most of the models were built based on homogeneous, older, and small data sets (e.g., COCOMO), which might result in underfitting or overfitting. Berlin et al. (2009), in turn, also employed an ensemble approach (ANN and linear regression) to the ISBSG benchmarking data with log transformation for effort estimation. In accordance with error measures for effort, their results were comparable to the results obtained in (Pospieszny et al., 2018). The results should also be compared with the software project effort ensemble models devised in (Pospieszny et al., 2015) using the ISBSG data set and the GLM, MLP, and CHAID decision tree algorithms. It shows that the approach adopted in the research of 2018 for preprocessing of data, including SVM in an ensemble model, and the use of cross-validation for result evaluation ensured more accurate estimates. Nonetheless, it must be noted that even with a similar data set and algorithms, the accuracy of the preliminary prediction might differ – depending in particular on the quality of data treatment and methods employed for this purpose.

To sum up, the described research achieved a high effectiveness of effort/cost estimation models through: (1) the use of the ISBSG data set and at the same time one based on the FSM methods approved by ISO/IEC, (2) appropriate data preparation, (3) application of three ML algorithms (SVM, MLP-ANN, and GLM) and their averaging, and (4) cross-validation. The obtained models are supposed to provide a decision-making support tool both for organizations developing and implementing software systems.

6 CONCLUSIONS AND FUTURE WORK

Incorrect cost estimation of software projects means that many of them are not completed or are completed at significantly greater costs than planned, which in turn not only reduces the success rate of software projects, but above all causes a waste of vast amounts of money invested in IT projects. Large losses in this respect are borne not only by companies, but most of all by state budgets, and therefore by us – citizens. The main purpose of this chapter was to demonstrate that ML-based software project cost estimation models can help minimize this problem, but the following conditions must be met: (1) the models are based on software FSM methods approved by ISO/IEC,

(2) the benchmarking data used by ML models must be of sufficient quantity and quality and hence their proper preparation needs to be prioritized.

The results of some ML algorithms, but primarily ensemble models for estimating effort/cost for software projects based on these assumptions, applied when they are needed most, i.e., at the initial stage of project life cycle, demonstrate that they are extraordinarily accurate compared with other approaches employed by researchers, experts, or software estimation enterprises. Nevertheless, still not many applications of such approaches can be found in practice – probably due to the emphasis on the effectiveness of individual algorithms and complexity of their construction and implementation procedures. However, the content of this chapter shows that it is worth using such models. Moreover, the accuracy of model prediction (effectiveness) for data specific to a given software development organization where the proposed approach would be employed, would be even greater due to homogeneity of data.

Models can be built using commercial or open-source software, but it is worth also considering their inclusion to the Enterprise Project Management software often existing in a software development organization in order to automate the entire process. They are supposed to support decision-making by an organization handling software development and/or implementation, in particular a large or medium-sized one, which should hold its own benchmarking data such that investors' funds are not wasted on systems that will never be implemented or their development and implementation costs will considerably exceed the planned budget. The application of the described models is the most desirable at the beginning of a project, when the uncertainty about the end product and about which variant of the system should be selected and delivered is the highest and any estimations are the most difficult, which is true regardless of the development methodology (waterfall, agile). This is because early estimation of the project budget and timeframe is often necessary to define the business case, set the functional boundaries of the product, and – consequently – have it approved by sponsors. Hence, they provide rational grounds for making the right business/investment decisions.

To recapitulate, it needs to be stated that, despite the fairly broad research conducted, in recent years, on the implementation of the state-of-the-art ML algorithms for the purpose of estimating effort/cost and duration of software system development projects, very few, if any, implementations can be found in software development organizations. And so, the information and knowledge contained in data is not extensively used in the practice of software system development, but some conceptual scientific studies illustrating the applications of predictive analytics methods in this area are promising. However, we can venture the statement that the practice of estimating the attributes of such projects will probably follow this direction in the nearest future. This is because, given the present dynamic environment of software development, the application of ML algorithms could improve the quality of project attribute estimation methods and contribute to a better allocation and engagement of resources. Prediction might be particularly useful in the initial phase of a project, when it might enable, among others, risk mitigation when estimating project attributes, minimization of costs and losses, and maximization of efficiency and organizational effectiveness of the contractor, or elimination of unnecessary requirements for a software system (Fatima, 2017). The knowledge derived from data could also effectively support contractors in reasonable task scheduling, subsequent cycles of the project, which ultimately increases the chance for delivering a complete software system within the set framework. It is not even impossible that, along with the technological progress, the process will be fully automated. At present, it is possible to model and simulate complex software construction processes as well as analyze various scenarios for the same project. Prediction techniques of this kind might definitely be at least an additional source of knowledge for the decisions to be made (Pospieszny et al., 2018).

As follows from the above discussions, the use of benchmarking data of an appropriate quality, relying on standardized FSM methods, is the right approach to the proper, i.e., reliable and objective, pricing of software systems. Research findings show that it is worth applying some of ML-based algorithms to this end, as they prove to be more effective than the traditional approaches to effort/cost estimation. As an effect, such an approach facilitates more the constraint on wasting the funds invested in software projects or, in broader terms, IT projects, by both businesses and the state budget. Therefore, it is worth beginning to apply them more extensively in practice, thus bridging the gap between the current research findings and implementation in organizations in practice.

This study has some limitations that determine the directions of further work. The main ones are: only generalized benchmarking data were taken into account, the usefulness of FSM methods was defined in a general way, without distinguishing them individually, and the verification was limited to ML algorithms only. Therefore, in the course of further research, attention should be given to the verification of ML models in several software development organizations, while taking into account organization-specific benchmarking data. Due to the fact that the software development effort/cost estimation accuracy depends directly on the explanatory variable, i.e., the estimated software system size, it is also worth attempting the identification of the FSM method that brings the best results from among the ones approved by ISO/IEC – the IFPUG and COSMIC methods are particularly deserving of being taken into consideration here. Moreover, the effectiveness of other data mining techniques is also worth verifying (see e.g., Stanek & Czarnacka-Chrobot, 2023).

REFERENCES

Albrecht, A.J. (1979). *Measuring Application Development Productivity*. Proceedings of IBM Application Development Symposium, Monterey, CA, 14–17.10.1979, 83–91.

Azzeh, M., Cowling, P.I., & Neagu, D. (2010). *Software Stage-effort Estimation Based on Association Rule Mining and Fuzzy Set Theory*. Proceedings of 10th IEEE International Conference on Computer and Information Technology (CIT-2010), 7th IEEE International Conference on Embedded Software and Systems (ICESS-2010), 249–256.

Ben-David, S., & Shalev-Shwartz, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 410.

Berlin, S., Raz, T., Glezer, C., & Zviran, M. (2009). Comparison of Estimation Methods of Cost and Duration in IT projects. *Information Software Technology, 51*, 738–748.

Cios, K., Pedrycz, W., Swiniarski, R., & Kurgan, L. (2007). *Data Mining: A Knowledge Discovery Approach*. Springer Science, 606.

Clarke, B., Fokoue, E., & Zhang, H.H. (2009). *Principles and Theory for Data Mining and Machine Learning*. Springer Science, 786.

COSMIC. (2008). *The COSMIC Functional Size Measurement Method, Version 3.0, Guideline for Sizing Business Application Software (Version 1.1).* Common Software Measurement International Consortium, 1–23.

Czarnacka-Chrobot B. (2009). The Role of Benchmarking Data in the Software Development and Enhancement Projects Effort Planning. In: H. Fujita, & V. Marik (eds.) *New Trends in Software Methodologies, Tools, and Techniques. Frontiers in Artificial Intelligence and Applications, 199*. IOS Press, 106–127.

Czarnacka-Chrobot, B. (2012). Software Development Processes Maturity versus Software Processes and Products Measurement. *International Journal of Computer and Systems Engineering, 6*(11), 20–26.

Fatima T. (2017). A Predictive Analytics Approach to Project Management: Reducing Project Failures in Web and Software Development Projects. *World Academy of Science, Engineering and Technology, International Journal of Materials and Metallurgical Engineering, 4*, 10–17.

Fernández-Diego, M., & González-Ladrón-De-Guevara, F. (2014). Potential and Limitations of the ISBSG Dataset in Enhancing Software Engineering Research: A Mapping Review. *Information Software Technology, 56*, 527–544.

Flasinski, M. (2007). *Zarządzanie projektami informatycznymi*. Wydawnictwo Naukowe PWN, 196.

Fronczak, E., & Michalcewicz, M. (2010). *Zastosowanie narzędzi eksploracji danych data mining do tworzenia modeli zarządzania wiedzą*. Polskie Stowarzyszenie Zarządzania Wiedzą, 89–100.

Galorath, D.D. (2004). *Effective Software Sizing. A Galorath Web Seminar*. Retrieved September 9, 2004. http://www.galorath.com/Effective_software_sizing.pdf.

García, S., Luengo, J., & Herrera, F. (2016). Tutorial on Practical Tips of the Most Influential Data Pre-processing Algorithms in Data Mining. *Knowledge-Based Systems, 98*, 1–29.

González-Ladrón-de-Guevara, F., Fernández-Diego, M., & Lokan, C. (2016). The Usage of ISBSG Data Fields in Software Effort Estimation: A Systematic Mapping Study. *Journal of Systems and Software, 113*, 188–215.

Han, J., Kamber, M., & Pei, J. (2006). *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 740.

Ho, T.K. (1998). The Random Subspace Method for Constructing Decision Forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*, 832–844.

Huang, J., Li, Y.F., & Xie, M. (2015). An Empirical Analysis of Data Preprocessing for Machine Learning-Based Software Cost Estimation. *Information Software Technology, 67*, 108–127.

Idri, A., Hosni, M., & Abran, A. (2016). Improved estimation of software development effort using classical and fuzzy analogy ensembles. *Applied Soft Computing, 49*, 990–1019.

IEC. (2023). The International Electrotechnical Commission. Retrieved January 10, 2023. https://iec.ch/who-we-are.

IEEE. (2023). The Institute of Electrical and Electronics Engineers. Retrieved January 10, 2023. https://www.ieee.org/about/vision-mission.html.

IEEE Std 610.12–1990: IEEE Standard Glossary of Software Engineering Terminology, The Institute of Electrical and Electronics Engineers. IEEE, 1–84.

ISBSG. (2005). *The ISBSG Report: Software Project Estimates – How Accurate Are They?* International Software Benchmarking Standards Group, 1–8.

ISBSG. (2012). *The ISBSG Report: Estimates – How Accurate Are They*. International Software Benchmarking Standards Group, 1–16.

ISBSG. (2016). *The ISBSG Special Analysis Report: The Use of Function Points in the Industry*. International Software Benchmarking Standards Group, 1–8.

ISBSG. (2020a). *ISBSG Corporate Presentation*. International Software Benchmarking Standards Group. Retrieved April 7, 2022. https://www.isbsg.org/wp-content/uploads/2020/02/2020.01.12-ISBSG-Corporate-Presentation.pdf.

