Volume 4 Number 4  December 2022

INSIGHTS INTO REGIONAL DEVELOPMENT*

* The journal was launched as a platform for dissemination of the project results, which has received funding from the European Union’s Horizon 2020 research and innovation programme European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme Marie Skłodowska-Curie Research and Innovation Staff Exchanges ES H2020-MSCA-RISE-2014 CLUSDEVMED (2015-2019) Grant Agreement Number 645730730
INSIGHTS INTO REGIONAL DEVELOPMENT ISSN 2669-0195 (online) is a peer-reviewed journal, which publishes original research papers and case studies. It is international journal published cooperating with universities, social companies, consultancies and associations. It is published quarterly.

Areas of research include, but are not limited to, the following:

- Conceptual/Practical Approaches and Methodologies towards Sustainable Regional Development
- Globalization, Internationalization and Solutions for Low-Carbon Economies of Scope or Scale
- Innovations and Technology Transfer Pilot Results Advancing Regional Development
- Information Technologies (IT) and Information Communication Technologies (ICT) for Regional Development, and Socio-Economic Innovations
- Environmental Engineering for Sustainability and Regional Development
- Smart Electricity Grid Solutions for Sustainable Regional Development
- Entrepreneurship for Regional Development
- Safety and Regional Development
- Threats to Society, Entrepreneurship and Regional Development
- Security and Safety of Countries, Regions and Society as Precondition of Sustainable Regional Development
- Boarder Guard Issues and Regional Development
- Migration and Sustainable Regional Development
- Terrorism, International Crime and Secure Regional Development
- Security and Safety Models for Sustainable Regional Development
- Emerging Business Drivers, Opportunities, or Constraints in Future Sectors of Current Markets
- Retrofitting Techniques and Technologies Near-Zero Energy Demand Buildings and Regional Development
- Urban Planning and Advanced Construction Materials for Energy-Efficiency or Smart Cities
- Contemporary Regional Development in the Context of Security, Comfort, or Legislation
- PPPs for Developments in Sustainable Entrepreneurship, Finance, and Investments
- Accounting for Sustainable Regional Development
- Sustainable Regional Development in the Context of Management or Life-Long Learning
- Strategic Management Practices and Creative Marketing Techniques for Sustainable Regional Development
- Organizational Studies and Sustainable Regional Development
- Regional Development in the Context of Marketing, ICT, and Creative Industries
- Intercultural Communication for Sustainable Regional Development
- Institutions and Regional Development
- Case Studies on Entrepreneurship and Sustainable Regional Development
- Social Cohesion and Social Innovations for Regional Development
- Business Models and Strategic Management for Sustainable, Secure and Safe Futures
- Corporate Social Responsibility and Sustainable Regional Development
- Regional Development in the Context of Psychology, Education, or Women
- Intuitions about Emerging Phenomena in Regional Development
- Start-Ups, Spin-Offs, SMEs in the Context of Market Growth and Regional Development
- Global Value Chain Management for Sustainable Regional Development
- Knowledge Management for Sustainable and Safe Regional Development
- Legal regulation peculiarities affecting regional development

Editorial correspondence including manuscripts and submissions:
E-mail: irdeditor@jssidoi.org
International Editorial Board

Editor-in-Chief

- **Prof. Salvatore Monni**, Roma Tre University, Italy

Editors

- **Prof. Manuela Tvaronavičienė**, Vilnius Gediminas Technical University, Lithuania
- **Prof. Wadim Strielkowski**, Prague Business School, Czech Republic
- **Prof. Dalia Štreimikienė**, Vilnius University, Lithuania
- **Dr. Waleed Mugahed Al-Rahmi**, Faculty of Social Sciences & Humanities, School of Education, Universiti Teknologi Malaysia, Malaysia
- **Prof. Yuri Bilan**, Rzeszow University of Technology, Poland
- **Prof. Gil Vieira**, Instituto Nacional de Pesquisas da Amazônia, Brazil
- **Dr. Keti Lelo**, Università degli Studi Roma Tre, Italy
- **Prof. Sebastian Kot**, Czestochowa University of Technology, Poland
- **Ms. Elisa Natola**, Brazilian National Council for the State Funding Agencies (CONFAP), Brazil
- **Dr. Rasmus Bøgh Holmen**, Institute of Transport Economics, Norway
- **Dr. Throstur Olaf Sigurjonsson**, School of Business, Reykjavik University, Iceland
- **Prof. Abel Femi Adekola**, Wilkes University, USA
- **Prof. Fan Lü**, Tongji University, China
- **Prof. Dr. Gilberto de Miranda Rocha**, Federal University of Pará, Brazil
- **Prof. Dr. Sibylle Heilbrunn**, Kinneret Academic College on the Sea of Galilee, Israel
- **Dr. Amr Radwan**, Academy of Scientific Research and Technology, Egypt
- **Prof. Armenia Androniceanu**, The Bucharest University of Economic Studies, Romania
- **Prof. Najiba El Amrani**, University Sidi Mohamed bin Abdellah, Morocco
- **Prof. Bora Aktan**, University of Bahrain, the Kingdom of Bahrain
- **Prof. Gunnar Prause**, Tallinn University of Technology, Estonia
- **Prof. Vera Komarova**, Daugavpils University, Latvia
- **Prof. Fernando García**, Polytechnic university of Valencia, Spain
- **Prof. Michal Fabuš**, School of Economics and Management in Public Administration in Bratislava, Slovak Republic
- **Prof. Sulphey, M. M.**, Prince Sattam Bin Abdulaziz Universtiy, Saudi Arabia
- **Prof. Mirjana Radović Marković**, Institute of Economic Sciences, Serbia
- **Prof. Ronaldo Lopes Rodrigues Mendes**, Federal University of Pará, Brazil
- **Dr. Andrea Bernardi**, Oxford Brookes University, UK
• Prof. Francesco Vigliarolo, Universidad Católica de La Plata, Argentina
• Dr. Laima Gerlitz, Wismar University of Applied Sciences, Germany
• Dr. Federico Tomassi, Agenzia per la Coesione Territoriale, Italy
• Dr. Kiran Javaria, University of Lahore, Lahore, Pakistan
• Prof. Maria da Gloria Gonçalves de Melo, Universidade do Estado do Amazonas, Brazil
• Prof. Mykola Sidak, University of Economics and Management in Public Administration in Bratislava, Slovak Republic
### CONTENTS

Tamas Somogyi, Rudolf Nagy.  
**SOME IMPACTS OF GLOBAL WARMING ON CRITICAL INFRASTRUCTURE PROTECTION - HEAT WAVES AND THE EUROPEAN FINANCIAL SECTOR**  
11

Matúš Grega, Pavel Nečas.  
**IMPLEMENTATION OF EFFECTIVE SOLUTIONS TO THE CRISIS’ TASKS AND ITS REGIONAL MANAGEMENT**  
21

Jereco Jims J. Agapito, Liza Lorena C. Jala, Rosemarie Cruz-Español.  
**PREDICTORS OF ENTREPRENEURIAL ATTITUDE AMONG ADMINISTRATORS OF SELECTED COMMUNITY COLLEGES AND UNIVERSITIES IN LEYTE**  
36

**WOMEN IN SCIENCE, TECHNOLOGY, ENGINEERING, AND MATHEMATICS (STEM) "EGYPTIAN CASE STUDY"**  
52

Alexandre Cabagnols, Ali Maalej, Pierre Mauchand, Olfa Kammoun.  
**THE DETERMINANTS OF ENTREPRENEURIAL INTENTION OF SCIENTIST PHD STUDENTS: AN APPROACH BASED ON THE THEORY OF PLANNED BEHAVIOR**  
63

Rasmus Bøgh Holmen.  
**PRODUCTIVITY IMPULSES FROM REGIONAL INTEGRATION: LESSONS FROM ROAD OPENINGS**  
83

Urim-Tummim Shipanga, Suzaan Le Roux, Jobo Dubihlela.  
**OPERATIONAL RISK FACTORS AND THE SUSTAINABILITY OF SMALL AND MEDIUM MANUFACTURING ENTERPRISES IN SOUTH AFRICA**  
126
K.G.N.U Ranaweera.
LOCATING CRIME IN THE FIELD OF SUSTAINABILITY: A BIBLIOMETRIC MAPPING ON "SUSTAINABILITY AND CRIME" FROM 1995 – 2022 140

Funmilayo Idowu Agbaje.
TOWARD QUERYING THE NATIONAL PERIL OF KIDNAPPING IN NIGERIA 155

Éva Hegedűs.
GEOPOLITICS AND GEOECONOMICS IN FUNDING HUMANITARIAN 166
SOME IMPACTS OF GLOBAL WARMING ON CRITICAL INFRASTRUCTURE PROTECTION - HEAT WAVES AND THE EUROPEAN FINANCIAL SECTOR

Tamas Somogyi ¹, Rudolf Nagy ²

¹ Doctoral School on Safety and Security Sciences, Óbuda University, Budapest, Hungary
² Donát Bánki Faculty of Mechanical and Safety Engineering, Óbuda University, Budapest, Hungary

E-mails: ¹ somogyi.tamas@phd.uni-obuda.hu; ² nagy.rudolf@bgk.uni-obuda.hu

Received 17 August 2022; accepted 17 October 2022; published 30 December 2022

Abstract. Global warming is one of the hottest topics today, impacting every aspect of our life on Earth. Surprisingly, the effects of climate change on critical infrastructure protection have not been studied extensively. This paper i) examines three challenges raised by heatwaves: managing environmental risks; failure of partners and supply chains caused by extreme weather; operation during heatwaves, and ii) gives an overview of the responses of the European financial sector. By contributing to the growing research area of global warming, this study provides an opportunity to understand some of the climate risks better and offers some insights into the climate strategy of the European banking industry. Research data is drawn from two primary sources. Relevant literature was explored, and current supervisory recommendations were examined along with the results of the European financial sector's first climate risk stress test.

Keywords: European Central Bank; banking industry; critical infrastructure protection; heat waves; global warming


JEL Classifications: Q54, M15, N24

1. Introduction

According to the widely accepted theory, the increasing amount of the so-called greenhouse gases in the atmosphere causes an increase in the global temperature by the adsorption of the heat that tries to leave the Earth after being heated by the Sun (Lechter, 2021). No doubt, the Earth is warming up in an accelerating way: in the last decades, it has been hotter than at any other time in the last millennium (McGuire, 2014). Among the effects of this global warming are rising sea levels, more frequent floods, severe storms, tornadoes and hurricanes, droughts and heatwaves (Fletcher, Smith, 2020).

The so-called heatwave is one of the most relevant extreme weather events due to its effects on health, society, agriculture and the environment. As the World Meteorological Organisation stated, 2015-2021 were the warmest years, with exceptional heatwaves in North America and Europe (World Meteorological Organization, 2022). For example, Greek data confirms the increasing trend in the frequency and length of heat waves between 1960 and
2019 (Founda et al., 2022). The latest report of the United Nations' Intergovernmental Panel on Climate Change (IPCC) forecasts that warming in Europe will continue to rise faster than the global mean (Intergovernmental Panel on Climate Change, 2022). Although hot weather jeopardizes all the regions of Europe, the Mediterranean is much more threatened (Molina, Sánchez, Gutiérrez, 2020). For the 2021–2050 period, heatwave amplitudes are expected to increase by 0.5–2 °C on average (and the maximum can reach 50 °C or more) in the warmest parts of the Mediterranean and southern Europe (Zittis et al., 2021). July 2022 was one of the three warmest Julys on record, and numerous July records for maximum temperature were broken in Western Europe, reported by the EU's Copernicus Climate Change Service (Copernicus Climate Change Service, 2022).

According to the World Health Organization, skin eruptions, heat fatigue, heat cramps, heat syncope, heat exhaustion, and heat stroke are classic heat-related illnesses (World Health Organization, 2004). The 2021 report of the Lancet Countdown on health and climate change has established a link between increased temperature and mortality (Romanello et al., 2021). It has been found that heat waves increase the number of deaths (Ye et al., 2012). Furthermore, it has been demonstrated in Spain that the vulnerability of the urban population to heatwaves is bigger than the vulnerability of the non-urban population (López-Bueno et al., 2021), which underpins that death is more likely to occur in high-income countries (Amirkhani et al., 2022), e.g., in the member states of the European Union.

Besides affecting our health and life, heatwaves also significantly affect physical security and fire safety: they create better conditions for fire in which opportunities may be seized. Fire can be used to cause damage to infrastructure and through it to an entire society, as it was experienced in 1991 when the Iraqi army set fire to the petroleum infrastructure in Kuwait (Moger, 2021). Moreover, arson has become a method of extremism, thus a weapon of terrorists threatening critical infrastructure and engaging security forces (Besenyő, 2017).

Thus, global warming and heat waves significantly affect our European society. The central questions of this study ask what heatwaves raise the challenges. How can we manage environmental risks and failures of partners and supply chains? How can the smooth operation of critical infrastructures be maintained during heat waves? A qualitative case study approach is used to answer these questions. This paper first examines how the European financial sector approaches the issue of global warming in general based on publicly available sector-specific data. It will then give an overview of the main challenges caused by heat waves. In the end, it attempts to identify areas for further research.

2. Critical infrastructure, financial sector

The financial sector also offers a wide array of services and products to individuals and corporations. Beyond dispute, any significant disruption of these services would have economic, social and perhaps political effects. A state's economic status is an essential condition of military and political security (Dumitru, Ferarau, 2018), hence the possible political effect. Taking into consideration the networks of parent companies and subsidiary companies in the European banking industry, these economic and political effects can easily cross the borders and impinge on other states as well. Therefore, financial services play an essential role in the economy's growth and well-being of people and are also vital for the states. This follows that banking industry services shall be considered essential (Nagy, Somogyi, 2021). Hence the importance of the infrastructure of the European financial system.
3. Global warming and the European financial sector

Having understood the existential threat posed by climate change, the European Union has adopted the regulation 2021/1119 of the European Parliament and of the Council of 30 June 2021, establishing the framework for achieving climate neutrality. According to Article 10, "the Commission shall engage with sectors of the economy within the Union that choose to prepare indicative voluntary roadmaps towards achieving the climate-neutrality objective" (Regulation (EU) 2021/1119). Within the European financial sector, the European Central Bank (ECB) and the national central banks are the sector-specific authorities, and one of their tasks is to promote the smooth operation of payment systems. In line with their legal and moral obligations, addressing global warming is a priority for the ECB and the national central banks. It has been confirmed that one of the critical factors of the secure and efficient activity of the European financial sector is supporting the sector members in their efforts to identify risks and reduce their impact on climate change (Sidak et al., 2020). In July 2021, an action plan was presented by ECB to include climate change considerations in its monetary policy strategy (European Central Bank, 2021).

One observer has already drawn attention to the paradox of securing critical infrastructure owned and operated by the private sector in a centralized way. The importance of involving economic and market actors in maintaining public security has been emphasized (Besenyő, Fehér, 2020). Nevertheless, sector-specific authorities are responsible for fostering the sector members' efforts to ensure the availability of essential services.

4. Heatwaves and critical infrastructure protection

The frequency and the length of heat waves are projected to increase, hence the importance of preparedness for extremely high temperatures. Three main areas affected by global warming will be discussed: managing climate change risks, outsourcing and supply chain failure caused by extreme temperatures, and daily operation during heatwaves. The final section summarises the main findings.

4.1 Enhanced risk management framework

In July 2022, further steps were taken by ECB. Besides reducing climate-related financial risks in Euro system credit operations, risk management practices have been enhanced with climate change-related risks (European Central Bank, 2022a). As a banking supervisor, ECB ensures that banks have an appropriate approach to identifying, assessing and managing climate-related risks. Believing that one of the main challenges for banks is addressing risks stemming from climate change, climate change-related risks have been defined as critical vulnerabilities and have become one of the supervisory priorities for 2022-2024 (European Central Bank, 2022b). In the first half of 2022, the first climate-related stress test was conducted in Europe (European Central Bank, 2022c) to assess the financial sector's preparedness for global warming hazards. The stress test focuses on risk management (European Central Bank, 2022d).

As generally accepted, risk management is the formal process utilized to quantify, qualify, and mitigate specific concerns an organization may discover or define (Broder et al., 2020). Without incorporating climate risks into the risk management framework, risks cannot be adequately identified, assessed and mitigated. Preparedness with appropriate mitigating actions and continuity plans depends on identifying and assessing risks. Hence the importance of enhancing risk management by adding climate risks to the existing risk portfolio, which has been found to reduce disaster risks by enabling the decision-makers to consider different types of extreme events (Urlainis et al., 2022).

---

1 see Article 2 and Article 3 of Protocol No 4 on the statute of the European system of central banks and the European central bank https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12016E/PRO/04
A Guide on climate-related and environmental risks has been issued by ECB in order to support financial institutions in their business strategy and risk management framework implementation (European Central Bank, 2020). According to this guide, the management must consider climate-related and environmental risks when developing the institution's overall business strategy. Moreover, institutions are suggested to consider the impact of climate-related and environmental events on their business continuity. Climate-related and environmental risks in the financial sector have come to be used for two categories: physical and transition risks. The former comprises risks from extreme weather events, such as heatwaves. The latter refers to possible financial loss that can result from the process of adjustment towards a lower-carbon and more sustainable economy. Therefore, physical risks caused by global warming are expected to be incorporated into the risk management framework of European banks.

4.2 Resiliency to the failure of partners and supply chains

Another area of critical infrastructure protection where considerable attention is needed is outsourcing and supply chains. The effects mentioned above of global warming affect everyone, including the partners. Although the banking industry is not a productive sector built on critical materials, the risks related to supply chains are worth mentioning. Extreme temperatures can make the supply chain fail by either causing failure in production or making transportation difficult. Hence the importance of proper supply change management. As suggested, multiple approaches are required to overcome a supply crisis (King, 2021), which depends on the particular sector and critical material. However, cooperation within the sectors can be suggested (Strauss, 2020), especially when it is coordinated by, e.g., the sector-specific authority. Addressing common barriers together and planning long-term crisis responses have been recommended to increase supply chains' resilience, especially after the coronavirus pandemic (Gebhardt et al., 2022).

The European Banking Authority has issued a guideline Regarding to outsourced ICT services (European Banking Authority, 2019a). According to this guideline, the contracts between members of the European financial sector and service providers have to include operational and security incident handling, thus raising the level of preparedness for service disruptions. Moreover, business continuity plans are made in the financial sector. According to the latest guideline on outsourcing arrangements issued by the European Banking Authority, banks maintain and periodically test appropriate business continuity plans for outsourced critical functions (European Banking Authority, 2019b). Considering that heatwaves can result in failures at the service providers, business continuity plans have been suggested for the outsourced services.

Moreover, risk assessment has to be performed regularly to assess the impact of the failure or inadequate services. The type of such risk assessment framework in the financial sector depends on the financial institution: the more significant the institution, the more sophisticated the risk assessment. It should be noted that evaluating cascading effects and interdependence between systems and services has been suggested as a methodology to focus on understanding the vulnerability of infrastructures to the impact of natural hazards (Krausmann et al., 2019).

---

2 Guides are not binding for the institutions but rather serve as a basis for supervisory dialogue
3 According to the Article 16 of EU Regulation No 1093/2010 authorities and financial institutions make every effort to comply with the EBA guidelines.
4.3 Daily operation and heatwaves

Facilities, mainly where essential services are being operated, should be prepared for the operation during heatwaves. Regarding extreme hot periods, two key factors can be mentioned: they can be predicted, and power outages are likely to occur (Watson et al., 2022). Early warning systems can help prepare, mainly if emerging technologies are used (Argyroudis et al., 2022) in data collection, procession and communication. However, avoiding the so-called alert fatigue caused by false alerts (Roberts et al., 2022) is essential since it may lead to indifference toward alerts. This information shows that global warming raises challenges. The significant difficulties faced by office areas (e.g., headquarters) and data centres will be discussed in this section.

4.3.1 Office areas

Health problems caused by heat are an increasing risk for indoor and outdoor workers, as has been found (Xiang et al., 2014). Besides affecting health, heatwave significantly affects productivity (Flourish et al., 2018). Since temperature is a crucial factor, air conditioning has been confirmed to reduce the impact of heat waves (Wang et al., 2016). However, a link between heat waves and electrical demand has been established (Agel et al., 2021). These results confirm that facilities should be prepared by appropriately being equipped with air conditioning systems and supplied with electricity. Undoubtedly, maintaining an appropriate temperature to reduce accidents and occupational health problems in the workplace is an obligation in the EU (see Council Directive 89/391/EEC). Considering the importance of the cooling ability, especially in critical infrastructures, passive and active cooling solutions can be suggested (Attia et al., 2021) that are robust and able to work under extreme temperatures.

Besides staying in cool places, drinking water is essential during heat waves (Kovats, Kristie, 2006). Providing enough drinkable cool water in office areas is undoubtedly an important task. Therefore it should be emphasized that becoming resilient to the failure of water dispensers is a must.

Creating a suitable environment for working is probably unaffordable for many people. Thus, being in the office instead of home-working is likely during heat waves. Moreover, it is almost certain that the energy demand of the office areas will be higher during heat waves. Planning for extreme heat in cities is poorly understood (Keith et al., 2019). However, cooperation between operators of essential services can be suggested: based on sector or locality, the possibility of a standard preparation should be investigated. It should be determined if it is possible to have one building where vital employees of companies providing essential services can work during heatwaves. This would mean that only one building should be equipped with complete air conditioning and an uninterruptible power system and supplied with enough cool drinkable water instead of many buildings in the same area.

One of the financial sector's specialities is that branches are operating countrywide beside the headquarters. However, there are larger branches with more services, e.g., in county towns. These branches can be considered critical branches with more advanced business continuity solutions. Keeping only the key branches open during extremely hot periods may simplify the preparation, save energy, and prevent heat illness.

4.3.2 Data Centres

Banks offer more and more digitalized services as Information and Communication Technology spreads. Therefore, the smooth operation of data centres is becoming increasingly important. Data centres are vulnerable to extreme temperatures and power outages. Thus, the sector-specific authority specifies the protection of data centres against natural hazards in the European financial sector. According to the Guidelines on ICT and security
risk management, physical security measures should be applied to protect the data centres from environmental hazards. Moreover, the applied measures should be commensurate with the importance of the buildings and the criticality of the operations or ICT systems in these buildings (European Banking Authority, 2019a). It is worth an overview of the main challenges heatwaves raised and the possible risk-mitigating actions within and outside the financial sector.

It is almost certain that the energy-efficient free cooling methods (Zhang et al., 2014) will fail at places hit by heatwave since the external air is not cool enough. Thus, appropriate cooling methods and air conditioning systems should be applied. However, a study has found inefficient air conditioning in more than half of the data centres in Europe (Ni, Bai, 2017). A method has already been proposed to enhance thermal awareness and identify the ineffective cooling system in a data centre to avoid hardware failure and wasting energy (Grishina et al., 2022). Nevertheless, developing air conditioning systems raises the problem of increased energy demand. Since air conditioning systems need energy, increased electricity demand is highly probable. Therefore, preparing for the increased energy demand during heatwaves is essential. It should be examined whether or not operators of essential services can cooperate in fulfilling the increased power demand of their data centres during heatwaves which are predicted to be more frequent. Cutting back the emission of greenhouse gases is a shared responsibility. Therefore, the solutions of switching to efficient cooling and renewable or clean energy in data centres should be explored by facilitating cooperation between sectors, including the financial sector.

Besides improving the cooling systems and being prepared for the increased energy demand, other risk-mitigating actions can be suggested. Optimizing the operation of not non-stop services and re-scheduling the operational tasks may save energy in data centres. It has been calculated that approximately 8-19% of the energy could be saved using the application signatures method (Salinas-Hilburg et al., 2022) when optimizing the scheduled running of tasks.

Furthermore, it has been demonstrated that the location of the data centres affects the energy spent on cooling. Due to the lower temperature in Northern Europe, less energy is needed for cooling the buildings (Avgerinou et al., 2017). Cooling is one of the crucial factors behind the emergence of the data centre industry in Nordic countries (Saunavaara et al., 2022). This information suggests several courses of action for the operators of essential services.

1. Northern Europe should be considered when either building new data centres or buying data centre services. The location of data centres supporting essential services should be considered very important, besides the redundancy and internationally recognized certificates on operation continuity of the course.

2. A data centre in Northern Europe seems to be an advantage for the operators of essential services. Reallocating virtual machines from a data centre to a Northern one before summer or the warmer periods may decrease energy consumption (Arshad et al., 2022) and mitigate the risk of significant disruptions caused by power outages.

One possible implication of these findings is that protecting data centres from global warming is the biggest concern within critical infrastructure protection. As pointed out, the most critical issues are cooling and providing the necessary energy. The following section summarises the key findings by discussing the three main areas affected by global warming (managing climate change risks, outsourcing and supply chain failure caused by extreme temperatures; daily operation during heatwaves).

---

4 According to the Article 16 of EU Regulation No 1093/2010 authorities and financial institutions make every effort to comply with the EBA guidelines.
5. Discussion & Conclusions

Global warming and heat waves significantly affect our European society by jeopardizing our health and infrastructure. Therefore, there is an urgent need to address the effects of global warming on critical infrastructure protection. Our research set out to assess the effect of heatwaves on critical infrastructure protection and provide an overview of the responses of the European banking industry. This study cannot encompass all the action plans and climate change-related strategies; the focus has been put on the risks related to the infrastructure rather than the financial risks.

The following conclusions can be drawn after examining the challenges raised by global warming and heat waves.

1. Assessing the sector's climate risk preparedness with the first climate risk stress test in Europe, ECB has found that banks do not yet sufficiently incorporate climate risks into their risk management. Preparing for the unique challenges raised by global warming is essential. Therefore operators of essential services are urged to enhance their risk management framework with climate risks.
2. The stress above test results showed the need for sector-wide cooperation to sharpen the focus on climate risks and create the best practice in the financial sector.
3. The strategy should be developed to manage possible supply chain failure or the outage/degradation of outsourced services due to extreme weather events. Sector-wide cooperation should be considered, or if it makes sense, cooperation based on locality rather than sector. Such cooperation could be fostered by an appropriate risk reduction programme driven by the EU.
4. Cutting back the emission of greenhouse gases is a shared responsibility. Therefore, the solutions of switching to efficient cooling and renewable or clean energy in data centres and office areas should be explored. Partnership with the academic world may enhance the development of efficient and green energy technologies.
5. Addressing the increasing vulnerability of data centres providing essential ICT services is fundamental. It has been shown that cooling data centres in Northern Europe need less energy than in areas hit by heatwaves. European operators of essential services may cooperate in moving some critical infrastructure elements to the North.
6. As heatwaves affect everyone, a comprehensive test could be organized for operators of essential services involving third-party service providers and relevant authorities.

The most prominent finding to emerge from this study is that climate risks must be addressed, and strategies must be developed to help our European society deal with the inevitable effects of global warming. No doubt, sectors providing essential services are also responsible for saving our planet.

References


17


**Author Contributions:** Conceptualization: Nagy, Somogyi; methodology: Nagy, Somogyi; data analysis: Somogyi; writing—original draft preparation: Somogyi, writing; review and editing: Nagy; visualization: Nagy, Somogyi. All authors have read and agreed to the published version of the manuscript.

Tamás SOMOGYI is a student at the Doctoral School on Safety and Security Sciences, Óbuda University, Budapest, Hungary.

**ORCID ID:** https://orcid.org/0000-0003-1397-697X

Colonel ret. Rudolf NAGY PhD is a lecturer, Donát Bánki Faculty of Mechanical and Safety Engineering, Óbuda University, Budapest, Hungary. He was a CBRN defence officer, and took part in industrial safety tasks. He gained experience as an operations officer in the NATO SFOR mission. After that, he became Deputy Head of the Emergency Management Department of Hungarian National Directorate General for Disaster Management. He has been teaching subjects of safety and security sciences since 2015.

**ORCID ID:** https://orcid.org/0000-0001-5108-9728
IMPLEMENTATION OF EFFECTIVE SOLUTIONS TO THE CRISIS TASKS AND ITS REGIONAL MANAGEMENT *

Matúš Grega ¹, Pavel Nečas ²*

¹ Armed Forces Academy, Demanová 393, 031 01 Liptovský Mikuláš, Slovakia
² Matej Bell University, Kurzmányho 1, 974 01 Banská Bystrica, Slovakia

E-mails: ¹ matus.grega@aos.sk; ² pavel.necas@umb.sk (Corresponding author)

Received 18 August 2022; accepted 20 October 2022; published 30 December 2022

Abstract. Because of unexpected negative social and natural phenomena, there is a growing need to prepare crisis structures of society, also due to technical and technological progress, which directly affect these facts. The facts defined in the article encourage the social requirement for practical, creative, and interactive training of individuals and teams in the public interest and public administration. The need for adequate preparation should reflect the logical and systematic analysis of dependencies and contexts in all procedures and processes of identifying negative phenomena to encourage and create thought stimuli of crisis staff, which were able to minimize losses and maximize opportunities. The potential of individual types of simulation technologies and their simulation tools is a possible means to achieve the already mentioned requirements for practical training of crisis staff.

Keywords: crisis management; crisis manager; education; simulation technology; financial calculation

Reference to this paper should be made as follows: Grega, M., Nečas, P. 2022. Implementation of effective solutions to the crisis tasks and its regional management. Insights into Regional Development, 4(2), 21-35. http://doi.org/10.9770/IRD.2022.4.4(2)

JEL Classifications: I18, J28, K10, K22, K37

Additional disciplines political sciences; sociology; information and communication; informatics

1. Introduction

Society and the individual, which is an integral part of it, have constantly encountered security conditions that affected him existentially throughout his existence. From a historical point of view, it is observable that each society was purposefully preparing itself, in a certain way, for already existing and potential threats. Society was able to face, to some extent, avoid, eliminate and even predict the specifics of security conditions. Hazards arising from natural patterns, such as windstorms with torrential rains and subsequent floods, devastating fires, earthquakes, etc., but also direct anthropogenic threats, such as war clashes, industrial cataclysms and catastrophic and apocalyptic manifestations thereof, were accompanying features when the individual and society learned to

* This work was supported by the Slovak Research and Development Agency under the contract No. APVV-20-0334.
manage and control the patterns arising from them, they were forced to face them directly. They had to take measures to ensure historically conditioned security. (Blažek, 2014)

Following the threats, the field of public administration has structurally and in terms of content created and gradually became an essential tool of management function. This management function was used and shared by the elements of social system management, which ensure the operation of the society in all its areas of existence and represent public administration and the public interest. The facts mentioned above directly affected and, in a way, created the conditions for actions and works associated with the management function of responsible individuals or groups to begin to fulfil the tasks that result from the specific functions of crisis management as a functioning unit. (Hašanová, Dudor, 2014)

Ensuring basic internal security processes and areas, such as defence, protection, external and internal security, and security of the social apparatus (state or individual) represents important attributes that activate, stimulate and thus directly affect the components of public administration. From a historical point of view, analysis of knowledge and experience and scientific research, the fundamental pillars of public administration - the authorities, in terms of opinion, and legislation, jointly or separately, participated in creating and ensuring the security of states, their citizens and other elements that make up the security environment. The development of a public administration system is an ongoing process in which ties are permanently managed, responsibilities updated, new relationships defined, competencies and implemented administrative national and international standards, the application of which is subject to a specific component of public administration, are strengthened. Despite the constant improvement of ties, procedures, and processes in matters of communication, decision-making, control, planning, executive, execution, etc., historical development in the area has failed to prevent the emergence of newly created crises, phenomena and situations that public administration authority has to face and for which they must be well prepared.

In these intentions, it is necessary to realize that crisis phenomena and situations have been and will continue to accompany individuals and communities. They are a particular representative of the interaction of the individual with society, the mutual exchange of societies and individuals, and humanity's interaction with its external environment. This is the impact of the adverse effects on the individual and community. Their specifics and common denominators are:
- complicated prediction,
- destabilization of social and natural balance,
- destabilization of system subjects and their subsystems,
- a direct threat to individuals, society, and their mutual ties,
- Destruction of cultural, social, and historical values.

Crisis phenomena directly but indirectly affect the economy, politics, and infrastructure of the community negatively; they also affect the lives of individuals, biosocial groups, their health, and material goods and, to a significant extent, negatively destabilize the natural and socio-social aspects of an individual's life. Experience shows that the invoice value of lost or degraded goods, health and capacity constraints of logistics infrastructure represent only a visible part of the costs and losses of companies. (Korecki, Adámková, 2020)

2. Crisis management as a complex system

The concept of crisis management in a regional public arena has resonated for a long time. It has been the subject of frequent scientific controversies and interpretations in the theory of law, politics, economics, and the community. One may consider the following entities as being relevant elements of the system.
Public administration

In general, an opinion has been settled that public administration consists of two primary components, namely state administration and self-governance, which realize public interest activities. Its main features are activity, initiative, determination, and focus on shaping the future while complying with the law and learning general tasks. Public administration can also be understood as a complex of methods, measures, ways and procedures that put the mechanism of economic policy in place, ensuring its objectives, regulation and development. This clearly shows the interrelationship between economic policy and public administration, as economic policy cannot exist without public administration and public administration loses its importance without economic policy. In this context, a more precise definition by the authors Strecková and Malý has a deeper meaning and substance: “The public administration is represented by the performance of self-administration and state administration, because the sector is financed from public finances and public finances are formed from sources of the population, there is no entity available for public sector management other than the public administration”. (Strecková, Malý, 1998, p.25).

Security system

The security system is representative of a complex of elements and means to ensure the safety and protection of lives, human health and property. It represents an integrator of elementary beams of security, institutional and system elements for its security. It is created based on risk identification, analysis, evaluation and taking the following measures to ensure security. Using the method of multitude quantitative risk analysis using mutual correlation, the dependency of the management structure of the selected critical infrastructure area is then determined, which is expressed by the mutual relation of identified risks. (Korecki, Adámková, 2018)

The created security system must represent a set of applications to have the qualities, which will enable it to ensure different levels of survival and functioning of the social whole in case of expected or unexpected threats. To ensure the required basic tasks, this system is designed to be open, flexible and evolving, with multiple interconnected networks of relationships in individual levels and levels of system and subsystems subjects, which allows it to respond adequately, to react to potential as well as natural threats (Ivančík, 2019). The primary qualitative attributes of a social security system include, for example, the ability to: analyze the security environment; identify and classify national security risks and threats and trends; identify procedures and measures to prevent or eliminate security risks and threats; ensure their continued readiness and ability to act (Ivančík, 2021); identify crisis solutions in line with current resources and capacities; ensure effective management of forces and aids in prevention, readiness to manage crises, adequate reactions and responses to the activated source of threat during an emergency, to mitigate and eliminate the consequences of situations and to create conditions for the renewal and development of society. Many of the tasks and requirements placed on the security system and its creation were often carried out in contrast to the situation, that the readiness of the system to act, to be adequately prepared for real as well as potential threats, could only be verified during an actual emergency or crisis. Experimenting with the security of society, states, and humanity is risky and possible only to a limited extent, so attention in this area has also focused on using crisis scenarios.

The gained experience and knowledge from personal crises of human society - natural disasters (floods, volcanic eruptions, earthquakes, extreme climatic fluctuations, etc.), technical and technological accidents (accidents of nuclear power plants, chemical plants, etc.), war conflicts (world and local war conflicts, the possibility to use weapons of mass destruction, etc.) were scientifically systematized. Attention was paid to the use of crisis scenarios in the crisis management system of public administration at all levels and areas.

The assumption that crisis scenarios are an effective and inexpensive tool for "managing" the adverse effects of crises, protection of lives, health and property of citizens and society, environment, defence and security interests has been confirmed. The growth of the close connection between the gained experience and scientific knowledge increased the importance of crisis scenarios for society. They became an essential part of theoretical and practical
activity in the security field, an important starting point for security policy and the creation of national security systems. In the current understanding, crisis scenarios result from recognizing the need to prepare society and individuals for threats and crises, the starting point for creating threats and problems identification tools, and mitigation or elimination of their adverse effects.

**Crisis scenario**

In the theory and practice of crisis management in public administration, we are increasingly facing the issue of crisis scenarios, which are the result of knowledge and risk assessment, as well as the development of the security environment. Crisis scenarios in public administration are based on historical experience and the degree of knowledge of their nature. The development of scientific knowledge in fields such as security risk theory, sociology, public administration theory, etc., enabled a deeper penetration into the essence of security, processes and mechanisms of ensuring it (security) as much as possible. At the same time, it enabled the transfer of scientific knowledge into crisis scenarios.

Throughout historical development, humans and their various social communities have encountered multiple forms of danger, threat, risk and the resulting crises. The historical stage of development of human society determined to what extent a particular community consciously prepared for natural as well as potential forms of threat and was able to prevent, confront or mitigate their harmful effects. Man, and human society, throughout their lives, have always encountered specific states of the security environment that have influenced it existentially as well as security, but also which he created by himself. In the historical development, emerged, structurally and in terms of content grew the field of public administration, which became a type of management activity by which the subjects of social system management carry out the course of the society in all areas of its existence and which constitutes public administration and management in the "public interest."

The attention of society was focused on the primary sources of danger, threats, risks, and the resulting crises, which it most frequently encountered in its historical development and whose destructive factors and extent was the greatest. Initially, supernatural forces were considered sources of threats depending on the development of science and human knowledge. They were gradually replaced by an approach that accepts the knowledge of natural and social pressures and their development patterns. Based on the gradual scientific penetration into their essence, three large groups of threat sources were created in the system of opinions: a congenial group of natural sources of threat, a congenial group of civilization sources of threat, and a disparate group of their combinations. We also encounter a simplified approach that presents the dangers, hazards, risks, and crises resulting from them in two large disparate groups, namely non-military and military threats. These are understood as:

- **Non-military threats** are real destructive potencies, the activation of which can lead to significant deterioration and weakening of the economic and security function of the state due to the action of natural, economic, internal, and international forces. Non-military threats must also be understood and analyzed in terms of mutual determination, accumulation, and the possibility of the emergence of an individual threats chain reaction, which may result in larger-scale threats. Non-military forces and means are used in the first place to eliminate them.

- **Military threats** represent such groups of activities of a state or coalition of states or organizations (especially terrorist ones) that endanger the security of coalitions of states or the state, comprising mainly the immediate use of military forces and means, as well as the mediated indirect use of military forces and means. The military threats associated with armed violence have become an area of exquisite growth in the historical development of humanity, even though it reflects a sad place in our history.
These groups of threats also correspond to crises:

- Non-military crisis - a crisis caused by the impact of natural disasters, technical and technological accidents, as well as the action of social forces. This type of crisis is in the legal norms of our state expressed in terms of an emergency state of emergency.

- Military crisis - a crisis that arose because of the threat of use or use of military force, which is expressed in our legal system in terms of the state of war, war.

Crises have different durations and intensities of harmful and destructive phenomena and affect different sizes of territories and numbers of people. They are the carriers of processes with a high degree of uncertainty for society and man. They hurt various areas of society; they can accumulate several random phenomena and trigger other destructive cycles. The core of the crisis consists of unexpected negative phenomena and their impact on the human, social system, social infrastructure, and the environment.

**Crisis management in public administration**

Crisis management in public administration represents a specifically hierarchical and fully functional system of public authorities and organizations, their ties, where their competencies and responsibilities are set. The output of this system is the implementation of approaches, opinions, experiences, recommendations, measures, and decisions. The general terminology about crisis management refers to solving various undesirable conditions. The first use of the term crisis management is highly debatable, but consensus prevails in the area where it was applied for the first time, and that is politics. In the 1970s, when the term crisis management became more widely used, the term began to be used in theories of business economics. However, its definitions differed, and the interpretation took different forms. Howard Chase and his colleague Barry Jones were the first, who try to define the term. They described crisis management as "an unsettled matter which is ready for decision" (Chase, 1984).

Today several definitions depending on the field in which it is applied (Bilczak, 2021; Rak et al., 2022). However, the essence remains unchanged. Crisis management represents a specific form of management that has a high degree of priority, and its main tasks are (Šenovský, Adamec, 2004):

- crisis state prevention,
- crisis state overcoming.

Prevention means undergoing critical processes using appropriate forecasting, analytical and management methods that allow the system to be protected and do not allow these essential processes to expand. Overcoming means a precisely oriented, thoughtful, and targeted reaction that recovers an incompetent, destroyed, or damaged system.

**Public administration versus crisis management**

In terms of universal definitions, it is possible to derive the meaning of crisis management for public administration. Crisis management in the context of public administration is understood as a set of tasks and measures that the administration performs independently or in cooperation with other organizations before, during and after crises to ensure the protection and security of the population. In a broader sense, these are:

- preservation and maintenance of public administration functionality,
- maintenance of population health,
- ensuring the availability of vital services to the population,
- ensuring vital material,
- securing private and public property,
- ensuring public order,
- preservation of cooperative activities of state rescue and security forces,
- ensuring informedness and
- provision of humanitarian aid.
Crisis management in public administration represents a specifically hierarchical and fully functional system of public authorities and organizations, their ties, where their competencies and responsibilities are set. The output of this system is the implementation of approaches, opinions, experiences, recommendations, measures, and decisions. Thus, crisis management represents a complex task whose primary goal is correction, prevention, contraction, reduction, and reconstruction. The problematic tasks and their goals should be based on a holistic point of view, i.e., that crisis state prevention and preparedness are both part of the process and partial tactical steps taken until the crisis occurs. This complex of tasks and their goals represents a closed circle because, after the elimination of a crisis state, there is a period of preparation for the future crisis state. (Figure 1)

![Fig. 1. Complex tasks and their goals by period](image)

A specific group of managers, namely crisis managers, was also identified by introducing and defining the term crisis management. Tasks connected to crisis management performed by a crisis manager are not usually a decision of one person but rather are served by several crisis managers (experts in specific fields) who are members of the crisis staff. The crisis staff implements complex tasks and measures before, during and after the crisis.

The security system, as a representative of a complex of elements and means in solving the tasks of crisis management in public administration, must have elementary attributes without which such a system would not work:

- Broad-spectrum and centrality define the participatory approaches of individual components. The components are in coherence with each other, where the structure, hierarchy and responsibilities are precisely defined, and these components are effectively influenced and guided.
- Legitimacy and legality, which are defined by the law order of the Republic, by respecting the results of regional and parliamentary elections, as well as by applying international standards and obligations.
Control and transparency exist where effective feedback mechanisms allow for evaluating the taken decisions and the planned decisions, especially in the competencies and responsibilities of crisis managers.

Versatility is representative of the continuous functioning of the system in both crisis and "non-crisis" states.

Functions of crisis management of public administration authorities functionally (functional dimension) carry out a complex of tasks. These functions represent another elementary framework that cannot be neglected and whose primary goal is:

- **Correction** - which means minimizing the resources or causes that precede the crisis state.
- **Prevention** - which means preparing for action in a crisis state. The need to apply preventive measures, renewal and revision of emergency, alarm, crisis plans, joint operational procedures, etc.
- **Counter(re)action** - active involvement in the emergence and gradation of the crisis state. The competent authorities need an erudite and practical approach based on early warning, notification, activation, analysis, and adoption of appropriate measures and decisions.
- **Reduction** - minimizing the outbreaks of crisis states and eliminating their adverse effects. The executive needs a coordinated approach and an active process by the managing authorities.
- **Reconstruction** means eliminating the consequences of a crisis state and returning to the state before the crisis state (stabilized condition). The need to stop the effects that caused a crisis state and take such decisions and measures that ensure that the recurrence of the crisis state is minimized.

The institutional dimension of crisis management of public administration authorities is presented as a competently integrated system of elements of crisis management. It is divided into management and executive components, which have a certain economic autonomy.

**Crisis staff and crisis manager**

The crisis staff is a working body of the statutory crisis management body serving their need to coordinate the intervening components and other elements in the cooperation of the components of the integrated rescue system in joint intervention and in resolving the crisis state. An inseparable assumption for ensuring the correct and efficient functioning of the crisis staff is its erudite and well-prepared staff.

The crisis manager, as a person, a member of the crisis staff, should be a representative of an effective, efficient and vigorous response, a representative of reception, analysis and correct evaluation of information and reports, as well as a representative of rationality, tolerance and complexity, when resolving a specific crisis state. The members of the crisis staff should form a group of erudite people who are efficient and united, where each member-element knows what his duties and competencies are. He is also aware that in resolving a crisis state, he must make quick and especially correct decisions under the stress factor of not only time. These reasons also represent a constant need for effective and efficient training for crisis managers. Training crisis managers is one of the essential elements to acquiring the necessary expertise and technical and technological readiness for solving potential crisis states and their management. (Grega, 2014)

The area crisis managers training should be prepared to prevent the occurrence of possible crisis states, ensure preparation for dealing with potential emergencies, ensure a well-founded solution to troubles within their competence at individual levels, solve tasks from superior authorities, also implementing measures to eliminate consequences of crisis states and the distribution of instructions to the executive elements involved in this process, is at a superficial level.
Crisis manager competencies
The abovementioned circumstances and facts point to the need for well-prepared staff. Nevertheless, what makes a crisis manager the right element in this system, what competencies should he have, and what negative factors should he eliminate when dealing with crises?

Decision-making is the most important function of managers at all levels of management, and all managerial procedures include decision-making. Decision-making, acceptance of a decision represents the principle of choice in the presence of several options by which a common goal can be achieved. Bad decisions by the crisis staff are often the result of an inability to consider all the contexts and doubts about the assumptions. The decision results in a consensus. However, it is acceptable only if it has not been adopted very quickly, given that unanimity does not prevail in the crisis staff because it suppresses the motivation to evaluate alternatives realistically.

The public administration crisis manager must consider the worst-case scenario when dealing with a crisis state. He must be adequately prepared for this state, and whereas he works in the public interest, he is the first to prevent the spread of panic, and he must be the first to inform the population and the media. The crisis manager takes preventive measures in the public interest to prevent the gradation of negative phenomena and must be able to evaluate what is a priority and paramount in a given situation.

3. Simulations

People want to be informed and prepared for the possible existence of incoming threats that could affect families, municipalities and cities, districts, or entire countries in any way. How to protect lives and properties, societies and their values? These denominators characterized the perception of security and readiness to respond correctly. The level of protection of specific fields, such as demography, the economy, the social and environmental field, and the political, reflects every society. In an unexpected, unplanned emergency, society must take appropriate measures to minimize negative phenomena (Cayirci, Marincic, 2009). The forces and means that must be used to protect population, property and social values are elements of the national security system. These elements must include all managerial, executive, operational and other components that must have different skills, such as the ability to communicate, make effective decisions, cooperate, cope with stress load, etc. Maintenance of these skills requires continuous and regular training using a variety of training methods based on simulation technologies. (Žentek, Nečas, 2020)

Simulations from the beginning, when implemented into the education and training environment, reflected the need to increase the effectiveness of staff training (from individual to coordinated staff) for events and situations that are very difficult to implement - to prepare in a natural environment. The benefits resulting from the implementation of individual types of simulations and their tools point to the growing interest and social need for their usage in the process of staff training in the field of management, coordination, and cooperation:

- improving the decision-making and planning process of staff (both civilian and military) in the preparation, planning and management of operations,
- obtaining alternative and backup solutions,
- risk reduction,
- saving staff capacity and time,
- minimizing adverse effects on the ecological environment, and
- saving costs and materials.

Any simulation of activities in terms of military and non-military operations is carried out on an appropriate, specific model based on statistical, mathematical, analytical methods, or a combination thereof, and on the widespread use of modern computer and information technologies. The method and methodological procedure of
its implementation depend mainly on the type of simulation used and the parameters of the created model at which the simulation is performed. Simulation has many advantages that encourage its constant development and expansion into new spheres and areas. Influenced by the development of the information-communication base, its potential is constantly enhancing. At present, simulation technologies have fully established themselves in the educational and training environment in the form of:
1. instrumented simulation (also referred to as live simulation),
2. virtual simulation,
3. constructive simulation and
4. distributed simulation, resp. blended simulation

Of course, there are many divisions and views on simulations. They can be classified according to approaches, according to the capacity of primary and secondary training elements (staff, tactical, operational, etc.), according to calculation methods (deterministic, stochastic, static, or dynamic), or also according to the end-user determination (military, police, medical, etc.). The simulation generally consists of five main parts (Cayirci, Marincic, 2009), resp. should consider the following aspects:
− simulation of performances that are performed in the simulation,
− models that provide real representations,
− input data that provide model definition,
− a visualization that shows the results of the simulation,
− communication that serves as a data interface.

Constructive simulation
Constructive simulation is the most used, in terms of time, the longest applied and the most used simulation, with a universal utilization and a wide range of uses. It is often referred to as the universal method. “The core of the constructive simulation is the usage of logical-mathematical models which are expressed usually by equations or systems of mathematical equations, inequalities (algebraic, differential, integral, etc.) and by algorithms or systems of closed or opened algorithms, while parameters (constant, variables) and interrelationships of these systems have deterministic or stochastic interpretation of material-energetic or thought-information objects and processes of armed encounter or conflict situation.“ (Rybár, 2000).

Today, computer technologies and software solutions are fully applied, where the relationships and dependencies of mathematical-logical interpretation are implemented. It is usually used in a distributed form, where several computing systems are connected via a computer network, or it can also be applied separately on a specialized hardware element. Constructive simulation is thoroughly used in training staff responsible for an operation or situation's planning and decision-making phase. In this way, the staff is trained in individual staff positions and management functions in the process of command, control, and verification of planning and decision-making activities.

A constructive simulation can also be specified as an artificial entity (model) representative of a man with actual behaviour, a natural technique, or specific units or wholes. Such a model is defined according to real data regarding its tactical-technical data, behaviour, and representation. In a constructive simulation, a real object (being, vehicle, technical device, system, living creature, etc.) is replaced by a model - an entity. It is, therefore, a simulation where synthetic beings, vehicles, systems, or technical devices move in a virtual environment and, depending on the simulated activities, also appropriately perform the assigned tasks. (Hubáček, Hausner, Vráb, 2013) Artificial entities behave according to defined algorithms, programmed in separate property classes (entity behaviour level). These models - entities based on interactions and defined semi-automatics can be divided into:
− deterministic - those that contain only precisely determining relationships (algorithms),
stochastic - those that have at least one relationship with the generated variable; other relationships are deterministic,

an entity with intelligence imitated by the properties of application programs, i.e., artificial intelligence.

Constructive simulation is based on mathematical methods and has many advantages that stimulate its constant development and expansion into new application areas in education and training. Like any other simulation method, constructive simulation has certain drawbacks and limitations. The main advantages of productive simulation methods include that (Grega, Bučka, 2013):

- it allows for simulating sources of threat and destructive processes of destruction during a crisis economically and efficiently,
- it allows the creation of an environment that is very similar to the real conditions of the emergency and the stages of the crisis,
- it will enable the simulation of a situation taking place in different geographical regions, in different climatic conditions, with various forces and means,
- constructive simulation models form important components also in other types of simulations and crisis management activities; they represent a universal unifying basis for all groups of simulations,
- reduce the damages caused by training from an ecological point of view, and the environment is not bothered by such training,
- its universality results from the applied mathematical basis,
- quantification of phenomena and processes enables their more accurate analysis in the decision-making activity of management authorities,
- it allows to model and simulation, at any level of similarity or with any degree of generalization, objects, phenomena and processes at the selected level (tactical, operational, strategic),
- it allows an objective expression of the influence of the terrain and other environmental factors on the processes of the crisis, as well as on the activity of crisis management,
- it gives the possibility of quality registration of the course of training and their use for a more objective evaluation of their results and formulation of conclusions and recommendations for theory and practice,
- the usability of constructive simulation in problem-solving crisis management is evident,
- it mediates the possibility of interconnecting constructive simulation systems with simulation systems of other categories,
- sets a measurable factor to determine the results of the operation of individuals and teams,
- it documents the course of solving the situation in 2D and 3D display,
- offers partial and complex statistical data,
- it records the planned and actual activity of the units,
- it allows searching for optimal solutions using forces and resources,
- it will enable crisis managers to take risks due to their own decisions.

As mentioned above, constructive simulation is usually implemented in a distributed form. Its basis is the implementation of the so-called Computer Assisted Exercise (CAX). Such form of training can be characterized as a sophisticated method of training commanders, managers, officials, and crisis staff members of the state, by which simulated processes and phenomena are potentially emerging or developing in carrying out the specified activity of the decision-making process, planning process or command and control process. Computers simulate processes and phenomena in the real or specified time and environment to give training entities the impression of performing real operations and activities as if they were performed in a natural setting and actual conditions. Constructive simulations and their computer-assisted exercises are excellent training and education in crisis management, mainly due to their versatility, broad application support and ease of implementation.
Simulation tools and their technologies are currently a global leader in training commanders, staff and participating units in their subordination and cooperation to solve crisis management tasks. There is no doubt that a well-prepared crisis staff and its members are the basis for the successful resolution of crises, primarily in matters of domestic crisis management and in cases of crisis management at the international level.

4. Training effectiveness and evaluation

The use of constructive simulation for staff and managers training will expand only if its usage is practical (especially more financially advantageous) compared to other training methods, by which the intervening components of crisis management are prepared to fulfil their tasks. The exact methodology for calculating the cost of conducting training by live simulation (training in natural conditions) and for the relevant calculation of the expenses of constructive simulation does not yet exist. The evaluation mentioned in (1) is focused on the content of the training and is not applicable to evaluate the financial demands of the training. However, a simplified model can be developed that can be used for comparative analysis in certain circumstances. The basis of the training effectiveness calculation model is the premise that in the analytical relationship for cost calculation are accepted identical variables characterizing a given phenomenon, process, or behaviour from both constructive and live simulations. (Žentek, Bučka, 2019)

The relationship for calculating the costs associated with the Ncv training can be expressed by equation (1):

\[ N_{cv} = N_a + N_b + N_t + N_p + N_n, \]

Wherein:
- \( N_a \) are the costs of using the machinery in training in [€],
- \( N_b \) are the costs of renting the premises for training in [€],
- \( N_t \) are the costs of using the technical means for intervention (action) in [€],
- \( N_p \) are the costs of supporting the training in [€],
- \( N_n \) are (for now) unspecified costs in [€].

The variable \( N_a \) represents the costs of using all types of machinery in training (passenger, transport, goods and special vehicles, aircraft, reconnaissance aircraft, etc.). Its value can be expressed by equation (2).

\[ N_a = \sum_{j=1}^{n} a_j \cdot l_j \cdot c_j \]

Wherein:
- \( a_j \) is the number of j machinery used in training,
- \( l_j \) is the number of kilometres the j machinery will cover (covered) during the training in [km],
- \( c_j \) are the costs of j machinery per 1 km of driving in [€/km].

Live simulation (training in real terrain) can take place on premises, the use of which is conditioned by payment for rent. The calculation of its value is based on the price per m² of the j area \( (c_j) \) multiplied by the total used j area \( (P_j) \), or the total invoiced value is substituted into equation 2.
The \( N_b \) calculation can be performed according to equation (3).

\[
N_b = \sum_{n=1}^{j} P_j \cdot c_j
\]

Wherein:
- \( P_j \) is the area of the \( j \) land used for training in \([m^2]\),
- \( c_j \) is the price per 1 m\(^2\) of the \( j \) land in \([€/m^2]\).

The costs of using technical means \( N_t \) represents the sum of funds, that must be spent for their use in training. The category of technical means refers to, for example, power plants for special power supplies and for powering other means (communication technology, computing technology, etc.). Equation (4) can be used to calculate these costs.

\[
N_t = \sum_{n=1}^{i} t_i \cdot c_i
\]

Wherein:
- \( t_i \) is the time of use of the \( i \) technical means for training in hours,
- \( c_i \) is the price of using the \( i \) technical means (operating costs) per 1 hour \([€/hour]\).

The costs for supporting the \( N_p \) training in Equation 4 represent the estimated or spent funds needed to support the training. This group can include, for example, the fees for training figurants, simulation system operators, etc., i.e., for people who are not trainers but whose work is necessary for the course of the training. These costs can be calculated using equation (5).

\[
N_p = \sum_{n=1}^{i} t_i \cdot c_i
\]

Wherein:
- \( t_i \) is the total time of deployment of \( i \) person to support the training in \([hours]\),
- \( c_i \) is the \( i \) person's hourly fee in \([€/hour]\).

The \( N_n \) costs represent unforeseen, unpredictable, additional costs that may arise during training. This category of costs can include costs for dealing with damage caused to materials, machinery, property of the population, etc., in direct and indirect connection with the training.

A coefficient \( k \) can increase the total cost of the training or its components \( N_x \), wherein \( k \) is the coefficient of tolerance of the population and its value is in the range \([k \geq 1]\). The coefficient \( k \) is dimensionless, and the value "1" means that the training is not perceived negatively by the population that does not participate in the training. The population participating in the training is included in the variable \( N_p \) (for example, in the category of figurants). If the population does not tolerate the training (because of restrictions, etc.), the coefficient \( k \) is greater than 1. The limit of the coefficient cannot be determined precisely yet, and its value will be estimated.
The overall effectiveness of the training is then evaluated as the ratio of the financial costs of live simulation training and training performed by constructive simulation. The verification of the described model was not exactly verified in the training conditions. Only an empirical comparison of training by constructive simulation with the equivalent of exercise performed in natural conditions (e.g., command and staff training in the military training area) was performed. The calculation showed that the training performed with the support of constructive simulation reaches the cost of about 10% of the cost of the training performed in real terrain. It can be assumed that in training on the components of crisis management using constructive simulation, a comparable value of training efficiency will be achieved. The costs associated with the catering for trainees and support elements, with their accommodation, the wage costs of trainees were not included in the calculations of the estimates, nor in the above model, as these costs are in many cases comparable regardless of whether the exercise is performed by constructive or live simulation. Accurate cost and training efficiency calculations are not yet possible because the data needed for the analysis are not available in 100% of cases.

Training using constructive simulation has been used by the armies of many countries for many years and brings results in the exercise of commanders and staff. Even in the case of this technology, it is becoming more and more widespread. In connection with this, its price is gradually falling, and the technology of constructive simulation is slowly beginning to be used outside the army. It is logical that the training of police and security forces, as well as rescuers, is another area into which this technology penetrates. Of course, the acquisition costs of simulation tools are not minor and will initially seem significantly higher than the cost of training in real conditions. Such activity cannot be replaced even by the perfect simulator, and the training of powerful elements must take place in a real environment with real equipment. However, only a limited number of people and machinery can participate in such training. Thus, lower management and performance elements are usually participating. A considerable complex training, with the participation of hundreds or thousands of intervening persons and dozens of special machineries, cannot be prepared in real conditions. Therefore, even the highest levels of management do not have adequate resources with an excellent response to their training. For these levels, the training using constructive simulation tools is tailor-made.

Individual training can be modified; it is possible to prepare various scenarios, move in space and not limit regular operation. For essential functions, live training remains the most important. Training using constructive simulation provides new opportunities in an almost real environment and expands training opportunities to an unprecedented extent. It allows preparing new members who can make mistakes during training without the risk of life-threatening and causing great material or other losses. The participants of such complex training can thus prepare to deal with real crises. The simulation can be repeated countless times with different staff, compare various solutions and learn from each other, gain new information, learn, look for weak points in the plans and improve before deploying. As in the case of the training of military rescuers, in future, the usage of constructive simulation can be expected for firefighters and civilian rescuers, including members of the crisis staff. Well-prepared control components can better coordinate rescue operations and prevent loss of life and material damage. As shown by the experience of floods in recent years, preparedness for crises is one of the primary conditions for successful interventions in a large area. However, it is better to gather this experience in a simulator than in a natural disaster at the cost of loss of life, destroyed houses, damaged infrastructure and other damage, especially if the actual training using the simulation is much cheaper than the same training in terrain. (Žentek, Grega, 2017)

Conclusions

The system of regional public administration is responsible for the human individual, their potential and the action of a particular societal group. It is the one which is the bearer of progress and modernization. But despite optimization procedures and methods, modern methods of work, the individual or the whole is unable to act
effectively in a matter of crises management unless it is qualified, well professionally erudite, technically and technologically prepared to deal with potential concerns (Andrassy, Grega, Nečas, 2018). The functionality of public administration in connection with organizational and managerial problems is integrated into the principal theoretical and practical demands, which are constantly discussed in the past and present. Interventions in this system, in its organizational structures, tend to be fundamental and substantial because their functional impact has a broad scope in the life of society in all its areas. (Šimák, 2006).

Modelling and simulation, aimed at optimizing activities using experimentation with a created computer model built into a real synthetic environment, provide an effective and efficient form of preparation and development of the characteristics and behaviour of individuals and groups. They appear to be a progressive means of training commanders and staff to find solutions to crises and acquiring habits and skills in individual and particular activities. Simulation technologies are therefore becoming a significant subject increasing the preparedness of individuals and groups in the frame of regional public administration, as a key to achieving training priorities because they can realistically simulate activities and prepare personnel to deal with complicated non-military and military operations of national as well as domestic crisis management.

References


Blažek, V. 2014. Východisková tvorby krízových scenárov pre potreby vzdelávania na APZ v Bratislave (The starting points for the creation of crisis scenarios for educational needs at APZ in Bratislava). Bratislava: Published by Academy of Police Forces.


Grega, M. 2014. Vzdelávanie krízových manažérov s podporou simulačných technológií (Education of crisis managers with the support of simulation technologies). In: Bezpečnost, spolehlivost a rizika. Liberec: Technical University.


https://doi.org/10.20858/sjsstst.2020.108.9


Šenovský, M., Adamec, V. 2004. Základy krizového manažmentu (Basics of crisis management). Ostrava: Published by SPBI.


Žentek, M., Grega, M. 2017. Simulation tools in the armed forces interfaced by the operational design. Revista Academiei Fortelor Terestre. ISSN 1582-6384. 22(2 (86)), 134-140. https://doi.org/10.1515/raft-2017-0019


Funding: This work was supported by the Slovak Research and Development Agency under the contract No. APVV-20-0334.

Author Contributions: The authors contributed equally.

Matúš GREGA, Dipl.Eng, PhD. is the researcher in Simulation Centre in Armed Forces Academy in Liptovský Mikuláš, where he deals with the issue of simulation tools and their application to the environment of education and training of crisis management personnel. He is the author and co-author of many scientific and professional monographs, articles, scripts and studies dealing with modeling and simulation, security and crisis management. He is the author of many research reports of national and international science and research projects. He was 6 years active in the Science and Technology Organization of NATO in the Modeling and Simulation Group.

ORCID ID: https://orcid.org/0000-0003-1065-2355

Pavel NEČAS, Dr.h.c. Prof. Dipl. Eng. PhD., MBA is professor at the Department of Security Studies and the Vice-Dean at the Faculty of Political Science and International Relations of Matej Bel University in Banska Bystrica. His research is focused on political-military strategic issues, particularly in defence and security policy, international relations, and international security studies. HE is an experienced European researcher and the author of several monographs, books, papers and articles related to the political sciences, security and defence published worldwide.

ORCID ID: https://orcid.org/0000-0001-7743-0453

Copyright © 2022 by author(s) and VsI Entrepreneurship and Sustainability Center
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/

Open Access
PREDICTORS OF ENTREPRENEURIAL ATTITUDE AMONG ADMINISTRATORS OF SELECTED COMMUNITY COLLEGES AND UNIVERSITIES IN LEYTE

Jereco Jims J. Agapito ¹*, Liza Lorena C. Jala ², Rosemarie Cruz-Español ³, Anthony G. Esquerra ⁴

¹,²,³,⁴ College of Business Administration, Graduate School of Business, University of the Visayas, Cebu City, 6000 Cebu Colon Street, Republic of the Philippines

E-mails: ¹jerecojims_agapito@uv.edu.ph ; ²ljlala@uv.edu.ph ; ³cbadean@uv.edu.ph ; ⁴anthonyesquerra71@gmail.com

Received 15 June 2022; accepted 18 October 2022; published 30 December 2022

Abstract. The school administrators adopted an “Entrepreneurial Attitude” approach to run the school profitably. Furthermore, administrators who embrace successful entrepreneur techniques were able to attain and establish an effective school, as the literature says. Using a descriptive survey method, this study determined the factors that affect the entrepreneurial attitudes of school administrators and faculty in selected universities and colleges in the 4th congressional district of Leyte through model generation. Hierarchical regression analysis showed that cognition is the most dominant and significant predictor of entrepreneurial attitude, which generated two models. An emergent model can enhance the lens to see a better picture of what school administrators should prioritize in improving their school management strategy. A simple model equation can be a foundation for enhancing someone's entrepreneurial attitude. A critical takeaway from the simple equation model was that enhancing someone's cognition can help school administrators transform their entrepreneurial mindset and attitude.

Keywords: Entrepreneurial attitude; school administrators; hierarchical regression; aptitude; boundaries; cognition; Leyte; Philippines

Reference to this paper should be made as follows: Agapito, J.J.J., Jala, L.L.C., Cruz-Español, R. 2022. Predictors of entrepreneurial attitude among administrators of selected community colleges and universities in Leyte. Insights into Regional Development, 4(4), 36-51. http://doi.org/10.9770/IRD.2022.4.4(3)

JEL Classifications: L16, L53, L31

Additional disciplines: political sciences, sociology, psychology, educology

1. Introduction

The economy has always been an unstable environment with evolving demands for human needs and wants. This demand provides opportunities for the community to develop and create innovative products and services. For the past years, the 4th congressional district of Leyte has experienced economic development with various emerging businesses such as the existing agriculture, energy supply, and other company offering essential commodities. Entrepreneurship is widely recognized as a crucial component in creating job opportunities, economic growth, and the strength of small businesses and academic institutions.
One of the industry's driving elements is entrepreneurship. In education, entrepreneurship is critical in encouraging all students to become more entrepreneurial (Hegarty et al., 2006). In the Philippines, the CHED views entrepreneurship education as a way to train people to become employers rather than workers when they graduate from college. It is envisioned that graduates would be educated and equipped with the required values and abilities to take business risks. However, the route to HEIs providing entrepreneurial education has been paved with numerous obstacles, including opposing philosophies and pedagogies. According to De Ocampo, Bagano, and Tan (2012), the ability to produce "entrepreneurs" has yet to be scientifically verified, but it will be felt in the following years. Furthermore, the future generation will be able to comprehend the context and significance of entrepreneurship.

Based on the study of Gibb and Hannon (2006), entrepreneurship education, or a narrow approach, i.e. education in starting and running businesses, and enterprise education, or a broad approach, with a focus on abilities that characterize entrepreneurs, such as opportunity seeking, initiative-taking, making things happen independently, problem-solving and risk-taking, commitment to work and tasks, and ability to cope with or enjoy uncertainty and opportunity seeking, initiative-taking. In general, the comprehensive approach is aimed at younger learners to provide opportunities for kids to acquire an entrepreneurial mindset. Entrepreneurship is defined broadly in Sweden and other Nordic countries, focusing on the process at individual and public levels (Sjøvoll and Pedersen, 2014). There seems to be widespread agreement in various nations that entrepreneurial education and training should encompass broad and limited knowledge (Leffler, 2009). In certain countries, such as the United States, the United Kingdom, and South Africa, the focus is mainly on limited knowledge, despite debates about giving more attention to the broader understanding (North, 2002).

In the Philippines, the innovation and the imperative of entrepreneurship in the educational setting are redefined in Ambisyon Natin 2040. Several countries share this vision, such as Sweden, where the New Swedish Education Act (2010) proposes that the concept of knowledge be given a broader meaning so that education can promote children's and student's overall personal development toward becoming active, creative, competent, and responsible individuals and citizens. That might include the ability to be creative and see opportunities, as well as the ability to take the initiative and put ideas into action, all of which are essential for establishing an entrepreneurial mindset. Entrepreneurship education programs foster an entrepreneurial mindset and the desire to pursue a career as an entrepreneur in the future (Souitaris, Zerbinati, & Al-Laham, 2007). The development of entrepreneurial vision should be at the heart of contemporary tertiary education for teachers, school officials, and students.

The individual who undertakes this is known as an "entrepreneur" (Martin & Osberg, 2007). While there are various definitions of entrepreneurship and entrepreneur in the literature, Richard Cantillon originated the two terms in 1755 and was generally acknowledged. Cantillon defined an entrepreneur as someone who accepts every risk after creating a purpose to make money and links the entrepreneur to the risk factor (Korkmaz, 2012).

Besides, Azim and Kahtani (2015) expressed that entrepreneurial people have different abilities, for example, correspondence, association, initiative, navigation, perceiving amazing open doors, organizing, using time effectively, and stressing the board, while Cooney (2012), then again, states that they have different abilities, for example, inside the discipline, risk-taking capacity, advancement, change direction, and assurance. These abilities add to the foundation that upholds the innovative capabilities of school administrators in authoritative terms. Simultaneously, they establish the main impetus of progress and advancement in light of drives (Fernald, Solomon, and Tarabishy, 2005). The reception of these perspectives and capabilities by school chairpersons is significant concerning expanding the entrepreneurial limit of their institutions. For instance, as per Kirkley (2017), enterprising school administrators are people who are fearless, aggressive, step up, and feel energy and excitement. Moving these characteristics to learning conditions and fields additionally adds to supportability given business and creation for schools. Simultaneously, the authority domain with entrepreneurial perspectives
makes the consciousness of seeing and getting the potential changes caused by difficult circumstances beforehand and defeating existing constraints (Xu, 2020).

As Sajkiewicz and Pashiardis (2020) discussed, as far as training, school directors rouse authoritative change and development through the navigation and show their innovative jobs by attempting new techniques. Along these lines, acting with their enterprising jobs, the school chief can produce livelihoods and assets that will give the school their imaginative and innovative activities and assess these assets concerning academic purposes. Consequently, the school chairperson showing enterprising perspectives and abilities by leaving their usual range of familiarity and facing challenges, and going to the obscure and brimming with opportunities can separate this school chief from different principals (Sharma and Dave, 2011).