ISBSG. (2020b). *ISBSG D&E – Corporate Release Demographics*, *ISBSG, R1 v1.0.* International Software Benchmarking Standards Group, 1–26.
ISO. (2023). The International Organization for Standardization. Retrieved January 10, 2023. https://www.iso.org/about-us.html.
ISO/IEC 14143:2002–2012 Information Technology – Software Measurement – Functional Size Measurement – Part 1–6. ISO/IEC, 1–6, 1–29, 1–23, 1–95, 1–27, 1–21.
ISO/IEC 14598:1998–2001 Software Engineering – Product Evaluation – Part 1–6. ISO/IEC, 1–19, 1–12, 1–16, 1–34, 1–35, 1–31.
ISO/IEC 15504:2003–2013 Information Technology – Process Assessment – Part 1–6. ISO/IEC, 1–19, 1–16, 1–54, 1–33, 1–196, 1–126.
ISO/IEC 19515:2019 Information Technology – Object Management Group Automated Function Points (AFP), 1.0. ISO/IEC, 1–28.
ISO/IEC 19761:2011 Software Engineering – COSMIC: A Functional Size Measurement Method. ISO/IEC, 1–14.
ISO/IEC 20926:2009 Software and Systems Engineering – Software Measurement – IFPUG Functional Size Measurement Method. ISO/IEC, 1–24.
ISO/IEC 20968:2002 Software Engineering – MkII Function Point Analysis – Counting Practices Manual. ISO/IEC, 1–93.
ISO/IEC 24570:2018 Software Engineering – NESMA Functional Size Measurement Method – Definitions and Counting Guidelines for the Application of Function Point Analysis. ISO/IEC, 1–70.
ISO/IEC 25000:2014 Systems and Software Engineering – Systems and Software Quality Requirements and Evaluation (Square) – Guide to Square. ISO/IEC, 1–27.
ISO/IEC 25001:2014 Systems and Software Engineering – Systems and Software Quality Requirements and Evaluation (Square) – Planning and Management. ISO/IEC, 1–13.
ISO/IEC 25010:2011 Systems and Software Engineering – Systems and Software Quality Requirements and Evaluation (Square) – System and Software Quality Models. ISO/IEC, 1–34.
ISO/IEC 25020:2019 Systems and Software Engineering – Systems and Software Quality Requirements and Evaluation (Square) – Quality Measurement Framework. ISO/IEC, 1–27.
ISO/IEC 25021:2012 Systems and Software Engineering – Systems and Software Quality Requirements and Evaluation (Square) – Quality Measure Elements. ISO/IEC, 1–37.
ISO/IEC 25023:2016 Systems and Software Engineering – Systems and Software Quality Requirements and Evaluation (Square) – Measurement of System and Software Product Quality. ISO/IEC, 1–45.
ISO/IEC 25030:2019 Systems and Software Engineering – Systems and Software Quality Requirements and Evaluation (Square) – Quality Requirements Framework. ISO/IEC, 1–46.
ISO/IEC 29881:2010 Information Technology – Systems and Software Engineering – FISMA 1.1 Functional Size Measurement Method. ISO/IEC, 1–16.
ISO/IEC/IEEE 12207:2017 Systems and Software Engineering – Software Life Cycle Processes. ISO/IEC/IEEE, 1–145.
ISO/IEC/IEEE 15288:2015 Systems and Software Engineering – System Life Cycle Processes. ISO/IEC/IEEE, 1–108.
ISO/IEC/IEEE 15939:2017 Systems and Software Engineering – Measurement Process. ISO/IEC/IEEE, 1–39.
ISO/IEC/IEEE 32430:2021 Software Engineering — Trial Use Standard for Software Non-Functional Sizing Measurements. ISO/IEC/IEEE, 1–76.
ISO/IEC/IEEE 90003:2018 Software Engineering – Guidelines for the Application of ISO 9001:2015 to Computer Software. ISO/IEC/IEEE, 1–69.
Johnson, J. (2021). *CHAOS 2020: Beyond Infinity.* The Standish Group International. Retrieved January 10, 2023. https://hennyportman.wordpress.com/2021/01/06/review-standish-group-chaos-2020-beyond-infinity/

Jørgensen, M. (2007). Forecasting of Software Development Work Effort: Evidence on Expert Judgment and Formal Models. *International Journal of Forecasting*, *23*(3), 449–462.

Kobylinski, A. (2005). *Modele jakości produktów i procesów programowych*. Oficyna Wydawnicza SGH, 258.

Larose, D.T., & Larose, C.D. (2015). *Data Mining and Predictive Analytics*. John Wiley & Sons, 824.

Linoff, G.S., & Berry, M.J.A. (2011). *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*. John Wiley & Sons, 896.

López-Martín, C. (2015). Predictive Accuracy Comparison between Neural Networks and Statistical Regression for Development Effort of Software Projects. *Applied Soft Computing, 27*, 434–449.

López-Martín, C., & Abran, A. (2015). Neural Networks for Predicting the Duration of New Software Projects. *Journal of Systems and Software, 101*, 127–135.

Lopez-Martin, C., Isaza, C., & Chavoya, A. (2012). Software Development Effort Prediction of Industrial Projects Applying a General Regression Neural Network. *Empirical Software Engineering, 17*(6), 738–756.

Mezouar, H., & El Afia, A. (2022). A Systematic Literature Review of Machine Learning Applications in Software Engineering. *Lecture Notes in Networks and Systems*, *489*, 317–331.

Minku, L.L., & Yao, X. (2013). Ensembles and Locality: Insight on Improving Software Effort Estimation. *Information and Software Technology, 55*(8), 1512–1528.

Mitchell, T.M. (1997). *Machine Learning*. McGraw-Hill, 421.

Mittas, N., Mamalikidis, I., & Angelis, L. (2015). A Framework for Comparing Multiple Cost Estimation Methods Using an Automated Visualization Toolkit. *Information and Software Technology, 57*, 310–328.

Moeyersoms, J., Junqué De Fortuny, E., Dejaeger, K., Baesens, B., & Martens, D. (2015). Comprehensible Software Fault and Effort Prediction: A Data Mining Approach. *Journal of Systems and Software, 100*, 80–90.

Myrtveit, I., & Stensrud, E. (2012). Validity and Reliability of Evaluation Procedures in Comparative Studies of Effort Prediction Models. *Empirical Software Engineering, 17*(1–2), 23–33.

Pai, D.R., McFall, K.S., & Subramanian, G.H. (2013). Software Effort Estimation Using a Neural Network Ensemble. *Journal of Computer Information Systems, 53*(4), 49–58.

Parthasarathy, M.A. (2007). *Practical Software Estimation: Function Point Methods for Insourced and Outsourced Projects*. Addison Wesley Professional, 426.

Port, D., & Korte, M. (2008). *Comparative Studies of the Model Evaluation Criterions MMRE and PRED in Software Cost Estimation Research*. Proceedings of the Second ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, 51–60.

Pospieszny, P., Czarnacka-Chrobot, B., & Kobylinski, A. (2015). Application of Function Points and Data Mining Techniques for Software Estimation – A Combined Approach. *Lecture Notes in Business Information Processing, 230*, 96–113.

Pospieszny, P., Czarnacka-Chrobot, B., & Kobylinski, A. (2018). An Effective Approach for Software Project Effort and Duration Estimation with Machine Learning Algorithms. *Journal of Systems and Software, 137*, 184–196.

Pujari, A.K. (2013). *Data Mining Techniques*. Orient Blackswan, 432.

Ruan, D., Chen, G., & Kerre, E.E. (2005). *Intelligent Data Mining: Techniques and Applications*, 5th ed. Springer Science & Business Media, 528.

Sage Technology. (2007). *Report on the SCUD Methodology Review*. Sage Technology, 1–7.

Sehra, S.K., Brar, Y.S., Kaur, N., & Sehra, S.S. (2017). Research Patterns and Trends in Software Effort Estimation. *Information and Software Technology, 91*(11), 1–21.

Shukla, R., Shukla, M., Misra, A.K., Marwala, T., & Clarke, W.A. (2012). Dynamic Software Maintenance Effort Estimation Modeling Using Neural Network, Rule Engine and Multi-regression Approach. ICCSA-2012. *Lecture Notes in Computer Science*, *7336*, 157–169.

Standish Group. (2018). *Project Resolution Benchmark Report*. The Standish Group International. Retrieved May 8, 2022. https://www.standishgroup.com/sample_research_files/DemoPRBR.pdf

Standish Group. (2022). The Standish Group International. Retrieved May 7, 2022. https://www.standishgroup.com/sample_research.

Stanek, M., & Czarnacka-Chrobot, B. (2023). Use of Artificial Neural Networks and Decision Trees for Cost Estimation of Software Projects – A Model Proposal. In: N. Thi Hong Nga, S. Agarwal, & E. Ziemba (eds.) *Analytics in Finance and Risk Management*. CRC Press, Taylor & Francis, (in print).

State Government of Victoria. (2000). *southernSCOPE. Reference Manual, Version 1.* Government of Victoria, Melbourne, 1–22.

Strike, K., El Emam, K., & Madhavji, N. (2001). Software Cost Estimation with Incomplete Data. *IEEE Transactions on Software Engineering, 27*, 890–908.

Wen, J., Li, S., Lin, Z., Hu, Y., & Huang, C. (2012). Systematic Literature Review of Machine Learning Based Software Development Effort Estimation Models. *Information and Software Technology*, *54*(1), 41–59.

17 The application of partial least squares structural equation modeling (PLS-SEM) algorithm to brand image and consumer loyalty at shoe shops

Cuong Tri Dam and Vijender Kumar Solanki

1 INTRODUCTION

Everyone views shoes to be a necessity, regardless of age, gender, or class. Shoes are part of the category of fashion products that also tends to consume a lot of resources. Moreover, shoes are not only for walking and protecting feet but also as fashion accessories, coordinated with other clothes and jewelry, increasing the aesthetics of each person. Therefore, practically each of us owns more than one pair of shoes. On average, each person will have at least two to three pairs of shoes. This number multiplies with a large customer base, which will open up enormous opportunities for shoe shops. Apart from the potential of the footwear business, however, many and even many other people are also trading this item, so they create competitive pressure to gain the market and get customers. On the other hand, purchasers in the present market hope to be happy with the item or brand they purchase; in any case, client satisfaction isn't sufficient to create a nonstop relationship with a brand. It is important to form a close-to-home bond beyond satisfaction to give buyer loyalty [1]. Brand loyalty is thoughtful as a fundamental factor for continuing in a competing situation and for the progress of any firm. It is an approach to staying with clients connected for a longer time and for acquiring a competitive advantage. When the same brand is purchased based on repeat purchases, brand loyalty is an effective approach for estimating the customer's propensity to repurchase a particular good and tends to have a larger market share. Due to customer devotion to a certain brand, many stores gain an advantage and capture

DOI: 10.1201/9780367854690-17

a significant market share. Brand loyalty implies clients are inclined toward one specific brand over the rivals' brands and also suggested others buy a similar item or brand [2]. Thus, shops need to create differentiation from competitors by offering products better than rivals to satisfy the needs of consumers. On the other hand, with the development and modernization of the retail field, the speed of progress of shoe shops has increased. Retailers need to keep close tabs on the brand image of their stores in clients' eyes to stay beneficial [3]. Likewise, a solid brand image assists with keeping the association in the psyche of shoppers. Client satisfaction was one of the primary elements of the shop; that results from the significant effect of the brand image and the products or services that they are giving to the clients. Thus, assuming the shop needs to compete with its rival, it ought to give quality items and services to make the client faithful to their items and services [4]. Furthermore, the shoe was a fundamental design extra for people, particularly the young ages. Picking a dependable, quality shoe shop was not easy for clients due to the many shoe stores [5]. Thus, making the right brand image to clients not similar to other shoe stores was crucial for senior footwear managers as clients routinely search for branded products to purchase [6]. Moreover, consumer satisfaction and faithfulness were the main concern for the achievement and greatness of the shop. Consumer loyalty had a positive and huge impact on creating client steadfastness. Client satisfaction was an issue that the shop is concerned about. With the end goal to meet consumer loyalty, shops should have the option to give the best quality items to clients [7]. Besides, shopper satisfaction as an intermediate variable is also ready to impact buyers' loyalty. The brand image additionally impacts purchasers' loyalty through consumer satisfaction as a mediating variable [8].

On the other hand, there are many studies on the influence of brand image on loyalty and satisfaction (e.g. [3, 8–10]). Only a small number of research, with customer satisfaction serving as an intermediary predictor at shoe stores, apply partial least squares structural equation modeling (PLS-SEM) algorithm for band image and loyalty. As a structural equation modeling technique, PLS-SEM enables the estimation of intricate cause-and-effect relationships in path models with latent variables. The numerous review studies that highlight the PLS-SEM algorithm's expanding application in a range of fields, including risk management, marketing, finance, etc., highlight its significance for both research and practice [11–14]. Therefore, this study applies the PLS-SEM algorithm to the connection between brand image and buyer loyalty with client satisfaction as an intermediate variable at shoe shops in Vietnam.

2 LITERATURE REVIEW

2.1 Consumer loyalty (CL)

CL is a select and/or continual client buy of the same brand or the same group of brands in a specific group of products/services [15]. Purchasers demonstrate loyalty by returning to the same shop and that is estimated by the proportion of the time spent at a solitary shop regardless of the offerings in competitive shops [16]. CL can be divided into two categories: one category is the clients' behavior, which is demonstrated by their preference for a brand, their repurchase of that brand, and their recommendation of that brand to others; the other category is the clients' attitude, which is CL's internal effect and perception component. Clients might show re-buy behavior because of restricted decision accessibility or inertia [17]. CL is firmly connected with the endurance and the more robust development of the organization. As a result, the profit ratio is more steady the higher the level of loyalty. It means vital to keep existing clients. Also, the size of the general market can be expanded by stimulating potential clients [18]. If a client who cherishes an individual's brand generally intends for the same brand when a person goes to a shop, an individual's loyalty can be viewed as high. This degree of loyalty is in the behavioral aspect and it is related to the quantity and frequency of the brand bought by the client [19]. Likewise, faithful clients often make repeat-purchases, which thus ensures a source of revenue for the organization, tend to purchase more, and are ready for greater expense, which will have an effect straightforwardly to the advantages of the organization [8].