Toward this path, school directors' administration furnished with entrepreneurial perspectives empowers these pioneers to see valuable changes before others and handle these changes in an inventive and imaginative structure (Pashiardis and Brauckmann-Sajkiewicz, 2020; Brauckmann-Sajkiewicz and Pashiardis, 2022). For instance, as Pashiardis and Baker (1992) did not entirely set in stone in their trial research, issues that emerge in schools can be transformed into an opportunity to the school executive. Nonetheless, to foster the entrepreneurial perspectives of the staff and faculty, these administrators can offer their schools adaptable environs that are accepted to make a growth opportunity for themselves and their staff by enduring the different enterprising mix-ups and disappointments of the staff (Pashiardis and Brauckmann-Sajkiewicz, 2020). These are extremely basic for school administrators to govern schools effectively.

Every part expects compelling schools of society, whether the part of the school's local area of understudies, educators, organizations, chiefs, and agents or the completely local location from a more extensive perspective (Rasmita and Shadma, 2009). According to Gronn (2003), the number and character of graduates who are prepared for graduate school can start new businesses and are not paid on a month-to-month basis are given special consideration by the school district.

The role of the administrator is to manage various academic and non-academic activities. This is not limited to merely collecting school in daily operational activity, but also administrator embodied entrepreneurial mindset. An entrepreneurial attitude is needed for people doing business/entrepreneurship. A business or company must have the essential entities Owner, Managers, Workers, Products, and Clients/Customers to carry on or operate. From an outer perspective of a university or college, the owner is either government or private, managers are the school administrators and principals, the workers are the faculties (Teaching and non-teaching staff), the products are the students, and lastly, the Customers are the Industries. From the inner perspective, the owner is either Government or the Private, the managers are the school administrators, the workers are the faculties, the products are the programs, and the customers are the students.

Therefore, agreeably, the school administrators running a university or a college are not different from those running a company. Desravines (2014) who stated that the CEOs of schools were school administrators. School directors are like CEOs; they work decisively to enlist the perfect individuals, foster compelling faculties at all levels of their association or department, guarantee that their representatives are upheld and considered responsible for outcomes, foster frameworks to advance productive tasks, and develop a positive culture zeroed in on their vision for progress and create an effective school.

2. Formulation of the Conceptual Framework

The entrepreneurial attitude was introduced by Robinson et al. (1991). They underlined that attitude is better to describe entrepreneurs than personality traits or demography. They focused on achievement, self-esteem, personal control, and innovation. However, the research was intended for students' entrepreneurial propensity. Some
studies didn't include demographic profiles either and used other factors such as the need for achievement, autonomy/independence, creative tendency, the calculated risk taken, drive, and determination in determining the entrepreneurial attitude (Abun et al., 2021). Other researchers like Rohitha Rosario & David Potts (2016) used factors such as innovation characteristics, opportunity seeking, and risk-taking. They included the demographic profile as a factor in determining entrepreneurial attitude; however, based on their results, it has no significant relationship or is not a determinant of an entrepreneurial mindset. Those studies tried to include the demographic profile using multiple regression and interpreted it individually. This study attempted to have the demographic profile as one of the factors of the entrepreneurial attitude of school administrators solving collectively simultaneously. In this study, the researcher focuses on school administrators of selected universities and colleges in the 4th congressional district of Leyte with the factors that may impact their entrepreneurial attitude. The factors, which are the byproduct of an entrepreneurial mindset, are the following:

Demographic profile - the school administrator's socio-demographic profile, such as age, gender, civil status, years of service, affiliation, and academic rank

Skills and knowledge – Entrepreneurial skills and knowledge school administrators should be based on the study of Carpenter (2008), which are the following: commitment, determination, and perseverance; the need for achievement; opportunity orientation; internal locus of control; tolerance to uncertainty; risk-taking; innovation and creativity; optimism; leadership; and ability to direct their future.

Motivation – In motivation, this study will adopt the Herzberg Theory. The faculty, as per the Herzberg hypothesis, has two degrees of requirements: "sterile" needs, which are related to fair treatment in regions like organization strategy, managerial practices, and oversight; and "persuasive" needs, mainly concerned with opportunities for achievement, recognition, responsibility, and advancement. Employee requirements are met by both "hygiene" and "motivator" variables, but "motivators" offer the motivating power that may lead to higher performance.

Cognition – The study of Inbar (1980) will be adopted to understand school administrators. The cognition preferences of school administrators have also termed "memory for specific facts or terms", namely, application, recall, principle, and questioning.

Organization boundaries — Administrators' boundary actions alternate between internal and exterior school-related concerns. Druskat and Wheeler's (2003) typology in the school setting to get a complete knowledge of school administrators' activities, including border management as a foundation for examining school administrators' internal and external border activities. Internal relating exercises, for example, perspectives, such as building trust among workforce staff and showing decency to school individuals in the dynamic interaction; exploring for data about qualities and shortcomings, as well as the troublesome issues personnel staff face; chairmen start correspondence with school staff to acquire data about inside occasions, encounters, requirements to distinguish and explain data that may be helpful to the dynamic interaction; convincing activities, which incorporate designating authority, practising adaptability with school staff choices, and training; and Empowering exercises, which incorporate assigning authority, practising adaptability with school staff choices, and instructing.

From the study of Kirby and DiPaola (2011), outside relating exercises are those where the school director addresses the school to outer partners to get sufficiently close to and secure assets and backing, as well as supervise the outside business matters for data collection or occasions that could impede or assist the school with accomplishing its targets. Building positive associations with outer partners, for example, understanding their power designs and game plans, as well as keeping up with positive associations with student guardians and the local area; collaborating and helping out essential networks beyond the school; limiting traversing exercises can safeguard the school from information excessiveness by exploring and looking for facts from outside partners to
distinguish significant ecological happenings and imparting this fact to employees. Practical exercises include getting outer help for the school by going about as a channel and facilitator.

Family and social boundaries – The family and social boundaries adopted from the "Work-Family Boundary" study of Allen et al. (2014). It gathered several studies suited for the integration and segmentation of family, social, and work-related aspects, such as the extent to which individuals need to keep work and family jobs discrete, inspiration to get across work and family areas (Matthews and Barnes-Farrell, 2010) and how much individuals permit breaks starting with one job then onto the next (Kossek et al. 2012).

Figure 1.

Figure 1 shows that age, gender, civil status, years of service, affiliation, and academic rank are in the Demographics set as Block 1, and motivation, cognition, skills, and knowledge are in the Aptitude set as Block 2. Organizational and family, and social care in Boundaries set as Block 3. They are all interconnected to the dependent variable, entrepreneurial attitude. This study determined the factors that impact the entrepreneurial attitudes of school administrators and faculty in selected universities and colleges in the 4th congressional district of Leyte, which generated models. To summarize, it is proposed that:

H_0: There were no factors associated entrepreneurial attitude of the school administrators of selected universities and colleges in the 4th congressional district of Leyte.

3. Methodology

Samples. To test our hypothesis, the data were collected from the few private and state universities and community colleges within the 4th District of Leyte. There are 6 municipalities and 1 City within the research area. There were 22 out of 25 school administrators (University/College Presidents, Vice Presidents, Board of Regents, and Campus directors) and 87 out of 100 faculty members, which yielded a total response rate of 87.2 per cent. This rate was higher than the previous response rate reported for an online web survey (see, e.g. Robinson, et al., 1991, Abun, 2021).
Methods. This study utilized a descriptive cross-sectional survey design to determine the factors that determine the entrepreneurial attitude of school administrators and faculty of selected universities in the 4th District of Leyte. Setting inclusion and exclusion criteria to narrow the participants of this study. All administrators in key positions of their respective institutions were included in this study. Those faculties who expressed willingness to participate were also included. However, if they signified not to participate were automatically excluded. This study contained five (5) questions about the factors (knowledge and skills, motivation, cognition, organization boundaries, and family and social boundaries) intended for the school administrator adopted from several articles. And the entrepreneurial attitude questionnaire with Thirty (30) questions for both faculty and school administrator has been adopted from the Personal Entrepreneurial Competencies (PECs) questionnaire.

Measures. These survey questionnaires underwent content validity and reliability testing to verify if the instrument measured what it intended. With a rating scale of four, from 4 = totally agree to 1= totally disagree. The dependent variable was the entrepreneurial attitude, and the independent variables were demographic, boundaries and Aptitude. Hierarchical regression was used to measure which among the independent variables is the significant factor.

Ethical Considerations. Before the research began, the protocol was reviewed and approved by the university's Research Ethics Committee (REC). The researcher ensured that all information from the respondents was treated with strict confidentiality. Respondents' names were not reflected in the file for data analysis. Instead, only the numbers corresponding to their names in the source code can be seen. The source code had the researcher only. The gathered data was only accessible to the researcher and statistician.

4. Results and Discussion

The data obtained were analyzed using jamovi 1.6.23. Aside from its convenience, the purpose of using jamovi 1.6.23 was to test the relationship pattern between the variables in the research model by estimating the parameters in the outer and inner models. Based on the result, school administrators were dominantly male, and most of them were married. This is only an indication that for someone to become an administrator, they need to count many years of service. School administrators are not for youngsters; the more mature the person is, the more they become bold enough to become an administrator. Most of the school administrators are teaching; this is reasonable because most do not want teaching to be left behind since that's their first job and passion, after all. The results also revealed that demographic profile contributes significantly as a factor or a predictor in entrepreneurial attitude collectively and individually.

Table 1. Administrators’ Opinion about Organizational Boundaries

<table>
<thead>
<tr>
<th>Statements</th>
<th>Mean ± SD</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communicate general objectives to the member of the organization</td>
<td>3.18 ± 0.96</td>
<td>Agree</td>
</tr>
<tr>
<td>Set general organizational objectives</td>
<td>3.18 ± 0.96</td>
<td>Agree</td>
</tr>
<tr>
<td>Empowering personnel in various organizational activities</td>
<td>3.27 ± 0.98</td>
<td>Strongly Agree</td>
</tr>
<tr>
<td>Provide fair opportunities in the decision-making process</td>
<td>3.14 ± 1.04</td>
<td>Agree</td>
</tr>
<tr>
<td>Establish a positive relationship between the organization and stakeholders through collaborative activities</td>
<td>3.18 ± 1.05</td>
<td>Agree</td>
</tr>
<tr>
<td>General WM</td>
<td>3.19 ± 1.00</td>
<td>Agree</td>
</tr>
</tbody>
</table>

Respondents rated Likert Scale from 1 to 5. 1.0 – 1.75 SD is weak agreement, 1.76 – 2.50 SD is liberal agreement, 2.51 – 3.25 A is moderate agreement, 3.26 – 4.00 SA is strong agreement.

Administrators strongly agree that there is a need to empower university personnel in various organizational activities (Table 1). All the rest agreed that general objectives must be set and communicated to them clearly, which in a way, provides opportunities for them to make their own decisions in a way that establishes suitable relational activities among stakeholders. Delegation and empowerment in the workplace go hand in hand, making
them understand what tasks they are responsible for; they take ownership of their work and thus will not create mediocre work but the best. This also allows the administrator to monitor their progress less because you have trusted their competency to get the job done. Although some might disagree with this because many thoughts that giving employees free rein to do whatever they want might mean that there is too much freedom. However, this can be improved when the administrator clearly defines project boundaries and the results expected of the employees to produce. This, in turn, help reinforce the standard of quality.

But they cannot be trusted fully on everything, too. The administrator must provide a clearly-defined pool of resources and tools and demonstrate to the workers how to utilize and access those resources as required. Furthermore, constantly promote open communication and inform them that the administrators are accessible for brainstorming sessions at any time. By outsourcing the task and establishing ground rules, they may begin to cultivate a more casual, informal atmosphere. If an employee needs to work from home or has a more flexible schedule, work with them to make it happen. This gives students the idea that the school is more concerned with their personal lives than nitpicking regulations.

Consequently, empowerment is not a gift that an employer may bestow or withdraw at any time. While little modifications are acceptable, workers will feel empowered if they know their autonomy might be removed at any time. While outstanding achievement should be acknowledged, all employees should have the same fundamental privileges to promote workplace unity and inclusion.

Table 2. Administrators' Opinion about Family and Social Boundaries

<table>
<thead>
<tr>
<th>Statements</th>
<th>Mean ± SD</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set a goal of accomplishing a task within working time before leaving the workplace</td>
<td>3.23 ± 0.97</td>
<td>Strongly Agree</td>
</tr>
<tr>
<td>Manage work and personal concerns and communication</td>
<td>3.14 ± 1.04</td>
<td>Agree</td>
</tr>
<tr>
<td>Allocate time for family and social gatherings</td>
<td>3.14 ± 1.08</td>
<td>Agree</td>
</tr>
<tr>
<td>Manage work and personal emergencies</td>
<td>3.05 ± 1.00</td>
<td>Agree</td>
</tr>
<tr>
<td>Establish work and personal life balance</td>
<td>3.05 ± 1.09</td>
<td>Agree</td>
</tr>
<tr>
<td>General WM</td>
<td>3.12 ± 1.04</td>
<td>Agree</td>
</tr>
</tbody>
</table>

1.0 – 1.75 SD  1.76 – 2.50  D  2.51 – 3.25 A  3.26 – 4.00 SA

School administrators strongly agree that setting goals for accomplishing tasks are imperative in thinning the boundaries between work and family (Table 2). Setting a plan for achieving a mission is a direct time management action; this item is ranked 1. According to Keiling (2021) Keeping track of what you need to do and when you need to do it may be easier if you stay organized. Being well-organized includes having an up-to-date calendar, quickly accessing particular files, keeping a clean environment, and taking exact, detailed notes. To be realistic about how much can be completed in a single day. Starting the day with an impossible-to-complete to-do list can only result in tension, irritation, and a lack of concentration. Setting priorities is also essential in accomplishing the task. Meeting deadlines or establishing one for oneself if a job lacks one. Otherwise, it's too easy to lose track of a project or let it slip through our fingers (see Table 3).
In terms of entrepreneurial attitude per dimension of personal entrepreneurial competencies such as opportunity seeking, persistence, commitment to work contract, risk-taking, demand for efficiency and quality, goal seeking, information seeking, systematic planning and monitoring, persuasion, and networking, there is no significant difference between faculty and school administrators. Many researchers suggest that an entrepreneurial attitude is meant for not only school administrators but also faculty.

According to Gibb and Hannon (2006), faculty with an entrepreneurial attitude play a critical role in establishing entrepreneurial ways of thinking and behaving in students or in encouraging the development of students' entrepreneurial competencies. To attain this purpose, it was recommended that faculty serve as entrepreneurial...
role models for learners. According to a previous study, entrepreneurial abilities and a good attitude toward entrepreneurship are required to improve entrepreneurial learning among students (Hytti & O’Gorman, 2004).

Table 4. Step 1-Model Coefficients - Entrepreneurial Attitude

<table>
<thead>
<tr>
<th>Predictor</th>
<th>B</th>
<th>p</th>
<th>β</th>
<th>R²</th>
<th>ΔR²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>0.732</td>
<td>0.732</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.0625</td>
<td>0.093</td>
<td>-0.7854</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>-0.2914</td>
<td>0.275</td>
<td>-0.1615</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Civil Status</td>
<td>-0.0470</td>
<td>0.855</td>
<td>-0.0327</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee Position</td>
<td>0.4196</td>
<td>0.044</td>
<td>0.3591</td>
<td></td>
<td></td>
</tr>
<tr>
<td>School Classification</td>
<td>0.6694</td>
<td>&lt;.001</td>
<td>0.8035</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of Service</td>
<td>0.0509</td>
<td>0.191</td>
<td>0.5710</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Age, gender, civil status, employee position, school classification, and years of service collectively explained 73.2% of the variance in entrepreneurial attitude (Table 4). School classification ($\beta = 0.8035, p < .001$) and employee position ($\beta = 0.3591, p = .044$) were significantly associated to entrepreneurial attitude. School classification has a big impact on the entrepreneurial attitude of school administrators.

State colleges and universities offering more programs than private and community colleges can trigger administrators to increase their entrepreneurial attitude since they need to accommodate more students to be a high-performing school. High-performing school is not an easy task. Administrators in comprehended schools will put more effort than lower ones (Gritter, 2006). The administrator position also has a significant impact on their entrepreneurial attitude. In the administrative domain, there are different positions. Some school administrators handle executive positions; however, their school rank is low, affecting their managerial performance. Others are assigned as administrators but do not enjoy the benefits of it. According to Zhao and Hughes (2020), school administrators can perform excellent tasks and effectively govern the school if they have a high or acceptable position.

Table 5. Step 2-Model Coefficients - Entrepreneurial Attitude

<table>
<thead>
<tr>
<th>Predictor</th>
<th>B</th>
<th>p</th>
<th>β</th>
<th>R²</th>
<th>ΔR²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 2</td>
<td>0.979</td>
<td>0.247</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.0236</td>
<td>0.070</td>
<td>-0.2960</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.0464</td>
<td>0.583</td>
<td>0.0257</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Civil Status</td>
<td>-0.0627</td>
<td>0.435</td>
<td>-0.0436</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employee Position</td>
<td>0.1135</td>
<td>0.108</td>
<td>0.0971</td>
<td></td>
<td></td>
</tr>
<tr>
<td>School Classification</td>
<td>0.1193</td>
<td>0.071</td>
<td>0.1433</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of Service</td>
<td>0.0217</td>
<td>0.106</td>
<td>0.2434</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organizational Boundaries</td>
<td>0.5234</td>
<td>0.005</td>
<td>0.5491</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family and Social Boundaries</td>
<td>0.2839</td>
<td>0.054</td>
<td>0.3082</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
When organizational boundaries and family and social boundaries were added to the model, the $R^2$ increased to 97.9% (Table 5). This means that the addition of the two predictors uniquely collectively explained an additional 24.7% of the variance in the dependent variable. Organizational boundaries is significantly correlated to entrepreneurial attitude ($\beta = 0.5491$, $p < .05$). There are things that need to have a consultative matter and permission to the other administrators and faculty from other departments of the school before take into action which act as a boundary because there are some personnel within the organization or even the organization that need to call in softy approach. School is an organization and as an organization, boundaries are always influenced by its surroundings. As Santos and Eisenhardt (2005) stated boundaries are the demarcation between an organization and its environment. Organizational boundaries have an impact on the entrepreneurial attitude of the school administrators because organizational boundaries offer a unique lens on how environments operate and relate to organizations.

Table 6. Step 3-Model Coefficients - Entrepreneurial Attitude

<table>
<thead>
<tr>
<th>Predictor</th>
<th>B</th>
<th>$p$</th>
<th>$\beta$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 3</td>
<td>0.987</td>
<td>0.008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.00728</td>
<td>0.604</td>
<td>-0.0915</td>
<td>0.5813</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>-0.03787</td>
<td>0.677</td>
<td>-0.0210</td>
<td>0.3004</td>
<td></td>
</tr>
<tr>
<td>Civil Status</td>
<td>-0.07987</td>
<td>0.320</td>
<td>-0.0556</td>
<td>0.1024</td>
<td></td>
</tr>
<tr>
<td>Employee Position</td>
<td>-0.05383</td>
<td>0.586</td>
<td>-0.0461</td>
<td>0.0819</td>
<td></td>
</tr>
<tr>
<td>School Classification</td>
<td>0.08532</td>
<td>0.294</td>
<td>0.1024</td>
<td>0.0819</td>
<td></td>
</tr>
<tr>
<td>Years of Service</td>
<td>0.00730</td>
<td>0.603</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Organizational Boundaries</strong></td>
<td>0.55401</td>
<td>0.078</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family and Social Boundaries</td>
<td>0.27667</td>
<td>0.149</td>
<td>0.3004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skills and Knowledge</td>
<td>-0.83242</td>
<td>0.188</td>
<td>-0.8615</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivation</td>
<td>0.10509</td>
<td>0.593</td>
<td>0.1094</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognition</td>
<td>0.76273</td>
<td>0.050</td>
<td>0.7849</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Aptitudes such as skills and knowledge, motivation, and cognition were added to the model; the $R^2$ increased to 98.7% (Table 6). Which collectively added 0.8% of the variance in the dependent variable on the entrepreneurial attitude. Cognition on the other hand ($\beta = 0.7849$, $p = .05$) was a positively significant predictor in relation to entrepreneurial attitude. This means that acquiring knowledge and skills in entrepreneurship is important in building an entrepreneurial attitude. Aptitude is a measurable aptitude for learning and skill in a given field or subject that may be learned or innate. Interest articulates Aptitude, shown in current performance, which is projected to increase with instruction over time (Rahman, 2014). In his experiment, the aptitude test of the faculty and administrator shows abstract reasoning, verbal reasoning, numerical reasoning, creativity, and critical thinking skills, all of which are aggregates of cognition, supporting the finding of this study. Kumar (2014) also concluded that teaching and administrative ability, especially cognition, are the strongest predictors of faculty and school principal performance. The methodological gap: based on the result of the study, collectively, demographic appeared to be a predictor. However, demographic is generally a predictor of entrepreneurial attitude (Robinson et al., 1991, Abun et al. 2021). It is intriguing that skills, knowledge, and motivation are not a predictor of the
entrepreneurial attitude of school administrators. It only implies that an individual must already obtain those traits before being an administrator. Obviously, in an administrative domain, all leaders should enhance their cognitive skills, such as logical reasoning, visual processing, and sustained and selective attention, as they manage an effective school. Adding more factors to the regression inherently improves the fit. Therefore, in the succeeding steps, there are predictors in the previous actions which no longer significant; it's just contributed a percentage to fit a new considerable variable.

**Conclusion**

In organizational boundaries, personnel empowerment is something on the distinction of any organization or institution. Although administrators understand the importance of personnel empowerment, there may be barriers to its implementation. The most imperative of all as far as demographic, boundaries, and Aptitude are concerned is cognition. In result of the study, two models were generated. First, it is emergent model, which can be used to enhance the lens to see better picture of what school administrators should prioritized in improving their school management strategy. In addition, a simple equation model that can be a foundation formula for improving someone's entrepreneurial attitude.

**Limitation of the Study**

This research used few factors in demographics, boundaries and in Aptitude as an independent variable. The limited independent variables used in this research, may not fully represent the perspectives in terms of entrepreneurial attitude of the school administrators. This research used only three major independent variables such as socio-demographic profile, boundaries, and Aptitude may not be able to explain the factor that influences the entrepreneurial attitude of the school administrators. Thus, limited independent variables may affect the finalized result of the study.

**Proposed Model for Entrepreneurial Attitude**

Based on the statistical outcome, we can generate an entrepreneurial attitude model of illustration for visualization of factors per block and a mathematical model to predict the entrepreneurial attitude of school administrators. When the demographic, aptitude, and boundaries are statistically treated by individual and by pair referring to the previous results and their individual factors (Figure 2).
We can create a Venn diagram that incorporated the demographics, boundaries, and Aptitude into one single figure. If demographic alone run by regression, it gives employee position and school classification as predictors. If Aptitude alone treated by regression, it gives cognition as predictor, and If boundaries, the organizational as the predictor. By doing it by pair, starting with demographic and boundaries, the gap "Organizational" emphasizes the significant predictor or the dominant predictor. If demographic and Aptitude, the gap tells void, meaning, there's no predictor. Since demographic does not have any connection to Aptitude, Aptitude can be obtained as long as you have the willingness to acquire it because Aptitude can be acquired. By pairing Aptitude and boundaries, organizational is still the predictor. If we merged it altogether, the demographic, boundaries and Aptitude, the dominant predictor is the cognition. That is why it placed at the center of the diagram. This emergent gave us a systematic understanding of associating those predictors with the school administrators into one single figure emergent (Model on entrepreneurial attitude/mindset based on the result of the study. In addition, emergent strategy is an action model that explains a corporate strategy and mindset that evolves over time as a company's aims and circumstances change. These methods occur when a company repeats a series of acts in order to build a habit pattern.

This model can also enhance the lens to see better picture of what school administrators should prioritized in improving their school management strategy. Other factors such as seminars and trainings can improve knowledge and skills in terms of entrepreneurial of school administrators as well as faculty. However, there is a specific aptitude which dominant. This specific aptitude, which can be seen at the center of the emergent model and can be acquired only by understanding through experience, thought and senses (Figure 3).

Figure 3. Emergent Model
Discussion of the Model

Mathematical models can be solved if they are correctly worded, which means that any dependent variable that has a significant association with entrepreneurial attitude may be expressed as a function of the independent variable, and therefore the assumptions' implications can be identified. The outcome of this study led to the formulation of Entrepreneurial Attitude Equation. A Simple Model of School Administrators Entrepreneurial Attitude. The model generated from the results of the study is simple and memorable, it just like reading the common formula function in linear algebra the “y is a function of x” or \( Y = f(x) \).

Here, we read, entrepreneurial attitude is a function of cognition or entrepreneurial attitude is directly proportional to cognition. It means that the higher cognitive skills a person has, the more entrepreneurial attitude he or she has.

\[
E = f(C)
\]

The equation demonstrates that entrepreneurial cognition differs from entrepreneurial attributes or other variables, which tend to seek for features that distinguish entrepreneurs from non-entrepreneurs, or school administrators from teachers, and then apply them to all scenarios. Entrepreneurial cognition focuses on the distinctive information processing, opportunity appraisal, thinking style, and decision-making processes of entrepreneurs in dynamic entrepreneurial contexts, and might explain why disparities in behavior outcomes occur at a deep level. According to Zhao et al., 2021, Thomas et al., 2019 and George et al., 2016 entrepreneurs' cognition will play an essential part in business model creation, and entrepreneurs will dictate how new enterprises choose their business models. In the part of school administration, the new ventures are the new programs. Today, “Industrial Revolution 4.0” time and thinking as well as innovation are fast. There are programs that are no longer marketable, or less marketable because of emerging new industries that need new specific skills. Our market or client is the industry that needs our product “the student”.

There should be a business model generation activities and analysis because, there are new openings of opportunities, and new emerging branches of skills. That is why schools are breaking down the main programs to catch up with the market needs. Like the three main programs of engineering (Electrical Engineering, Mechanical Engineering and Civil Engineering), REE (Electronic, Communication, IT, Software Developer, Robotics), RME (Biotechnology, Automotive, metallurgical, Nautical, Marine Engineering). There are school that even merge Electrical and Mechanical, the Electromechanical Engineering), Civil Engineering (Water engineering, environmental engineering, geodetic engineering).

Problems nowadays are very tough, on the part of administrators; they do endless efforts in managing quality education generating quality students as a product. Improving their entrepreneurial attitude by amplifying their cognition is one of the solutions in solving school related problems. The applications of their entrepreneurial cognition are determining the relevant programs and curriculum that are in need to be changed or abolished, creating plan for equitable educational opportunities, develop of rigorous and coherent curriculum, instruction and assessment systems. This mathematical model can be a foundation for the formulation of other models pertaining to predictors of the entrepreneurial attitude. A unique entrepreneurial environment drives the process of entrepreneurial cognition and reasoning that leads to the selection of a new venture's business model (Martins et al., 2015). Hence, the outcome of this study is not questionable at all. One important takeaway from this model is that enhancing someone's cognition can help them transform their entrepreneurial mindset and attitude.
References


Data Availability Statement: More information and data can be obtained from the authors on a reasonable request.

Author Contributions: The authors contributed equally. All authors have read and agreed to the published version of the manuscript.

Jereco Jims J. AGAPITO  
ORCID ID: https://orcid.org/0000-0003-4084-6740

Liza Lorena C. JALA  
ORCID ID: https://orcid.org/0000-0002-8233-4201

Rosemarie CRUZ-ESPAÑOL  
ORCID ID: https://orcid.org/0000-0001-5039-566X

Anthony G. ESGUERRA  
ORCID ID: https://orcid.org/0000-0002-3599-5146

Make your research more visible, join the Twitter account of INSIGHTS INTO REGIONAL DEVELOPMENT: @IntoInsights

Copyright © 2022 by author(s) and VsI Entrepreneurship and Sustainability Center  
This work is licensed under the Creative Commons Attribution International License (CC BY).  
http://creativecommons.org/licenses/by/4.0/
WOMEN IN SCIENCE, TECHNOLOGY, ENGINEERING, AND MATHEMATICS (STEM) "EGYPTIAN CASE STUDY"

Mohamed Ramadan A. Rezk 1, Breyhan Shafai 2, Leonardo Piccinetti 3, Nahed Salem 4, Shaimaa ElBanna 5, Amr Radwan 6, Mahmoud M. Sakr 7

1,4,6,7 Academy of Scientific Research and Technology (ASRT), 101 Elkasr Alini St, 11516, Cairo, Egypt
2,5 British Council, 192 El Nil Street, Cairo, Egypt
3 Europe for Business Ltd, Warrington United Kingdom

E-mails: 1 mramadan79@gmail.com; 2 breyhan.shafai@britishcouncil.org.eg; 3 l.piccinetti@e4business.eu; 4 nahedelashkar.asrt@gmail.com; 5 Shaimaa.AlBanna@britishcouncil.org.eg; 6 Radwan.Amro@gmail.com; 7 msakr@asrt.sci.eg

Received 15 August 2022; accepted 20 October 2022; published 30 December 2022

Abstract. The women representation in science, technology, engineering, and mathematics (STEM) fields in Egypt is spanning hundreds and even thousands of years back. Yet, there is efforts needed to show women's contributions in recent history and current time. This paper highlights the role of women in STEM and their role as main partners in the science, technology and innovation system, participation of women in science for both decision-making and the scientific community. The survey has been designed to look at and support the participation and progression of women in STEM professionals and to encourage more girls and women to continue their studies and practical life in the STEM Fields. Literature, and particularly analytical literature, available around women in STEM in Egypt is limited; it is often tackled within the scope of women’s employment, women’s economic empowerment, or education. The paper reviews many challenges faced by women in science, technology and engineering and mathematics but from a practical point of view, where the questions are designed to reflect the scientific and leadership background as well as entrepreneurship and the relationship with the industry and to identify the obstacles that women face in being a business leader and the obstacles that women face in dealing with industry. Through the case study we unpack and examine the multiple thresholds of women and girls in higher education and employment in STEM fields in Egypt. The case study sample shows that while women are engaging in STEM at the tertiary education level, it is the integration into the labor force which makes it difficult for there to be a more equitable distribution of women across all STEM-related sectors. Egypt has a strong tradition of female empowerment and development but remains patriarchal in many aspects of progress and opportunity; the number of women in STEM education does not translate to those in the work force.

Keywords: Women; STEM; Egypt; STEM survey; entrepreneurship


JEL Classifications: Q20, Q25

* EU-funded WomenUP project ENI/2017/394-091
1. Introduction

Women's engagement in the fields of Science, Technology, Engineering, and Math (STEM) is crucial for expanding knowledge, ensuring economic growth, fostering prosperity, and improving societal well-being (Jiang, 2021). The gap of gender differences in the fields of STEM can be shown in the unequal representation of women in publications, salary, senior positions, annual output, and the distribution of resources (Abuwatfa et al., 2021; Huang et al., 2020). The underrepresentation of women in science, technology, engineering, and mathematics STEM fields is a global phenomenon (Ceci and Williams, 2011). Women are less likely than men to enrol in STEM programmes and pursue jobs in STEM fields, even though the number of women joining these fields is expanding (Wang and Degol, 2016; Charlesworth and Banaji, 2019). There are several disparities in women's participation that cause their underrepresentation in STEM fields may be due to the influence of male-dominated environments that have sown the idea that STEM fields are masculine (Sarseke, 2017). However, a recent study found that Arab women's contributions to the scientific community have inevitably grown to be substantial and that their participation in STEM fields has become increasingly obvious (Darwish et al., 2020).

In ancient Egypt, records do not only show that women enrolled in science education but that there were around 100 female pioneers in the STEM domain. A prominent female doctor at that time is Merit Ptah, the first-ever named female doctor as well as the first woman in the pharmaceutical field. Another significant figure of Ancient Egyptian women in science is a doctor named Cleopatra, who was an expert in the field of obstetrics and gynaecology, whose work on pregnancy, labour, and women’s well-being were studied for more than a century after (Khalil et al., 2017). Despite significant improvement over the past few decades, there are extremely few female scientists working across the globe. According to the UNESCO Institute for Statistics, the average proportion of female researchers worldwide was just 29.3 percent in July 2019. The gap grows with seniority. Women have only received 3% of the scientific Nobel prizes given out to date. Even in tertiary education, there is a shortfall in STEM-related professions. Only 35% of students enrolling in STEM-related fields of study are women at this point, when specialization starts, and students decide which topics to take (UN Women, 2020).

According to the UNESCO Institute for Statistics, there is an estimated 45.3% of women engaged in scientific research in Egypt. (UIS-UNESCO, 2019). In Egyptian Science, Technology and innovation (ESTIO) report (2019) it estimated that 41% of women are in the field of natural sciences, 29% are in the field of engineering and technology, 49% in the field of medical science, and 35% are in the field of agricultural science. According to the CAPMAS 2017 bulletin of employment, wages, and working hours, female specialists in the fields of natural sciences, mathematics and engineering sciences working full time are only 18.44% of the public sector workforce. The number of working hours varies considerably from the public to the private sector; the average number of working hours per week in the public sector was 47 (from 5 institutions surveyed), compared to 52 hours per week in the private sector (from the 20 organisations surveyed). This could go some way to explaining the statistic that women working in STEM fields prefer to do so in the public sector (21%) over the private sector (11%), however the numbers change depending on the specific field (CAPMAS, 2018).

Many studies have been trying to know the impact of work environment on women working in STEM (Saxena et al., 2019; Fouad et al., 2016), showed that, negative affects toward women such as a severe criticism, belittling ideas, or intentionally directed derogatory comments frequently may take the shape of undermining women in the STEM workplace (Duffy et al., 2002), the female scientists' work does not get as much recognition as that of men, that appears in the number of men who receive scholarly awards and prizes compared to women despite efforts to increase the nominations of women (Lincoln et al., 2012).
There are many initiatives that have appeared in the recent period directed at women in the STEM field such as Women-UP initiatives, is an Egyptian initiative funded by the EuropeAid and coordinated by the Academy of Scientific Research and Technology in cooperation with SEKEM Foundation and Techno Khair Foundation. women up program aims to empower women social entrepreneurs and female households as job creators through dual participatory approach (women-to-women loop) as well as empower women social entrepreneurs and female households as job creators through dual participatory approach (women-to-women loop). Women Up program worked to create forums and interactive discussion through the Women Entrepreneurs Platform. In addition to it provided technical support for the Women Entrepreneurs in Egypt (Training sessions and capacity building) and offered helpdesk for financial support and guidance for getting funding. Also, The ASRT manages several state prizes for women that recognize individual efforts such as Nile awards, state merit awards, state awards of excellence, as well as the state encouragement awards, which cover STEM filed specializations.

There are some international programs that support women in science, such as L’Oréal-UNESCO For Women in Science Egypt Young talents program fellowships, that aims to promote the participation of young women in STEM coming from Egypt, this program identifies and rewards talented young female scientists in the STEM field. USAID supports women entrepreneurs in developing their businesses and improving productivity through training, business incubators, career fairs, and corporate matchmaking events. This article highlights the barriers women face in STEM fields that create an unsupportive environment for women. The focus will be on female researchers in the STEM to provide recommendations for best practices to encourage women in research, development, and entrepreneurship, and to create an appropriate social environment for women's work in Egypt.