2.2 Brand image (BI)

BI refers to the comprehensive understanding and emotional connection that consumers have with brands, playing a pivotal role in shaping their purchasing behavior [20]. It encapsulates the collective perception of a brand, which is cultivated through various brand interactions and experiences [21]. The dynamic interaction between brand cues and user interpretations is encompassed by BI, giving rise to a continuous cycle of communication. This process, driven by perception, originates from the evaluation of brand consumers and holds the potential to develop into a central determinant of purchasing decisions [22]. Moreover, a fundamental role is played by BI in shaping the distinctive identity and perceived value of a product [23]. For the customers to get a brand picture, they don't need to purchase an item or service, but they had an experience. BI could be created from the shop's impressions that buyers got from different resources connected with a brand [24]. BI similarly included customer feedback on the product's name, sign, impression, and picture. BI was similarly viewed as a collection of assets and liabilities connected to the name and symbol of the brand, which might raise or lower an enterprise's value [25]. BI isn't found in the specialized attributes of the original item. However, BI is something that brings out by advertisements, promotions, etc. At the point when a client utilizes the item and shows his/her opinions about the item before others. That is building a BI of that item. BI is something about an item that includes traits, design, packaging, and the attributes of that brand or item that makes them unique and extraordinary from different items [4]. BI was additionally demonstrated to be an insight and conviction remembered for clients' recollections as a portrayal of the association contained in clients' minds [26]. There is a quality that causes a BI to include internal features, such as the essential advantage provided and the beliefs, sentiments, and connections related to a company's brand. When it comes to the client's knowledge of the brand, BI tends to the exceptional circumstance of all impressions. A brand is a key tool for an organization to create a positive impression in the client's mind, which translates to creating a devoted client base and maintaining the company's market share. The brand-faithful client will purchase and suggest the brand to the following likely clients [27]. Previous research stated that BI has an antecedent of CL and it had positively affected CL [9, 10]. Moreover, an organization or item that has a good enough picture in the marketplace can be to get a better position in the marketplace and guarantee that they are ready to compete with rivals. Also, BI plays as the lifelike features that show what the brand offers and satisfied what needs of the customers. Some scientists confirmed that BI had a predictor of client satisfaction (CS) [7, 8].

Therefore, we provide the following hypotheses:

H1: There is a connection positively between BI and CL
H2: There is a connection positively between BI and CS

2.3 Client satisfaction (CS)

CS was from the item/service assessments as shown by the client experiences and the general assessment of the user experience [28]. CS is a feeling of delight or displeasure perceived by clients of the performing buying of products/services desired [21]. Satisfaction has been depicted as a shopper's post-buy assessment of products/services, given pre-buy assumptions [26]. CS is a huge component to appreciate satisfying clients about what they need. It is implied by the word "satisfying" that satisfaction either boosts enjoyment or lowers resentment or anxiety. The client's presumptions define what is satisfying. Also, if the perceived help execution doesn't meet client assumptions, the probable outcome is disappointment [29, 30]. One of the two approaches for separating satisfaction is as transaction-specific satisfaction [31] or as satisfaction with cumulation/post-utilization [15]. After the 1990s, many scientists consider satisfaction as clients' cumulation, after buying, and in the general assessment of buying behavior [32, 33]. CS is an essential factor in servicing delivery since understanding and fulfilling clients' needs can induce expanded market share from repeat buys and recommendations [34]. Overall client satisfaction came from the ability of the service to meet the client's expectations, presumptions, and needs in line with the service. CS is appraised exceptionally as a strategic purpose, as it influences directly the retention of clients and profits [33]. CS will influence repeat buy intentions in the future and the clients will impart their positive experiences to different customers [29]. Former studies stated that CS was a predictor of CL and influenced positively CL [21, 29, 35]. On the other hand, Simatupang and Purba [8] recommended that CS was a mediating factor in the relation of BI to CL; and the finding showed that BI affected positively CL through the CS mediating variable. Another study such as Putra et al. [21] suggested that at the shoe shop; however, the result disclosed that BI did not influence CL through CS.

Thus, we provide the following hypotheses:

H3: There is a connection positively between CS and CL
H4: CS as mediating variables in the linking between BI and CL

3 RESEARCH METHODOLOGY

3.1 Sample approach

For quantitative research, the non-probability method was used to gather the population sample. According to Fan et al. (2016) [36] in structural equation analysis, the minimum sample size is from 100 to 200 responses. In order to guarantee the validity of the study, 253 respondents were gathered from shoe shops in Ho Chi Minh City, Vietnam. The statistic of the sample features is shown in Table 17.1:

As presented in Table 17.1, the sample had 147 male customers (58.1%) and 106 female customers (41.9%), making male customers the majority (58.1%) of the sample. Regarding ages: 18–25 (31.2%), 26–40 (51.8%), 41–55 (12.7%), and >55 (4.3%); thus, the sample with ages from 18 to 40 accounts for the majority (83.0%).

3.2 Measurement

Applying a five-point Likert with one (totally opposed) and five (totally agreed) to measure items, the study applied four items of BI from [37], four items of CL, and four items of CS from [35].

3.3 Analytical approach

A model that is built between one or more endogenous variables and one or more exogenous variables can be tested using a set of statistical techniques known as structural equation modeling (SEM). Each endogenous and exogenous variable may be expressed as a latent variable or as a construct composed of some manifest or indicator variables. In general, a type of SEM based on covariance and variance is known as partial least squares (PLS). The PLS-SEM algorithm is a powerful and flexible analysis tool because it can be used with a wide range of data measurement scales (nominal, ordinal, interval, and ratio), can be applied with small sample sizes, and is independent of normal multivariate distribution [11, 38].

TABLE 17.1
Sample features

Characteristics	Classifications	Frequency	Percent
Gender	Male	147	58.1
	Female	106	41.9
	Total	253	100
Age	18–25	79	31.2
	26–40	131	51.8
	41–55	32	12.7
	>55	11	4.3
	Total	253	100

TABLE 17.2
PLS-SEM algorithm

Stage 1	**Iterative appraisal of loads and latent variable points**
	Starting at stage #4, rehash stages #1 to #4 until achieving convergence
#1	Inward loads (here acquired by utilizing the variable weighting plan)
	$V_{ji} = \begin{cases} \operatorname{cov}(Y_j; Y_i) \\ 0 \end{cases}$ if Y_j and Y_j are adjoins otherwise
#2	Inner estimate
	$\widetilde{Y}_j = \sum_i b_{ji} Y_i$
#3	Outer loads; resolve for
	$\widetilde{Y}_{jn} = \sum_{kj} \widetilde{W}_{kj} X_{kjn} + d_{jn}$ in a Mode A block
	$\widetilde{X}_{kjn} = \widetilde{W}_{kj} \widetilde{Y}_{jn} + e_{kjn}$ in a Mode B block
#4	External estimate
	$Y_{jn} = \sum_{kj} \widetilde{W}_{kj} X_{kjn}$
Stage 2	Estimate of external loads, outer loadings, and path coefficients
Stage 3	Estimate of location parameters

The PLS-SEM method is being utilized more frequently in business research [39]. PLS-SEM is crucial for research and practice, as shown by the numerous review studies that show its expanding application across a range of disciplines, including risk management (e.g. [12]), marketing (e.g. [13]), finance (e.g. [14]), etc. In path models with latent variables, the PLS-SEM algorithm, an SEM method, permits the estimation of complex cause-and-effect connections [12]. A novel statistical tool for the analysis process is the PLS-SEM method [40]. The PLS-SEM algorithm was presented by Lohmoller [41] (see Table 17.2).

4 RESULT AND DISCUSSION

4.1 Partial least squares structural equation modeling (PLS-SEM) algorithm

4.1.1 Reliability and validity

Applying the PLS-SEM algorithm with SmartPLS software showed that α (Cronbach's alpha) & CR (Composite reliability) values of the predictors were above 0.70. Hence, the consistent reliability of these factors was supported. The outer loadings and AVE (average variance extracted) were over 0.50. Thus, the convergent validity of these factors was good (see Table 17.3).

Moreover, using the Fornell–Larcker criteria [42] to estimate the discriminant validity. Table 17.4 pointed out that AVE's square root values (in bold – from 0.758–0.831) were higher than the correlation between each concept with any other concepts. So, these findings support the discriminant validity of the factors.

4.1.2 Hypotheses testing

PLS-SEM algorithm for hypotheses testing, as demonstrated in Table 17.5, the R^2 coefficient of the general model was 0.526, below 0.67, which is considered a moderate impact [43]. As such, BI and CS explained a 52.6% variance of CL; we saw BI has a more critical effect (0.476) than CS (0.408). Besides, BI depicted an 11.6% difference in CS. Likewise, Table 17.5 demonstrated the assumption testing's outcomes. To estimate parameters, we applied the 5,000-re-sampling bootstrapping technique. The outcomes showed that H1, H2, H3, and H4 had correspondent path coefficients of 0.479, 0.340, 0.408, and 0.139 in a 95% confidence interval (CI). That means all parameter measures of these relationships were statistically valued at the 5% level. Thus, these assumptions were verified.

4.2 Discussion

This study has analyzed the BI's effect on CL with the mediating role of CS at shoe shops in Ho Chi Minh, Vietnam, using the PLS-SEM algorithm. The result

TABLE 17.3
The outcomes of factors

Constructs and measurement scales	Outer loadings	α	CR	AVE
Brand Image		0.849	0.899	0.690
1. IMAGE1	0.827			
2. IMAGE2	0.777			
3. IMAGE2	0.805			
4. IMAGE2	0.909			
Customer Loyalty		0.746	0.840	0.574
1. LOY1	0.794			
2. LOY2	0.774			
3. LOY3	0.871			
4. LOY4	0.556			
Customer Satisfaction		0.787	0.860	0.606
1. SAT1	0.686			
2. SAT2	0.817			
3. SAT3	0.793			
4. SAT4	0.810			

TABLE 17.4
Discriminant validity

	BI	CL	CS
BI	0.831		
CL	0.615	0.758	
CS	0.340	0.570	0.778

TABLE 17.5
Direct & indirect effect results

Paths	Hypotheses	Path coefficients	95% CI	Factor	R^2
BI → CL	H1	0.476	[0.364–0.577]	CS	0.116
BI → CS	H2	0.340	[0.233–0.449]		
CS → CL	H3	0.408	[0.315–0.500]	CL	0.526
BI → CS → CL	H4	0.139	[0.083–0.208]		

of this investigation affirmed that the proposed four hypotheses in the study were acknowledged.

The examination finding supported that BI has measurably significant, and connected positively with CL. The BI was a predecessor of CL. The previous review affirmed this examination [23, 24]. Additionally, the finding likewise reinforced that BI has genuinely critical, and emphatically connected with CS. The BI was an antecedent of CS. This study was confirmed by previous research [7, 8].

Besides, the results additionally affirmed that CS has genuinely significant, and decidedly affected CL. CS was a forerunner of CL. The earlier review affirmed this exploration [16, 26, 32].

Finally, the outcomes also checked that CS was a mediator variable in the connection between BI and CL. This outcome was illustrated by the early study [8].

5 CONCLUSION AND LIMITATIONS

This assessment scrutinized the PLS-SEM algorithm in the link between BI and CL with CS as an intermediate variable at shoe shops in Vietnam. The primary contribution of the study is to offer a theoretical framework for brand image and loyalty, using the PLS-SEM algorithm to mediate customer satisfaction for shoe businesses in the Vietnamese market. Moreover, the findings of this research can aid shoe shop administrators in giving advertising plans that work on clients' perspectives toward BI and CS to upgrade CL.

This assessment helps shoe-shopping managers comprehend the significance of BI on CS and CL. Like this, senior footwear supervisors ought to lay out a fair association of BI in the client's memory. In case clients perceive a good BI, they will satisfy and unwavering ness while shopping at the shoe store.

The assessment results in like manner show that CS is a crucial factor affecting CL. Similarly, CS takes an intermediate into the association between BI and CL; this helps managers see the significance of CS to CL. Therefore, senior shoe managers sought to create and execute promoting tasks to build CS through clients' experience and address customers' shopping issues, prompting extended CL.

This examination has a few limitations. This exploration clarifies 52.6% of CL's variance by two factors (BI and CS). Along these lines, future assessments should add components to boost descriptions for variance in CL. This concentration additionally can't be generalizable, so the forthcoming ought to be analyzed in different enterprises.

REFERENCES

1. Unal, S., Aydın, H.: An Investigation on the Evaluation of the Factors Affecting Brand Love. *Procedia - Soc. Behav. Sci.* 92, 76–85 (2013). https://doi.org/10.1016/j.sbspro.2013.08.640.
2. Iqbal, U., Rizwan, M., Zafar, A., Khan, H., Usman, M., Iqbal, D.-S.: Determinants Uncovering the Brand Loyalty: A Signaling Effect of Price on Quality Perception. *J. Basic Appl. Sci. Res.* 3, 212–221 (2013).
3. Silva, T.S., Giraldi, J.d.M.E.: The Influence of Store Image on Customer Satisfaction: A Case Study of a Shoe Store. *Brazilian Bus. Rev.* 7, 60–77 (2010). https://doi.org/10.15728/bbr.2010.7.2.4.
4. Abbas, U., Islam, K.M.A., Hussain, S., Baqir, M., Muhammad, N.: Impact of Brand Image on Customer Loyalty with the Mediating Role of Customer Satisfaction and Brand Awareness. *Int. J. Mark. Res. Innov.* 5, 1–15 (2021). https://doi.org/10.46281/ijmri.v5i1.987.
5. DNVNTC: Top 10 Most Beautiful and Quality Footwear Stores in Ho Chi Minh City (Vietnamese). https://dvt.vn/doanh-nghiep/top-10-cua-hang-giay-dep-dep-va-chat-luong-nhat-tp-hcm.htm, last accessed 2019/06/12.
6. Neupane, R.: The Effects of Brand Image on Customer Satisfaction and Loyalty Intention in Retail Super Market Chain UK. *Int. J. Soc. Sci. Manag.* 2, 9–26 (2015). https://doi.org/10.3126/ijssm.v2i1.11814.
7. Diputra, I.G.A.W., Yasa, N.N.: The Influence of Product Quality, Brand Image, Brand Trust on Customer Satisfaction and Loyalty. *Am. Int. J. Bus. Manag.* 4, 25–34 (2021).
8. Simatupang, P., Purba, F.: The Brand Image and Its Effect on Consumer Loyalty and Satisfaction as a Variable Intervening of Aqua Mineral Water Product (Study on Undergraduate Student of Management Study Program, Universitas Simalungun). *Budapest Int. Res. Critics Inst.* 3, 1902–1910 (2020).
9. Azizan, N.S., Yusr, M.M.: The Influence of Customer Satisfaction, Brand Trust, and Brand Image towards Customer Loyalty. *Int. J. Entrep. Manag. Pract.* 2, 93–108 (2019). https://doi.org/10.35631/ijemp.270010.
10. Durmaz, Y., Cavusoglu, S., Ozer, O.: The Effect of Brand Image and Brand Benefit on Customer Loyalty: The Case of Turkey. *Int. J. Acad. Res. Bus. Soc. Sci.* 8, 524–536 (2018). https://doi.org/10.6007/ijarbss/v8-i5/4140.
11. Tohari, A., Faisol, F., Rahmad, A.: A the Use of Partial Least Squares Modeling in Finance Business Partnering Research. *J. Ilm. Kursor.* 11 (2021). https://doi.org/10.21107/kursor.v11i1.256.
12. Kassem, M.A.: Risk Management Assessment in Oil and Gas Construction Projects Using Structural Equation Modeling (PLS-SEM). *Gases.* 2, 33–60 (2022). https://doi.org/10.3390/gases2020003.

13. Amarullah, D., Handriana, T.: Ewom Credibility, Trust, Perceived Risk, and Purchase Intention in the Context of E-Commerce: Moderating Role of Online Intention in the Context of E-Commerce: Moderating Role. *J. Ekon. Bisnis dan Kewirausahaan*. 11, 61–83 (2022).
14. Jain, N., Raman, T. V.: A Partial Least Squares Approach to Digital Finance Adoption. *J. Financ. Serv. Mark*. 27, 308–321 (2022). https://doi.org/10.1057/s41264-021-00127-8.
15. Oliver, R.L.: *Satisfaction: A Behavioral Perspective on the Consumer*. The McGraw-Hill Companies, Inc., New York (1997).
16. De Wulf, K., Odekerken-Schröder, G.: Assessing the Impact of a Retailer's Relationship Efforts on Consumer's Attitudes and Behavior. *J. Retail. Consum. Serv.* 10, 95–108 (2003). https://doi.org/10.1016/S0969-6989(02)00013-9.
17. Bloemer, J.M.M., Kasper, H.D.P.: The Complex Relationship between Consumer Satisfaction and Brand Loyalty. *J. Econ. Psychol.* 16, 311–329 (1995). https://doi.org/10.1016/0167-4870(95)00007-B.
18. Kim, M.K., Park, M.C., Jeong, D.H.: The Effects of Customer Satisfaction and Switching Barrier on Customer Loyalty in Korean Mobile Telecommunication Services. *Telecomm. Policy*. 28, 145–159 (2004). https://doi.org/10.1016/j.telpol.2003.12.003.
19. Odin, Y., Odin, N., Valette-Florence, P.: Conceptual and Operational Aspects of Brand Loyalty: An Empirical Investigation. *J. Bus. Res.* 53, 75–84 (2001). https://doi.org/10.1016/S0148-2963(99)00076-4.
20. Zhang, Y.: The Impact of Brand Image on Consumer Behavior: A Literature Review. *Open J. Bus. Manag.* 03, 58–62 (2015). https://doi.org/10.4236/ojbm.2015.31006.
21. Putra, R., Yeni, F., Fitri, H., Hutapea, N.: The Role of Consumer Satisfaction in Mediating Service Quality and Brand Image on Consumer Loyalty. *Sentralisasi*. 9, 108 (2020). https://doi.org/10.33506/sl.v9i2.970.
22. Wu, C.S., Chen, T.-T.: Building Brand's Value: Research on Brand Image, Personality and Identification. *Int. J. Manag. Econ. Soc. Sci.* 8, 299–318 (2019). https://doi.org/10.32327/IJMESS.8.4.2019.19.
23. Meenaghan, T.: The Role of Advertising in Brand Image Development. *J. Prod. Brand Manag.* 4, 23–34 (1995).
24. Keller, K.L.: Conceptualizing, Measuring, and Managing Customer-Based Brand Equity. *J. Mark.* 57, 1–22 (1993). https://doi.org/10.2307/1252054.
25. Magid, J.M., Cox, A.D., Cox, D.S.: Quantifying Brand Image: Empirical Evidence of Trademark Dilution. *Am. Bus. Law J.* 43, 1–42 (2006). https://doi.org/10.1111/j.1744-1714.2006.00010.x.
26. Kotler, P., Keller, K.L.: *Marketing Management*. Pearson Education, Inc., Upper Saddle River, NJ (2016).
27. Septiani, D., Chaerudin, R.: The Effect of Customers' Price Perception, Perceived Quality and Brand Image Toward Purchasing Intention in Bandung Local Shoe Brand. *KnE Soc. Sci.* 2020, 1242–1254 (2020). https://doi.org/10.18502/kss.v4i6.6674.
28. Joewono, T.B., Kubota, H.: User Satisfaction with Paratransit in Competition with Motorization in Indonesia: Anticipation of Future Implications. *Transportation (Amst).* 34, 337–354 (2007). https://doi.org/10.1007/s11116-007-9119-7.
29. Li, J.: Factors Affecting Customer Satisfaction and Customer Loyalty towards Belle Footwear Company in Lanzhou City, Gansu Province of the People's Republic of China. *IOSR J. Bus. Manag.* 14, 41–48 (2013). https://doi.org/10.9790/487x-1424148.
30. Oliver, R.L.: A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *J. Mark. Res.* 17, 460–469 (1980). https://doi.org/10.2307/3150499.
31. Olsen, L.L., Johnson, M.D.: Service Equity, Satisfaction, and Loyalty: From Transaction-Specific to Cumulative Evaluations. *J. Serv. Res.* 5, 184–195 (2003). https://doi.org/10.1177/1094670502238914.

32. Tu, Y.-T., Li, M.-L., Chih, H.-C.: An Empirical Study of Corporate Brand Image, Customer Perceived Value and Satisfaction on Loyalty in Shoe Industry. *J. Econ. Behav. Stud.* 5, 469–483 (2013).
33. Jones, T.O., Sasser, W.E.: Why Satisfied Customer Defect. *Harv. Bus. Rev.* 73, 88–99 (1995).
34. Barsky, J.D.: Customer Satisfaction in the Hotel Industry: Meaning and Measurement. *J. Hosp. Tour. Res.* 16, 51–73 (1992). https://doi.org/10.1177/109634809201600105.
35. Leninkumar, V.: The Relationship between Customer Satisfaction and Customer Trust on Customer Loyalty. *Int. J. Acad. Res. Bus. Soc. Sci.* 7, 450–465 (2017). https://doi.org/10.6007/ijarbss/v7-i4/2821.
36. Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S.R., Park, H., Shao, C.: Applications of Structural Equation Modeling (SEM) in Ecological Studies: An Updated Review. *Ecol. Process.* 5, 5–19 (2016). https://doi.org/10.1186/s13717-016-0063-3.
37. Lien, C.H., Wen, M.J., Huang, L.C., Wu, K.L.: Online Hotel Booking: The Effects of Brand Image, Price, Trust and Value on Purchase Intentions. *Asia Pacific Manag. Rev.* 20, 210–218 (2015). https://doi.org/10.1016/j.apmrv.2015.03.005.
38. Hair, J.F., Black, W.c., Babin, B.J., Anderson, R.E.: *Multivariate Data Analysis*. Pearson Education Limited, London (2014).
39. Sarstedt, M., Ringle, C.M., Henseler, J., Hair, J.F.: On the Emancipation of PLS-SEM: A Commentary on Rigdon (2012). *Long Range Plann.* 47, 154–160 (2014). https://doi.org/10.1016/j.lrp.2014.02.007.
40. Henseler, J., Hubona, G., Ray, P.A.: Using PLS Path Modeling in New Technology Research: Updated Guidelines. *Ind. Manag. Data Syst.* 116, 2–20 (2016). https://doi.org/10.1108/IMDS-09-2015-0382.
41. Lohmoller, J.-B.: *Latent Variable Path Modeling with Partial Least Squares*. Springer-Verlag, Heidelberg, Berlin (1989). https://doi.org/10.1007/978-3-642-52512-4.
42. Fornell, C., Larcker, D.F.: Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *J. Mark. Res.* 18, 39–50 (1981).
43. Garson, G.D.: *Partial Least Squares: Regression & Structural Equation Models*. Statistical Associates Publishers, Asheboro, NC (2016).

18 Effect of the general government fiscal deficit on the inflation rate

OECD countries with the upper middle income

Semra Aydoğdu Bağci

1 INTRODUCTION

The phenomenon of inflation is a dilemma that is difficult to be solved in many countries, especially in developing countries. This study examines the effect of fiscal deficit (FD) and GDP on IR. In this study, the effects of the FD, GDP, and expansionary money supply (MS) on the IR are investigated by utilizing the obtained data over the period 1990–2020 for Colombia, Costa Rica, Mexico, and Turkey, which are among the upper middle-income countries of the OECD country group. The objective of the research is to examine whether the FD, especially those upper middle-income level countries in the OECD country group have a problem, leading to IR. Many academic studies researched the effect of the IR on the FD, but there has been a lack of academic studies researching the effect of the FD on the IR. None of them studies the effect of the FD on the IR for the upper middle-income countries of the OECD country group. Therefore, this study is relevant. In the first section of the manuscript, theoretical views on the FD and empirical literature are introduced. In the second section, the IR and FD development in those countries are mentioned. In the third section, such an effect is analyzed econometrically.

2 LITERATURE REVIEW

In this part, the theoretical and empirical literature is introduced. In theoretical literature, FDs are discussed in terms of the Classical, Monetarist, Keynesian, Neoclassical, Ricardian, and New Keynesian views. In the part on the empirical literature, the studies and their findings are also included.