2. Methodology

2.1. Survey Design
The paper relied on data collected via an online survey. To get a fair reflection of the women researchers in Egypt, the survey is designed by Egyptian Science, Technology, and Innovation Observatory (ESTIO) affiliated to the Academy of Scientific research and Technology (ASRT), in cooperation with the British Council (BC) in Cairo. The target population for the Women in STEM Survey is targeted to women researchers in Egyptian research centers and universities working in the areas of science, technology, engineering, and mathematics (STEM). The sample was chosen randomly from the ESTIO database.

The survey was evenly distributed among the 30 research centers and universities covering all Egypt. It specifically targeted women researchers who have or are currently working in STEM fields. The questionnaire was designed to reviews examples of women who have inspired researchers to work in STEM, the efforts of women researchers to establish start-up or join in incubators and the obstacles that women have faced as entrepreneurs, in addition review any support received by women during their work in STEM. It also discussed the funding for research projects, available opportunities that women must exploit to lead in STEM and the challenges facing women in STEM community.

Also, the paper used some of data extracted from a woman in STEM case study of British Council in Cairo (British Council, 2021), that rely on the twenty-two participants were interviewed face to face, in small focus groups or through online meeting platforms. In addition to this, an online survey was widely distributed through various network channels. Of the 63 eligible survey responses, 32 were STEM students while 31 were employed in an area of STEM, the study based on the British Council in Egypt commissioned Pivot Global Education and Mariam Mecky a gender researcher to conduct the study on women in STEM.
2.2. Characteristics of the Participants in the Survey
The survey covers 411 women researchers in 16 governmental research centers and 14 state and private universities in STEM fields, the 30% of the respondents work in natural sciences, while about 27% work in medical and engineering sciences and 16% of the respondents work in agricultural sciences (figure 1). In the survey the women researchers were asked whether they had completed their higher education (bachelor’s degree) inside or outside Egypt. The results showed that most of the respondents to survey (95.6%) had completed their higher education inside Egypt and only 4.4% had completed their higher education outside Egypt. Meanwhile, the results showed that most of respondents obtained their masters and PhD degrees from Egypt, where only 2.9% and 8% of all respondents received their master and PhD degrees respectively from international universities.

The largest number of respondents to the survey were from researcher/lecturer degree, representing 36.5% of all respondents, followed by less than researcher/lecturer (24.1%), professor and professor emeritus (20.9%) and assistant professor (18.5%) respectively. The survey was targeting to get more information from young women; therefore, most of the respondents of the survey were from the age group between 25 and 40 years representing 70.3%. Only 8% of the respondents are experts over 60 years old (figure 2).
3. Results and Discussion

3.1. Leadership Positions and Historically Influential Figures in STEM
The respondents were asked if they had held any leadership positions within their organization or in another organization. The results showed that 40% of all women researchers held leadership positions during their work, the leadership positions varied by researcher degree across universities and research centres (Table 1). The results showed that there is a direct relationship between the academic degrees of female researchers and occupying of leadership positions, 85% of Professor/Professor Emeritus had held leadership positions. The women took over various leadership positions (head of a research center, the dean of a faculty, president and vice President of a university, head of a research department etc.

<table>
<thead>
<tr>
<th>Researcher Degree</th>
<th>Women Researchers That have Leadership Positions</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Less than Researcher/Lecturer</td>
<td>18</td>
<td>81</td>
</tr>
<tr>
<td>Researcher/Lecturer</td>
<td>40</td>
<td>110</td>
</tr>
<tr>
<td>Assistant Professor</td>
<td>32</td>
<td>44</td>
</tr>
<tr>
<td>Professor/Professor Emeritus</td>
<td>73</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>163</td>
<td>248</td>
</tr>
</tbody>
</table>

Source: Results of the survey

The role models, such as teachers, classmates, public figures, and teaching assistants, have been shown in the literature to be useful in attracting more women to STEM disciplines (Herrmann et al., 2016; Canaan and Mouganie, 2019; Mansour et al., 2021), When the women researchers were asked if they have a role model in STEM. The answers varied between the selection of historical and scientific personalities and among the eminent women scientists in the fields related to the researchers, but it was noticeable that for all women researchers they have one or more role models.

3.2. Women Entrepreneurship
In terms of creating new jobs and growing the gross domestic product (GDP), women can contribute significantly to entrepreneurial activity and economic development (Noguera et al., 2013; Kelley et al., 2017; Hechevarra et al., 2019), which can have a positive impact on lowering poverty and social exclusion (Langowitz and Minniti, 2007). The study of women entrepreneurs has grown significantly in recent years, gaining academic acceptance and, most importantly, helping in the knowledge of all the aspects that contribute to women's challenges in starting their own businesses (Cardella et al., 2020). Startups establishment is one of the top challenges facing women in STEM, as all researchers focus on publishing research, registering patents, and solving industrial challenges, but when they are moving from research to start a startup based on their innovations, there are many challenges facing them.

To understanding the challenges for women researchers to be entrepreneurs, the respondents were asked if they created Startups, or they have been trying to create start-ups and what challenges they are facing during creating the start-ups. The results showed that only 6.3% of women researchers have created start-ups through incubators, exist in their universities, research institutions, or through incubators funded by funding agencies such as the
Academy of Scientific Research and Technology, while the majority (93.7%) have no activities related to entrepreneurship. (Figure 3).

Only 4.1% of women researchers were trying to establish Startups but their attempts were not successful due to many barriers. The main barrier that is facing the women researchers to be entrepreneurs is the difficulty in dealing with the surrounding environment that includes; the lack of training for academic and university professors, society's traditional view of women, women’s family responsibilities, social and cultural restrictions, lack of knowledge of marketing methods, absence of women's awareness of entrepreneurship, lack of awareness among women about the procedures and the nature of projects that they can start with, difficulty in obtaining funds, difficulty in accessing the market, administrative and cultural issues and social duties.

Also, the results showed that 44.3% of women researchers were facing some difficulties in the communication with industry. When the women researchers were asked to mention the top barriers that prevent them from working or collaborating with industry, the main barriers include; limited occupations to women, the difficulty of women presence in industrial areas, the lack of information about industrial partners, traveling abroad and being absent for long periods of time from home country, industrial work environment and location of industrial installations, the absence of communication channels between scientific research and industry and the difficulty of having women working in hard industries such as cement and petroleum explorations.

3.3. Social and Professional Impact of Women in STEM

The women respondents were asked to choose the degree of impact on their participation in STEM fields (negative, no change and positive), weather the impact on their family life, or social life, or professional life. The results showed that for impact on family life, 59.5% of women researchers consider their participation in STEM fields have a positive impact on family life, while 17.6% indicate that it has a negative impact and 23% consider has no change on family life (Figure 4).
The results of survey on the impact of participation of women in STEM fields on the social life showed that 59.6% of women researchers consider that participation in STEM have positive impact on social life, while 13.6% indicate that participating in STEM has a negative impact and 13.6% consider that participating in STEM has no change on social life. As the family and social impact, the impact of the participation of women in STEM on professional life have the similar results showed that 89.3% consider that participation of women in STEM have a positive impact, 2.4% have a negative impact and 8.3% have no change on professional life.

3.4. Support to Women in STEM

Women’s employment is largely governed by policies and legal frameworks that can aid or hinder their work, such as childrearing and maternity leave as well as legal structures that largely dictate the laws around gender. The Egyptian Constitution guarantees equality between women and men with Article 9 affirming the state’s commitment to equal opportunities for all. More notably, Article 11 of the Egyptian Constitution of 2014 stipulates that “the State shall ensure the achievement of equality between women and men in all civil, political, economic, social, and cultural rights in accordance with the provisions of this Constitution”. Other policies, or government strategies to support women in STEM, include the Strategy for Empowerment of Egyptian Women 2030 (Egyptian Constitution, 2014).

The women researchers were asked were asked if they received any type of support in STEM fields such as Moral support from colleagues at work, managers, or from outside work (family, friends), or if they received financial support from their organizations, at the government level, or from outside Egypt or other bodies. The results showed that 86% of women researchers have received support in STEM fields while 14% didn’t receive any kind of support, 67% of women receive support from people outside the field from family and partners and the rest receive support from colleagues. The results also showed that 81.5% of women researchers in STEM received financial support, including 37.2% from their work, 18.5% from the government and 8% from other sources (European Union, USA, etc.).

In the case of support networks, the women researchers in STEM fields mentioned affiliation with women specific networks, such as the National Committee of Women in Science, Women in Science for Developing Countries, Women in Science without Borders, and other international and/or informal women’s networks. Most respondents mentioned connections with informal networks formed in their departments, with colleagues, or simply using their families and friends as their support networks. However, many of those women did mention the
importance of having more opportunities for networking and connecting with other people. There is a demonstrable need for women to connect professionally to share their work practices, as well as finding support and possible solutions for some of the struggles they face while pursuing their professional development.

Data from the literature and the interviews, demonstrate that the strategies in place hold major potential to promote women’s equal access to opportunities in the STEM field (Keune et al., 2019; Dasgupta and Stout, 2014). Moreover, the structure of certain STEM fields, such as ICT, is an opportunity to promote women’s inclusion in STEM. Women surveyed and interviewed believe that there is a need for better and more transparent policies in the workplace, both in industry and academia, for job hiring and promotions.

3.5. Challenges Facing Women in STEM Fields

The participants were asked if they are facing any difficulties in integrating into STEM community. The results showed that 20% are facing some difficulties in integration into STEM community while 80% do not have any difficulties. Also, the respondents were asked also to choose the most important difficulties they are facing through the following options: difficulties to get a promotion at work, difficulties in dealing with other research agencies, difficulties to get a research fund, difficulties to get training, and difficulties in dealing with industry

![Fig. 5. Percentage of Women Facing Difficulties in STEM Field](image)

Source: Results of the survey

The results showed that, the majority of women (50.6%) have difficulties to get a research fund from different institutions, while 15.6% of women researchers are facing difficulties to get a promotion at work, 31.4% have difficulties to get training to increase their capacity building, 34.5% have difficulties in dealing and collaborating with industry and 21.4% have difficulties in dealing with other research agencies (Figure 5).

Conclusions

While there are several structural and systemic challenges to women’s participation in the STEM field, largely concentrated on gender dynamics, there are multiple opportunities to be capitalised on to support women in STEM and overcome hurdles. The major one of these is the high number of female students going into STEM subjects. Support to keep them engaged and to increase awareness at the early education stage may be the best way to further grow this demographic. What is more difficult is changing the mindsets of those inside the workplace and the gender biases that are prevalent, not only in Egypt but around the world. Hence, any proposed interventions should be multi-faceted. Despite the existence of the financial or moral support for women in
STEM fields such as the supportive initiatives for women or government programs offered to them, but there is a clear deficit in entrepreneurship and cooperation with industry, that need more efforts from governmental to encouraging women in these directions.

The women in the interviews proposed access to professional development opportunities, such as leadership and commercialization workshops, entrepreneurship courses, and more practical courses on communication skills, preparation of curriculum, and transferable skills, as well as the development of concrete support programmes for women in STEM at the different stages of their careers. This should include mentorship and leadership programmes by other female leaders in their fields. Partnership with other organizations. There are significant numbers of young women engaging in STEM activity, despite there being considerable barriers to female engagement; there are considerable opportunities available to women through scholarships and policies, and yet structures and cultural norms are holding back true access and participation. These dichotomies speak to opportunity, rather than pure constraint. There exists an underlying foundation to support increased female participation and success in STEM and this needs to be further supported and grown to the point of institutional reality.

Acknowledgment

We extend our thanks to the British Council in Egypt, Academy of Scientific Research and Technology, Europe for Business Ltd, as well as the support from the EU-funded WomenUP project ENI/2017/394-091 in analysis of this paper.

Data Availability Statement: More information and data can be obtained from the authors on a reasonable request.

Author Contributions: The authors contributed equally. All authors have read and agreed to the published version of the manuscript.

References


Egyptian Constitution. 2014 at: https://www.sis.gov.eg/newvr/theconistitution.pdf [accessed 18 September 2022]


Mohamed Ramadan A. REZK
ORCID ID: https://orcid.org/0000-0002-7677-3072

Breyhan SHAFAI
ORCID ID: https://orcid.org/0000-0003-1844-2491

Leonardo PICCINETTI
ORCID ID: https://orcid.org/0000-0002-7861-5668

Nahed M. SALEM
ORCID ID: https://orcid.org/0000-0003-3536-8502

Shaimaa ELBANNA
ORCID ID: https://orcid.org/0000-0002-4512-4942

Amr RADWAN
ORCID ID: https://orcid.org/0000-0002-5567-6708

Mahmoud M. SAKR
ORCID ID: https://orcid.org/0000-0001-9467-9250

Register for an ORCID ID:
https://orcid.org/register

Make your research more visible, join the Twitter account of INSIGHTS INTO REGIONAL DEVELOPMENT: @IntolInsights

Copyright © 2022 by author(s) and VsI Entrepreneurship and Sustainability Center
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/
THE DETERMINANTS OF ENTREPRENEURIAL INTENTION OF SCIENTIST PHD STUDENTS: ANALYTICAL VS EMOTIONAL FORMATION OF THE INTENTION

Alexandre Cabagnols 1, Ali Maalej 2*, Pierre Mauchand 3, Olfa Kammoun 4

Polytech Clermont-Ferrand, University of Clermont-Auvergne, CLerMA Lab (EA3849) - 63178 Aubière, France
University of Sfax, Laboratoire de recherche en Économie et Gestion, LR18ES27, Tunisia
Université de Lyon, UJM Saint-Étienne, Département STAPS, F-42023 Saint-Étienne, France
Université de Tunis, ESSECT, LR18ES02 LARIME, 1089, Montfleury, Tunisia

E-mails: 1 alexandre.cabagnols@uca.fr; 2* ali.maalej@enis.tn (Corresponding author); 3 pierre.mauchand@univ-st-etienne.fr; 4 olfakammoun@yahoo.fr

Received 11 July 2022; accepted 10 October 2022; published 30 December 2022

Abstract. We use the theory of planned behavior (TPB) to investigate the determinants of the entrepreneurial intention (EI) of two samples of PhD students in science in France (N=101) and Tunisia (N=308). We performed PLS-PM estimations of a structural path model. The survey design makes possible an explicit distinction between the direct and indirect measures of the Attitude, Norm and Control for the prediction of the EI. In a situation of “analytical” decision making process we would expect a strong connection between direct and indirect measures which not observed in the case of Attitude. In both countries the main direct explanatory variables of EI are the Perceived Behavioral Control (PBC) and the Attitude (ATT) whereas the Subjective Norm (SN) has no direct significant impact. We also observe a very strong relationship between the SN and the PBC and between PBC and ATT. It raises questions concerning the real causal relationship between these explanatory variables which could have important consequences in terms of entrepreneurship promotion strategies. In contrast with SN and PBC we observe that the indirect measures of ATT are not strongly connected to its direct measure. It may indicate in the case of SN and PBC that the evaluation process is analytic whereas it would be more “emotional” for ATT. We discuss the implications of that distinction. Our results indicate that the manipulation of believes related to PBC is certainly the most relevant strategy of behavioral change.

Keywords: entrepreneurial intention; Theory of Planned Behavior; academic entrepreneurship; science and technology policy; rational decision-making; PhD student entrepreneurship

* The dissemination of the research supported by European Union's Horizon 2020 project H2020-MSCA-RISE-2014 CLUSDEVMED, Grant Agreement Number 645730730
1. Introduction

In economies increasingly focused on knowledge and the strengthening of the linkages between governments, science and industry (Etzkowitz and Leydesdorff, 1995), researchers and doctoral students are integral intermediaries in the diffusion of academic knowledge towards the business world (Neumann and Tan, 2011; Autant-Bernard et al., 2007; Wright et al., 2008). The specific contribution of PhD students to this process can be a part of the traditional framework for being hired at the end of their degree. It can also be through the formal or informal relations they have maintained with companies during their research. They, therefore, carry a double contribution (Wallgren and Dahlgren, 2007; Thune, 2009; Mangematin and Robin, 2003). Their contribution to the dissemination of knowledge can also involve the creation of companies stemming from their research, which is commonly classified into the category ‘academic entrepreneurship’ (Shane, 2004; Cantu-Ortiz et al., 2017; Miranda et al., 2017a). The entrepreneurship of PhD students can serve the collective interest and at the same time mitigate the risks of unemployment and over-education (Horta et al., 2015). At least until the COVID19 epidemic, in most developed countries, the number of the doctoral graduates was structurally greater than the vacancies of academic researchers (Auriol et al., 2013a; Neumann and Tan, 2011). Young PhDs were obliged to seek jobs outside the academic sphere which, usually, was not their primary career intention. They had not only the problems of identifying and matching their skills with the needs of the available companies (Durette et al., 2016; Couston and Pignatel, 2018; Mangematin, 2000) but also of cohering their career aspirations and the career opportunities offered to them outside the academic sphere (Agarwal and Ohyama, 2012). The consequence was a slower job placement for doctoral students than that of the lower-level graduates (Auriol et al., 2013b).

Consequently, PhD student entrepreneurship is seen as a vector of economic development leading to a more coherent use of the competencies of the PhDs, to higher job creation and finally resulting in higher levels of competitiveness and innovation in the market (Maalej, 2013, 2022; Doanh and Bernat, 2019; Barba-Sanchez et al., 2022). Given the high inertia of doctoral trajectories (Hayter and Parker, 2018; Mangematin, 2000), rapid demonstration of entrepreneurial intent can help doctoral students to guide the development of their knowledge and skills in order to limit the costs of a possible future reorientation of their career towards entrepreneurship (Lee et al., 2010). It accordingly increases the chances of success of their projects. Business creation can either be an open ‘option’ or become (or be from the beginning) the primary goal of their research. In both cases a deliberate strategy can be developed, one that assumes the existence of a minimum level of entrepreneurial intent that possibly results in an entrepreneurial behaviour.

Maalej and Cabagnols (2020) maintain that any entrepreneurial action is preceded by the intention to undertake. This intention expresses a person’s determination to start their own business. This determination can be accounted for by individual skills of the entrepreneur, their social environment or even by cultural specificities (Maalej, 2013). Our work aims to identify, more precisely, the factors that explain the intention of the PhD students to start a business. We deal with this topic by using Ajzen's theory of planned behavior (TPB) (Ajzen, 1991). It offers a strong methodological starting point on which many empirical studies have already been conducted in the field of entrepreneurship with numerous sub-populations. That large empirical literature makes possible the identification of the specific attribute of the PhDs.

Our work is characterized by a strict application of the standard methodological procedure described by Ajzen, only a few studies in entrepreneurship follow these steps. In particular, we performed an initial belief survey and
made a clear-cut distinction between direct and indirect measures of the determinants of the intention. Although we use the TPB methodology, the distinction between direct and indirect measures allows us to highlight results whose interpretation refers to the situated action (Sassetti et al., 2018; Suchman, 1987) and effectual theories (Sarasvathy, 2001). It leads to a specific discussion concerning the place of the “analytical” vs “emergent” “contextual” and maybe “emotional” nature of the decision-making processes behind the formation of the intention. The possible cultural sensitivity of our results is assessed with the help of a bi-cultural sample made of French and Tunisian PhD students.

2. Literature and Theoretical Background

This work is based on Ajzen's Theory of Planned Behavior (TPB) (Ajzen, 2012, 1991). Its explicit use in the field of entrepreneurship goes back to the work of Krueger & Carsrud, 1993). It is presented as a theoretical framework that federates a few entrepreneurial researches conducted during the 1970s and 1980s, such as the one of Shapero, 1975, Shapero & Sokol, 1982 et Bird, 1988). This theory is particularly applicable to the case of business start-up behaviors because they usually occur after a ‘rational’ decision-making phase of the project initiators who engage themselves in an ‘intentional’ and ‘prospective’ approach (Krueger and Carsrud, 1993). This model of behavioral prediction has shown its effectiveness in many areas of ‘planned’ decision-making both in the laboratory (by studying game strategy choices, for example) and in the field by studying voting behavior (Ajzen and Fishbein, 1980), the choice to breastfeed (Manstead et al., 1983), the use of abortion (Smetana and Adler, 1980), the birth of another child (Vinokur-Kaplan, 1978) or physical exercise (McAuley et al., 2001) and the use of university sports facilities (Sniehotta, 2009); drinking Alcohol (Ajzen et al., 2011). Starting from a traditional TPB approach our work will specifically focus on two questions relevant for entrepreneurship theory: firstly “how do we measure the factors that explain the intention? (“directly” and/or “indirectly”)) and, secondly, “what kind of structural model should we take into account?”.

2.1. Direct vs. indirect measures: Analytical vs emotional rational decision-making

In the light of a set of experimental tests, the TPB postulates that in the case of planned behaviors, intention to act is the most immediate determinant of behavior. The effect of other explanatory factors on behavior is systematically mediated by intention. Ajzen's work shows that the intention would be directly determined by three main variables that are themselves the consequence of the individuals' beliefs about different issues (see Figure 1 below in which the black arrows represent the causal structure of the model of Ajzen):

- The attitude consists in the globally favorable note or unfavorable evaluation of an individual with respect to the realization of a behavior (here the behavior of creation of company) or not. That global evaluation will be referred as the “direct measure” of the intention.
  In a rational decision-making perspective, the attitude of the individual towards entrepreneurship is supposed to be the result of a set of conscious subjective beliefs about the favorable / unfavorable consequences of business creation. The inventory of these believes is made during a preliminary semi-directive interview with a small sample. Thereafter, these “behavior beliefs” are used as question items during a second stage of large-scale surveying. For each belief, two questions are asked. The first question measures the intensity of the belief, the second one measures the expected impact of the belief in the formation of the global attitude. When these weighted behavioral believes are combined they form an “indirect measure” of the attitude.
- The subjective norm reflects the overall perception by the individual of the most salient social pressures regarding the realization a behavior or not. That overall perception will be later referred as the “direct measure” of the subjective norm. In a rational decision-making perspective, the “direct measure” of the subjective norm is supposed to be the result of the opinion that the subject has regarding the opinions of the various members of his social network. ‘Normative beliefs’ represent the beliefs of the individual about the favorable / unfavorable expectations of specific social network regarding the behavior of business creation. The ‘motivation to comply’ measures the importance that the individual associates with the opinion of the
social network. Thus, the subjective norm is considered in the TPB as the consequence of the normative beliefs of the individual about each member of his/her surroundings considered weighted by the motivation to conform to each of them. Combined these normative believes form the “indirect measure” of the subjective norm.

- Perceived behavioral control refers to the overall perception of the ease or difficulty that the individual expects to encounter in the achievement of the behavior being studied. It will be referred as the “direct measure” of the perceived behavioral control.

TPB presupposes that perceived control results from the combination of two elements: on the one hand, the intensity of the obstacles and facilitating elements that the subject predicts to encounter in the realization of the behavior (the ‘control beliefs’) and on the other hand, the level of control he/she believes to have on each of them (‘evaluation of obstacles and facilitator’). Combined the control believes give an “indirect measure” of the perceived behavioral control.

The direct measures represent the global perception of the individual concerning the situation whereas the indirect measures reflect the intensity of the rational arguments that are supposed to back his global evaluation. As such, we will consider that the indirect measures reflect an “analytical” evaluation whereas the direct measure reflect a more global “emergent”, “contextual and maybe “emotional” evaluation (Colombetti, 2010). In an usual rational decision-making process, we could expect a strong causality running from the indirect measures to the direct measures: Global evaluations are the result of taking into account all relevant elements in the judgement from the evaluator’s point of view. On the contrary, a low correlation between the direct and indirect measures would be the sign of an evaluation process that is not directly related to the explicit criteria of judgement but to more holistic mechanisms such those described by the theory of situated action and effectuation (Sarasvathy, 2008, 2001; Sassetti et al., 2018). In these theories the evaluation of the situation and the behaviour are not only guided by plans but emerge during the action in relation to the context. Feelings and “emotions” may strongly contribute to the judgement (Colombetti, 2010). It doesn’t mean that the decision-making process is not “rational” but that it is not grounded on an analytical and summative decontextualized evaluation process. In a behavioral change perspective, the identification of the evaluation process behind the decision-making process of the PhDs is important since it could lead to quite different strategies either devoted to change believes (in an analytical perspective) or devoted to change “emotions” with the help of “contextualization” and “enaction” tools.

2.2. Choice of a causal structure

The causal structure of the TPB model has generated many theoretical debates as indicated in the review article by Liska (Liska, 1984): anteriority of the intention before the behavior, the causal independence of attitude, norm, control and the anteriority of attitude, control and norm with respect to intention etc... In the field of entrepreneurship, the hypothesis of causal independence is probably not verified as indicated by different studies (Alonso-Galicia et al., 2015; Liñán et al., 2011; Liñán and Chen, 2009; Miranda et al., 2017b; Obschonka et al., 2012; Trivedi, 2017; Doanh and Bernat, 2019; Barba-Sanchez et al., 2022). Accordingly, we introduced three relationships that are represented by gray arrows in Figure 1. The relationship that goes from the subjective norm to the behavioral control is justified by the fact that the perceived social support intervenes potentially as a factor of performance since the social network of project promoters is often considered as a key factor of success (Fernández-Pérez et al., 2014; Liñán and Chen, 2009; Maes et al., 2014; Sun et al., 2017; Trivedi, 2017). The relationship between the social norm and the attitude can be explained by a phenomenon of internalization of the social norm in the perspective of the theories of dissonance which is a function of the motivation to conform to individuals as it has already been suggested in several studies (Giger, 2008; Liñán and Chen, 2009; Maes et al., 2014; Miniard and Cohen, 1981; Ryan, 1982; Sun et al., 2017; Trivedi, 2017). Be out of a strict TPB framework, the relationship between perceived behavioral control and attitude has already been explored in studies of attitudes toward computer use and learning to read (Gardner et al., 1993; Morgan and Fuchs, 2007; Vaknin-Nusbaum et al., 2018). These works reveal a causal relationship that seems to go from control to attitude.
rather than vice versa. In entrepreneurship, Obschonka et al. (2012) statistically highlight a correlation between these two constructs without considering a causal relationship whereas Sun et al. (2017) explicitly test it and obtain a low but positive and slightly significant path coefficient. Barba-Sanchez et al. (2022) show that attitude mediates the relationship between perceived behavioral control and the university students’ entrepreneurial intention. Therefore, in this article, we will assume a relationship that goes from perceived behavioral control to attitude.

![Figure 1. Schematic representation of the theory of planned behavior (TPB) (Ajzen, 1988)](image)

3. Methodological choices, Survey and data

As indicated by Liñán and Fayolle (2015) in their review of literature since the founding article of Krueger & Carsrud, 1993, the TPB model and its adaptations represent a very important part of the research dedicated to the study of the explanatory factors of entrepreneurial behavior. The method has been applied to many subpopulations (students, engineering students, researchers, and in some part to the general population) and in contexts of intercultural comparisons. Most of these studies validate the existence of one or more significant effects associated with attitude, subjective norm, or perceived control over intent (Iakovleva et al., 2011; Kautonen et al., 2015). Nevertheless, the magnitude of the estimated effects of these three variables varies greatly from one study to another depending on the samples and their exact methodologies which compromise any generalization as to the exact value of the parameters obtained (J. Boissin et al., 2009; J.-P. Boissin et al., 2009; Emin, 2004).

To circumvent the biases due to our methodological choices we decided to make a “strict” implementation of the TPB approach of Ajzen in the way we built the survey and organized data collection. Firstly, we conducted a pre-investigation that aimed at collecting the salient believes. Secondly, after identification of the salient believe we performed an internet survey in which a clear distinction was made between indirect and direct measures.

In addition as it is pointed out by Antonioli et al. (2016), most of the existing works in the field of academic entrepreneurship deal with “academic researchers” in general. They do not focus exclusively on doctoral students who are likely to have marked specificities compared to other academic subjects given their unconventional position between “student” and “researcher”. That heterogeneity in the samples may blur the results. Therefore, we have decided to focus exclusively on a sample of PhDs in both stages of pre-investigation and large-scale final survey diffusion.

† They conducted a SEM path analysis.
3.1. Pre-investigation and the Construction of the Belief Survey

In order to obtain the relevant basic material for constructing the questionnaire (i.e., the population's strong beliefs about the behavior studied), 17 doctoral students in science (7 men and 10 women) from the Universities of Clermont-Ferrand (N = 4), Saint-Etienne (N = 4), Sfax (N = 5) and Tunis (N = 4) were surveyed in one-to-one interviews. Questions about the pros and cons expected by the subject particularly when he or she actually achieves the proposed behavior were asked in order to probe into the behavioral beliefs (e.g., ‘What benefits would you see in starting a business?’). As such probing into normative beliefs, the questions focused on identifying those who may be supportive or unfavorable to the subject performing the same behavior (e.g., ‘Who would approve of starting a business?’). Finally, at the level of control beliefs, it was a question of identifying the personal and external factors that, according to the subject, facilitated or complicated the actual performance of the behavior (e.g., ‘What personal factors could encourage you to start a business?’)

To make the questionnaire more representative, only a subset of the beliefs of each type was retained. According to (Ajzen and Fishbein 1980), it is necessary to choose as many beliefs as necessary to arrive at a representative percentage of the subjects consulted. A minimum of 75% seems reasonable.

In the case of behavioral beliefs, to reach 75% of doctoral students consulted, 4 beliefs were included in the questionnaire. Normative beliefs were retained in their entirety, given the fact that they were few (3). Regarding the control beliefs, there are four in the structure of the questionnaire.

In conjunction with the questions about the beliefs (indirect measures), we have, according to the recommendations of Ajzen and Fishbein, developed questions dealing with the other variables (direct measures): attitude, subjective norm, perceived control and intention.

3.2. Sample of the Final Investigation

Based on the beliefs index developed in the pre-survey phase, we conducted an online survey of students enrolled in PhDs in science at four different universities. For France, these are University Clermont-Auvergne in Clermont-Ferrand and Jean-Monnet in Saint-Etienne. For Tunisia, the universities of Sfax and Tunis were targeted. The sampling frames were created from the e-mail lists of the students enrolled in the PhD in these universities at the time of the launch of the survey in mid-October 2016. Apart from Saint-Etienne, after the first sending process, we sent 3 reminders 10 days apart to each of those who had not explicitly asked not to be relaunched. The final response rate is 27% in France and 37% in Tunisia. It indicates a certain interest of the students for this subject.

In order to avoid a possible selection bias linked to the reluctance of some PhDs when it comes to entrepreneurship, the survey was described as a "Survey about the professional insertion of doctoral students". The questionnaire was also structured around the notion of employability. Entrepreneurship was presented as a possible alternative among others in order not to discourage the respondents who were not interested in entrepreneurship.

The questionnaire had to be completely filled in so that the answers could be used in this study which represents 15% of the original French sample and 16% of the Tunisian one (the response rates were respectively 27% and 37%). In the end, 409 observations will be used: 101 in France and 308 in Tunisia. The analyze will then be conducted by bringing together on one side the two French universities and on the other the two Tunisian ones in order to highlight any national specificities and stabilize the results.

The main differences between the French and Tunisian samples come from the distribution of men and women (79% of women in Tunisia against 53% of men in France); age structure (27.31 in France vs. 29.07 in Tunisia);
the average duration of the doctoral registration (2.31 in France vs. 2.9 in Tunisia). The evaluation that PhD students make of the applied nature of their thesis does not differ significantly between countries (5.01 in France and 5.31 in Tunisia on a 10 point scale). The high overrepresentation of women in the Tunisian stems from a large structural imbalance between men and women in the Tunisian higher education sectors which are even more marked in the scientific field. The French and Tunisian sample also strongly differ in terms of field of research. 71% of the French sample is composed of PhD students that are carrying researches in the field of ‘Mathematics, physics, chemistry and computer sciences’ whereas the Tunisian sample is constituted 61% of PhD students enrolled in ‘Biology and life sciences’.

Regarding the professional projections of doctoral students (Table 1), it should be noted that the French are more likely to think of being salaried than researchers while the opposite is the case in Tunisia. Business creation comes last in the professional perspectives of the French as well as the Tunisian. The gap between the least attractive alternative and entrepreneurship does not seem to be more pronounced in the face of entrepreneurship in France (5.09-3.70 = -1.39) compared to Tunisia (5.39-4.07 = -1.32). The intentions of starting a business just after the doctorate and ‘later’ are on the other hand much less strong in France than in Tunisia. It indicates a less attractive on the event of the creation of company in France compared to Tunisia.

Table 1. The professional perspectives of the PhD students

<table>
<thead>
<tr>
<th>Variables measured with a 7-points Likert scale</th>
<th>FR</th>
<th>TN</th>
<th>t-value FR-TN</th>
</tr>
</thead>
<tbody>
<tr>
<td>After the PhD, the future profession I imagine is ...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>… researcher</td>
<td>5.09 (1.88)</td>
<td>5.39 (1.93)</td>
<td>-1.35</td>
</tr>
<tr>
<td>… salaried</td>
<td>5.31 (1.51)</td>
<td>4.77 (1.94)</td>
<td>2.52</td>
</tr>
<tr>
<td>… business creator</td>
<td>3.72 (1.83)</td>
<td>4.07 (1.92)</td>
<td>-1.61</td>
</tr>
<tr>
<td>I intend to create a business:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>… just after my PhD</td>
<td>2.62 (1.73)</td>
<td>3.90 (1.96)</td>
<td>-5.62</td>
</tr>
<tr>
<td>… later after or during my professional life</td>
<td>4.25 (1.88)</td>
<td>5.20 (1.64)</td>
<td>-4.77</td>
</tr>
</tbody>
</table>

3.3. Measurements and Descriptive Analysis of the TPB Variables

The final survey was conducted exclusively online. The order of the questions was fixed but the response items were presented in random orders.

3.3.1. Measure of the entrepreneurial intent (I)

Entrepreneurial intent is the dependent variable of the model. It is measured by using three items. Two items at 7 levels {very unlikely, very likely}: ‘I intend to start a business just after obtaining my Ph.D.’ and ‘After the thesis, the professional future I imagine is a business creator’. A three-level item {never, sometimes, often}: ‘I’ve already thought about starting a business’: three-level item. Table 2 summarizes the results obtained on the two samples. On average, Tunisian doctoral students seem to be more inclined to consider the creation of companies at the end of the thesis, as indicated by the very significant difference observed for the item ‘intention_ct’.

—

‡ In 2014-2015, the proportion of women enrolled in doctorates in Tunisia was 67% for all sectors, it reaches 73% at the University of Sfax (Bureau des études, de la planification et de la programmation, 2016).
Table 2. Measure of entrepreneurial intention

<table>
<thead>
<tr>
<th></th>
<th>Scale</th>
<th>FR*</th>
<th>TN*</th>
<th>p-value FR-TN**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Future: I have already think of setting up a business</td>
<td>0-3</td>
<td>0.82 (0.75)</td>
<td>0.80 (0.73)</td>
<td>0.81754</td>
</tr>
<tr>
<td>Future crea: After my PhD, I think of creating a business in my professional life</td>
<td>1-7</td>
<td>3.72 (1.83)</td>
<td>4.07 (1.92)</td>
<td>0.09927</td>
</tr>
<tr>
<td>Intention: I intend to create a business just after setting my PhD</td>
<td>1-7</td>
<td>2.62 (1.73)</td>
<td>3.90 (1.96)</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

* in brackets are reported the standard deviations
* * p-values from the comparison test of the French and Tunisian averages under the hypothesis of heteroscedasticity

3.3.2. Direct measures of attitude, subjective norms and behavioral control

The direct measures of Attitude, Subjective Norm and Behavioral Control were conducted by using a single question giving rise to a measured response on unipolar scales ranging from 1 to 7. The scores were not displayed on the scales.