2.1 Theoretical literature

According to the standard view, the substitution of the FD with current taxation causes an increase in aggregate consumer demand. Meaning that national savings

DOI: 10.1201/9780367854690-18

decrease as private savings increase less than tax deductions. In a closed economy, this leads to a rise in the expected real interest rate (RIR) to improve the equity between demand for investment and national savings. Investments are crowded out by an increase in the RIR and production capital is reduced in the long run. Thus, Modigliani (1986) perceived public debt as the reason for future generations to attain less capital. As Martin Feldstein (1974) mentioned, social security programs enhance aggregate demand for goods, thus, causing a rise in the RIR and a decline in the productive capital stock. In autarkic economies, the host country's substitution of the FD with current taxes causes foreign borrowing. Thus, the FD causes the current account deficit (CAD) to increase. The FD excludes domestic investment in the host country, causing a decrease in domestic investment, and a decline in the capital stock in the long run. The rise in the CAD also reduces social welfare in the long run and leads foreigners to complain about borrowing (Barro, 1989: 37–38).

Sargent and Wallace (1981) stated that in the first case where monetary policy gained predominance over fiscal policy, for instance, the monetary authority acted independently of monetary policy in declaring the growth rate of the current and future period of the base money. In doing so, monetary policy determines the level of revenue supplied by the fiscal authority, which is subject to the constraint determined by the demand for bonds, offsetting an FD using a combination of the monetary authority's seigniorage revenues and bond sales to the public sector. The monetary authority constantly controls IR in the monetary economy since it is free to determine the base money. In the 2nd case, where fiscal policy gained predominance over monetary policy, the fiscal authority constituted its budget independently by declaring all current and future FDs and surpluses which, in turn, determined the amount of revenue that needed to be increased through the sale of bonds and seigniorage. In this case, the monetary authority is subject to the constraint determined by the demand for public bonds, and such demand is crucial for the monetary authority to determine whether or not it should constantly have control over IR. In particular, upon demand for public bonds, the interest rate of the bonds is thought to be higher than the growth rate. Once the fiscal authority has a deficit in the budget, the monetary authority can control neither the economic growth that constitutes the monetary base nor the IR. Although the monetary authority can constantly control IR, in the 2nd case the monetary authority is weaker than in the 1st one. In the 2nd case, the fiscal authority's FD cannot be offset solely by the sale of new bonds since the monetary authority would have difficulty in creating money and tolerating additional inflation (Sargent and Wallace, 1981: 1–2). Friedman (1948) stated that if the prices of the final products were not set correctly and the factors of production were inelastic, the condition of full employment could not be met. It was also stated that cyclical fluctuations in output and employment could not be fully resolved. Leaving the authority to control the amount of money and supply money to offset FDs to the government may lead to irresponsible governmental behavior and inflation. Therefore, the state's control over the amount of money should be eliminated (Friedman, 1948: 263–264).

Instead of assuming the deficit budget was wrong, Keynes advocated the fiscal policies implemented in accordance with the economic fluctuations against the conjuncture. For instance, Keynesian economists advocated deficit-based public expenditures on labor-intensive infrastructure projects throughout a recession. Keynesian economists increased taxes and prevented inflation to alleviate the economy when demand-side growth was high. Monetary policy practices tend to enhance investment, for instance, by lowering interest rates to stimulate the economy. The exceptional case arises with the liquidity trap leading to an increase in output and employment. Keynes argued that the government was effective in solving problems in the short run. The theory of Keynesian economists became predominant throughout the period following WWII until the 1970s. Keynesian economists could not find any solution to the phenomenon of stagflation, in which slow growth was accompanied by high levels of IR (Jahan et al., 2014: 2).

According to the Neoclassical view, forward-thinking individuals tend to plan their consumption throughout their life span. FDs are increased by imposing taxes on future generations and making lifetime total consumption. Upon utilizing economic resources for full employment, enhanced consumption compulsorily reduces savings. Thus, the interest rate should be increased to fulfill the capital market equilibrium. Persistent FDs reduce private capital accumulation via the crowding-out effect (Bernheim, 1989: 55).

According to the Ricardian view, overlapping generations are altruistic through the voluntary transfer of resources. Under certain conditions, consumption becomes a function of the total resources of taxpayers and future generations of taxpayers. With future generations paying taxes, the FD would be devolved on future generations without affecting the total resources (Bernheim, 1989: 56).

New Keynesian Economics Woodford mentioned that the public budget deficit led to public borrowing. It was found that the unexpected rise in the public's primary budget deficit increased IR, real GDP, and the NIR. The capital loss of the nominal public debt due to the rise of the IR is not sufficient to prevent the current public debt from increasing at the beginning of the 1st year. In the coming years, real public debt would become stagnant, nonetheless, with the rise in seigniorage revenues and the decline in the RIR, the public debt may be rolled over (Woodford, 1996: 18).

2.2 Empirical literature

The empirical literature is sorted from the most recent to the most outdated. Upon examining the empirical literature, some of the studies (Ahmad and Aworinde, 2019; Danlami et al., 2019; Duodu et al., 2022; Durguti et al., 2020; Eita et al., 2021; Jalil et al., 2014; Lin and Chu, 2013; Maraş and Dumrul, 2019; Myovella and Kisava, 2017; Ssebulime and Edward, 2019) asserted that FD positive affected IR; some studies (Ezeabasılı et al., 2012; Güneş, 2020) found negative impacts in that regard; and some other studies concluded that the FDs did not affect IR (Karadeniz, 2021; Olaniyi, 2020; Tiwari et al., 2012). In some studies (Catao and Terrones, 2005; Doğan and Günel, 2021; Kaur, 2021; Olubiyi and Bolarinwa, 2018) FDs were detected to affect IR, either positively or negatively, according to the maturity date, country, or analysis methods used.

Catao and Terrones (2005) examined the impact of FD on IR for 107 developed and developing countries using the MG and PMG estimators over the period 1960–2001. IR was chosen as the dependent variable, whereas the MS/GDP, the central

FD/GDP, the trade deficit, and petroleum prices were used as independent variables. According to the MG estimator result, it was determined that the FD positively affected the IR in all countries and country groups. According to the PMG estimator, FD had a negative impact on the IR in developing countries, whereas a positive influence in developing countries.

Tiwari et al. (2012) examined the factors affecting the FD, IR, MS, and public expenditures in India over the periods 1970–1971 and 2008–2009 by conducting the Granger Causality analysis, standard Granger Causality analysis, and VAR analysis developed by Dolado and Lütkepohl (DL) (1996). As a result of the Granger Causality analysis developed by DL; unilateral causal relationships running from public expenditures and MS to FD were found. As a result of the standard Granger causality analysis, unilateral causal relationships running from the 1st difference in the FD to the 1st difference in the MS, from the 1st difference in public expenditures to the 1st difference in the FD, and from the 1st difference in the MS to the 1st difference in the FD were detected. The VAR analysis revealed an inverse causal relationship from the 2nd lag of the FD to the MS was found. There was a positive causality running from the 1st lag of public expenditures to the FD and public expenditures, whereas a negative causality from public expenditures to the 2nd lag of MS existed. A positive causal relationship from the 1st lag of the IR to the IR was found. Negative causal relationships from the 1st lag of the MS to public spending and from the 2nd IR of the MS to the MS were detected. It was determined that the FD did not affect the IR.

Ezeabasılı, Mojekwu, and Herbert (2012) explicated the impact of the FD on IR in Nigeria over the years 1970–2006 by conducting a cointegration analysis. IR was used as the dependent variable; whereas FD, MS, GDP, depreciation of the national currency, real exchange rate (RER), and the last period's IR were used as the independent variables. Although the 1st lag of the FD and the 1st lag of depreciation of the national currency negatively affected the IR; the 1st and 2nd lags in the MS enhanced the IR. Although the rise in GDP decreases the IR, the 3rd lag of GDP increased the IR. A rise in the 1st lag of the IR increased the IR.

Lin and Chu (2013) analyzed the impact of FD on IR in 91 countries over the years 1960–2006 by employing the Dynamic Panel Quantile Regression (DPQR) method. IR was chosen as the dependent variable; whereas the FD/MS, the FD/GDP, the growth rate of MS, IR of gasoline prices, and trade deficits were used as the independent variables. It was asserted that the FD caused inflation. It was determined that such impact was higher in countries with high and moderate IR, whereas lower in countries with low IR.

Jalil, Tariq, and Bibi (2014) investigated the impacts of FDs, interest rates, trade openness, RER, petroleum prices, public sector borrowings, private sector borrowings, real money demand for real MS, wheat prices, import price index, and the lag in IR on IR in Pakistan over the period 1972–2012 by employing the ARDL bounds approach. It was detected that the FD had a positive influence on the IR, and the budget imbalance caused inflation.

Myovella and Kisava (2017) investigated the impact of the FD on IR over the period 1970–2015 using the ARDL Bounds Test for Tanzania. It was detected that the FD positively affected IR in the long run.

Olubiyi and Bolarinwa (2018) investigated the impact of the FD on IR over the period 1994–2015 by performing the ARDL Boundary Approach test for Nigeria, South Africa, Mali, Kenya, and Egypt. IR was used as the dependent variable, and the 1st lag of the IR, the RER, the 1st lag of the RER, the expansionary MS, the foreign debts, and the 1st lag of the foreign debts were used as independent variables. External debts increased the IR for South Africa, Mali, and Nigeria, whereas foreign debts decreased the IR for Egypt. The 1st lag of external debt increased the IR in Kenya. The 1st lag of the IR negatively affected the IR in Egypt and Nigeria, whereas the 1st lag of the IR in South Africa and Mali positively affected the IR in the short run. Expansionary MS negatively affected the IR in Nigeria; whereas it had a positive impact in Mali, Kenya, and Egypt. The RER positively affected the IR in Kenya, Nigeria, and Egypt. In the short run, the 1st lag of the RER increased the IR in Nigeria. In the long run; although foreign debts increased the IR in Nigeria, external debt negatively affected the IR in Kenya and Egypt. In the long run, expansionary MS positively affected the IR in all countries. The RER increased the IR in Nigeria, South Africa, Mali, and Kenya.

Danlami et al. (2019) investigated the influence of FD on IR in Nigeria over the years 1970–2016 performing the ARDL bounds test approach. The 1st lag of IR, GDP, RER, and FD were chosen as dependent variables; whereas IR was chosen as the independent variable. At the end of the analysis, the rise in the 1st lag of IR and the rise in the FD increased the IR in the short run, whereas the rise in GDP and RER reduced the IR. The rise in the FD increased the IR; whereas the rise in GDP and RER reduced the IR.

Maraş and Dumrul (2019) analyzed the association between FD and IR for Turkey by performing the ARDL bounds test approach with the monthly data obtained between 2006:01 and 2018:10. IR was used as the dependent variable; whereas M3, the TL/USD rate, and FD were used as the independent variables. It was revealed that the budget balance caused a decline in the IR in the long run, and therefore, the IR was increased with the rise in the FDs. The MS and the TL/USD rate negatively affected the IR.

Ssebulime and Edward (2019) estimated the impact of the FD on the IR in Uganda over the period 1980–2016 performing a cointegration analysis. IR was used as the dependent variable, whereas the change in the MS, trade balance, GDP change rate, and NIR were used as the independent variables. The 1st difference of the FD, the 2nd difference of the MS, the trade balance, and the 2nd difference of the trade balance were detected to have positive impacts on the IR. The rate of change in GDP negatively affected the IR.

Ahmad and Aworinde (2019) examined whether or not the FD caused IR in 12 African countries utilizing the quarterly data obtained over the period 1980–2018 by employing the TAR and the M-TAR. It was found that the FD had a positive impact on the IR.

Olaniyi (2020) explicated the relationship between FD and IR in Nigeria using the quarterly data obtained over the period 1981:Q1–2016:Q4 by conducting both symmetrical and asymmetrical causality analyses developed by Hatemi-J and El-Khatib (2016). As a result of the analysis, neither symmetrical nor asymmetrical causality relationships were detected.

Güneş (2020) examined the impact of FDs on IR for 28 OECD countries over the period 1995–2018 by conducting Vector Autoregressive (VAR) and Panel Granger Causality Analyses. Although a causal relationship running from IR to FD was determined; no causality from FD to IR was detected. The VAR analysis revealed that the FD had a negative impact on the IR.

Durguti et al. (2020) explicated the impact of FD on IR in six Western Balkan countries over the period 2001–20017 by employing the VECM. IR was the dependent variable; whereas FD, the government debt/GDP, RER, and unemployment rate were used as independent variables. Although it was detected that the FD and public debts increased the IR; it was detected that the RER and the unemployment rate decreased the IR.

Doğan and Günel (2021) investigated the impacts of FDs on IR in eight Balkan countries employing the panel ARDL method and using the data obtained between 1999 and 2019. It was detected that the FD increased IR in the short run, whereas the FDs and the 4th difference of the FD negatively affected the IR.

Kaur (2021) examined the effect of the FD on the IR in India by employing the ARDL boundary approach method and using the monthly data obtained over the period 1996–1997 and 2016–2017. IR was the dependent variable; whereas the total budget deficit/GDP, M3, RER, gasoline (energy) index, and GDP were used as independent variables. The FD and M3 positively affected the IR; whereas the RER and gasoline (energy) index negatively affected the IR. In the short term; although the lag of the IR and the 5th lag of the interest rate as well as the 5th, 6th, and 7th lags of the FD negatively affected the IR; the M3 positively affected the IR. In the short run, the gasoline (energy) index negatively affected the IR, whereas the 1st lag of the gasoline (energy) index positively affected the IR.

Karadeniz (2021) measured the impact of FD on CAD, economic growth, and IR in 14 developing countries over the period 1994–2019 employing the Panel Average Group Estimator (AMG) method. In the 1st model, the CAD was used as the dependent variable, whereas the FD and growth rate were the independent variables. In the 2nd model, the growth rate was used as the dependent variable, while the FD and IR were used as the independent variables. In the 3rd model, the IR was the dependent variable, whereas the FD and MS were the independent variables. It was detected that the rise in the FD and growth rate in the 1st model increased the CAD; in the 2nd model, the decrease in the FD and the IR increased the growth rate; and in the 3rd model, the rise in the MS increased the IR. In the 3rd model, it was determined that the FD did not have a statistically significant impact on IR.

Eita et al. (2021) examined the effect of FD on IR in Namibia over the period 2002:Q2–2017:Q2 by employing the ARDL and Granger causality methods. Namibia's IR was the dependent variable; whereas the FD/GDP, IR, and interest rate of South Africa were used as independent variables. The rise in the FD and South Africa's IR increased Namibia's IR. The 1st lag in Namibia's IR, South Africa's IR, and the rise in the FD increased the IR. However, a causality from FD to IR was detected, but no causality from IR to FD could be found.

Duodu et al. (2022) investigated the influence of FD on IR with the Granger causality analysis and VECM over the periods 1999:Q1–2019:Q4 in Ghana. IR was the dependent variable; while the rate of change in MS, foreign trade balance, the FD/IR, GDP, and nominal exchange rates were used as independent variables. Neither the FD caused IR nor IR caused FD. According to the VECM result, the FD led to a rise in IR.

3 DEVELOPMENT OF IRS AND FDS IN OECD COUNTRIES WITH UPPER MIDDLE INCOMES

The high IRs and high FDs in these countries account for considering the upper middle-income countries for the study sample.

Figure 18.1 illustrates the IR development of Colombia, Costa Rica, Mexico, and Turkey over the period 1990–2020.

Upon examining Figure 18.1; it is seen that the countries with the highest IRs are Turkey Mexico, Colombia, and Costa Rica, respectively. Inflation in Turkey was on the rise over the period 1990–1992, it decreased in 1993, and climaxed as of 1994. The IR, which generally decreased over the period 1995–2005, has been stable since 2006. The IR in Mexico, which decreased over the period 1990–1994, climaxed in 1995 and declined over the period 1996–2007, and has, in general, remained stable since 2008. The IR in Colombia, which has been on the rise since 1990, climaxed in 1991. The IR, which fell over the period 1992–2007, increased in 2008 and has followed a stable trend since 2009. In Costa Rica, the IR increased over the period 1990–1991 and climaxed in 1991. The IR, which fell until 1993, increased until 1995, and followed a downward trend from 1996 onwards, except for the periods 2004–2005 and 2007–2008.

Figure 18.2 illustrates the development of FD.

It is seen that the countries with FDs are Turkey, Colombia, Costa Rica, and Mexico, respectively. For Turkey, the FD was high during the 1990–2001 period, this deficit decreased over the period 2002–2008, the FD increased during the 2009–2010 period, the FD decreased over the period 2011–2016, and the FD, in general,

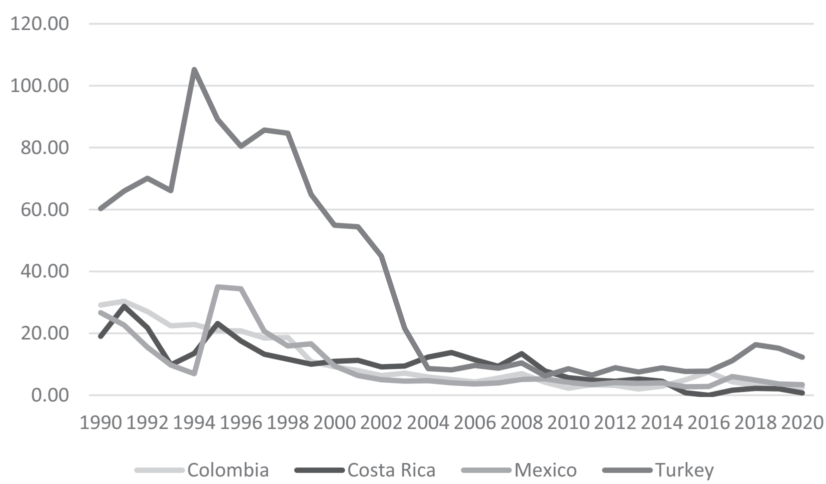


FIGURE 18.1 Development of IR.

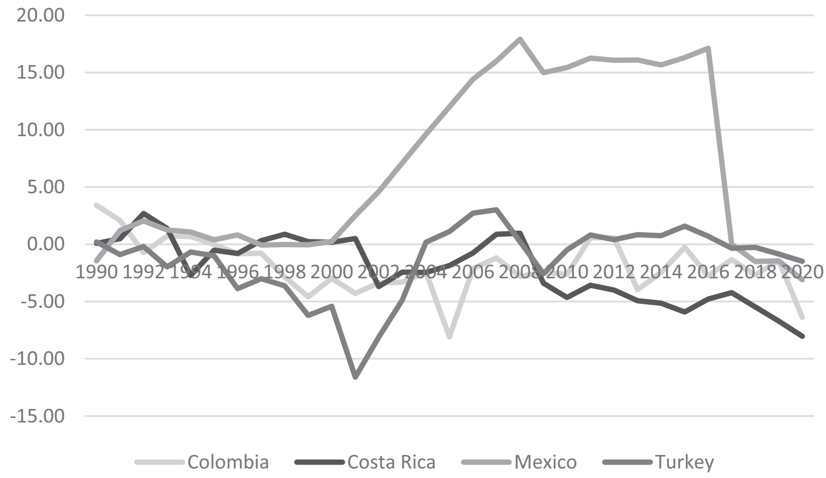


FIGURE 18.2 Development of FD.

increased in 2017. For Colombia, the FD decreased during the 1990–1994 period, but increased over the period 1995–2010, decreased over the period 2011–2012, and has been increasing since 2013. For Costa Rica, the FD increased over the periods 1994–1996, 2002–2006, and 2009–2020, whereas decreased during other periods. In Mexico, the FD was high over the periods 1991–1996 and 2000–20017, but decreased in 1990, 1997–1999, and 2018–2020.

4 PROPOSED WORK

The study covers the annual data of Colombia, Costa Rica, Mexico, and Turkey which are the OECD countries with upper-middle incomes over the period 1990–2020. The data of IR and FD variables are obtained from the Government Finance Statistics section, which is the public database of the IMF; whereas the data of GDP and MONEY variables are obtained from the WDI database. Stata 14 software is utilized for the analysis.

Upon examining Table 18.1, the Annual Percentage Change of the IR is seen as the dependent variable, and the Government's Fiscal Deficit/GDP (FD) is seen as the independent variable.

4.1 Examination of descriptive statistics of the panel data model

Descriptive statistics of the panel regression model are examined. The descriptive statistics results are presented below.

Upon examining Table 18.2, it is seen that a huge difference exists between the minimum and maximum values of the IR, GDP, and MONEY variables. Therefore, the differences are reduced by taking the logarithm of this variable. Thus, the model;

$$LOGIR_{it} = \alpha_{it} + \beta_{1it}FD_{it} + \beta_{2it}LOGGDP_{it} + \beta_{3it}LOGMONEY_{it} + u_{it} \quad (18.1)$$

TABLE 18.1
Variables used in analysis and their explanations

Variables	Explanations
IR	Annual percentage change in the consumers' price index (IR)
FD	Ratio of the government's fiscal deficit to GDP
GDP	Real GDP per capita
MONEY	Rate of change in expansionary money supply

TABLE 18.2
Descriptive statistics

Variables	Mean	Std. dev.	Min.	Max.	# of obs. (N)/# of observed groups
IR	16.48241	20.8948	0.0174789	105.215	124/31
LOGIR	0.9624123	0.5020158	−1.757487	2.022078	124/31
FD	2.533304	2.221358	0.0198232	11.60721	124/31
GDP	7687.083	2390.256	3639.719	12755.17	124/31
LOGGDP	3.863211	0.144209	3.561068	4.105686	124/31
MONEY	25.28749	28.79917	−45.47297	154.3908	124/31
LOGMONEY	1.237892	0.3965632	−0.0262785	2.188621	124/31

4.2 Determining the estimation method of the model

The study involves a panel dataset since it includes data from four countries and long periods (31 years). Although studies on panel data were first conducted by Hildreth (1950), Kuh (1959), Grunfeld and Griliches (1960), Zellner (1962), Balestra and Nerlove (1966), and Swamy (1970), the panel data model has been employed since the 1990s (Yerdelen Tatoğlu, 2018: 3). A linear panel data model:

$$Y_{it}= \beta_{0it} + \beta_{1it}X_{1it} + \beta_{2it}X_{2it} + \cdots + \beta_{kit}X_{kit}+ u_{it},\ i =1, \ldots , N;\ t = 1, \ldots, T \quad (18.2)$$

Briefly;

$$Y_{it} = \beta_{0it} + \sum_{k=1}^{K} \beta_{kit}X_{kit} + u_{it},\ i = 1, t\ldots, N;\ t = 1,\ldots, T \quad (18.3)$$

Here, i is the unit size, and t is the time dimension. β_{0it} is the constant term; β_{kit} is a vector of parameters with K × 1 dimension; X_{kit} is the value of the kth explanatory variable at time i for the value of the ith unit; Y_{it} is the value of the dependent variable for the ith unit at time t. If both the constant and slope parameters are fixed with respect to the unit and time, the Classical model is involved. If the slope parameter is fixed and the constant parameter is variable with respect to units, there is a unit

TABLE 18.3
Results of the within-group estimators

Probability value of within-group estimator for the unit effect	Probability value of within-group estimator for the time effect
0.0000	0.0000

effects model. This model is called the one-way model. If the slope parameter is fixed and the constant parameter is variable with respect to both units and time, the unit and time effects model is involved. This model is also called the two-way model (Yerdelen Tatoğlu, 2018: 37–40).

To decide whether the research model is classical or not, the F test results in Table 18.3 are considered. Since the p-value is lower than 0.05, the model is determined not to be classical. Thus, the presence of time and unit effects in the model should be tested with the within-group estimator (WE).

The null hypothesis, which implies that no unit and time effect exist, is rejected at the 95% confidence interval. Accordingly, unit and time effect exists. Then, it is determined whether such an effect is fixed or random.

Although it is assumed that no correlation exists between unit effects and explanatory variables in the random effect model (REM), this correlation is assumed to be different from zero in the fixed effect model (FEM) (Yerdelen Tatoğlu, 2018: 79).