Table 3 shows that the levels of Subjective Norm and Attitude of French and Tunisian PhD students do not differ significantly at the 5% level. On the other hand, the perceived behavioral control of Tunisians (4.92) is much higher than that of French (3.95). It indicates that the Tunisian doctoral students generally feel more able to create than French PhD ones.

Table 3. Direct measures of attitude, subjective norm and behavioral control

<table>
<thead>
<tr>
<th></th>
<th>Scale</th>
<th>FR*</th>
<th>TN*</th>
<th>p-value FR-TN**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ad: Attitude: ‘My attitude concerning the creation of a company is …’</td>
<td>1-7</td>
<td>4.267 (1.6964)</td>
<td>4.627 (1.7885)</td>
<td>0.07010</td>
</tr>
<tr>
<td>Nd: Subjective Norm: ‘My entourage would approve that I create a company …’</td>
<td>1-7</td>
<td>4.792 (1.3514)</td>
<td>4.623 (1.5716)</td>
<td>0.29768</td>
</tr>
<tr>
<td>Cd: Behavioral control: ‘I feel able to create a business …’</td>
<td>1-7</td>
<td>3.951 (1.8296)</td>
<td>4.929 (1.5964)</td>
<td>0.00000</td>
</tr>
</tbody>
</table>

* in brackets are reported the standard deviations
* * p-values from the comparison test of the French and Tunisian averages under the hypothesis of heteroscedasticity

3.3.3. Indirect measurements

Indirect measurements were made using 7-point Likert scales ranging from 1 to 7. The scores were not displayed on the scales so as not to encourage the effects of ‘social desirability’ in the responses. In accordance with the procedure suggested by Ajzen and Fishbein and for each indirect measure (i) the subjects were first asked about the level of their beliefs (Bi) and then asked to evaluate the importance they attached to each of them (Ei). The final score associated with each item i (Si) is calculated as the level of the belief considered multiplied by its evaluation. To carry out this calculation, we have retained the formula Si = (Bi × Ei)^0.5

With the help of a PLS-regression theses score will be combined to produce 3 latent variables:
- Ai: ‘indirect measure of attitude’
- Ni: ‘indirect measure of subjective norm’
- Ci: ‘indirect measure of perceived behavioral control’
Indirect measure of attitude (Ai)

Four statements were used that evoke possible consequences of business start-up behavior: ‘being independent’ (indep), ‘being your own boss’ (boss), ‘having a good salary’ (wage), and ‘creating jobs’ (Job). The subjects first had to indicate their beliefs about the proposed consequences on a scale from ‘totally disagree’ to ‘totally agree’ (e.g., ‘If I started a business, it would be a good way for me ... ‘to be independent’?). They then had to indicate the value they gave to each of these consequences (e.g., ‘Personally, what value do you give to the fact ... of being independent’?). Table 4 shows that the scores (Si) of the indirect measures of attitude differ sharply between the two samples: Tunisian doctoral students have a significantly more favorable view on entrepreneurship.

Indirect measurement of subjective norm Ni

In the first place, the doctoral students indicate their normative beliefs. In other words, for each referent that is envisaged (family, friends or colleagues and thesis director), they evaluate the probability (on a 7-point Likert scale from ‘very improbable’ to ‘very likely’) that the referent approves the behavior studied. Secondly, on a 7-point Likert scale ranging from ‘very weak’ to ‘very strong’, they evaluate the importance given to the opinion of each referent. Table 4 reports higher scores (Si) for the Tunisian sample mainly regarding the family and the thesis director.

Indirect measurement of perceived behavioral control Ci

Behavioral beliefs are evaluated on four factors from the prompt ‘After I get my doctorate, if I want to start a business ... I will find the necessary funding; I will have a network of relationships; I will have the necessary knowledge in economic and / or legal matters; I will meet a favorable socio-economic context’. The importance associated with each factor is assessed by using the prompt: ‘To start a business, how much importance would you give to finding the necessary funding, having a sufficient relationship network, having the necessary knowledge in the socio-economic context). Table 4 shows that the scores (Si) associated with the indirect measures of behavioral control are significantly lower in France than in Tunisia (except for the financing aspect).
Table 4. Indirect measures of attitude (Ai), subjective norm (Ni) and perceived behavioral control (Ci)

<table>
<thead>
<tr>
<th></th>
<th>FR*</th>
<th>TN*</th>
<th>p-value FR-TN**</th>
</tr>
</thead>
<tbody>
<tr>
<td>S_i=(Belief_i × Evaluation_i)^0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indirect measures of Attitude (Ai) : “If I started a business it would be a good way of…”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S_indep</td>
<td>5.129</td>
<td>5.537</td>
<td>0.00656</td>
</tr>
<tr>
<td>… being independent</td>
<td>(1.2594)</td>
<td>(1.3847)</td>
<td></td>
</tr>
<tr>
<td>S_boss</td>
<td>5.035</td>
<td>5.561</td>
<td>0.00049</td>
</tr>
<tr>
<td>… being my own boss</td>
<td>(1.2797)</td>
<td>(1.3273)</td>
<td></td>
</tr>
<tr>
<td>S_wage</td>
<td>4.352</td>
<td>5.759</td>
<td>0.00000</td>
</tr>
<tr>
<td>… earing a good salary</td>
<td>(1.3166)</td>
<td>(1.0974)</td>
<td></td>
</tr>
<tr>
<td>S_job</td>
<td>4.967</td>
<td>5.866</td>
<td>0.00000</td>
</tr>
<tr>
<td>… creating jobs</td>
<td>(1.3317)</td>
<td>(1.2537)</td>
<td></td>
</tr>
<tr>
<td>Indirect measures of subjective norm (Ni) : “Who think that you should start a business…”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S_family</td>
<td>4.258</td>
<td>4.847</td>
<td>0.00070</td>
</tr>
<tr>
<td>… my family</td>
<td>(1.4824)</td>
<td>(1.5003)</td>
<td></td>
</tr>
<tr>
<td>S_friends</td>
<td>4.368</td>
<td>4.64</td>
<td>0.09432</td>
</tr>
<tr>
<td>… my friends or colleagues</td>
<td>(1.3986)</td>
<td>(1.4408)</td>
<td></td>
</tr>
<tr>
<td>S_dr</td>
<td>3.877</td>
<td>4.418</td>
<td>0.00233</td>
</tr>
<tr>
<td>… my PhD supervisor</td>
<td>(1.4773)</td>
<td>(1.683)</td>
<td></td>
</tr>
<tr>
<td>Indirect measures of behavioral control (Ci) : “After obtaining my PhD, if I want to start a business…”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S_finance</td>
<td>4.537</td>
<td>4.358</td>
<td>0.23940</td>
</tr>
<tr>
<td>… I will find the necessary finance</td>
<td>(1.1843)</td>
<td>(1.6745)</td>
<td></td>
</tr>
<tr>
<td>S_net</td>
<td>4.463</td>
<td>4.911</td>
<td>0.00294</td>
</tr>
<tr>
<td>… I will have an adequate social net work</td>
<td>(1.2247)</td>
<td>(1.4958)</td>
<td></td>
</tr>
<tr>
<td>S_knowledge</td>
<td>4.009</td>
<td>5.072</td>
<td>0.00000</td>
</tr>
<tr>
<td>… I have developed the necessary knowledge</td>
<td>(1.2325)</td>
<td>(1.5289)</td>
<td></td>
</tr>
<tr>
<td>S_context</td>
<td>4.15</td>
<td>4.795</td>
<td>0.00004</td>
</tr>
<tr>
<td>… the socio-economic context will be favorable</td>
<td>(1.2485)</td>
<td>(1.582)</td>
<td></td>
</tr>
</tbody>
</table>

* we report the mean of each variable and in parenthesis the standard deviation
** p-values resulting from a test of comparison between the French and Tunisian means (taking account of a possible heteroscedasticity)

4. Final Model, Estimation Method and QUALITY ASSESSMENT

Figure 2. Measurement model and structural model
Figure 2 schematically represents the model that we estimated. The estimation of the measurement model and the structural model were made by using a PLS regression carried out with the R PLS-PM package (Sanchez, 2013). This method has the advantage of being ‘robust’ in the presence of small samples and when the variables are not multi-normal. The analysis of the results will enable us to test the magnitude of the ‘direct’ effects associated with the TPB variables (coefficients g, h, i) in addition to the indirect effects when the effect of one variable passes through another to calculate the ‘total’ effect of each variable. Significance tests were performed by bootstrap. We will also test the existence of differences between France and Tunisia.

We work on to estimate the value of the linear coefficients [a ... i] (path coefficient) which link the different variables of interest of the TPB to each other. Note that latent variables associated with indirect measures are all considered exogenous. We also note that direct measurements are only made with one measure. This means the latent variables Ad, Nd, Cd are confused with their measures. Therefore, the loadings of their manifest variables will be equal to 1.

The structural model can also be written as follows:

\[
\begin{align*}
Nd &= a \cdot Ni + \varepsilon_N \\
Cd &= b \cdot Nd + c \cdot Ci + \varepsilon_C \\
Ad &= d \cdot Nd + e \cdot Cd + f \cdot Ai + \varepsilon_A \\
I &= g \cdot Nd + h \cdot Cd + i \cdot Ad + \varepsilon_I 
\end{align*}
\]

Where the \( \varepsilon \) represent the prediction errors related to the existence of omitted variables in the model.

### 4.1. Quality assessment

#### Unidimensionality and convergence

All the measurement variables are considered reflective of the latent variables they are associated to. In other words, we will assume that their level is the consequence of the latent variables they measure. If these manifest variables measure the same latent underlying variable, a high level of correlation between them should be observed (unidimensional criterion evaluated by Cronbach's Alpha). Moreover, the level of association of the manifest variables with the latent variable with which they are associated should be greater than with the other latent variables (criterion of specificity evaluated by cross-loadings).

In table 5 for France and Tunisia, we check that:

- Manifest variables are relatively homogeneous as indicated by Cronbach's alpha (\( \alpha_c \)) greater than 0.7 except in two cases where it is 0.697 and 0.698.
- The Dillon-Goldstein rho and the Cronbach alpha measure the convergence of the manifest variables but considering the weight of each variable in the construction of the latent variables obtained by PLS (the loadings). This indicator is therefore more reliable than Cronbach's alpha for PLS (Chin, 1998; Vinzi et al., 2010). The rho's are all well above 0.7 indicating that the relationship between latent and manifest variables are good.
- The analysis of Eigen value from the matrix of correlations between manifest variables gives a final indication of the reliability of the measurements: the first Eigen value (VP1) are all much greater than 1. The second ones (VP2) are all less than 1. This difference in value informs us of the importance of the ‘polarization’ of the manifest variables which all point in the same direction.
Table 5. Verification of the unidimensionality of the manifest variables on the whole sample

<table>
<thead>
<tr>
<th>Latent variables</th>
<th>FR+TN N= 409</th>
<th>FR N= 101</th>
<th>TN N= 308</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normei</td>
<td>αc 0.738</td>
<td>ρ 0.851</td>
<td>VP1 1.97</td>
</tr>
<tr>
<td></td>
<td>0.599</td>
<td>0.830</td>
<td>0.898</td>
</tr>
<tr>
<td></td>
<td>2.24</td>
<td>0.494</td>
<td>0.698</td>
</tr>
<tr>
<td></td>
<td>0.833</td>
<td>1.87</td>
<td>0.646</td>
</tr>
<tr>
<td>Normed</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Controlei</td>
<td>0.887</td>
<td>0.922</td>
<td>2.99</td>
</tr>
<tr>
<td></td>
<td>0.488</td>
<td>0.876</td>
<td>0.915</td>
</tr>
<tr>
<td></td>
<td>2.92</td>
<td>0.476</td>
<td>0.895</td>
</tr>
<tr>
<td></td>
<td>0.927</td>
<td>3.05</td>
<td>0.438</td>
</tr>
<tr>
<td>Controled</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Attitudei</td>
<td>0.829</td>
<td>0.886</td>
<td>2.64</td>
</tr>
<tr>
<td></td>
<td>0.637</td>
<td>0.784</td>
<td>0.861</td>
</tr>
<tr>
<td></td>
<td>2.43</td>
<td>0.686</td>
<td>0.831</td>
</tr>
<tr>
<td></td>
<td>0.888</td>
<td>2.66</td>
<td>0.641</td>
</tr>
<tr>
<td>Intention</td>
<td>0.734</td>
<td>0.850</td>
<td>1.96</td>
</tr>
<tr>
<td></td>
<td>0.626</td>
<td>0.856</td>
<td>0.913</td>
</tr>
<tr>
<td></td>
<td>2.33</td>
<td>0.439</td>
<td>0.697</td>
</tr>
<tr>
<td></td>
<td>0.832</td>
<td>1.87</td>
<td>0.667</td>
</tr>
</tbody>
</table>

As part of a PLS estimation, the convergent validity of the measurements is based on the examination of the correlations (or loadings) between the manifest variables and the latent variables associated with them. An informal rule used by researchers is to consider that the correlation coefficient must be greater than 0.7 implying that there is more than 0.7²≈50% of shared variance between the manifest variable and the latent variable (Fernandes, 2012; Hair et al., 2016). Referring to Table 6 below, it is verified that the ‘loadings’ of the variables manifested with the latent variables that they reflect are greater than or equal to 0.7 in all cases (shaded boxes). The correlations of the manifest variables with the latent variables to which they are not associated (the ‘cross-loadings’) are on the other hand always lower than those obtained for their latent variables of membership. It can, therefore, be concluded that the manifest variables used in this study are sufficiently correlated to their latent variables of attachment and are specifically related to them.

Table 6. Loadings and cross-loadings on the full sample (France and Tunisia simultaneously)

<table>
<thead>
<tr>
<th>Normei</th>
<th>Normed</th>
<th>Controlei</th>
<th>Controlle</th>
<th>Attitudei</th>
<th>Attitudeu</th>
<th>Intention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normei</td>
<td>S_family 0.885</td>
<td>0.551</td>
<td>0.335</td>
<td>0.508</td>
<td>0.370</td>
<td>0.362</td>
</tr>
<tr>
<td></td>
<td>S_friends 0.816</td>
<td>0.379</td>
<td>0.274</td>
<td>0.364</td>
<td>0.288</td>
<td>0.248</td>
</tr>
<tr>
<td></td>
<td>S_dr 0.711</td>
<td>0.317</td>
<td>0.383</td>
<td>0.369</td>
<td>0.285</td>
<td>0.279</td>
</tr>
<tr>
<td>Normed</td>
<td>MDN 0.536</td>
<td>1.000</td>
<td>0.279</td>
<td>0.429</td>
<td>0.303</td>
<td>0.416</td>
</tr>
<tr>
<td>Controlei</td>
<td>S_finance 0.307</td>
<td>0.223</td>
<td>0.780</td>
<td>0.325</td>
<td>0.198</td>
<td>0.286</td>
</tr>
<tr>
<td></td>
<td>S_net 0.383</td>
<td>0.297</td>
<td>0.888</td>
<td>0.470</td>
<td>0.309</td>
<td>0.343</td>
</tr>
<tr>
<td></td>
<td>S_knowledge 0.326</td>
<td>0.206</td>
<td>0.882</td>
<td>0.458</td>
<td>0.279</td>
<td>0.320</td>
</tr>
<tr>
<td></td>
<td>S_context 0.366</td>
<td>0.238</td>
<td>0.903</td>
<td>0.428</td>
<td>0.277</td>
<td>0.284</td>
</tr>
<tr>
<td>Controlle</td>
<td>MDC 0.524</td>
<td>0.429</td>
<td>0.492</td>
<td>1.000</td>
<td>0.473</td>
<td>0.635</td>
</tr>
<tr>
<td>Attitudei</td>
<td>S_indep 0.273</td>
<td>0.261</td>
<td>0.204</td>
<td>0.355</td>
<td>0.815</td>
<td>0.321</td>
</tr>
<tr>
<td></td>
<td>S_boss 0.252</td>
<td>0.260</td>
<td>0.153</td>
<td>0.326</td>
<td>0.829</td>
<td>0.275</td>
</tr>
<tr>
<td></td>
<td>S_wage 0.364</td>
<td>0.221</td>
<td>0.317</td>
<td>0.442</td>
<td>0.814</td>
<td>0.306</td>
</tr>
<tr>
<td></td>
<td>S_job 0.386</td>
<td>0.244</td>
<td>0.331</td>
<td>0.407</td>
<td>0.793</td>
<td>0.312</td>
</tr>
<tr>
<td>Attitudeu</td>
<td>MDA 0.373</td>
<td>0.416</td>
<td>0.358</td>
<td>0.635</td>
<td>0.375</td>
<td>1.000</td>
</tr>
<tr>
<td>Intention</td>
<td>Intention 0.492</td>
<td>0.386</td>
<td>0.481</td>
<td>0.610</td>
<td>0.388</td>
<td>0.511</td>
</tr>
<tr>
<td></td>
<td>Future_crea 0.336</td>
<td>0.248</td>
<td>0.323</td>
<td>0.506</td>
<td>0.231</td>
<td>0.484</td>
</tr>
<tr>
<td></td>
<td>Future_ 0.274</td>
<td>0.314</td>
<td>0.222</td>
<td>0.441</td>
<td>0.228</td>
<td>0.498</td>
</tr>
</tbody>
</table>

Nb: We also verify that for each country individually the convergence rule was satisfied.
The measurement model seems generally correct. The unidimensionality, the convergence of the measurements and their specificity are verified on the two samples considered jointly and separately. It should be noted, however, that the quality of the measures seems better for the French sample than for the Tunisian one.

### Predictive power of the structural model and global quality

As it is reported in Table 7, the overall predictive power seems much higher in France than in Tunisia (GOF = 0.59 in France against 0.49 in Tunisia). It should be noted that the intention is particularly well predicted in France with a $R^2$ of 75% against 41% in Tunisia. In both cases the predictive powers are correct. This, nevertheless, raises the question of the possible heterogeneity of the Tunisian sample composed of 2 cities: Sfax and Tunis. It would also seem that in the Tunisian context of profound institutional and economic change, it would be difficult to obtain a reliable predictive model of intention (Maâlej, 2013).

**Table 7.** $R^2$ of the equations of the structural model and goodness of fit of the full model (GOF)

<table>
<thead>
<tr>
<th>Equation</th>
<th>FR+TN</th>
<th>FR</th>
<th>TN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norm: Nd = a . Ni + εN</td>
<td>0.287</td>
<td>0.315</td>
<td>0.311</td>
</tr>
<tr>
<td>Control: Cd = b . Nd + c . Ci + εC</td>
<td>0.334</td>
<td>0.431</td>
<td>0.314</td>
</tr>
<tr>
<td>Attitude: Ad = d . Nd + e . Cd + f . Ai + εA</td>
<td>0.433</td>
<td>0.536</td>
<td>0.408</td>
</tr>
<tr>
<td>Intention: I = g . Nd + h . Cd + i . Ad + εI</td>
<td>0.501</td>
<td>0.750</td>
<td>0.417</td>
</tr>
</tbody>
</table>

**Goodness Of Fit (GOF):**

<table>
<thead>
<tr>
<th></th>
<th>FR+TN</th>
<th>FR</th>
<th>TN</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOF</td>
<td>0.5147</td>
<td>0.5982</td>
<td>0.4938</td>
</tr>
</tbody>
</table>

### 4.2. Analysis of the estimated path coefficients

**Figure 3.** Results of the structural PLS estimations when the French and Tunisian samples are used separately (France / Tunisia)
Analytical vs “emotional” evaluations

The intensity of the connection between the indirect and direct measures provides an indication of the strength of the connection between the analytical evaluations made by the PhDs and their corresponding overall evaluations of attitude, norm and control. The higher that connection is, the higher may be the impact of a manipulation of their beliefs on their intention.

Concerning subjective norm and perceived behavioral control, the connection between the indirect and direct measures is significant and relatively strong in France and in Tunisia (the path coefficients are between .32 and .56) (See Figure 3). Firstly, it indicates that the global evaluation made by the PhDs concerning these two items is more clearly related to an analytical process than to an “emotional” one. Secondly it means that the manipulation of the main accessible believes should have an impact on their global perception of behavioral control and their subjective evaluation of the social norms.

On the contrary, we observe that the connection between the indirect and direct measures of attitude is not significant in Tunisia and significant and quite low in France. Two explanations can be suggested. Firstly, it may result from the omission of relevant believes that have not been elicited during the preliminary interviews. Secondly, we think that it may indicate that the attitudes concerning entrepreneurship are not primarily driven by an analytical evaluation process but by a more “emotional” evaluation that can neither be easily described in term of salient believes nor reduced to a summation of believes about the possible outcomes of the behavior. Such a result raises questions concerning the proper strategies that could influence such direct “emotional” evaluations.

Core TPB path coefficients

Following the literature about the intercultural dimensions of entrepreneurship, we expected that some of the core TPB coefficients\(^1\) may exhibit significant differences between France and Tunisia (Franco et al., 2010; Léger-Jarniou, 2008; Thurik and Dejardin, 2012; Trivedi, 2017). However, we notice that two out of three TPB’s core path coefficients are similar in France and Tunisia:

- The lowest coefficient in both countries is obtained for the path ‘Subjective Norm → Intention’: it is not significant for France and very low in Tunisia (0.11). In most of the existing researches (Doanh and Bernat, 2019; Barba-Sánchez et al., 2022) conducted in the field of entrepreneurship that coefficient is lower than the other and often not significant.
- The path perceived behavioral control → Intention is positive and significant and very close in France (0.39) and Tunisia (0.36).

The only statistically significant difference (at 5%) that we obtain between France and Tunisia is associated to the path coefficient that goes from Attitude to Intention. We observe that the French path coefficient is significantly higher than the Tunisian one (with a 10% p-value threshold). In fact, it is the highest of the three TPB’s core path coefficients for the French sample. Several entrepreneurial studies report similar highest coefficients for that Attitude → Intention path (Alonso-Galicia et al., 2015; Ambad and Damit, 2016; Barba-Sánchez and Atienza-Sahuquillo, 2017; Goethner et al., 2012; Hesse and Brünjes, 2018; Liñán et al., 2011; Liñán and Chen, 2009; Miranda et al., 2017b; Yurtkoru et al., 2014)**.

In Tunisia, the attitude → intention path coefficient is lower than in France. Firstly, it may be explained by a problem of measure since we also observe that the R-square associated to the Attitude equation is 13% lower in

\(^1\) Core TPB path coefficients: Norm→Intention; Control→Intention; Attitude→Intention

\(^{**}\) Some studies report lower but still positive and significant coefficients (Delanoë and Brulhart, 2011; Devonish et al., 2010; Obschonka et al., 2015, 2012).
Tunisia than in France. It indicates a lower predictive power of the Tunisian dataset compared to the French one. Secondly it could be explained by the fact that French culture encourages people to behave according to their attitude and realize themselves, while in the Tunisian culture several other factors come into play such as for example the destiny and self-confidence that are less pregnant in the French culture.

**Likelihood of the hypothesis of causal independence**

One key question was the existence of ‘lateral’ relationships between the TPB explanatory variables of the intention that is referred to as ‘the hypothesis of causal independence’ (Liska, 1984).

**Norm: a larger impact on control than on attitude**

The previous researches that examined the paths Norm → Attitude and Norm → Control reported positive significant coefficients for both but no significant differences between each other†† (Alonso-Galicia et al., 2015; Liñán et al., 2011; Liñán and Chen, 2009; Maes et al., 2014; Obschonka et al., 2012; Sun et al., 2017; Trivedi, 2017). In the case of doctoral students, our results show that in France and in Tunisia the Subjective Norm tends to have a stronger impact on Control than on Attitude (respectively in France 0.38*/0.03ns, in Tunisia 0.33*/0.18*). This is consistent with several studies that highlight the importance of the relational capital of PhD students in the achievement of their doctoral and entrepreneurial projects (Benmore, 2016; Bienkowska et al., 2016; Bienkowska and Klofsten, 2012).

**Control and Attitude: a strong connection**

The explicit hypothesis of a link from Control to Attitude in the field of entrepreneurship has been studied by (Sun et al., 2017). They report a slightly positive and significant coefficient. In our case, we observe a strong and very significant connection between these two variables: it is the highest estimated path coefficient for the French sample and the second highest one for the Tunisian sample. In line with the experimental researches conducted in the field of education sciences (Gardner et al., 1993; Morgan and Fuchs, 2007; Vaknin-Nusbaum et al., 2018), it indicates that the perceived behavioral control may also be a major explanatory variable of the attitude of PhD student toward entrepreneurship.

**Synthesis of the results in comparison with the previous literature**

![Figure 4. Synthesis of the structural relationships (arrows proportional to the intensity of the relationships)](image)

†† The studies based on PLS models report coefficients between .25 and .4 (Alonso-Galicia et al., 2015; Liñán et al., 2011; Liñán and Chen, 2009; Obschonka et al., 2012).
Figure 4 box A represents the usual results reported in most of the previous studies about the TPB determinants of the entrepreneurial intention. Boxes B and C summarize the results obtained with our two samples of PhD students. Roughly speaking, we observe the same kind of structural relationship in France and Tunisia:

- The subjective norm positively impacts the perceived behavioral control
- The perceived behavioral control results in more positive attitude
- Attitude and perceived behavioral control directly determine the intention whereas the impact of the subjective norm is only indirect.

5. Conclusions and Discussion

5.1. Analytical vs emotional decision-making
From a theoretical point of view these results indicate that the distinction between “direct” and “indirect” measures may contribute to a better understanding of the evaluation process that is carried out by the decision maker. In the case of PhDs, we observe that the subjective norm and the perceived behavioral control are certainly the consequence of an analytical process of evaluation whereas it may not be the case for the attitude. Additional investigations are necessary to assert that finding and to understand how not-analytical evaluations (“emotional”) can be incorporated in a rational decision-making framework. In a dynamic perspective it may lead to a distinction between two kinds of process: the process of the believes formation relevant for analytical evaluations and the process of “emotional experiencing” when the evaluation is not analytic. Both would be relevant for the prediction of the intention as it is the case in our survey. Indeed, we observe in France and Tunisia a strong impact of the direct measure of the attitude on the entrepreneurial intention whereas that in direct measure is not strongly connected (indirect measures). On the basis of a situated action perspective developed by Suchman (1987), Maâlej and Cabagnols (2020) attempt such a kind of combination between the TPB and the more context dependent and emotional perspective.

5.2. Lateral connections between Control and Attitude
We observe a very strong connection between the direct measures of control and attitude. When people consider that they have enough ability and control to run a business, their attitude towards this goal improves, causing a positive effect on entrepreneurial intention (Zhao et al., 2015; Barba-Sanchez et al., 2022). From the point of view of the entrepreneurship theory, the possible existence of relationships between attitude, norm and control is very important since it could explain why the probability to start a business is growing over time and reaches a peak between 40 years-45 years before declining both at the macroeconomic and individual levels (Lévesque and Minniti, 2011; Azoulay et al., 2018). We may think that the accumulation of knowledge during the career leads to an increase of the perceived behavioral control which results in a more favorable attitude and an evolution of the intention and of the probability of entrepreneurial triggering during the first part of the career whereas in its second part the opportunity cost may become too high and reduce the attitude in a proportion that exceeds the positive impact of knowledge accumulation (Amit et al., 1995). This hypothesis of causality has yet to be confirmed. Does it result from measurement problems? What is the direction of the causalities? To clarify such matters, it would be necessary to carry out longitudinal cohort monitoring and, in the best case, to carry out experiments in which each component of the TPB model would be controlled to elucidate the causalities and a possible phenomenon of reinforcement of the entrepreneurial intention trough time.

5.3. Strategies of behavioral change for PhDs
In practical terms, the TPB approach leads to the identification of the major determinants of the intention and of the salient believes that are associated to them. The manipulation of these believes is thus considered as a way of influencing decision making in a chosen direction. Concerning PhD entrepreneurship, our results indicate that the main direct determinants of the entrepreneurial intention of PhDs are their perceived behavioral control and attitude whereas the impact of the subjective norm is mainly indirect (via its impact on the perceived behavioral
control. Consequently, the perceived behavioral control and the subjective norm may be first two candidates for manipulation. However, we have shown that the salient believes associated to the attitude are not strongly correlated to the overall attitude of PhDs whereas the connection between salient beliefs of control (indirect measures) and the direct measure of the perceived behavioral control is much stronger. Consequently, it seems more reliable to envisage strategies of influence that are based on the manipulation of the beliefs of control than of the attitude. In addition, our results and the existing literature indicate that the perceived behavioral control may also have a very strong impact on the attitude which in turn should reinforces the intention. If the manipulation of the salient believes associated to the perceived behavioral control is difficult a second and easier strategy may be the use of the PhD supervisor advices. Indeed, we observe a significant connection between the normative believe concerning the PhD supervisor and the global evaluation of the subjective norm.

5.4. Opening
That research raises points out that the traditional TPB can still be relevant for further researches in as much as that methodology is strictly applied. It seems important to perform a systematic assessment of the salient believes in a pre-survey and to establish a clear distinction between direct and indirect measures of attitude, norm and control. Such simple methodological steps could open fruitful analysis concerning the analytic vs emotional background evaluations involved in the decision-making process that underly the entrepreneurial decision.

References


**Funding:** The dissemination of the research supported by European Union's Horizon 2020 project H2020-MSCA-RISE-2014 CLUSDEVMED. Grant Agreement Number 645730730

**Data Availability Statement:** The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

**Author Contributions:** AC, AM, PM and OK had an equal contribution to that research. All authors read and approved the final manuscript.

Alexandre CABAGNOLS, PhD in economics, University of Clermont-Auvergne, France

ORCID ID: https://orcid.org/0000-0003-4891-6344

Ali MAÂLEJ, PhD in economics, University of Sfax, Tunisia,

ORCID ID: https://orcid.org/0000-0003-1011-7512

Pierre MAUCHAND, PhD in psychology, Jean Monnet University, France

ORCID ID: https://orcid.org/0000-0003-4679-7273

Olfa KAMMOUN, PhD in economics, University of Tunis, Tunisia

ORCID ID: https://orcid.org/0000-0003-4894-9496

Make your research more visible, join the Twitter account of INSIGHTS INTO REGIONAL DEVELOPMENT: @IntoInsights
PRODUCTIVITY IMPULSES FROM REGIONAL INTEGRATION: LESSONS FROM ROAD OPENINGS*

Rasmus Bøgh Holmen

Institute of Transport Economics, Gausdalallén 21, 0349 Oslo, Norway

E-mail: rbh@toi.no

Received 10 July 2022; accepted 10 October 2022; published 30 December 2022

Abstract. In recent years, assessment of wider economic impacts has become an integrated part of transportation appraisal in many developed countries. The practices have also spread to sparsely populated countries, for which the empirical evidences for such impacts remain thin. In this paper, we conduct a multi-level examination on productivity impulses of regional integration caused by road constructions in Coastal Southern Norway. We measure market access in the national road network by power and exponential distance decay, using local estimates for the distance decay parameters from Holmen (2022a) in our baseline specifications. Our endogeneity test and earlier studies suggest that productivity analyses of impulses from Norwegian road constructions do not suffer from reverse causality. Still, we operate with buffer zones of twenty traveling kilometers around each receiver of impulses from market access, where traveling times are held constant. Total factor productivity is pre-estimated, before the impacts of increased market access are assessed at firm and industry level. We find some indications of more commuting and regional industry restructuring subsequent to road openings. The most striking evidences are nevertheless that the openings neither appear to have enhanced productivity growth at firm level nor induced welfare-enhancing reallocation of factor inputs within or between local industries.

Keywords: urban economics; rural economics; infrastructure, productivity; wider economic impacts; market access; road constructions; transportation appraisal


JEL Classifications: D24, F15, R11, R12

Additional disciplines: economics, transportation planning

1. Introduction

There is a growing consensus in the economic geography literature that there may be causal linkages from economic density to productivity (see for instance Graham et al. 2010 and Behrens et al. 2014). Theoretical rationales for such linkages include direct transportation costs savings, production agglomeration and competition

* The paper was funded by the Research Council of Norway, partially through the program, the ‘Industrial PhD-Program’, and partially through the research project ‘Impact’ on wider economic impacts.
effects. Yet, causal identification of impulses on productivity has proved challenging (e.g. Graham et al. 2010, Straub 2011 and Melo, Graham and Brage-Ardao 2013). This has become an important focus in recent research on agglomeration and productivity.

For mobility, traveling speed constitutes a direct substitute for physical proximity in most contexts. In the recent literature, proximity is often interpreted in broader terms than the pure physical sense, also involving traveling time reductions (e.g. Rice, Venables and Patacchini 2006 and Graham et al. 2010). While new major road constructions often have limited impact on traveling distances, they may reduce traveling time substantially.

By this broader interpretation of proximity, studies of road constructions may shed light on the relationship between productivity and economic density. For an economic actor, decreased traveling time and transportation costs to surrounding areas imply more potential profitable economic transactions with other firms and individuals, as well as better access to public goods and fiercer competition. Moreover, new major road constructions in populated areas may increase local value creation by growing the market access for local economic actors and strengthen the local competition. Yet, infrastructure projects for transportation may be motivated by or related to productivity concerns, so the causality in the decision-making should be addressed in such studies.

Until recently, many of the studies in this literature have been macro-oriented. Furthermore, the empirical strategies aimed to determine causal linkages from economic density to productivity has not always been convincing (confer Melo, Graham and Noland 2013 for a review).

In recent years, the research field of wider impacts from infrastructure projects has become more micro-oriented and more focused upon causal impacts from increased market access induced by transportation infrastructure explicitly. Holl (2016) finds support for productivity effects from proximity to highways in the Spanish manufacturing sector. She exploits 1760 postal routes and the old Roman roads as sources to exogenous variation, also controlling for human geographic, geologic and historic circumstances. Another study that applies historical routes for instrumentation is carried out by Duranton, Morrow and Turner (2014). They find that American cities with more highways specialize in sectors producing heavy goods. Some other studies also instrument road expansions on historical routes (e.g. Baum-Snow 2007 and 2010, Michaels 2008, Duranton and Turner 2011 and 2012, Garcia-López, Holl and Viladecans-Marsal 2015 and Baum-Snow et al. 2017) or physical geography (e.g. Banerjee, Duflo and Qian 2012, Faber 2014, Jebwad and Moradi 2016 and Alder 2019).

Utilizing a market access index based on traveling time and employment figures, Gibbons et al. (2019) find that new road infrastructure in the United Kingdom provides positive impulses on local employment, number of firms and productivity in small-scale geographical areas. To address possible reverse causality challenges related to road constructions, the authors hold traveling times constant over time within buffer zones. Gibbons et al. conclude that new transport infrastructure attracts transport-intensive firms to the local area from other areas. Studying US highways outside metropolitan areas, Chandra and Thompson (2000) find evidence that positive impulses for service industries in regions with new highways partly come at the expense of rural regions nearby. Their identification strategy is to consider areas that ‘just happen to be along the new route’ between two major

---

† Direct transportation costs savings involve cost reductions related to actual transportation processes (e.g. Shirley and Winston 2004 and Venables 2007). Production agglomeration involves synergies in production for firms and individuals being near each other. Important production agglomeration arguments include sharing of product markets, factor markets and common goods; more efficient matching of factor inputs and learning in terms of knowledge exchange (confer Duranton and Puga 2004 and Rosenthal and Strange 2004 for overviews, both building on Marshall 1890). Competition effects include firm selection, disciplinary competition effects and impacts on market power exploitation (e.g. Fujita 1988, Melitz and Ottaviano 2008 and Behrens Duranton and Robert-Nicoud 2014). We refer to Holmen and Hansen (2020) for a general review of impacts of transportation measures.
cities, an identification strategy later adapted by several other studies (e.g. Holl 2004, Melo, Graham and Noland 2010, Ghani, Goswami and Kerr 2016 and Ahlfeldt and Feddersen 2018).

The new empirical evidence has contributed to justification of supplementary quantitative analysis for wider economic impacts in ex ante transportation appraisal, carried out before project implementation. Moreover, conduction of such analyses has over the last decade become standard in transportation appraisal in many Western countries, including several ones that are more sparsely populated (e.g. Wangsness, Rødseth and Hansen 2017 and Holmen, Biesinger and Hindriks 2020). These studies often indicate wider economic impacts of tens or hundreds for road construction projects located in rural countries (e.g. Tveter and Mørkrid 2018), but the empirical foundation of such effects in rural areas remains thin.

Admittedly, some studies apparently find support for productivity impulses from road constructions in more rural areas. Notably, Börjesson et al. (2019) and Tveter (2021) find indications of higher wages subsequent to expansions in the road network, assessing Mid-Sweden including Stockholm and most of Norway’s municipalities respectively. Yet, neither of these studies control for industry-year fixed effects, and both simply assume that there is no reverse causality associated with the investments. The later assumption was made based on previous findings suggesting that net benefit of road construction projects do not affect the selection of Norwegian and Swedish road construction projects, as long as the net benefits were positive in Sweden and even regardless of the sign of the net benefit in Norway (Eliasson et al. 2015). In case of Tveter (2021), the study takes place over the financial crisis from 2006 to 2009, where oil and gas supply industry had very strong development near some of the road projects implemented (e.g. the areas surrounding the Eiksund Connection and European Route 18 in Agder). Overall, the empirical evidence for wider economic impacts in rural areas are at least questionable.

In this paper, we study impulses from regional integration on productivity in a rural business sector, exploiting changes in traveling time caused by major road constructions. Utilizing the richness of Norwegian firm and panel data of the road network between Norwegian and zip codes and municipalities in the neighboring countries from 2004 to 2014, we conduct detailed investigations on the heterogeneity related to impulses from road constructions. Earlier studies suggest that productivity studies of Norwegian road constructions do not suffer from reverse causality, a finding that is confirmed by our endogeneity tests. Still, we operate with buffer zones of twenty kilometers traveling kilometers (measured in the initial year) around each receiver of impulses from market access to adjust for potential endogeneity, as suggested by Donaldson and Hornbeck (2016) and Gibbons et al. (2019). Here, the traveling times are held constant, both to adjust for potential exceptions from exogeneity and to limit the influence of noise in our data.

As our study region, we focus on Coastal Southern Norway, which has three advantages for identification. First, the timing of the concrete road constructions and the implementation of Norwegian road constructions more generally are relatively detached from concerns about productivity potential (confer subsection 2.1 and appendix A for elaboration on this matter). Second, the geographic structure is rather transparent with relatively homogenous industry composition for a given urbanization level with most regional economic activities located along the 250 kilometers of coastline. The fact that the region is relatively small and concentrated enables us to explore composition effects of road openings somewhat systematically. Third, the region has experienced a relatively large number of road openings, causing substantial traveling time reductions over just a decade. Since the turn of the millennium, no other Norwegian road openings inducing more than five minutes’ traveling time reductions have concerned more inhabitants within half an hour’s reach on both sides of the road segments in question than the segment at European Route 18 between Kristiansand and Grimstad. Two other construction projects in the region induced some of the highest traveling time reductions in Norway after the turn of the
millennium (i.e. European Route 39 between Flekkefjord and Lyngdal and County Route 465 in Vest-Agder between Farsund and Kvinesdal).

We address both productivity impulses at firm level and composition effects related to factor allocation, which both contribute to higher productivity at a regional level. Hereby, we investigate how increased market access impacts productivity differently depending on industry, firm size and entrepreneurial status. We pre-estimate total factor productivity by Wooldridge and Levinsohn-Petrin’s production estimation procedures at firm and industry level, which are state-of-the-art procedures for causal productivity studies in large panel dataset. To control for firm-specific and industry-specific price developments, we introduce a new control for terms of trade in our firm regressions. As for robustness, we also explore how our results are affected by the use of alternative neoclassical and control function approaches to production estimation, as well as use of alternative control groups beyond our study region. To be able to distinguish between local displacement effects and partial national effects (before potential general equilibrium mechanisms come into play), we choose to operate both a core study region, consisting of Coastal Southern Norway, and an extended study region, also involving municipalities in neighboring regions not located next to major road openings (confer subsection 2.1 for details).

Agglomeration synergies for an economic actor tend to increase with the number of possible economic interactions with other firms and individuals in surrounding areas. Similarly, improved market access to product and factor markets will tend to increase competition locally. To capture the key features of regional integration, we apply two alternative market access measures – one with power distance decay and one with exponential distance decay. For both functions, we apply the distance decay parameter values estimated by non-linear estimation techniques for municipalities in the Southern parts of Norway in Holmen (2022a), controlling for annual growth trend, capital intensity and industry composition.

Our study relates to the empirical literature on composition effects of market enlargement. Combes et al. (2012) study regional densification and productivity in France with a nested model for selection and agglomeration synergies. They find that firm selection is insufficient to explain regional productivity differences. Combes et al. also find that regional productivity is determined not only by productivity at the enterprise level, but also by the development of employment in companies and regions. Several other authors also find that highways attract economic activities, thereby increasing the local density of economic activities (confer Redding and Turner 2015 and Combes and Gobillon 2015 for general reviews).

2. Empirical Strategy

In this section, we present our empirical strategy for addressing the productivity impact of regional integration caused by new road constructions. First, we address our identification of regional integration through road constructions. We account for the geography in the study region and argue that the timing of local road constructions is largely detached from productivity concerns. Second, we present a framework suited for estimating market access measures that capture the agglomeration decay over space, controlling for differences in capital intensity, industry composition and economic growth trend. We also adjust the market access measure to design hypothetical market measures suited for placebo tests. Third, we account for the estimation techniques applied to estimate total factor productivity. In particular, we focus on techniques that address the simultaneity biases related to endogenous factor inputs and controls suited for catching industry-specific developments. To control for firm-specific productivity developments, we introduce a new firm-specific control for terms of trade. Fourth, we present our framework for studying causal productivity impulses of increased market access on TFP and factor usage. Factor usage is also included at this point, considering that composition effects contribute to aggregate TFP developments.
2.1 Identification of Regional Integration through Road Investments

Although economic impacts of transportation measures in secondary markets – so-called wider economic impacts – often come out with substantial magnitudes in empirical research, they are commonly omitted from cost-benefit analyses. For that reason, several researchers have addressed the importance of and potential issues with obtaining reliable estimates for wider economic impacts in recent years (e.g. Venables 2007, Vickerman 2007, Banister and Thurstain-Goodwin 2011, Graham and Gibbons 2019 and Holmen, Biesinger and Hindriks 2020). In the context of the related empirical investigations, there has been a debate in the literature to what extent implemented road construction projects’ effect on productivity can be considered as exogenous and thereby suited for causal identification (e.g. Baum-Snow 2010, Funderburg et al. 2010, Crescenzi and Rodríguez-Pose 2012, Leduc and Wilson 2012, Redding and Turner 2015, Holl 2016 and Gibbons et al. 2019). If road decision processes leading up to road investments paid little attention to productivity potential and associated conditions, road openings may be a valid source of exogenous variation in economic density without further instrumentation. At least the endogeneity challenge would be less severe.

Throughout our study period, economic appraisal including cost-benefit analysis was mandatory for major Norwegian road investments (i.e. above NOK 1 billion in current prices), which are carried out by the public sector. As accounted for in subappendix A.1, studies on Norwegian road investment decisions indicate that these are relatively detached from concerns about net benefits in general and productivity potential in particular. Instead, the project selection is largely influenced by regional horse-trading and solid state finances, as well as maintenance and safety concerns. Considering that Norwegian road construction projects are commonly motivated by other factors than productivity potential, they appear as a good case for studies of regional integration caused by road openings, although instrumentation may still be needed. Rather than studying the country as a whole with all its regional heterogeneity, we choose to turn our attention to a transparent region with relatively homogenous industry composition. Most road openings after the turn of the millennium either occurred in close proximity to the capital area of Oslo parallel to the implementation of other infrastructure projects or in form of new mainland connections and fjord crossings in more sparsely populated areas along the Western and Northern Coast. In contrast to this, Coastal Southern Norway constitutes a relatively homogenous, transparently structured and populated Norwegian region with major road constructions at different points in time. Coastal Southern Norway also distinguishes itself from other Norwegian regions by being subject to relatively many road openings after the turn of the millennium. As briefly reviewed in the introduction in section 1, there are some notable studies on productivity impulses from increased market access caused by expansion of the road network.

To study the productivity impact of regional integration, we focus on Southern Norway in the period from 2004 to 2014. Our study region is rather streamlined with main traffic arteries along a coastline of 250 kilometers and limited population in non-coastal parts of the region. The regional capital, Kristiansand, is located in the middle of the region’s coastline with comparable population bases on each side. The municipalities east of Kristiansand are located in Aust-Agder county, while Kristiansand and the municipalities west of Kristiansand are part of Vest-Agder county. The terrain is relatively similar along the coast of Southern Norway. The geographic structure means that it becomes easier to isolate the impact of decreased traveling times. Another advantage with our study case is that the industry composition for a given urbanization level is relatively homogeneous across

---

1 In the 2000s, major road openings along the Western coast of Norway included National Route 519 Finnfast, County Route 553 in Rogaland by the T-Connection, National Route 13 Hardanger Bridge, National Route 635 Eiksund Connection and National Route Atlantic Tunnel, while European Route 10 Lofast constituted the largest road opening in Northern Norway by far. In the capital area of Oslo, new road openings occurred incrementally along European Route 18 south of the city on both sides of the Oslofjord, while European Route 6 and the Oslo Airport at Gardermoen were gradually expanded north of the capital.

2 January 1. 2020, Aust-Agder county and Vest-Agder county were merged to form Agder county.
municipalities in the region. Thus, potential shocks in the business sector should be rather evenly distributed over the region, limiting industry-related measurement errors that are not captured by firm- and industry-specific controls. A clear geographic structure also makes it easier to keep track of regional composition effects related to the spatial allocation of labor.

The municipalities along the coastline and the remaining suburban municipalities bordering on Kristiansand accounted for 89.9 percent of the inhabitants in Southern Norway in 2004, and 90.3 percent of the population in 2014. In our study, we focus on these municipalities and neglect the more peripheral non-coastal municipalities, since they are somewhat different by nature. We also include the inland municipalities which surround Kristiansand. With a population of 264 200 inhabitants in 2004 and 297 100 inhabitants in 2014, the region is relatively small compared to most regions previously studied in the agglomeration literature (confer section 4 for descriptive statistics). In Fig. 1 below, we have provided a map over the study region with indications of population level in the start of our study period, as well as the location of major road openings and road construction packages during our study period.

The clear structure of our study region also enables us to track the decision processes related to all main arteries between the municipalities in our study, in addition to Kristiansand. Four road packages cover all main roads between municipalities in our study. The renewal of European Route 18 in Southern Norway covers the municipalities located east of Kristiansand, while the renewal of European Route 39 covers the municipalities located west of Kristiansand. Furthermore, the Lister Package covers the study region’s western municipalities, some of which are located along European Route 39, while the Kristiansand Package covers Kristiansand and its suburbs. The Setesdal Region, located in the Northern non-coastal part of Aust-Agder, was not covered by any of these packages, but instead granted an own package (i.e. the Setesdal Package).

The location of European Route 18 and 39 in Southern Norway does to a large extent follow the Norwegian main road from Oslo to Stavanger from the 1800s (Irgens 1978). Much of today’s route segments were built in the postwar era, when parts of them were placed inland instead of going through the coastal towns. A comprehensive overview of the changes in the road network during our study period, as well as some historical context, is provided in subappendix A.2.

Three major road construction projects were realized in Southern Norway in the years between 2004 and 2014 – one in connection with European Route 18 (east in our study region) and two in connection to the Lister package (west in our study region). 30th of August 2006, a renewal of European Route 39 between Flekkefjord and Lyngdal in connection with the Lister package reduced the traveling time at the route segment by 17 minutes, from 46 to 29 minutes. 26th of August 2009, the traveling time between Kristiansand and Grimstad on European Route 18 was reduced by 12 minutes, from 34 to 22 minutes. Since the turn of the millennium, no other major road constructions in Norway (inducing traveling time reductions of at least five minutes) have concerned more people (within 30 minutes from the reduction sites) on both sides of the renewed route segment. Also connected to the Lister Package, County Route 465 in Vest-Agder between Farsund and Kvinesdal was renewed in the period from 2009 to 2012, reducing the traveling time between the municipalities by 18 minutes, from 52 minutes to 34 minutes. Of these, 5 minutes’ traveling time reduction was induced 24th of November 2009, while the reduction of the 13 remaining minutes was realized 19th of October 2012.
Fig. 1. Overview over the core study region of Coastal Southern Norway with highlighting of municipal populations in 2004 and new major road constructions (light green arrows) and road packages (indicated by ovals of dashed lines) from 2004 to 2014.

Beyond these three projects, the road network in our study region was affected by the renewal of European Route 18 between Risør and Gjerstad 20th of August 2004 (reducing traveling time by 6 minutes), gradual increases in speed limits on County Route 751 in Vest-Agder between Kvinesdal and Hægebostad from 2008 to 2013 (reducing traveling time by 5 minutes) and the renewal of County Route 43 in Vest-Agder between Farsund and Lyngdal (reducing traveling time by 2 minutes). There were no major road constructions within the immediate vicinity outside our study region, so other road constructions in the neighboring regions are unlikely to have had substantial direct impacts on the economic outcomes along the coastline of Southern Norway (but they may of course have induced general equilibrium mechanisms).

Reviewing the decision documents for the major road projects in Southern Norway during our study period, we do not find conclusive evidence that rules out that productivity potential could been decisive for their realization. In the public decision documents, the road constructions are primarily motivated by traffic mobility and need for maintenance and improved road standards without mentioning productivity potential. The timing of the different projects also appears to be rather random beyond maintenance considerations. Yet, direct transportation costs savings and housing and labor market region integration are also presented as arguments in the decision documents, particularly in the most recent ones. Although the argumentation in the project documents may not have influenced project selection in practice, it is reasonable to question whether the road investments can be considered exogenous to economic outcomes. Thus, we conduct endogeneity tests to address this matter and adjust our empirical identification strategy accordingly, a matter that we will address in subsections 2.2 and 5.1.

Instead of applying a general equilibrium framework, we conduct a flexible and detailed econometric investigation on productivity impulses of increased market access caused by road investments. The downside is that we neither capture potential national displacement nor ripple effects. In order to shed light on the distinction
between local displacement effects and partial national effect, we operate with an extended study region in addition to our core study region of Southern Norway. Our extended control region includes the two counties of Southern Norway (i.e. Aust-Agder and Vest-Agder), their bordering counties (i.e. Rogaland and Telemark) and their bordering counties again (i.e. Buskerud, Hordaland and Vestfold). Together, our extended study region covers the most southern parts of Norway. Compared to the rest of Norway, these regions are relatively comparable to Southern Norway in terms of somewhat similar industry structure and degree of urbanization.

In Fig. 2, we have illustrated our extended study region, where our core study region of Coastal Southern Norway is colored in blue and the rest of the extended study region is colored in green. We have utilized dark colors for municipalities with road openings inducing traveling time reductions of at least five minutes within 30 minutes’ reach during our study period and light colors for other municipalities.

Since our focus is on productivity impulses in Southern Norway, we omit the municipalities next to road openings in the extended study region. Municipalities are only excluded when the concerned road renewals constitute their fastest traveling path to surrounding destinations. Overall, our extended study region (including Coastal Southern Norway) consists of 142 municipalities, of which 29 are taken out as a consequence of our exclusion criterion. Compared to the road openings in Southern Norway, the concerned constructions either involve traveling time reductions of just above five minutes within 30 minutes realized incrementally over time (i.e. European Route 18
in Eastern Norway), realization in the end of our study period (i.e. the T-Connection) or relatively rural areas on one side of the construction route segment (i.e. Finnfast) or a combination of these characteristics (i.e. the other highlighted connections outside Southern Norway).

2.2 Market Access Measures

In order to study how productivity depends on spatial configuration, we need to apply a market access measure that captures proximity to and the magnitude of places with economic activity nearby into account. Market access functions should also capture how impulses from economic activities diminish over space, a phenomenon known as ‘agglomeration decay’ or ‘distance decay’.

In our main estimation, we will make use of market access functions with power distance decay and exponential distance decay, exploiting the parameter estimated estimated by Holmen (2022a). He estimates the market access functions with power distance decay and exponential distance decay for the Southern parts of Norway by nonlinear estimation techniques at municipal level, controlling for capital intensity, annual growth trend and industry composition. Holmen’s findings suggest a relative sharp distance decay in rural areas compared the urban, which are more commonly addressed in the literature.

Formally, let \( D_{r,t} \) denote a matrix consisting of traveling times between region \( r \) and all regions \( s \) at time \( t \), while \( d_{r,s,t} \) is the minimum traveling time between municipality \( r \) and municipality \( s \) at time \( t \). Furthermore, let \( N_s \) be a vector of the labor stock in all municipalities at time \( t \), while \( N_{s,t} \) is number of employees with workplace in municipality \( s \). Note that \( N_{s,t} \) could be considered as a proxy for potential market connections at given locations, while \( d_{r,s,t} \) could be considered as a proxy for frictions in connectivity. We denote the distance decay parameter by \( \delta_{\text{pow}} \) in case of power distance decay and \( \delta_{\text{exp}} \) in case of exponential distance decay.

We calculate estimate logarithm of market access, \( g_{r,t} \), for each municipality at each time, which are applied in our further empirical investigations:

\[
(1) \quad g_{r,t}^{\text{pow}} (\mathcal{N}_t, D_{r,t}) = \ln \left( \sum_{s=1}^{S} \frac{N_{s,t}}{d_{r,s,t}^{\delta_{\text{pow}}}} \right), \quad g_{r,t}^{\text{exp}} (\mathcal{N}_t, D_{r,t}) = \ln \left( \sum_{s=1}^{S} \frac{N_{s,t}}{\exp(d_{r,s,t}^{\delta_{\text{exp}}})} \right)
\]

where \( \delta_{\text{pow}} \) and \( \delta_{\text{exp}} \) is set equal to the values equal to 2.3 and 0.07 respectively, in accordance with the findings of Holmen (2022a). Further investigations with lower decay parameter values are briefly pursued in appendix A.

In our study, we focus on productivity impulses from increased market access caused by changes in the road network. Urbanization patterns related to employment and settlement will also constitute important sources for increased market access, but these may be subject to even more severe endogeneity challenges. Thus, local developments in employment and economic performance are interlinked. In our investigations, we therefore utilize market access measures, where employment is held constant at the initial level in 2004 (i.e. \( \mathcal{N}_{t,0} \)):

\[
(2) \quad g_{r,t}^{\text{pow}} (\mathcal{N}_{t,0}, D_{r,t}) = \ln \left( \sum_{s=1}^{S} \frac{N_{s,t,0}}{d_{r,s,t}^{\delta_{\text{pow}}}} \right), \quad g_{r,t}^{\text{exp}} (\mathcal{N}_{t,0}, D_{r,t}) = \ln \left( \sum_{s=1}^{S} \frac{N_{s,t,0}}{\exp(d_{r,s,t}^{\delta_{\text{exp}}})} \right)
\]

In line with the advice of Holmen (2022a), we measure potential market connections by the average of employment by residence and employment by workplace, as these measures’ relevance differs from market to market. Furthermore, we will keep employment constant to the initial level to avert endogeneity issues.
Consistent with the empirical strategy of Gibbons et al. (2019), and Donaldson and Hornebeck (2016), we fix the traveling times within buffer zones surrounding each location, in our case fixed to twenty traveling kilometers (measured in the initial year). While Gibbons et al. use buffer zones as a way of obtaining causal identification, an additional justification for us has been to limit noise associated with small changes in traveling time in our data. In this regard, it is worth noting that our market access measures pass the reverse causality test with regard to productivity (as explained further in subsection 5.1).††

In addition to designing instruments, we want to construct suitable market access measures for placebo tests. In our study case, the road opening that affected by far the most people took place in 2009, but road openings also induced substantial traveling time reductions in 2006 and 2012 (confer subsection 2.1 with elaboration in subappendix A.2). In our placebo investigation, we let the period from 2004 to 2006 be the pretreatment placebo period with traveling times and employment figures for 2004, and the period from 2007 and 2008 the posttreatment placebo period with traveling times and employment figures of 2014. We omit the municipalities next to the construction in 2006 (i.e. Flekkefjord and Lyngdal) from this exercise. When designing the placebo measure for the road construction, we hold the population constant at 2004 levels for all year and use 2004 traveling time spans for the pretreatment period and 2014 traveling time spans for the posttreatment period (i.e. $D_{t}^{f}(D_{r}^{f}; N_{t}^{p^{f}})$).

While the estimation of the decay parameters in Holmen (2022a) is carried out at municipal level, we apply the parameter estimates at zip code levels. This raises a concern about small traveling times below the ones typically observed between municipalities and averagely within municipalities. Especially the exponential distance decay function is very sensitive to small values. In coastal municipalities along Southern Norway, the average internal traveling time in the municipalities ranges from nine to 13 minutes except for one outlier on each side. Robustness checks carried out by Holmen suggest that setting the minimum traveling time equal to ten minutes has little impact on the distance decay parameters. The geographical detail level of which zip codes are defined varies substantially between municipalities.

In addition, our traveling time involves some random noise, while the market access functions by construction appear to be oversensitive to changes in traveling time, when the traveling time is low. To limit the noise associated with changes in low traveling time, we set the minimum traveling time to ten minutes in our further

** To deal with endogenous scheme location, Gibbons et al. (2019) operate with buffer zones of 10 to 30 kilometers from the project sites in their study of economic impulses from expansion of the road network the modern United Kingdom. By the same token, Donaldson and Hornbeck (2016) operate with buffers from 10 to 100 miles from railways and waterways in their study of expansion of the rail network in the United States in the 1800s.

†† We have also explored an alternative identification strategy with two sets of instrument candidates, inserting alternative measures for distance friction and potential market connection into the market access functions in equations Error! Reference source not found.. In the design of the first set of instruments, we replaced traveling times with air distances as distance measure, also adjusting for average traveling time reductions in the road network. In the design of the second set of instruments, we replaced employment with severe victims in accident (i.e. fatalities or severely injured) per road surface during the past five years on a preexisting road as measure for potential market connections, once again adjusting for average traveling time reductions in the road network. The adjustment of aggregate changes in traveling time was carried out at county level in addition to national level, since several studies over time suggest that the distribution of road funding between Norwegian counties has been stable over time (e.g. Strand 1983 and 1993, Nyborg and Spangen 1996 and Strand et al. 2015 – confer subappendix A.1). For this purpose, we employed data on victims with severe injuries, fatalities and road surface from Statistics Norway, as well as air distances calculated based on coordinates and the Earth’s curvature. Yet, we abandoned the instrumentation strategy, as our market access measure with constant employment and buffer zones passed the endogeneity test (confer subsection Error! Reference source not found.), and because the instrument candidates turned out to have limited internal validity within the study region under investigation. Yet, this identification strategy could still be useful in other contexts. Similarly, potential market connections (e.g. employment) might be instrumented in studies of market access growth related to urbanization trends and industry developments.
empirical investigations. Robustness checks carried out by Holmen suggest that this assumption would not change the estimates for the distance decay parameters much. Note that our specifications with buffer zones and fixed traveling time beyond some limit in principle still allows for differences in market access beyond ten minutes, but they will be captured by the regressions fixed effects.

2.3 Total Factor Productivity Estimation

We will now account for how we estimate total factor productivity over industries at firm level and municipality level in our main analyses, as well as at municipal level in the preparation of the market access measure. Other components of technical productivity, such as scale efficiency and allocative efficiency, are briefly assessed in appendix B. In our empirical analyses, we first estimate total factor productivity as the regression residual of the production function specification. Then, we assess the impact of increased market access on total factor productivity (TFP) in line with the next subsection. Note that is standard to apply a two-stage procedure in non-frontier productivity studies on panel data (e.g. Graham et al. 2010 and Holl 2016). The approach implies that TFP implicitly will be a function of market access:

\[ a_{i,t} = a_{i,t} (g_{i,t}^f) \]

where \( a_{i,t} \) is total factor productivity (in natural logarithmic form) in firm or regional industry \( i \) at time \( t \) with corresponding gravity measure \( g_{i,t}^f \) with distance decay assumption \( f \).

While the first estimation step is all about obtaining the most reliable TFP estimates, the second estimation step addresses the endogeneity issues related to changes in market access, as explained in the next subsection. As long as firm does not move, changes in market access will be outside the firms’ control, so the associated variables are not considered as control variables in firms’ profit optimization in our main specifications. If the few cases where firms choose to move to another municipality, it will be accounted for as an entry and exit in our study.

In our productivity investigations, we control for the development in terms of trade for each observation unit (i.e. firms or industries). Comparison of development in price margins over industries and municipalities reveals

---

11 An alternative would be to add a constant, which could be interpreted as a fixed traveling time regardless of actual transportation process. This alternative assumption did however provide less robust results in the empirical estimation of the market access function, as rendered in Error! Reference source not found. in case of no such assumptions, so we abandoned it. Note that our specifications with buffer zones and fixed traveling time beyond some limit in principle still allows for differences in market access beyond ten minutes, but they will be captured by the regressions fixed effects.

88 Theoretically, TFP accounts for effects in total output growth relative to the growth in factor inputs for a given production technology. In measurement, growth in TFP corresponds to the portion of output not explained by the amount of input factors, expressed by the Solow residual in the production function and the error term. This implies that total factor productivity estimates will not only capture technological progress and possibly factor efficiency in the strict sense, but also omitted variables (e.g. infrastructure and institutional factors). In addition, the TFP estimates will be affected by simultaneous biases related to the factor input, measurement errors and specification errors. Improvements’ scale and allocative efficiency will not be captured by TFP measures.

111 Admittedly, our applied approach involves an efficiency loss, since we estimate TFP and the impact of increased market access on TFP growth in two separate steps. An alternative procedure would be to estimate TFP with market access as a control variable in a single step. This would be to assume that firms adapt their factor utilization directly to changes in market access. A one-step approach would neither allow us to correct for observation unit fixed effects, industry-year developments nor clustering of standard errors on observation units, which corresponds to the approach suggested in the following subsection. As the control variables for terms of trade include regional factor prices for buildings and land area, the specification might also suffer from simultaneity, if these prices are affected by the increase in market access. We therefore stick with a two-step procedure as our main specification, but report the main results for a one-step procedure in appendix A. Note that a two-step procedure recently has been undertaken by several notable contributions, including the ones by Gibbons et al. (2019) and Holl (2016).
substantial differences in the terms of trade development within our study region. Firms maximize profits in current prices and not in fixed prices, so this should be accounted for in our TFP estimation. Profit maximization suggests that a beneficial price development may make the firm willing to accept a lower marginal productivity.††† We assess the net value added-based terms of trade for an observation as the ratio between output-based price deflators for net value added and input-based price deflators for components of net value added. While the output deflators capture the product price development after intermediates are paid, the input deflators capture the price development for the internal factor inputs. Since we have access to disaggregate deflators for different key economic components over industries, we have the option to calculate both output-based and input-based deflators by utilizing price contribution from different components. By this methodology, the net value added-based terms of trade $T_{it}^s$ for study unit $i$ at time $t$ can be calculated by Error! Reference source not found. below:

$$T_{it}^s = \frac{\frac{X_{it}^{FP\text{ output}}}{W_{it}^{FP\text{ input}}} - \frac{M_{it}^{FP\text{ input}}}{W_{it}^{FP\text{ input}}} - \frac{D_{it}^{FP\text{ input}}}{W_{it}^{FP\text{ input}}}}{\frac{X_{it}^{def\text{ output}}}{W_{it}^{def\text{ input}}} - \sum_{j=1}^{J} \frac{M_{jt}^{def\text{ input}}}{X_{jt}^{def\text{ input}}} - \sum_{c=1}^{C} \frac{D_{ct}^{def\text{ input}}}{K_{ct}^{def\text{ input}}}}$$

(4)

where $V_{it}, X_{it}, W_{it}, R_{it}, M_{it}, D_{it}$ and $K_{it}$ are net value added, gross production, labor costs, net operational profits, intermediates, capital depreciations and fixed capital services for observation unit $i$ at time $t$ respectively. We utilize that all mentioned variables could be measured in current prices, fixed output prices and fixed input prices with corresponding output deflators and input deflators. At last, we have two forms of intermediates $\sum_{j=1}^{J} M_{jt}$ and ten types of fixed capital $\sum_{c=1}^{C} K_{ct}$ and $\sum_{c=1}^{C} D_{ct}$ with (confer subsection 3.1 for elaboration on subgroups). Note that the somewhat messy fraction involving fixed capital services and net operational profits implies that the net operational revenues are deflated with a capital deflator weighted over capital forms in line with each capital stock’s annual user contribution.

Terms of trade in period $t$ as expressed in equation Error! Reference source not found. are not suited as a control in the production estimation directly, since the input factors might be adjusted in period $t$, inducing a simultaneous bias in case of inclusion of the variable. We solve this challenge by instead applying the volumes from period $t-1$ as proxies for the volumes in current period, while valuing the volumes in the prices at period $t$:

$$T_{it}^s = \frac{\frac{V_{it}^{FP\text{ output}}}{W_{it}^{FP\text{ input}}} - \sum_{j=1}^{J} \frac{M_{jt}^{FP\text{ input}}}{X_{jt}^{FP\text{ input}}} - \sum_{c=1}^{C} \frac{D_{ct}^{FP\text{ input}}}{K_{ct}^{FP\text{ input}}}}{\frac{V_{it}^{def\text{ output}}}{W_{it}^{def\text{ input}}} - \sum_{j=1}^{J} \frac{M_{jt}^{def\text{ input}}}{X_{jt}^{def\text{ input}}} - \sum_{c=1}^{C} \frac{D_{ct}^{def\text{ input}}}{K_{ct}^{def\text{ input}}}}$$

(5)

††† The impact of terms of trade on total factor productivity has not been devoted much attention in the economic productivity literature, particularly not at micro level. The topic is studied by inter alia Dievert and Morrison (1985), Fox and Kohli (1998) and Kehoe and Ruhl (2008).
Lagged volume variables are not always available at firm level, as firms are entering and exiting their markets over time. In this case, we use lagged industry- and municipality-specific volumes (i.e. A64 second revision, confer subsection 3.1) instead. In cases where an industry is absent from a municipality, we apply lagged volumes at national level instead.

We estimate TFP by applying non-frontier techniques to production estimation. Our main analyses are carried out with the estimation procedures developed by Wooldridge (2009). In the robustness analyses in appendix B, we utilize a wide range of control function and neoclassical estimation procedures. The purpose is both to address the results’ robustness and to illuminate how different key features in the production estimation affect the results. We apply net value added as our production measure, which captures the ability to transform technological change into income and final demand.

Control function techniques are semi-parametric non-frontier production estimation procedures that apply control functions to handle simultaneous biases originating from estimating output at the same time as factor inputs, particularly fixed capital stock. Following Olley and Pakes (1996), some new non-frontier and semi-parametric approaches to total factor productivity have emerged to deal with the endogeneity challenge with help of control functions. These methods may also address selection biases related to exits. In their paper, Olley and Pakes suggest investments as a proxy to handle the simultaneous bias related to fixed capital. Considering that investments often are zero, Levinsohn and Petrin (2003) suggest intermediates as an alternative proxy to investments. They apply a somewhat more general data generating process than Olley and Pakes do (controlling for heteroskedasticity and autocorrelation in of the error terms).

Various modifications in these approaches have been suggested to address potential dependence challenges. Two important contributions in this regard are Wooldridge (2009) and Ackerberg, Caves and Frazer (2015). Further developing Levinsohn and Petrin’s estimation procedure with intermediates as proxy, Wooldridge (2009) proposes to jointly estimate the elasticities for employment and capital in a single step. He computes standard errors with general method of moments rather than bootstrapping, such that efficiency increases, and computation of standard errors becomes easier. Rather than inverting unconditional input demand functions as Olley and Pakes and Levinsohn and Petrin (2003) and Wooldridge (2009) do, Ackerberg, Caves and Frazer (2015) condition the inverted input demand function on choice of labor input. This does in turn allow for a more general data generating process, where Blundell and Bond’s (2000) dynamic panel estimator is exploited. The authors’ own

---

There are some advantages to using value added as production measure instead of gross production. First, we avoid the issue with double counting of intermediates over firms at aggregate level, which makes the results vulnerable to consolidation. Second, we do not need to address that providers of intermediate inputs do not have to be local. Third, we avoid more complex estimation procedures. On the other hand, value added-based TFP measures ignore the magnitude of intermediate inputs. Thereby, it assumes no surplus extraction between providers of intermediates and providers of internal factor inputs. What is more, zero or negative value added may be a problem when value added is chosen as production measure. Given zero pure profit and Cobb-Douglas technology, the relationship between gross production-based TFP estimates and gross value added-based TFP estimates could be derived by applying the ratio between gross production and value added, known as the Dolmar factor (Dolmar 1961). We refer to Balk (2009) for a review of the relationship between gross production-based and value added-based TFP estimation. We choose to measure value added in net terms rather than gross terms for two reasons. First, depreciation is a cost and therefore does not belong in the output measure. Second, we have capital depreciation figures with sufficient quality to do so.

Potential collinearity issues (as pointed out by Bond and Söderbom 2005) are particularly important. This challenge is more severe when gross production is applied as production measure rather than value added (which we use). Gandhi, Navarro and Rivers (2020) show that control function methods are not identified non-parametrically in the case of flexible inputs. They suggest utilizing non-parametric information contained in the first order conditions for the flexible inputs to resolve this issue. We refer to Van Beveren (2012) for a practical review and comparison of new non-frontier and semi-parametric approaches to TFP estimation.
Monte Carlo experiments reveal that more robustness in Ackerberg, Caves and Frazer’s approach does come at an efficiency cost.