By performing the Hausman test developed by Hausman in 1978, it is determined whether the model is FEM or REM. In the Hausman test, it is tested whether a difference occurs between the REM estimator $\left(\hat{\beta}_{GLS}\right)$ and the FEM estimator $\left(\hat{\beta}_{FE}\right)$ for the model below:

$$Y_{it} = X_{it}\beta + \mu i + \varepsilon_{it},\ i = 1, \ldots, N;\ t = 1, \ldots, T \tag{18.4}$$

$$\hat{q} = \left(\hat{\beta}_{GLS} - \hat{\beta}_{FE}\right) \tag{18.5}$$

The hypotheses (Hausman, 1978: 1261–1263):

$$H_0: \text{corr}(\mu_i, X_i) = 0 \tag{18.6}$$

There is no difference between the REM and FEM estimators. The REM is efficient.

$$H_1: \text{corr}(\mu_i, X_i) \neq 0 \tag{18.7}$$

There is a difference between the REM and FEM estimators (FEM is efficient).

If the calculated $p < 0.05$ according to the hypotheses, H_0 is rejected; whereas it is accepted if $p > 0.05$. Accordingly, if the p-value is lower than 0.05, the REM would be considered, and if the p-value exceeds 0.05, the FEM would be applicable. The result of the Hausman estimator is presented below (Table 18.4).

TABLE 18.4
Hausman estimator for the unit and time effects

Hausman test for the unit effects			Hausman test for the time effects		
Chi^2	Probability value (p)	Estimation method	Chi^2	Probability value (p)	Estimation method
47.85	0.0000	FE	30.30	0.0000	FE

Upon examining the above table, since the probability values of the Hausman test for unit and time effects are lower than 0.05, a two-way FEM is involved.

4.3 Testing the assumptions of the model

4.3.1 Normality assumption

Spiegel and Stephens (2011) and Oral Erbaş (2008) asserted that; according to the central limit theorem, in samples of N selected from a finite population, for quite large values of N ($N \geq 30$), the sampling distributions of the means exhibited an approximately normal distribution regardless of the population. The unit size (n) representing the number of countries is four, whereas the time dimension (t) representing years is 31. Thus, to express the sample number in another way, since the number of observations is $N = n \times t = 124$, the mean value of the samples exhibits a normal distribution.

4.3.2 Multicollinearity assumption

In multiple regression models with more than one independent variable, the relationship between two, some, or all of the independent variables is known as multicollinearity. With the variance inflation factor (VIF), it is determined how far the variances of parameter estimates diverge from their actual values due to multicollinearity:

$$\mathrm{VIF_i} = \frac{1}{1 - R_i^2}, \; i = 1, 2, \ldots, k \tag{18.8}$$

Here, R_i^2 is the value of the models in which the independent variables are the dependent variable one by one and the others are the independent variables. Starting from the model with *k* independent variables below, the auxiliary regression models are estimated *k* times, and the VIF value is calculated from all of them.

$$Y_{it} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + u$$

$$X_1 = \beta_0 + a_1 X_1 + a_2 X_2 + \cdots + a_k X_k + u, \; \mathrm{VIF_1} = \frac{1}{1 - R^2_{X_1 X_2 X_3 \ldots X_k}} \tag{18.9}$$

$$X_k = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + v,\ VIF_k = \frac{1}{1 - R^2_{X_k X_1 X_2 \ldots X_{k-1}}} \tag{18.10}$$

When the above criteria are examined;

VIF = 1 if $R_i^2 = 0$, and no multicollinearity exists.

The VIF value ranges between 0 and 5 if $0 < R_i^2 < 0.5$, and no multicollinearity exists.

The VIF value ranges between 5 and 10 if $0.50 < R_i^2 < 0.80$, and the multicollinearity is moderate.

The VIF value exceeds 10 if $0.80 < R_i^2 < 1$, and strong multicollinearity exists.

It is called 1/VIF tolerance number and is used as a criterion providing information about multicollinearity. When VIF > 10, the multicollinearity problem is quite strong (Yerdelen Tatoğlu, 2020a: 111, 115). Table 18.5 presents the result of the VIF criterion by which multicollinearity is measured.

In Table 18.5, it is seen that the VIF values of the model are lower than five and no multicollinearity problem exists.

4.3.3 Testing the omitted variable in the model

Ramsey (1969) proposed a model to test whether the model specification was correct and this model has been the most performed test to determine the specification error in the literature. Based on the following model in the Ramsey Reset test,

$$Y = X\beta + u \tag{18.11}$$

X is of dimension NT × k and u is normally distributed with a zero mean.

$E[u|X] = \xi \neq 0$ if Model (18.11) has a specification error. In the Ramsey Reset test, ξ is considered as $Z\theta$.

In the following model:

$$Y = X\beta + Z\theta + u \tag{18.12}$$

The null hypothesis:

H_0: $\theta = 0$ (the model specification is correct)

TABLE 18.5
Multicollinearity test results

	VIF criterion	
Variables	**VIF value**	**1/VIF value**
FD	1.20	0.835360
LOGGDP	1.13	0.884931
LOGMONEY	1.08	0.928215

Mean VIF: 1.13

The constrained F test is performed to test constrained model specification error in Model (18.11), and unconstrained model specification error in Model (18.12) as follows:

$$F = \frac{R_{UR}^2 - R_R^2 / p}{1 - R_{UR}^2 / NT - k - p} \tag{18.13}$$

R_{UR}^2 denotes the salience coefficient of the unconstrained model in Equation (18.9); whereas R_R^2 denotes the salience coefficient of the constrained model in Equation (18.12). The calculated F value, p, fits the F distribution with NT − k − p degrees of freedom.

If H_0 is rejected, it is accepted that a specification error exists (Yerdelen Tatoğlu, 2020a: 308). The result of this test, called Ramsey Reset, is presented in Table 18.6.

Since the p-value of the Ramsey Reset test exceeds 0.05 in Table 18.6, it is determined that there is no omitted variable in the model.

4.3.4 Autocorrelation assumption

The fact that the error terms are correlated with the error terms of other periods is described as autocorrelation (Yerdelen Tatoğlu, 2020a: 130). For the Model, which is FEM; autocorrelation is tested with the Durbin–Watson test proposed by Bhargava, Franzini, and Narendranathan and LBI tests proposed by Baltagi-Wu. The test result is presented in Table 18.7.

Upon examining Table 18.7, the Durbin–Watson test suggested by Bhargava, Franzini, and Narendranathan for the unit effect and the LBI test suggested by Baltagi-Wu for the unit effect indicates that values are lower than 2, and thus, a

TABLE 18.6
Ramsey reset test

Testing the omitted variable in the model with the Ramsey reset test	
Test value	**Probability value (p)**
2.12	0.1016

TABLE 18.7
Autocorrelation test

Determining autocorrelation by performing Durbin–Watson and LBI tests			
For the unit effect		**For the time effect**	
Test name	**Test value**	**Test name**	**Test value**
Durbin–Watson	1.2235021	Durbin–Watson	2.3696565
LBI	1.264315	LBI	2.7405054

1st-order autocorrelation exists (Yerdelen Tatoğlu, 2018: 225). It is seen that there is no autocorrelation problem.

4.3.5 Heteroscedasticity assumption

If the conditional variance of the error term remains the same depending on the independent variable, there is homoscedasticity, and if it varies depending on the independent variable, there is heteroscedasticity (Gujarati and Porter, 2012: 365). The results of heteroscedasticity tested with the modified Wald test for the unit and time effects for the FEM are shown in Table 18.8.

Upon examining Table 18.8, it is determined that a heteroscedasticity problem exists for the time and unit effects, according to the results of the heteroscedasticity assumption test for the FEMs.

4.3.6 Cross-sectional dependence (CSD) assumption

After determining the estimation method of the model, the assumption of CSD should be examined.

To test the CSD, Breusch Pagan's (1980) LM test is performed when T is large and N is small. Here, Breusch Pagan's (1980) LM test is performed since T (time dimension) is 31 and N (unit size) is 4. Table 18.9 presents the test results.

Upon examining Table 18.9, it is concluded that there is autocorrelation according to the test results for the model.

4.3.7 Unit root test

The measurement of stationarity in time-series and panel datasets is made by unit root tests (Yerdelen Tatoğlu, 2013: 199). If there is autocorrelation, the 2nd-generation tests, otherwise, the 1st-generation tests are performed (Yerdelen Tatoğlu, 2020b: 21). Since autocorrelation exists, the 2nd-generation tests should be performed. The 2nd-generation tests are categorized into three groups. The 1st group consists of Levin, Lin, and Chu (LLC); Harris and Tzavalis (HT); Breitung, Hadri, Im, Pesaran,

TABLE 18.8
Heteroscedasticity test

Testing heteroscedasticity for the unit effect by performing the modified Wald test		Testing heteroscedasticity for the time effect by performing the modified Wald test	
Test value	**Probability value (p)**	**Test value**	**Probability value (p)**
80.98	0.0000	2916.57	0.0000

TABLE 18.9
Breusch Pagan's (1980) LM test results

Test statistic	Probability value (p)
37.56800	0.0000

and Shin (IPS); Fisher ADF; Fisher Philips and Perron (Fisher PP); and Choi Fisher ADF panel unit root tests. In the 2nd group of 2nd-generation tests, there are the Multivariate Augmented Dickey–Fuller (MADF) and the Seemingly Unrelated Regression Augmented Dickey–Fuller (SURADF) panel unit root tests. Panel unit root tests in the 3rd group of 2nd-generation tests are Moon and Perron (2004); Cross-Sectional Augmented Dickey–Fuller (CADF); Augmented Cross-Section Im, Pesaran and Shin (CIPS); Panel Analysis of Nonstationarity in Idiosyncratic and Common (PANIC); Augmented Sargan and Bhargava (CSB); and PANICCA (Yerdelen Tatoğlu, 2020b: 67–100).

Here, one of these tests, the IPS panel unit root test is performed.

Table 18.10 presents the IPS panel unit root test results. According to this result, it is seen that all variables are stationary at the 1st difference (I(1)).

4.3.8 Testing the homogeneity of slope parameters

To test the Random Coefficients Model (RCM), the difference between the Ordinary Least Squares (OLS) estimators that ignore the panel structure of the data and the weighted average matrices of the WE can be considered. If no statistically significant difference exists, the parameters are homogeneous. The hypothesis to be tested would be established as follows:

$$H_0\!: \beta_i = \beta \tag{18.14}$$

Statistics in the Swamy S test, which was derived by Swamy (1970) and is a Hausman type, would be written as follows:

$$\hat{S} = X^2_{k(N-1)} = \sum_{i=1}^{N} \left(\hat{\beta}_i - \bar{\beta}^*\right)' \hat{V}_i^{-1} \left(\hat{\beta}_i - \bar{\beta}^*\right) \tag{18.15}$$

Here, $\hat{\beta}_i$ denotes the OLS estimators obtained from the regressions according to units, $\bar{\beta}^*$ represents the weighted WE estimator, and $\hat{V}_i$ denotes the difference

TABLE 18.10
IPS panel unit root test results

Variables	Test level	Constant/ with trend	Statistic value of the test	p-Value	Decision
LOGIR	Level	Constant	−0.1789	0.4290	I(1)
	1st Diff.	Constant	−8.2251	0.0000	
FD	Level	Constant	−4.3700	0.1510	I(1)
	1st Diff.	Constant	−10.2179	0.0000	
LOGGDP	Level	Constant	0.8963	0.8150	I(1)
	1st Diff.	Constant	−5.5870	0.0000	
LOGMONEY	Level	Constant	−1.5405	0.0617	I(1)
	1st Diff.	Constant	−15.4885	0.0000	

TABLE 18.11
Swamy S homogeneity test results

Chi^2 test statistic	p-Value
100.62	0.0000

TABLE 18.12.
Results of the Driscoll–Kraay's AR(1) linear regression with residuals

$LOGIR_{it} = \alpha_{it} + \beta_{1it}FD_{it} + \beta_{2it}LOGGDP_{it} + \beta_{3it}LOGMONEY_{it} + u_{it}$

Dependent variable (LOGIR)	R-sq	F test value	Prob. value (p)
	0.5682	52.63	0.0000
Independent variables	**Coefficient**	**t test value**	**Prob. value (p)**
FD	–0.0217854	–1.55	0.124
LOGGDP	–0.5929577	–2.67	0.009
LOGMONEY	0.8188152	9.86	0.000

between the variances of the two estimators. The test statistic exhibits a X^2 distribution with K(N − 1) degrees of freedom. The parameters are heterogeneous if the test statistic exceeds the critical value; whereas homogeneous if the test statistic is lower than the critical value (Yerdelen Tatoğlu, 2020b: 247). Table 18.11 presents the Swamy S homogeneity test results.

Upon considering Table 18.11, according to the Swamy S test results, H_0 is rejected, it is accepted that the parameters are not homogeneous and tend to vary from unit to unit, and are heterogeneous.

5 RESULT AND DISCUSSION

Upon examining the basic assumption tests, it is seen that the autocorrelation problem exists for the unit effect; whereas heteroscedasticity and CSD exist for the unit and time effects. Driscoll and Kraay's (1998) estimator can accurately estimate the parameters using the POLS method under the assumption that the error term is heteroscedastic, autocorrelated, and CSD. Furthermore, for fixed effects, Driscoll–Kraay has derived an AR(1) linear regression with residuals model, where residual 1st-order autocorrelation follows a regressive process (AR(1) correlation, which is the case with 1st-order autocorrelation) (Yerdelen Tatoğlu, 2018: 276, 279).

Accordingly, the result of Driscoll–Kraay's AR(1) linear regression with residuals model for the Model with the number (18.1) is presented in Table 18.12.

Upon examining Table 18.12, it is seen that the entire model and all independent variables of the model are statistically significant at the 5% level, according to the p-value, which determines the significance of the overall model. Although

LOGIR variable positively affects LOGIR, LOGMONEY variable negatively affects LOGIR. A 1% increase in LOGGDP decreases the LOGIR variable by about 0.6%; whereas a 1% rise in LOGMONEY enhances LOGIR by almost 0.82%. FD is not statistically significant.

6 CONCLUSION AND FUTURE WORK

Price stability, which is expressed as the continuation of the change in the general level of prices at an acceptable rate, is one of the main economic objectives. Monetary and fiscal policies are the two main policy tools in ensuring the aforementioned stability. Various debates and different views on the effectiveness of these policy tools have continued in the literature from the past to the present. According to the Classical view, the substitution of the FD with current taxation in autarchic economies increases the total consumer demand, and thus, the RIR increases and investment decreases. Accordingly, the production capital declines in the long run. In open economies, the main country's substitution of the FD with current taxes causes foreign borrowing, hence, the CAD. The FD also causes a decrease in domestic investment, while a decrease in the capital stock in the long run by excluding domestic investments.

In the Monetarist view, two situations are considered in terms of the predominance of monetary and fiscal policies.

In the case of the predominance of monetary policy on the fiscal policy along with an FD, such a deficit is offset by the fiscal authority using a combination of the monetary authority's seigniorage revenues and bond sales to the public sector. The monetary authority constantly controls IR in the monetary economy since it is authorized to freely determine the monetary base. In the 2nd case, where fiscal policy predominates over monetary policy, the fiscal authority creates its budget independently by declaring all current and future FDs and surpluses which, in turn, determines the amount of revenue that needs to be increased through the sale of bonds and seigniorage. In the event that the fiscal authority has a deficit in the budget, the monetary authority can control neither the economic growth that constitutes the monetary base nor the IR.

Instead of accepting the deficit budget as wrong, Keynes advocated the fiscal policies implemented against the conjuncture in accordance with the economic fluctuations. Keynesian economists increase taxes and prevent inflation to soothe the economy when demand-side growth is high.

According to the Neoclassical view, FDs are increased by imposing taxes on future generations and making lifetime total consumptions. Increasing consumption, by definition, reduces saving. Thus, the interest rate should rise to bring the capital market into equilibrium. Permanent FDs reduce private capital accumulation by creating a crowding-out effect.

According to the Ricardian view, overlapping generations are altruistic about resource transfer through voluntary bonds. With future generations paying taxes, the FD is passed on to future generations without altering the total resources of taxpayers and their descendants.

New Keynesian Economics Woodford mentioned that the public budget deficit caused public borrowing, and found that the unexpected increase in the public's primary budget deficit increased IR, real GDP, and the nominal interest rate (NIR).

The impacts of FD, GDP, and expansionary MS on the IR were investigated for Colombia, Costa Rica, Mexico, and Turkey, which are OECD-member countries in the upper middle-income group, over the period 1990–2020. Since heteroscedasticity, autocorrelation, and CSD problems were detected as a result of the assumption tests, Driscoll and Kraay's (1998) estimator method was employed because it estimated the parameters correctly even if those problems were present. It was detected that the FD did not have a statistically significant impact on the IR. It was found to be consistent with the findings of Karadeniz (2021), Olaniyi (2020), and Tiwari, Tiwari, and Pandey (2012). According to this result, policymakers should pay attention to the MS, which affects the inflation phenomenon, instead of considering the FD problem upon examining IR. The expansion in the MS encourages both households and producers to consume more. Policymakers can increase the MS mainly in the expansion and the peak phase and decrease the MS mainly in the contraction and recovery phases. FD is seen more due to the increase of the IR. Because of this reason future studies can be expanded by the inclusion of different country groups and periods, and even import and export data to be used as the independent variables.

REFERENCES

Ahmad, A. H. & Aworinde, O. B. (2019). Are fiscal deficits inflationary in African countries? A new evidence from an asymmetric cointegration Analysis. *The North American Journal of Economics and Finance*, 50, 1–23.

Balestra, P. & Nerlove, M. (1966). Pooling cross section and time series data in the estimation of a dynamic model: The demand for natural gas. *Econometrica*, 34(3), 585–612.

Barro, R. J. (1989). The Ricardian approach to budget deficits. *Journal of Economic Perspectives*, 3(2), 37–54.

Bernheim, B. D. (1989). A neoclassical perspective on budget deficits. *Journal of Economic Perspective*, 3(2), 55–72.

Breusch, T. S. & Pagan, A. (1980). The Lagrange multiplier and its applications to model specification in econometrics. *Review of Economic Research*, 47(1), 239–253.

Catao, L. A. V. & Terrones, M. E. (2005). Fiscal deficits and inflation. *Journal of Monetary Economics*, 52, 529–554.

Danlami, I. A., Hidthiir, M. H. B. & Hassan, S. (2019). Dynamic analysis of the effect of fiscal deficit on inflation in Nigeria. *Academic Journal of Economic Studies*, 5(2), 159–165.

Doğan, B. S. & Günel, T. (2021). Avrupa'nın yükselen ekonomilerinde bütçe açığının enflasyon üzerindeki etkisi: Panel ARDL yaklaşımı. *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, 32, 17–30.

Dolado, J. J. & Lütkepohl, H. (1996) Making Wald tests work for cointegrated VAR systems. *Econometric Reviews*, 15 (4), 369–386.

Driscoll, J. & Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent data. *Review of Economics and Statistics*, 80, 549–560.

Duodu, E., Baidoo, S. T., Yusif, H. & Frimpong, P. B. (2022). Money supply, budget deficit and inflation dynamics in Ghana: An empirical investigation. *Cogent Business & Management*, 9, 1–23.

Durguti, E., Kryeziu, N. & Gashi, E. (2020). How does the budget deficit affect IR-Evidence from Western Balkans. *International Journal of Finance & Banking Studies*, 9(1), 1–10.

Eita, J. H., Manuel, V., Naimhwaka, E. & Nakusera, F. (2021). The impact of fiscal deficit on inflation in Namibia. *Journal of Central Banking Theory and Practice*, 1, 141–164.

Ezeabasılı, V. N., Mojekwu, J. N. & Herbert, W. E. (2012). An empirical analysis of fiscal deficits and inflation in Nigeria. *International Business and Management*, 4(1), 105–120.

Feldstein, M. S. (1974). Social security, induced retirement, and aggregate capital accumulation. *Journal of Political Economy*, September/October, 82, 905–926.

Friedman, M. (1948). A monetary and fiscal framework for economic stability. *The American Economic Review*, 38(3), 245–264.

Grunfeld, Y. & Griliches, Z. (1960). Is aggregation necessarily bad? *Review of Economics and Statistics*, 42(1), 1–13.

Gujarati, D. N. & Porter, D. C. (2012). *Basic Econometrics*. New York: Mc Graw Hill. Güneş, H. (2020). Tanzi ve Patinkin Etkisi çerçevesinde bütçe açığı ile enflasyon arasındaki ilişkinin incelenmesi. *AÇÜ Uluslararası Sosyal Bilimler Dergisi*, 6(2), 1–11.

Hatemi-J A. & El-Khatib Y. (2016). An extension of the asymmetric causality tests for dealing with deterministic trend components. *Applied Economics*, 48(42), 4033–4041.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6), 1251–1271.

Hildreth, C. (1950). Combining cross-section data and time series. *Cowles Commission Discussion Paper*, No. 347.

Jahan, S., Mahmud, A. S. & Papageorgiou, C. (2014). What is Keynesian economics? *Finance & Development*, September, 1–2.

Jalil, A., Tariq, R. & Bibi, N. (2014). Fiscal deficit and inflation: New evidence from Pakistan using a bounds testing approach. *Economic Modelling*, 37, 120–126.

Karadeniz, Y. (2021). Mali kural uygulayan bazı gelişmekte olan ülkelerde bütçe açığının cari açık, ekonomik büyüme ve enflasyon üzerine etkisi. *Maliye Araştırmaları Dergisi*, 7(1), 51–58.

Kaur, G. (2021). Inflation and fiscal deficit in India: An ARDL approach. *Global Business Review*, 22(6), 1553–1573.

Kuh, E. (1959). The validity of cross-sectionally estimated behaviour equations in time series applications. *Econometrica*, 27, 197–214.

Lin, H. Y. & Chu, H.-P. (2013). Are fiscal deficits inflationary? *Journal of International Money and Finance*, 32, 214–233.

Maraş, D. & Dumrul, G. (2019). Türk ekonomisinde bütçe açığı ve enflasyon ilişkisinin sınır testi yaklaşımı ile eş-bütünleşme analizi. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 54, 314–342.

Modigliani, F. & Sterling, A. (1986). Government debt, government spending, and private sector behavior: Comment. *The American Economic Review*, 76(5), 1168–1179.

Moon, H. & Perron, B. (2004). Testing for a unit root in panels with dynamic factors. *Journal of Econometrics*, 1, 81–126.

Myovella, G. A. & Kisava, Z. S. (2017). Budget deficit and inflation in Tanzania: ARDL bound test approach. *Journal of Business, Economics, and Finance*, 7(1), 83–88.

Olaniyi, C. O. (2020). Application of Bootstrap Simulation and Asymmetric Causal Approach to fiscal deficit-inflation nexus. *Global Journal of Emerging Market Economies*, 12(2), 123–140.

Olubiyi, E. A. & Bolarinwa, M. A. (2018). Fiscal deficit and IR in selected African Regional Blocs: A comparative analysis. *Turkish Economic Review*, 5(1), 89–101.

Oral Erbaş, S. (2008). *Olasılık ve istatistik*. Ankara: Gazi Kitabevi.

Ramsey, J. B. (1969). Tests for specification errors in classical linear Least-Squares regression analysis. *Journal of the Royal Statistical Society*. Series B (Methodological), 31(2), 350–371.

Sargent, T. J. & Wallace, N. (1981). Some unpleasant monetarist arithmetic. *Federal Reserve Bank of Minneapolis Quarterly Review*, 5(3), 1–18.

Spiegel, M. R. & Stephens, L. J. (2011). İstatistik (S. Çelebioğlu, Çev.). Ankara: Nobel Akademik Yayıncılık.

Ssebulime, K. & Edward, B. (2019). Budget deficit and inflation nexus in Uganda 1980–2016: A cointegration and Error Correction Modeling approach. *Journal of Economic Structures*, 8(3), 1–14.

Swamy, P. A. (1970). Efficient inference in a random coefficient regression model. *Econometrica*, 38(2), 311–322.

Tiwari, A. K., Tiwari, A. P. & Pandey, B. (2012). Fiscal deficit and inflation: What causes what? The Case Of India. *Journal of International Business and Economy*, 13(1), 57–81.

Woodford, M. (1996). Control of the public debt: A requirement for price stability? *NBER Working Paper*, No. 5684, 1–35.

Yerdelen Tatoğlu, F. (2013). İleri panel veri analizi stata uygulamalı. İstanbul: Beta Basım Yayım Dağıtım A.Ş.

Yerdelen Tatoğlu, F. (2018). *Panel veri ekonometrisi Stata uygulamalı.* İstanbul: Beta Basım Yayım Dağıtım A.Ş.

Yerdelen Tatoğlu, F. (2020a). *Ekonometri stata uygulamalı*. İstanbul: Beta Yayınları.

Yerdelen Tatoğlu, F. (2020b). *Panel zaman serileri analizi*. İstanbul: Beta Yayınları.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of American Statistical Association*, 57, 348–368.

Index