Wooldridge’s (2009) estimation procedure is chosen in our baseline empirical investigations due to relatively efficient optimization algorithms, the advantages by using intermediates as proxy instead of investments, efficient implicit selection correction and decent small sample properties. A one stage procedure also makes computation of standard errors easier, although it requires non-linear estimation for a larger set of parameters than the other approach. Wooldridge’s procedure is somewhat more complex one-step estimation by general method of moments, so we are aware of possible estimation challenges related to estimation on small samples. The other control function estimation procedures applied estimate the production function in two steps and thereby involve efficiency losses in the estimation. The possible collinearity challenge in Levinsohn and Petrin’s procedure associated with intermediates as proxy for productivity shocks may cause a challenge in case of this approach, although it is less severe when value added is applied as production measure instead of gross production. Ackerberg, Caves and Frazer’s estimation approach has relatively poor small sample properties, while Olley and Pakes’ (1996) approach is somewhat more basic than the other mentioned control function approaches and relies on positive investments as a proxy to investment shocks. Due to the mentioned weaknesses, these estimation approaches are only applied in our robustness checks.

To implement the control function estimation procedures, we apply the Stata program ‘prodest’, developed by Rovigatti and Mollisi (2018). We utilize positive gross investments as proxy**** in Olley and Pakes’ procedure and intermediates as proxy in the other procedures. Prodest also allows us to apply implicit selection correction related to exits, which we apply at firm level.

Production estimation may also suffer from other biases than simultaneous biases (e.g. endogenous factor inputs) and selection biases (e.g. in the case of entry and exits). These include measurement biases (e.g. imprecise price deflators or errors in the financial reporting) and specification biases (e.g. omitted variables or limitations of the Cobb-Douglas technology assumption). To address the robustness of our results with regard to these possible biases and illuminate the results of more basic approaches to production estimation, neoclassical estimation techniques come in handy. As simple benchmarks in our productivity estimation in appendix B, we estimate total factor productivity with Cobb-Douglas technology with ordinary least squares, both controlling for terms of trade development and not. Furthermore, we estimate total factor productivity in a panel data linear least square effects fixed effect framework to decrease the measurement errors by controlling for year and observation units fixed effects, as well as autocorrelation and heteroscedasticity.

In addition, we assess the specification bias by detaching ourselves from the Cobb-Douglas technology assumption about an elasticity of substitution equal to one. Instead, we assume a Constant Elasticity of Substitution (CES) technology, following the estimation procedures suggested by Kmenta (1967). First, we apply the nonlinear least squares estimation method, seeing that the CES production function cannot be monotonically transformed into a linear function. In this case, we assume constant return to scale and skip fixed effect dummies to obtain convergence. Second, we a panel data linear least square with year and observation unit fixed effects to estimate Kmenta’s Taylor approximation of the CES-function, controlling for heteroskedasticity and autocorrelation. Note that this Taylor approximation collapses to a variant of the translog representation suggested by Berndt and Christensen (1973).

**** When looking into negative investments in our data, it appears that large negative investments in line with large positive investments predict a reactive actor and improved productivity. Thus, the relationship between gross investments and predicted productivity appears to be nonlinear around zero. We therefore ignore negative investments rather than creating a gross investment index in our implementation of Olley and Pakes’ and Ackerberg, Caves and Frazer’s estimation procedures, being aware that the methods only allow for one proxy.
2.4 Measuring Impact of Increased Market Access

Our object is to explore the impact of increased market access on total factor productivity and input factor usage. We aim to find out how impulses on firm level productivity and factor allocation contribute to higher social wealth. In order to do so, we make use of a panel data least square framework with year fixed effects, observation unit fixed effects (i.e. firm, industry or municipality), and standard errors clustered on observation units.

Our starting point for impact measurement is a panel data least square regression framework with fixed effects that assess the impact of increased market access on a given economic key variable. We let \( x_{i,t} = \{a_{i,t}, l_{i,t}\} \) be log of total factor productivity (\( a_{i,t} \)) or labor input (\( l_{i,t} \)) for observation group \( r \) at year \( t \).†††† Our regression equation is then given by:

\[
(6) \quad x_{i,t} = \beta_0 + DJT_{i(t),t} + DU_{i} + \beta_1 g_{r(i),t}(N_{r_{t_0}}, D_{r_{t_0}}) + \epsilon_{i,t}
\]

where \( DJT_{i(t),t} \) is an industry-year dummy for observation unit \( i \)’s industry \( j \) (i.e. A64 Rev. 2 industry, confer subsection 3.1) at time \( t \) and \( DU_{i} \) is a dummy for observation unit \( i \). Note that \( DJT_{i(t),t} \) collapses to year dummies in regressions are municipal industries and municipalities. \( \epsilon_{i,t} \) is the error term corrected for spherical correlation patterns, and \( \beta_0 \) and \( \beta_1 \) are regression coefficients. \( \beta_1 \) can be interpreted as an agglomeration elasticity. The key to assessing the possible impact of the road opening lies of course in whether \( \beta_1 \) is significant different from zero or not.

\( g_{r(i),t}(N_{r_{t_0}}, D_{r_{t_0}}) \) represents the market access function for firm \( i \) in region \( r \) at time \( t \), given distance decay assumption \( f \). As accounted for in subsection 2.2, we hold employment matrix \( (N_{r_{t_0}}) \) constant in the market access function to adjust for urbanization trends. Furthermore, we let the distance matrix \( (D_{r_{t_0}}) \) vary freely over time beyond buffer zones of twenty traveling kilometers (measured in the initial year) from each location (as endogeneity tests and background literature suggest that causal effects can be captured without instrumentation, confer subsection 5.1 and appendix A respectively). By this specification of approach, we only capture changes in market access related to decreased traveling times and not changes related to other labor market developments.

In case of production estimation, selection biases related to firm exits are implicitly taken into account in the control function approaches. When investigating average impact on firm size in a supplementary analysis, we instead utilize Heckman’s correction procedure (confer Heckman 1979) to correct for this bias, still applying firm fixed effect, industry-year interaction dummies and clustering of standard errors on firms. This is further accounted for in subappendix B.3.

Utilizing the placebo measures for market access established in subsection 2.2, we will now construct a framework for placebo testing to enable revelation of potential underlying trends. Again, we utilize a panel data least square framework with industry-year interaction dummies (simplifying to year dummies on aggregated observation level), observation unit group fixed effects and clustering standard errors on observation units.

†††† In appendix B, we extend the vector of outcome variables \( x_{i,t} = \{a_{i,t}, l_{i,t}, k_{i,t}, F_{i,t}\} \) to include fixed capital services \( (k_{i,t}) \) and firm size in terms of the number of persons engaged (i.e. employees and self-employed) per firm \( (F_{i,t}) \) for observation group \( r \) at year \( t \).
We use 2004 to 2008 as our placebo test period, since only one substantial road construction opening took place in this period. This opening was between Lyngdal and Flekkefjord municipalities, which therefore are taken out of the sample in the placebo tests. In our placebo test framework, we pretend that the road constructions took place in the beginning of 2007, thereby applying 2004 to 2006 as our placebo pretreatment period and 2007 to 2008 as our placebo posttreatment period. In the placebo measure for market access, we hold the labor stocks constant at 2004 levels and apply 2004 traveling time spans for the placebo pretreatment period and 2014 traveling time spans for the placebo posttreatment period (i.e. \( g_{r(i),x}(N_{i,t}, D_{i,t}^{pl}) \)).

\[
(7) \quad x_{i,t} = \beta_0^{pl} + D_{i,t}^{pl} J_{r(i),x} + D U_{i,t} + \beta_1^{pl} \sigma_{r(i),x}(N_{i,t}^e, D_{i,t}^{pl}) + \varepsilon_{i,t}^{pl}
\]

where \( D_{i,t}^{pl} \) is an industry-year dummy for observation unit \( i \)’s industry \( J \) at time \( t \) and \( D U_{i,t} \) is a dummy for observation unit \( i \). Furthermore, \( \varepsilon_{i,t}^{pl} \) is the error term corrected for heteroskedasticity and autocorrelation, while \( \beta_0^{pl}, \beta_1^{pl} \) and \( \beta_2^{pl} \) are regression coefficients. Note that a \( \beta_1^{pl} \) significantly different from zero or of the same magnitude as \( \beta_2 \) indicates there might be an underlying trend. A significant placebo test could of course also be an expression of a temporary random deviation from a more ordinary development path. It could even involve lagged impulses of market access, or responses to the announcement of a road opening rather than the road opening itself (probably more relevant for factor utilization than productivity). Yet, it does in any case indicate that corresponding potential significant impulse estimates may be questionable. Comparison of estimates for \( \beta_1 \) and \( \beta_1^{pl} \) can also be useful when considering potential impacts of the road constructions.

3. Data Sources

To be able to study regional integration impacts from major road constructions, we have combined and processed various sources for firm data and traveling data. These data are accounted for in the following.

3.1 Firm Data

In our analyses, we apply accounting data from the Norwegian Register of Business Enterprises at the Brønnøysund Register Centre (Norwegian Register of Business Enterprises in short) from 2004 to 2014, where data for 2003 is used in addition to estimate terms-of-trade control. The Norwegian Register of Business Enterprises covers about 95 percent of the business sector employment in Norway including public firms (Holmen 2022b). For some firm branches in some years, the zip codes are missing or did not exist in our study period. In this case, we exploit our own mapping of historical zip codes, time series of reported zip codes, correspondence between postal box codes and physical zip codes and other reported geographical information to determine geographical location by zip code, when possible. When monetary values are stated in foreign currency, we convert the monetary values to NOK, using the Norwegian Central Bank’s statistics for historical exchange rates. As justified in subsection 2.1, we limit our study to 16 municipalities along the coastline of Southern Norway.

Applying European enterprise register data involves a branch challenge, since the geographic distribution of economic activities beyond employment in firms with branches at multiple geographic locations is not
Consequently, we limit ourselves to single-branch firms in analyses concerning other economic key variables than employment, since we do not have geographic identification for activities within multi-branch firms.\textsuperscript{‡‡‡‡} To avoid possible biases related to single-branch firms becoming multi-branch firms over the study period and vice versa, we classify all firms that are multi-branch firms at least once during our study period as multi-branch firms. For each year, we also omit firms without employment or fixed capital holdings, as such firms are rather peculiar, and both factor inputs are essential in neoclassical production functions.

In some of our analyses on employment composition effects, we have approximated the distribution of labor costs over branches in multi-firm by the help of employment shares and industry belonging, as previously done in Holmen (2022b). In cases where firms are registered with more employees than its branches, the surplus of employees are allocated to their headquarters. Formally, labor costs $w_{b,t}$ in branch $b$ at time $t$ is approximated as:

\begin{equation}
(8) \quad w_{b,t} = \frac{\tilde{\rho}_{i(b),t} n_{b,t}}{\sum_{b=1}^{B_f(b)} \tilde{\rho}_{i(b),t} n_{b,t}} \cdot w_{f(b),t}
\end{equation}

where $\tilde{\rho}_{i(b),t}$ and $n_{i(b),t}$ are the labor costs and employment for single-branch firms in the industry $i(b)$ of branch $b$ at time $t$ respectively, while $w_{f(b),t}$ and $B_f(b)$ are labor costs and the number of branches in firm $f$ of branch $b$ respectively. Put differently, $w_{f(b),t}$ is the labor costs to be distributed, while the fraction corresponds to branch $b$’s share of the variable at firm level. Single-industry ratios are obtained from the enterprise register, when available, and from the national accounts otherwise.

In our study, we utilize the industry classification standard applied in European national accounts, A64 second revision. We have grouped the industries defined by the A64 second revision industry classification into ten industries, as accounted for in the Table 1. We do not consider resource industries, oil and gas extractors, construction, finance and insurance and non-market-oriented industries. These industries are excluded because of measurement and identification issues, as they are either relatively volatile, not profit-driven, strongly driven by international prices and local natural resources, strongly regulated, requiring alternative measurement of value added, directly affected by road construction processes or poorly captured by the enterprise register. Higher market access might induce productivity impulses in these industries as well, but since they are likely to create noise in our analyses, we exclude them.

In addition, we omit the oil and gas suppliers, since these firms generally have had a stronger development than mainland suppliers across NACE codes, strongly dependent on the development in the oil and gas markets (e.g. Grünfeld et al. 2013). This industry sector is also characterized by quality competition, resulting in relatively heterogeneous performance, which is unlikely to be related to expansion of the road network. Since the NACE system largely constitutes an activity-based rather than a market-oriented industry classification, these firms

\begin{itemize}
\item \textsuperscript{‡‡‡‡} Firms with reporting duty to the Norwegian Register of Business Enterprises are only obliged to report industry belonging, number of employees, head quarter status and geographic location are reported at branch level.
\item \textsuperscript{§§§§} Studying the Canadian manufacturing sector, Rigby and Brown (2015) investigates how agglomeration gains differ between single-branch and multi-branch firms. Their empirical results suggest that single-branch firms obtain stronger productivity gains from the matching of workers and knowledge spillovers, whereas multi-branch obtain stronger productivity gains from the presence of upstream input suppliers. In general, they find that single-branch firms obtain more agglomeration gains than multi-branch firms do. To the contrary, Rosenthal and Strange (2003) find little evidence that aggregate agglomeration impulses differ for subsidiaries and independent units. Studying six manufacturing industries in the United States, they only find significant differences in agglomeration benefits for the fabricated manufacturing industry.
\end{itemize}
cannot be identified by the traditional NACE system. Instead, we have applied the oil and gas firm population developed by International Research Institute of Stavanger and Menon Economics with regional collaborators (Blomgren et al. 2015). It contains firms that mainly deliver goods to the petroleum extractors’ value chain. The firm population is also utilized in some supplementary investigations in subappendix B.5.

Table 1. Industry classification

<table>
<thead>
<tr>
<th>Industry sector</th>
<th>NACE codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer good manufacturing</td>
<td>NACE 8 to 15, 18 and 31 to 32</td>
</tr>
<tr>
<td>Material manufacturing</td>
<td>NACE 16 to 17 and 19 to 25</td>
</tr>
<tr>
<td>Technological manufacturing</td>
<td>NACE 26 to 30 and 33</td>
</tr>
<tr>
<td>Wholesale and vehicle trade</td>
<td>NACE 45 to 46</td>
</tr>
<tr>
<td>Retail trade</td>
<td>NACE 47</td>
</tr>
<tr>
<td>Transportation</td>
<td>NACE 49 to 52 except 49.5, 50.101 and 50.201</td>
</tr>
<tr>
<td>Tourism</td>
<td>NACE 55 to 56 and 70</td>
</tr>
<tr>
<td>Information services</td>
<td>NACE 58 to 63</td>
</tr>
<tr>
<td>Knowledge consultancy</td>
<td>NACE 69 to 75</td>
</tr>
<tr>
<td>Business support services</td>
<td>NACE 77 to 78 and 80 to 82</td>
</tr>
</tbody>
</table>

We apply processed fixed figures (i.e. figures for fixed capital services, gross investments, depreciation in fixed prices and industry- and asset-specific capital deflators), obtained from Holmen (2022b). We will operate with ten different forms of fixed capital: ‘land area’, ‘buildings’, ‘construction’, ‘machinery’, ‘equipment’, ‘small means of transportation’, ‘large means of transportation and mobile production units’, ‘research and development’, ‘patents’ and ‘cultivated biological resources’. Furthermore, we operate with two forms of intermediates – ‘commodity purchases’ and ‘service purchases’, which include electricity purchases, but exclude financial costs. Commodity and service deflators are also obtained from Holmen (2020), which are based on the Norwegian national accounts’ gross production and intermediate deflators and industry input-output matrixes. Note that regional price developments are accounted for in the capital deflators for buildings and land area and the service purchase deflators (which is affected at a rate proportional to share of services purchases related to real

**** Regional collaborators include Bodo Science Park, Center for Economic Research at Norwegian University of Science and Technology, Eastern Norway Research Institute and Impello Management.

†††† Holmen (2022b) estimates fixed capital figures by the perpetual inventory methods with industry-specific capital deflators and depreciation rates, utilizing firm data from the Norwegian Register of Business Enterprises and macro data derived from inter alia the Norwegian National Accounts. In case of real estate, the deflators are also regional. For asset forms without capital depreciation, the author uses book value adjustment instead. We utilize fixed capital services as our capital measure, which adjusts the capital utilization for capital composition by weighting each type of asset in accordance to their stock volume and annual required return (confer Jorgenson and Griliches 1967 and Christensen and Jorgenson 1969 for details on the concept). We refer to Holmen (2022b) for details on the practical implementation of the fixed capital services concept.
In addition, we apply industry-specific deflators for gross production deflators and use industry-specific development in labor costs per person employed to deduce labor costs input deflators.

3.2 Spatial Data

In our research, we utilize annual traveling time data between all Norwegian zip codes (represented by post offices) from 2004 and 2014. The data are constructed by the Institute of Transport Economics and Menon Economics based on Geographic Information Systems data from the Norwegian Mapping Authority’s ‘Elvegdata’, which contains periodical shapefiles of the Norwegian route network. The calculations are carried out with the Dijkstra’s (1959) algorithm, using the application ‘Network Analyst’ in ArcGis. Traveling times are based on speed limits, implying that road quality that does not alter speed limits or distance (e.g. measures that reduce congestion or variations in waiting time at ferry crossings) is left unaccounted for. The same goes for more generalized transportation costs (e.g. road tolls and logistical optimization). Moreover, traveling time reductions related to waiting and more generalized transportation costs often partly depend on local economic outcomes and would in that case be endogenous. Ferry transitions are assigned five minutes boarding time and a speed of 15 kilometers per hour.

The basic data from these calculations involve snapshot shots of the road network at different points in time. We have recalculated them into annual averages, where we have taken into account what time of the year major road construction projects were implemented. In this regard, all road connections with a traveling time reduction above five minutes within thirty minutes in the dataset were quality assured and corrected in case of inaccuracies. In the quality assurance of the Norwegian traveling data, the Norwegian Public Road Administration’s project database was utilized. A few temporary breaches in the road network (e.g. due to winter closed roads or ferry crossings) were identified and corrected. Data from Gule Sider and Google Maps have been used to connect geographical destinations to the road network in a few incidences of missing data. To account for traveling data to neighboring countries, the Institute of Transport Economics has collected traveling times to all municipalities in Denmark, Sweden, Northern Finland (i.e. Kajanaland, Lapland and Pohjois-Pohjanmaa counties) and Murmansk Oblast from all border crossing from Norway, using Open Street Map. Note that this includes all foreign municipalities within six hours’ reach from Norway including travel by ferry.

The division for population statistics at Statistics Norway has calculated the built-up areas of each zip code based on maps over buildings and population in Norway. Based on these areas, we have approximated internal traveling distances, assuming that the average traveling time equals half of the radius of a circle with similar area. In line with assumptions made by the division for population statistics at Statistics Norway, we assume a speed limit of 60 kilometers per hour within zip codes. We do not allow internal traveling times within zip codes to change over time, but we set them equal to their approximated value in the initial year (i.e. 2004). We refer to the acknowledgement section towards the end of this paper for an overview of the people, who have contributed to the construction of the traveling dataset.

Employment at residence and workplace and population figures are obtained at zip code level and municipal level from Statistics Norway. The geographical location information of between 1.5 to 3 percent of employment figures at municipal level is not identified at zip code level. We have distributed municipal employment figures without known zip code location proportionately to the annual employment shares within the municipalities. Employment and population figures for municipalities in the other neighboring countries are collected from Statistics Denmark.

‡‡‡‡‡ In the fixed capital and deflator data applied, it is distinguished between regional price differences related to real estate capital and real estate purchases, based on processed price data from Statistics Norway and Eiendomsverdi (confer Holmen 2020 for details). The other price deflators are national, implying that we will not be able to adjust for local price variations for other economic components beyond what follows from differences in industry composition and differences in the relative strengths of the ground components.
Statistics Finland and Statistics Sweden. For Murmansk Oblast, we approximate employment and population data by combining municipal data from City Population with regional data from Russian Federal State Statistics Service and national data from the World Bank and the International Labour Organization. Here, we have stipulated the local development based on the aggregate development and figures, and otherwise assumptions about geometric growth in years between the reported observations.

Data utilized to shed light on descriptive statistics, including statistics on income and population, are collected from Statistics Norway. We apply dwelling prices harmonized over quality at municipality level, which are processed by Holmen (2022b), based on price statistics from Eiendomsverdi and Statistics Norway. In this data, the municipal composition of dwelling types is assumed to be the same as at national level.

4. Descriptive Statistics

Before carrying out our empirical investigations, it is worth taking a closer look at the development patterns in our study region. In the following, we will present the main regional economic development of Coastal Southern Norway compared to the rest of Norway and look at status and development for market access in the region.

4.1 Economic Development in the Study Region

In Fig. 3 below, we depict some important development patterns for our study region of Coastal Southern Norway, benchmarked against the other municipalities in our extended study region. A peculiar characteristic of our study region is that it has built up a large supply industry for the petroleum sector over our study period. Accordingly, we have excluded petroleum suppliers from our sample, as accounted for in subsection 3.1. We see relatively weak developments in employment by working place and in the unemployment rate in Coastal Southern Norway late in the study period. Similarly, the housing price development was relatively weak in Coastal Southern Norway in the last years of our study period. Throughout our study period, Coastal Southern Norway had a stronger population growth for people at working age than the rest of our extended study region, while the development in average personal income follows about the same development path.

Our data selection contains 6,134 firms in the core study region and 26,602 firms in the extended study region (including our core study region) over 11 years. In 2004, these firms employed 18,154 persons in the core study region and 81,472 persons in the extended study region, compared to 21,291 and 97,372 in 2004 respectively. As accounted for in section 3.1, we have neglected certain industries (i.e. industries with developments dominated by natural resources, heavy capital investments and non-market-oriented industries) and multi-branch firms to achieve a cleaner identification. Per 2004, 58.2 and 65.0 percent of the employees registered in the Norwegian Enterprise Register worked in the selected industries in the core study region and the extended study region respectively, whereof 49.6 and 49.8 percent worked in single-branch firms. Per 2014, the share of enterprise employment working in the selected industries had fallen to 51.8 percent in the core study region and to 59.1 percent in the extended study region. Yet, the single-branch firms’ share of employment in the selected industries remained nearly the same, at 48.7 percent in the core study region and 48.3 percent in the extended region in 2014.
Fig. 3. Development for the whole study region compared to other municipalities with intermediate urbanization level with 2008 as base year. Money terms are stated in current prices.

In Fig. 4, we provide more comparative statistics on Coastal Southern Norway and the rest of our extended study region. We see that overall development in productivity and net value added is not very strong, which must be seen in relation to changes in the industry composition and omission of industry developments from our study (e.g. the rise of the oil and gas supply sector). Overall, the in-sample developments in the core study region and the rest of the extended region follow each other rather closely across variables over our study period.
Fig. 4. Development for our data sample in the study region compared to other municipalities with intermediate urbanization level estimated by 2008 as base year. Money terms are stated in fixed prices at average price level from 1999 to 2014.

4.2 Status and Development in Market Access

In the following, we will account for status development in market access measures with power and exponential distance decay, measured by the framework established in subsection 2.2. In line with Holmen (2002a), we have set the power distance decay parameter to $\delta^{pow} = 2.3$ and the exponential distance decay parameter to $\delta^{exp} = 0.07$. In both cases, we hold employment (the applied source of potential market connections) constant to the initial year and operate with constant traveling time within buffer zones of twenty traveling kilometers from each location. We will focus on our core study region, as we have excluded municipalities close to road openings from our extended study region and the remaining municipalities only are subject to low or relatively moderate changes in market access.

In Fig. 5, we depict the measured market access in 2004 by the two measures compared to Kristiansand – the regional largest city (marked by dark blue in the maps). As expected, the alternative measures show resembling market access patterns. We see that the differences in market access between Kristiansand and the second most central municipalities in the region are a bit larger in case of power distance decay measure than in case of exponential distance decay measure. In case of exponential distance decay, measured market access is closely related to proximity to Kristiansand, while the pattern is somewhat weaker in case of power distance decay.
Fig. 5. a) Market access in our study region in 2004 measured with a) power distance decay function with $\delta^{\text{Pow}} = 2.3$ (top) and b) exponential distance decay function with $\delta^{\text{Exp}} = 0.07$ (bottom). Measured relative to the market access in Kristiansand in 2004.

In Fig. 6, we have illustrated the municipal increases in market access during our study period from 2004 and 2014. The highest increase in market access by both measures occurred in connection with the renewal of European Route 18 east of Kristiansand in August 2009, through Lillesand to Grimstad. The highest relative increase in market access by both market access measures occurred in Lillesand (marked in dark blue), which became more integrated with both of its neighboring municipalities. Northeast of our study region, Risør and Tvedestrand were also affected by another road opening at European Route 18 north towards Gjerstad municipality and Telemark county, which was implemented already in August 2004. Note that market access is measured on an annual basis.

The increase in market access in the Lister Region (capturing the four municipalities west in our study region) were subject to two investments, which triggered modest increases in market access. Although the investments induced substantial traveling timesavings, the traveling times between the municipalities remained relatively high. The first of these projects was realized at European Route 39 along the regional west-east axis of the Lister Region in August 2006, whereas the second was implemented incrementally from November 2009 and October 2012 along the regional north-south axis of the Lister Region at County Route 465 in Vest-Agder. The regional increase in market access in the Lister region was more prominent, when exponential distance decay was applied, than when power distance decay was applied, given our pre-estimated distance decay parameters. This is partly
due to lower initial market access levels in case of exponential distance decay. The highest increase in market access in this subregion occurred in Flekkefjord (the municipality located farthest west in our study region).

**Fig. 6.** a) Increase in market access in our study region from 2004 to 2014 measured with a) power distance decay function with $\delta^{\text{pow}} = 2.3$ (top) and b) exponential distance decay function with $\delta^{\text{exp}} = 0.07$ (bottom). Measured relative to the market access in Kristiansand in 2004

5. **Empirical Investigations**

In this section, we study the productivity impact of regional integration caused by road constructions empirically. We start by verifying the endogeneity of Norwegian road investments. Then we consider productivity impulses at firm level, before we turn to composition effects. Composition effects from regional integration affect the aggregate productivity developments through changes in factor allocation and heterogeneous impacts over firms. They may occur within or between industries and regions. Various robustness checks are provided in appendix B. We distinguish between three types of composition effects. First, we consider intra-industry composition effects, which reflect that regional integration may have different impacts at firm level and at industry level. This may be due to heterogeneous impact over firm sizes or restructuring within the industry under investigation. Second, we assess composition effects between industries at a given location. These reflect local reallocation of factor inputs,
possibly towards industries with higher factor return. Third, we address regional composition effects. These reflect how the regional factor allocation between municipalities is affected.

5.1 Endogeneity of Norwegian Road Investments

As briefly discussed in subsection 2.1 and elaborated on in subappendix A.1, earlier studies suggest that concerns about wider economic impacts and net benefits more generally play a rather limited role in the practical decision processes that lead up to new road construction projects. To ascertain that road construction does not suffer from reverse causality with respect to productivity in our study sample, we have performed a simple endogeneity test, by considering last period’s labor productivity impact on current period’s road-related increases in market access. We consider three regional delimitations with municipal level over ten years – the coastal municipalities in Southern Norway (i.e. our core study region), the Southern counties of Norway (i.e. the extended region applied in our estimation of the distance decay parameters) and the whole country. For each municipality, we considered the weighted traveling time to all 3,196 geographical zip codes in Norway as well as 465 municipalities in the neighboring countries (confer subsection 3.2 for details).

As depicted in Table 2, the test indicates that Norwegian road construction can be considered exogenous to productivity, regardless of the choices of regional delimitation and distance decay pattern. For our study region, the p-values range from 45 to 55 percent, so no further testing is needed. We report the results on municipal level, as economic activity is rather uneven between zip codes, but it can be noted that the results also hold on regressions at zip code level.

Table 2. Impact of lagged net value added per person engaged on market access from 2004 to 2014, estimated on municipal level. (* for p < 0.1, ** for p < 0.05 and *** for p < 0.01)

<table>
<thead>
<tr>
<th>Market access (logarithm)</th>
<th>Core study region</th>
<th>Extended study region</th>
<th>Whole country</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Industry structure control</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Capital intensity control</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Specification with power distance decay</td>
<td>Lagged net value added per person engaged</td>
<td>0.018</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.020)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Specification with exponential distance decay</td>
<td>Lagged net value added per person engaged</td>
<td>0.014</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.022)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>No. of observations</td>
<td>160</td>
<td>160</td>
<td>1,420</td>
</tr>
</tbody>
</table>

5.2 Genuine Productivity Effects at Firm Level

Our baseline results for productivity impulses at firm level are shown in
Table 3. The results provide little evidence of firm level TFP impulses from expansion in the road network. Within Coastal Southern Norway, there are weak indications that the local tourism industry has had relatively high productivity growth in areas close to major road openings, but this result disappears when we apply our broader study region. By a similar token, there are some indications of productivity improvements in information service firms close to road openings, but this result is clearly induced by an underlying productivity trend, when we consider the broader study region.

In the manufacturing sector in Coastal Southern Norway, consumer material manufacturing firms close to road openings seem to have escaped a negative underlying trend, while the opposite is the case for technological manufacturing firms. Yet, neither of these results hold when the extended study region is applied. In case of technological manufacturing, the development may have been influenced by the rise of the regional oil and gas supply manufacturing industry (which we have omitted from our study, confer subsection 3.1) in the study region over the study period. For the other industries investigated, our main regressions at firm level produce no significant results. Overall, the point results are influenced considerably by what market access function is applied, but the overall patterns are the same.

Supplementary investigations to these analyses are reported in subappendix B.1. Breakdown of firms lasting over our study region into subsamples based on firm size gave few indications of productivity impulses from increased market access. Our supplementary analyses suggest that small firms in Coastal Southern Norway within tourism, material manufacturing and wholesale and vehicle trade have had a stronger productivity growth than other firms in the region. Yet, these results do no longer hold when we consider a broader control region.

With regard to robustness, alternative estimation procedures on firms lasting over our study period and less aggressive distance decay patterns produce somewhat different results, occasionally involving slightly stronger indications of productivity impulses for consumer manufacturing firms and weaker indications for tourism firms within Coastal Southern Norway. In particular, our baseline results involving positive productivity impulses for tourism firms at firm level do not appear very robust, while positive productivity impulses on firms within consumer manufacturing appear to be subject to an underlying trend in case of application of the extended study region. Both firms within tourism and consumer manufacturing seemingly receive positive productivity impulses from increased market access when less aggressive distance decay patterns are assumed, but only when our regressions are limited to our core study region. Exclusion of our regression control for terms of trade from our baseline regression and implicit estimation of the market access impulses in the production estimation also involve some weak indications of productivity impulses, but the differences from the baseline regressions are likely to be related to lack of control for industry-specific developments.

Moreover, the patterns remain the same with few indications of productivity impulses when a broader study region is applied. In the subappendix B.1, we also consider impulses from higher market access through road openings on other components of technical productivity than TFP at firm level, including scale efficiency and allocative efficiency. These analyses clearly suggested that no such impulses were in play.

Overall, our firm level investigation does not provide any convincing and unequivocal evidence of firm productivity impulses from regional integration caused by major road openings. In some industries, there are weak indications that regional firms close to road openings have experienced stronger productivity growth than other regional firms prior to openings. Yet, these results are no longer significant when a broader study region is applied.
Table 3. Impacts and placebo impacts of increased market access caused by traveling time reduction on TFP over industries at firm level, estimated by linear panel data regressions using industry-specific year dummies, firm fixed effects and clustered standard errors on firms. Market access is measured with power distance decay ($\delta^{pow} = 2.3$) and exponential distance decay ($\delta^{exp} = 0.07$). TFP is pre-estimated by Wooldridge’s estimation procedure with terms of trade as control variable and correction for firm exits. (* for $p < 0.1$, ** for $p < 0.05$ and *** for $p < 0.01$)

<table>
<thead>
<tr>
<th>Impact on TFP</th>
<th>Study region</th>
<th>Core</th>
<th>Impact</th>
<th>Extended</th>
<th>Placebo</th>
<th>Extended</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry \ distance decay</td>
<td>Power</td>
<td>Exp.</td>
<td>Power</td>
<td>Exp.</td>
<td>Power</td>
<td>Exp.</td>
</tr>
<tr>
<td>Consumer manufacturing</td>
<td>0.614</td>
<td>0.867</td>
<td>-0.047</td>
<td>0.108</td>
<td>-6.433***</td>
<td>-3.318**</td>
</tr>
<tr>
<td></td>
<td>(0.460)</td>
<td>(0.599)</td>
<td>(0.260)</td>
<td>(0.331)</td>
<td>(0.951)</td>
<td>(1.370)</td>
</tr>
<tr>
<td></td>
<td>n = 1,139</td>
<td>n = 1,139</td>
<td>n = 4,770</td>
<td>n = 4,770</td>
<td>n = 510</td>
<td>n = 510</td>
</tr>
<tr>
<td>Material manufacturing</td>
<td>1.289</td>
<td>0.752</td>
<td>0.002</td>
<td>-0.121</td>
<td>0.190</td>
<td>0.235</td>
</tr>
<tr>
<td></td>
<td>(0.947)</td>
<td>(0.700)</td>
<td>(0.131)</td>
<td>(0.131)</td>
<td>(0.313)</td>
<td>(0.387)</td>
</tr>
<tr>
<td></td>
<td>n = 1,348</td>
<td>n = 1,348</td>
<td>n = 6,291</td>
<td>n = 6,291</td>
<td>n = 651</td>
<td>n = 651</td>
</tr>
<tr>
<td>Technological manufacturing</td>
<td>0.439</td>
<td>0.662</td>
<td>1.092</td>
<td>0.857</td>
<td>1.471***</td>
<td>3.089***</td>
</tr>
<tr>
<td></td>
<td>(0.407)</td>
<td>(0.533)</td>
<td>(1.047)</td>
<td>(1.047)</td>
<td>(0.494)</td>
<td>(1.013)</td>
</tr>
<tr>
<td></td>
<td>n = 841</td>
<td>n = 841</td>
<td>n = 3,787</td>
<td>n = 3,787</td>
<td>n = 387</td>
<td>n = 387</td>
</tr>
<tr>
<td>Wholesale and vehicle trade</td>
<td>0.985</td>
<td>1.500</td>
<td>0.067</td>
<td>0.079</td>
<td>0.285</td>
<td>0.336</td>
</tr>
<tr>
<td></td>
<td>(0.625)</td>
<td>(1.110)</td>
<td>(0.121)</td>
<td>(0.123)</td>
<td>(0.234)</td>
<td>(0.365)</td>
</tr>
<tr>
<td></td>
<td>n = 3,793</td>
<td>n = 3,793</td>
<td>n = 17,113</td>
<td>n = 17,113</td>
<td>n = 1,709</td>
<td>n = 1,709</td>
</tr>
<tr>
<td>Retail trade</td>
<td>0.078</td>
<td>-0.355</td>
<td>0.112</td>
<td>0.061</td>
<td>-0.109</td>
<td>0.828</td>
</tr>
<tr>
<td></td>
<td>(0.327)</td>
<td>(0.829)</td>
<td>(0.152)</td>
<td>(0.216)</td>
<td>(1.054)</td>
<td>(1.926)</td>
</tr>
<tr>
<td></td>
<td>n = 5,014</td>
<td>n = 5,014</td>
<td>n = 20,832</td>
<td>n = 20,832</td>
<td>n = 2,306</td>
<td>n = 2,306</td>
</tr>
<tr>
<td>Transportation</td>
<td>0.116</td>
<td>0.038</td>
<td>0.035</td>
<td>0.060</td>
<td>4.187</td>
<td>2.942</td>
</tr>
<tr>
<td></td>
<td>(0.682)</td>
<td>(0.869)</td>
<td>(0.096)</td>
<td>(0.101)</td>
<td>(2.753)</td>
<td>(1.863)</td>
</tr>
<tr>
<td></td>
<td>n = 1,572</td>
<td>n = 1,572</td>
<td>n = 8,770</td>
<td>n = 8,770</td>
<td>n = 670</td>
<td>n = 670</td>
</tr>
<tr>
<td>Tourism</td>
<td>10.85*</td>
<td>19.05***</td>
<td>1.174</td>
<td>0.191</td>
<td>-1.090***</td>
<td>-1.457***</td>
</tr>
<tr>
<td></td>
<td>(5.738)</td>
<td>(7.171)</td>
<td>(4.222)</td>
<td>(3.534)</td>
<td>(0.372)</td>
<td>(0.643)</td>
</tr>
<tr>
<td></td>
<td>n = 1,968</td>
<td>n = 1,968</td>
<td>n = 8,530</td>
<td>n = 8,530</td>
<td>n = 821</td>
<td>n = 821</td>
</tr>
<tr>
<td>Information services</td>
<td>0.313*</td>
<td>0.563*</td>
<td>0.111*</td>
<td>0.124*</td>
<td>-0.172</td>
<td>-0.250</td>
</tr>
<tr>
<td></td>
<td>(0.188)</td>
<td>(0.288)</td>
<td>(0.068)</td>
<td>(0.072)</td>
<td>(0.261)</td>
<td>(0.313)</td>
</tr>
<tr>
<td></td>
<td>n = 1,434</td>
<td>n = 1,434</td>
<td>n = 6,191</td>
<td>n = 6,191</td>
<td>n = 623</td>
<td>n = 623</td>
</tr>
<tr>
<td>Knowledge consultancy</td>
<td>0.232</td>
<td>0.138</td>
<td>0.143</td>
<td>0.074</td>
<td>0.234</td>
<td>0.132</td>
</tr>
<tr>
<td></td>
<td>(0.234)</td>
<td>(0.247)</td>
<td>(0.092)</td>
<td>(0.101)</td>
<td>(0.661)</td>
<td>(0.700)</td>
</tr>
<tr>
<td></td>
<td>n = 5,068</td>
<td>n = 5,068</td>
<td>n = 20,271</td>
<td>n = 20,271</td>
<td>n = 2,039</td>
<td>n = 2,039</td>
</tr>
<tr>
<td>Business support</td>
<td>0.132</td>
<td>0.043</td>
<td>0.510</td>
<td>0.301</td>
<td>0.671</td>
<td>1.499*</td>
</tr>
<tr>
<td></td>
<td>(0.366)</td>
<td>(0.416)</td>
<td>(0.495)</td>
<td>(0.270)</td>
<td>(0.626)</td>
<td>(0.791)</td>
</tr>
<tr>
<td></td>
<td>n = 1,203</td>
<td>n = 1,203</td>
<td>n = 5,747</td>
<td>n = 5,747</td>
<td>n = 477</td>
<td>n = 477</td>
</tr>
</tbody>
</table>

5.3 Productivity Effects at Industry Level

Firm level productivity impulses generally do not coincide with the industry productivity at municipal level. First, productivity impulses may be uneven distributed over firm size, where unweighted firm level regressions like ours will put relatively much weight on small firms. To the contrary, our municipal level regressions will weight each municipality equally, regardless of their industry size provided that they are active. Second, factor inputs many be reallocated to more productive firms, inducing industry productivity gains even without firm productivity impulses.
As we do not find any clear indications of productivity impulses at firm level, intra-industry composition effects between firms appear to be plausible explanations for potential industry productivity impulses of increased market access caused by new road constructions. Furthermore, industrial observation units at municipal level imply that industries in municipalities with low activity are weighted equally to major municipal industry clusters. This weighing will also affect the results, if factor inputs of an industry are reallocated between relatively urban and relative rural areas subsequent to a major road opening.

In our baseline regressions at industry level, we utilize municipal industries as observation units. Intra-industry composition effects over regions are investigated further in relation to regional composition effects in subsection 5.5. Our baseline results for industry productivity are depicted in
Table 4.

The three manufacturing industries and wholesale and vehicle trade all show signs of productivity impulses from the road openings, but in all four cases the placebo tests reveal underlying development trends. The remaining service industries show few signs of productivity impulses, although in some of them, both positive and negative significant underlying development trends seem to have changed (i.e. retail trade, transportation industry and information services).

Supporting examinations on potential productivity effects at industry level are to be found in subappendix B.3. Results from alternative specifications to our baseline regressions reported in
Table 4 produces similar results without any notable different conclusions (i.e. involving estimation on disaggregated industry aggregates, alternative approaches to TFP estimation and less aggressive distance decay).

In our investigation of potential intra-industry composition effects contributing to higher productivity at industry level, we have also decomposed each industry into masses of lasting firms and entering and exiting firms. These investigations give some indications of a restructuring of the retail industry next to road constructions, but not for other industries.

In line with similar investigations at firm level, we neither find signs of increased scale efficiency nor allocative efficiency. For retail trade and to a somewhat lesser extent business support, we find signs of retail industry restructuring towards larger multi-branch firms in areas close expansions in the road network, subsequent to the corresponding road openings.
Table 4. Impacts and placebo impacts of increased market access caused by traveling time reduction on TFP over industries at municipal level, estimated by linear panel data regressions using year dummies, municipality fixed effects and clustered standard errors on municipalities. Market access is measured with power distance decay ($\delta^\text{pw} = 2.3$) and exponential distance decay ($\delta^\text{exp} = 0.07$). TFP is pre-estimated by Wooldridge’s estimation procedure with terms of trade as control variable and correction for industry closures. (* for $p < 0.1$, ** for $p < 0.05$ and *** for $p < 0.01$)

<table>
<thead>
<tr>
<th>Impact on TFP</th>
<th>Distance decay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Core</td>
</tr>
<tr>
<td></td>
<td>Power</td>
</tr>
<tr>
<td>Consumer manufacturing</td>
<td>4.551***</td>
</tr>
<tr>
<td>(1.392)</td>
<td>n = 174</td>
</tr>
<tr>
<td>Material manufacturing</td>
<td>5.082**</td>
</tr>
<tr>
<td>(2.180)</td>
<td>n = 176</td>
</tr>
<tr>
<td>Technological manufacturing</td>
<td>10.32*</td>
</tr>
<tr>
<td>(5.280)</td>
<td>n = 162</td>
</tr>
<tr>
<td>Wholesale and vehicle trade</td>
<td>2.192***</td>
</tr>
<tr>
<td>(0.483)</td>
<td>n = 176</td>
</tr>
<tr>
<td>Retail trade</td>
<td>-0.553</td>
</tr>
<tr>
<td>(0.371)</td>
<td>n = 173</td>
</tr>
<tr>
<td>Transportation</td>
<td>-0.689</td>
</tr>
<tr>
<td>(1.189)</td>
<td>n = 164</td>
</tr>
<tr>
<td>Tourism</td>
<td>0.925</td>
</tr>
<tr>
<td>(2.214)</td>
<td>n = 164</td>
</tr>
<tr>
<td>Information services</td>
<td>-1.041</td>
</tr>
<tr>
<td>(1.235)</td>
<td>n = 163</td>
</tr>
<tr>
<td>Knowledge consultancy</td>
<td>0.480</td>
</tr>
<tr>
<td>(0.390)</td>
<td>n = 176</td>
</tr>
<tr>
<td>Business support</td>
<td>-0.592</td>
</tr>
<tr>
<td>(1.857)</td>
<td>n = 161</td>
</tr>
</tbody>
</table>

### 5.4 Inter-Industry Factor Composition Effects

At a given location, aggregate productivity impulses from regional integration may occur due to inter-industry composition effects, even when industry level productivity impulses are absent. Such composition effects between industries necessarily involve reallocation of factor inputs to more efficient factor utilization. We will focus mainly on reallocation of labor, as the aggregate labor’s share of net value added is nearly 90 percent, and as fixed capital either tends to be mobile (e.g. machinery and equipment, means of transportation and immaterial capital) or immobile (e.g. buildings and construction) subsequent to investments.
In principle, there is a distinction between industries that provide constructions and services for the local community and industries producing commodities and services which are primarily targeted towards export outside the region. Industries dominated by export will be more flexible with regard to adjustment in activity level than industries dominated by the provision of goods consumed locally. In practice, the distinction is not clean cut, as some industries may produce goods for both local and external consumption, where the degree of exporting activities depends on competitiveness and composition of their deliveries. Firms in some industries may also deliver customized deliveries to dominate firms in their local industry, which in turn are tied up in global value chains. Overall, a strengthening of the regional trade balance is in principle not tied to particular industries in the long run, making the activity levels of the associated industries considerably more flexible than the ones associated with local service industries. For all industries, however, improvements in the work distribution between industries within or across municipalities could contribute to higher productivity at macro level.

In our investigation of composition effects between industries, we will carry out regression analyses at municipal level for each industry, but also consider the aggregate developments in our core study region and the rest of our extended study region. As mentioned in the previous subsection, usage of municipalities as unweighted observation units implies that municipal industry observations – with limited, but positive, activities – are weighted equally as municipal industries representing larger industry clusters.

In Table 5, we shed light on the overall movement of labor across sectors over the study period. As we have employment figures for branches in multi-branch firms in addition to single-branch firms, we include these in our baseline investigation. On the left hand side of the table, we have shown the labor costs in each industry in the core study region and in the rest of the extended study region. Differences in labor costs do of course partially reflect differences in human capital, but still give an indication of the industries’ labor compensation capability. In the middle of the table, we depict the employment distribution over industries in the initial study year of 2004. On the right hand side of the table, we show the growth in employment shares measured in percentage points from 2004 and 2008, and further from 2008 to 2014. As the largest changes in market access take place in 2009, the period of 2004 to 2008 could roughly be considered as the pretreatment period, while 2009 to 2014 can be roughly considered as the treatment period for this exercise.

Overall, Table 5 shows no clear indications that employment distribution has moved towards industries with relatively high labor cost levels (i.e. labor cost level above average). The manufacturing sector held a relatively strong position in Coastal Southern Norway over our study period, which is actually even stronger than the table suggests, as we have omitted oil and gas suppliers from our study sample. Prior to 2009, the employment shares of high-remuneration industries were mostly growing more (measured in percentage points) in the extended study region than along the coastline of Southern Norway, with technological manufacturing as a notable exception.

Among the service industries with relatively low remuneration (i.e. labor cost level below average), the employment shares of Coastal Southern Norway grew relatively much in the tourism industry and the retail industry and relatively little in the business support industry.
Table 5. Annual growth in employment shares in single- and multi-branch firms across industries measured by percentage points with additional depiction of average relative labor costs and initial employment shares. The shares are limited to our industry selection. Labor costs are deflated with basis in average cost levels from 1999 to 2014. Comparison between Coastal Southern Norway (core study region) and municipalities in neighboring counties. Industries with higher labor costs or growth in employment share than the benchmark region are indicated by green color.

<table>
<thead>
<tr>
<th>Industry / study region</th>
<th>Impact on employment</th>
<th>Relative labor cost level</th>
<th>Share of employment in single- and multi-branch firms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Average 2004-2014</td>
<td>Status 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Core</td>
<td>Extended except core</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Core</td>
<td>Core</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Core</td>
<td>Extended except core</td>
</tr>
<tr>
<td>Consumer manufacturing</td>
<td></td>
<td>0.9685</td>
<td>0.0688</td>
</tr>
<tr>
<td>Material manufacturing</td>
<td></td>
<td>1.2066</td>
<td>0.1563</td>
</tr>
<tr>
<td>Technological manufacturing</td>
<td></td>
<td>1.4534</td>
<td>0.1208</td>
</tr>
<tr>
<td>Wholesale and vehicle trade</td>
<td></td>
<td>1.0852</td>
<td>0.1259</td>
</tr>
<tr>
<td>Retail trade</td>
<td></td>
<td>0.6090</td>
<td>0.1968</td>
</tr>
<tr>
<td>Transportation</td>
<td></td>
<td>0.9975</td>
<td>0.0866</td>
</tr>
<tr>
<td>Tourism</td>
<td></td>
<td>0.5516</td>
<td>0.0695</td>
</tr>
<tr>
<td>Information services</td>
<td></td>
<td>1.3768</td>
<td>0.0483</td>
</tr>
<tr>
<td>Knowledge consultancy</td>
<td></td>
<td>1.3224</td>
<td>0.0617</td>
</tr>
<tr>
<td>Business support</td>
<td></td>
<td>0.7365</td>
<td>0.0654</td>
</tr>
</tbody>
</table>

In the period from 2008 to 2014, the Coastal Southern Norway’s relatively development patterns in employment share of remained largely the same in the period from 2004 to 2008. Among industries with relatively high wage levels, Coastal Southern Norway’s relative growth in employment shares measured by percentage points rose compared to the previous growth period in some industries (e.g. material manufacturing, wholesale and vehicle trade and information services) and fell in others (e.g. technological manufacturing and knowledge consultancy). By the same token, the relative developments in employment composition for Coastal Southern Norway show no clear patterns of restructuring away from low-remuneration industries (e.g. the growth in employment shares measured by percentage points becomes relatively weaker for tourism and business support, but stronger for retail trade).

Although regional developments show no clear signs of factor improvements between industries caused by road openings, restructuring between industries over municipalities within Southern Norway could potentially have contributed to higher regional value creation. In
Table 6, we consider how municipal employment shares (given our industry selection) have been affected by increased market access through new road constructions, including both single- and multi-branch firms.

Our results suggest that the tourism industry has experienced relatively high employment growth in municipalities close to road openings, at least within Coastal Southern Norway. A similar development after adjustment for a negative underlying trend is seen in the retail industry, where the results once again are stronger within our core study region. On the other hand, the municipal knowledge consultancy industries located close to new road construction appear to have had a negative development subsequently to the corresponding openings.
Table 6. Impacts and placebo impacts of increased market access caused by traveling time reduction on municipal employment shares in single- and multi-branch firms over industries, estimated by linear panel data regressions using municipality fixed effects and clustered standard errors on municipalities. Market access is measured with power distance decay ($\delta_{\text{pow}} = 2.3$) and exponential distance decay ($\delta_{\text{exp}} = 0.07$). Core region: Coastal Southern Norway. Extended region: Coastal Southern Norway and municipalities in neighboring counties not located next to road openings. (* for $p < 0.1$, ** for $p < 0.05$ and *** for $p < 0.01$)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry \ distance decay</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumer manufacturing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.008</td>
<td>-0.060</td>
<td>0.045</td>
<td>0.017</td>
<td>0.010</td>
<td>-0.018</td>
<td>-0.192</td>
<td>-0.135</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.042)</td>
<td>(0.072)</td>
<td>(0.066)</td>
<td>(0.057)</td>
<td>(0.100)</td>
<td>(0.094)</td>
<td>(0.132)</td>
<td>(0.095)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 174</td>
<td>n = 174</td>
<td>n = 1,079</td>
<td>n = 1,079</td>
<td>n = 80</td>
<td>n = 80</td>
<td>n = 500</td>
<td>n = 500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Material manufacturing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.132</td>
<td>0.057</td>
<td>-0.394***</td>
<td>-0.288**</td>
<td>1.263**</td>
<td>0.964</td>
<td>0.229</td>
<td>0.189</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.187)</td>
<td>(0.230)</td>
<td>(0.126)</td>
<td>(0.113)</td>
<td>(0.582)</td>
<td>(0.570)</td>
<td>(0.194)</td>
<td>(0.161)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 176</td>
<td>n = 176</td>
<td>n = 1,151</td>
<td>n = 1,151</td>
<td>n = 80</td>
<td>n = 80</td>
<td>n = 535</td>
<td>n = 535</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technological manufacturing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.024</td>
<td>0.090</td>
<td>-0.051</td>
<td>0.005</td>
<td>-0.070</td>
<td>-0.083</td>
<td>0.197</td>
<td>0.108</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.146)</td>
<td>(0.198)</td>
<td>(0.148)</td>
<td>(0.137)</td>
<td>(0.271)</td>
<td>(0.251)</td>
<td>(0.161)</td>
<td>(0.116)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 172</td>
<td>n = 172</td>
<td>n = 1,026</td>
<td>n = 1,026</td>
<td>n = 76</td>
<td>n = 76</td>
<td>n = 467</td>
<td>n = 467</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wholesale and vehicle trade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.049</td>
<td>-0.051</td>
<td>0.091</td>
<td>0.075</td>
<td>-0.177</td>
<td>-0.101</td>
<td>-0.078</td>
<td>-0.098*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.064)</td>
<td>(0.083)</td>
<td>(0.144)</td>
<td>(0.102)</td>
<td>(0.184)</td>
<td>(0.185)</td>
<td>(0.065)</td>
<td>(0.052)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 176</td>
<td>n = 176</td>
<td>n = 1,159</td>
<td>n = 1,159</td>
<td>n = 80</td>
<td>n = 80</td>
<td>n = 528</td>
<td>n = 528</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retail trade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.170**</td>
<td>0.034</td>
<td>-0.018</td>
<td>-0.090</td>
<td>-0.966**</td>
<td>-0.812**</td>
<td>-0.363***</td>
<td>-0.280**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.080)</td>
<td>(0.134)</td>
<td>(0.158)</td>
<td>(0.121)</td>
<td>(0.332)</td>
<td>(0.297)</td>
<td>(0.120)</td>
<td>(0.114)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 176</td>
<td>n = 176</td>
<td>n = 2,156</td>
<td>n = 2,136</td>
<td>n = 80</td>
<td>n = 80</td>
<td>n = 559</td>
<td>n = 559</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transportation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.085*</td>
<td>-0.081</td>
<td>0.045</td>
<td>-0.008</td>
<td>-0.134</td>
<td>-0.112</td>
<td>0.354*</td>
<td>0.490*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.043)</td>
<td>(0.073)</td>
<td>(0.133)</td>
<td>(0.094)</td>
<td>(0.367)</td>
<td>(0.299)</td>
<td>(0.199)</td>
<td>(0.257)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 176</td>
<td>n = 176</td>
<td>n = 1,138</td>
<td>n = 1,138</td>
<td>n = 80</td>
<td>n = 80</td>
<td>n = 514</td>
<td>n = 514</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tourism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.084***</td>
<td>0.076**</td>
<td>0.139</td>
<td>0.155*</td>
<td>0.047</td>
<td>0.033</td>
<td>0.014</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.026)</td>
<td>(0.032)</td>
<td>(0.097)</td>
<td>(0.078)</td>
<td>(0.159)</td>
<td>(0.144)</td>
<td>(0.098)</td>
<td>(0.114)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 164</td>
<td>n = 164</td>
<td>n = 1,112</td>
<td>n = 1,112</td>
<td>n = 72</td>
<td>n = 72</td>
<td>n = 497</td>
<td>n = 497</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.057</td>
<td>-0.069</td>
<td>0.015</td>
<td>-0.010</td>
<td>0.011</td>
<td>0.003</td>
<td>-0.075</td>
<td>-0.111</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.052)</td>
<td>(0.048)</td>
<td>(0.049)</td>
<td>(0.049)</td>
<td>(0.050)</td>
<td>(0.042)</td>
<td>(0.088)</td>
<td>(0.087)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 168</td>
<td>n = 168</td>
<td>n = 956</td>
<td>n = 956</td>
<td>n = 74</td>
<td>n = 74</td>
<td>n = 434</td>
<td>n = 434</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge consultancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.088***</td>
<td>-0.093**</td>
<td>0.025</td>
<td>0.031</td>
<td>0.046</td>
<td>0.050</td>
<td>0.072</td>
<td>0.068</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.021)</td>
<td>(0.032)</td>
<td>(0.054)</td>
<td>(0.038)</td>
<td>(0.106)</td>
<td>(0.096)</td>
<td>(0.078)</td>
<td>(0.067)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 176</td>
<td>n = 176</td>
<td>n = 1,195</td>
<td>n = 1,195</td>
<td>n = 80</td>
<td>n = 80</td>
<td>n = 546</td>
<td>n = 546</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.128</td>
<td>0.151</td>
<td>0.052</td>
<td>0.062</td>
<td>0.165</td>
<td>0.166</td>
<td>-0.088</td>
<td>-0.037</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.078)</td>
<td>(0.106)</td>
<td>(0.084)</td>
<td>(0.076)</td>
<td>(0.108)</td>
<td>(0.111)</td>
<td>(0.106)</td>
<td>(0.113)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 166</td>
<td>n = 166</td>
<td>n = 953</td>
<td>n = 953</td>
<td>n = 73</td>
<td>n = 73</td>
<td>n = 412</td>
<td>n = 412</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In subappendix B.4, we report supporting analyses to the investigations on industry composition effects carried out in this subsection. The most striking finding of these investigations is the lack of agreement of the results for only single branch firms and all firm branches, including branches of firms that were multi-branch firms at least once over our study period. This mismatch supports the hypothesis that the road constructions in Coastal Southern Norway have not induced substantial composition effects between industries at either municipal or regional level. Another difference is that the employment development for business support services becomes relatively negative, when employment stocks replaces employment shares as outcome variable in the baseline regressions. Considering changes in fixed capital, our supportive analyses suggest that increased market access through expansion of the road network had no clear impact on the fixed capital intensity.
5.5 Regional Factor Composition Effects

Regional integration through new road construction may contribute to higher factor return on regional level, even in the absence of positive productivity impulses and factor composition effects at local level. Such effects could occur through improvements in factor composition effects between different locations, where factor inputs are reallocated from regions with relative higher factor return after the implementation of a transportation measure. In Fig. 7, we have considered how the regional commuting into and out from Kristiansand city has developed over our study period. To make the figure easier to read, we have colored the municipalities in Aust-Agder county – which to a varying degree were affected by the opening of European Route 18 between Kristiansand and Grimstad August 2009 – in blue. By the same token, we have colored the municipalities in Vest-Agder county – on the other side of Kristiansand – in green.

![Fig. 7. a) Outgoing commuting rates (l.h.s.) and b) incoming commuting rates (r.h.s) into and out from Kristiansand in 2004, 2008 and 2014 with depiction of county (Aust-Agder in blue and Vest-Agder in green)](image)

We see that all municipalities located in Aust-Agder experienced increased commuting towards Kristiansand from 2008 to 2014. The largest increases are seen in Lillesand, Grimstad and Arendal, which are the three municipalities affected the most by the renewal of European Route 18 in addition to Kristiansand. On the other hand, the commuting rates to Kristiansand from other municipalities in Vest-Agder remained the same or even declined in this period, such as in Songdalen and Sogne. We also note that Flekkefjord – which experienced traveling time reductions from August 2006 through the renewal of European Route 39 – increased its commuting to Kristiansand from 2004 to 2008, although the levels were low, and the rates decreased again from 2008 to 2014.

If we look at incoming commuting to the other municipalities of Coastal Southern Norway from Kristiansand, we still see clear increases in the commuting rates concerning Aust-Agder. Yet, the municipalities on the other side of Kristiansand, particularly Vennesla and Sogne, also experienced increased commuting from the city. Thus, we are reluctant to make categorical conclusions on the role of the opening of European Route 18. The increase of commuting between Kristiansand and Lillesand – the municipality measured to be subject to the highest increase in market access – could be seen in conjunction with the Norwegian Communications Authority’s move from...
Oslo to Lillesand May 2007. With the new directorate followed 140 workplaces, which corresponded to 4.2 percent of the workplaces in Lillesand at the time.

In subappendix B.5, we carry out supplementary investigations on regional composition effects. Here, we examine the changes in commuting patterns within Southern Aust-Agder in relation to the regional road opening at European Route 18 and within the Lister Region in relation to the regional road openings at European Route 39 and County Route 465 in Vest-Agder. In this analysis, we also find that commuting into the more urban municipalities has increased, subsequent to the openings. When conducting a regression analysis on outgoing commuting after the increase in market access caused by expansions in the road network, the results become insignificant, although they are significantly different from the placebo trend. This weakens the hypothesis that increased market access improves regional social wealth considerably through stimulating commuting somewhat, even though alternative investigations and various explanations could be put forward in the opposite direction. Investigating the potential gains of commuting, it appears as if both the highest levels and increases in outgoing commuting shares, and possible gains for commuting, occur from the municipalities in Coastal Aust-Agder to Kristiansand in relation to the opening at European Route 18. Yet, we approximate these gains to be relatively modest in the big picture, amounting to less than NOK 10 million in fixed prices at average price level from 1999 to 2014.

Looking more closely at the regional work distributions within each industry, we do not find any clear indications of a welfare-improving reallocation of labor between Kristiansand and its eastern suburbs (defined as Lillesand, Grimstad and Arendal) subsequent to the road opening at European Route 18. Some weak tendencies are still seen in the data, inter alia including a weakening of the retail industry in Kristiansand’s eastern suburbs.

At last, we conduct regression analyses on how various measures for employment and fixed capital holdings are affected by increased market access through road constructions at macro level, controlling for local oil and gas supply activities. In these regression analyses, we do not find any sign of increased factor input utilization in municipalities next to new road constructions, subsequent to their opening.

6. Discussion and Conclusions

Urban areas are on average more productive than peripheral areas. In the economic geography literature, there is a growing consensus that firms in principle may benefit from high mobility to surrounding areas. While traveling distances seldom change much, new major road construction projects may reduce traveling time substantially. This suggests that studies of road constructions might shed light on the relationship between productivity and economic density.

Some recent papers provide microeconomic evidence on productivity impulses from increased market access through expansion in the road network (e.g. Duranton, Morrow and Turner 2014, Hall 2016 and Gibbons et al. 2019). Yet, productivity impulses in rural areas in developed countries and composition effects are seldom addressed. Despite of limited of empirical evidence, supplementary quantitative analyses on wider economic impacts constitute an integrated part of the transportation appraisal practices in many sparsely populated countries (e.g. Wangsness, Rødseth and Hansen 2017 and Holmen, Biesinger and Hindriks 2022).

As a sparsely populated country with solid state finances, Norway has invested relatively much in road infrastructure over the last decades. Coastal Southern Norway constitutes a transparent and streamlined region subject to several major road constructions over a brief period of time. Thus, it emerges as a good case study suited to providing new insights to the literature on productivity impulses from market access and transportation networks with focus on productivity, composition effects and rural areas. Since the turn of the millennium, no
road openings have reduced traveling time between each other for more a larger number of Norwegians with each other (through a traveling time reduction of at least five minutes within 30 minutes’ reach) than European Route 18 at in Coastal Southern Norway. If regional integration caused by road constructions were to boost total factor productivity at firm level in sparsely populated areas of developed countries, Coastal Southern Norway therefore appears as a reasonable place to look for such effects.

In this paper, we have studied productivity impulses from increased regional integration through road constructions in the business sector of Coastal Southern Norway from 2004 to 2014. We assess both productivity impulses at firm level and composition effects within and between industries and locations, which possibly contribute to higher factor return for the region as a whole. Admittedly, road constructions may be motivated by or related to productivity, so the causality in the decision-making must be addressed in such studies. Both our own causality checks and previous studies in the literature (reviewed in more detail in appendix A) suggest that the implementation of Norwegian road constructions is exogenous with respect to productivity, we proceed to our analyses on productivity impulses from increased market access without instrumentation of the latter variable. Still, we operate with buffer zones of twenty traveling kilometers (measured in the initial year) around each receiver of impulses from market access, where the traveling times are held constant, both to handle potential local endogeneity and data noise.

We measure market access under assumptions of power distance decay and exponential distance decay, exploiting the estimates for the Southern parts of Norway from Holmen (2022a) in our baseline regressions. As we are only interested in increases in market access caused by expansions in the road network and not the ones that are intertwined with urbanization developments, we hold our measure for potential market connections (i.e. mean of employment by residence and workplace) constant in our operational implementation of the market access measure.

As our baseline framework for industry-specific empirical examinations, we apply a panel data linear least square specification with industry-year dummies and observation units (i.e. firms or municipal industries) as basis for fixed effects and clustering of standard errors. We pre-estimate total factor productivity by Wooldridge's estimation procedure, controlling for developments in terms of trade. To be able to distinguish between local displacement effects and more direct contributions to higher social wealth (before potential displacement effects and ripple effects through general equilibrium mechanism), we conduct our analyses both on our core study region alone (i.e. Coastal Southern Norway) and a broader study region, also consisting of comparable municipalities in neighboring counties (except the ones next to new road openings). As actual impulses from expansions in the road network may easily be confused with underlying development trends, we also introduce a framework for placebo testing, where we pretend that road openings were realized at an earlier point in time. To alter robustness and obtain a deeper insight into the study phenomena, various further investigations and robustness tests were carried out in addition, as reported in appendix B.

Overall, our empirical investigations neither provide robust support to productivity impulses from road constructions at firm or industry level. Furthermore, welfare-enhancing composition effects between industries do not appear to be prevailing. Our results provide some support to the hypothesis that industry restructuring has occurred in areas next to new road constructions, inter alia within the retail trade industry. However, these tendencies do not provide traceable productivity impulses that could be traced to higher aggregation level in a robust manner and may be interlinked with underlying development patterns. In principle, fiercer competition for some industries (e.g. retail trade and transportation) or more efficient regional industry organization could have led to higher consumer surplus for households rather than increased producer surplus. Yet, our study does not provide supporting evidence for such hypotheses.
Lastly, we consider regional composition effects. There are clear indications that regional integration through new road constructions contributes to higher commuting to the rural integrated area to the urban areas. This is particularly the case for the commuting from Coastal Aust-Agder to the regional city of Kristiansand along European Route 18, which represents the route that induced the highest increase in market access in our study. Nevertheless, we approximate the gain from commuting due to the road opening to be relatively low. Furthermore, we are also unable to relate these developments to a particular reallocation of industry activities. We do not find signs of increased factor usage subsequent to road openings.

Moreover, the economic appraisal associated with the three major road openings assessed in our study involves a negative net benefit of NOK 269.4 million. We estimate the regional increases in market access caused by expansion in the road network to be 7.6 percent in case of power distance decay and 9.6 percent in case of exponential distance decay. Given our rough estimates for the regional GDP in our study region, compensation for the negative net costs of the project would have required an agglomeration elasticity in the range from 0.003 to 0.004. Even though these estimates are relatively low compared to other agglomeration elasticities found in the literature, our empirical analyses suggest that the agglomeration elasticity is unlikely to be anywhere close to this level of magnitude. Overall, our study does not support the hypothesis that expansions in the road network produce substantial productivity impulses in rural areas of developed countries. This suggest that possible inclusion of wider economic impacts in transportation appraisal should be done with caution, particularly in less inhabited areas. Moreover, our study provides little support to the current transportation appraisal practices on wider economic impacts, applied in some sparsely populated countries.

For further research, we encourage more studies on wider economic impacts in rural areas and how such impacts vary over the degree of urbanization. Impulses form increased market access induced by other infrastructure or urbanization should also be explored further. Measurement of market access should also be investigated further, possibly involving choices of functional form and contents, as well as heterogeneities concerning industries and geographic configuration. Furthermore, our empirical results suggest that road openings may have affected the regional organization of some industries, which constitutes a topic that may be explored further, either taking an even more micro-oriented econometric approach or exploring general equilibrium mechanisms. At last, our results on regional factor composition effects suggest that commuting may contribute some to higher regional value creation. This topic should be investigated properly with data on individuals, for instance by matching new commuters with similar individuals in other regions that do not start to commute.

Eventually, the net benefits for the road investments are estimated to be NOK minus 692.9 million for the investment at European Route 18 between Kristiansand and Grimstad, NOK 255.8 million for the investment at European Route 39 between Flekkefjord and Lyngdal and NOK 167.7 million for the investment at County Route 465 in Vest-Agder between Farsund and Kvinesdal (Norwegian Ministry of Transport and Communications 2002 and 2004, and Transportation and Communications Committee at the Storting 2002 and 2005). All this figures are recalculated to fixed prices at average price levels from 1999 to 2014 with analysis year in the middle of this period (i.e. between 2006 and 2007) based on the concurrent appraisal methodology with a discount rate of 0.06, a deflation rate of 0.02 and an appraisal period of 25 years (Norwegian Ministry of Finance 2005 and Norwegian Public Road Administration 2006).

In this calculation, we approximated the regional GDP by two alternative approaches, which led to estimates of roughly the same magnitude. In the first approach, we took basis in Statistics Norway’s regional accounts and approximate our study region’s share of GDP in Aust-Agder and Vest-Agder’s counties by help of employment statistics from Statistics Norway. Second, we took basis in our enterprise data and added value added estimates for firms without reporting duty and public sector firms based on employment and industry statistics from Statistics Norway. In both cases, we also utilized the national accounts deflators.
References


Holmen, R. B. 2022a. Agglomeration Decay in Rural Areas. *Insights into Regional Development*, 4(3), 139-155. [https://doi.org/10.9770/IRD.2022.4.3(9)](https://doi.org/10.9770/IRD.2022.4.3(9))


Irgens, J. B. 1978. Around the main route of Western Norway (in Norwegian). Published by Historielaget for Dypvåg, Holt og Tvedestrand. Contact information for access: janoaslaksen@gmail.com


Nyborg, K., & Spangen, I. 1996. Political decisions about investment in roads, Interviews with members of the Transportation and Communications Committee at the Storting (in Norwegian), TOI note 1026:1996. URL: https://www.toi.no/getfile.php?mmfileid=11400


Rovigatti, G., & Mollisi, V. 2018. PRODEST: Stata module for production function estimation based on the control function approach. Statistical Software Components. URL: [https://ideas.repec.org/b/cod/s458239.html](https://ideas.repec.org/b/cod/s458239.html)


Strand, A. 1983. The county distribution for national route grants (in Norwegian). Samferdsel, 5/1983. Contact information for access: hse@toi.no

Strand, A. 1993. Commitment to transportation - urban policy or rural policy? (in Norwegian). Regionale trender, 1/93. URL: [https://www.nb.no/nbsok/nb/ee283d6a3fb37c4bf62a7255b6093c15](https://www.nb.no/nbsok/nb/ee283d6a3fb37c4bf62a7255b6093c15)


Funding: The paper was funded by the Research Council of Norway, partially through the program, the ‘Industrial PhD-Program’, and partially through the research project ‘Impact’ on wider economic impacts.

Data Availability Statement: In principle, the data applied in this study is available from the data providers reported in the article, including the Norwegian Enterprise register at Brønnøysund, Statistics Norway and the Norwegian Mapping Authority and other data providers of regional economic figures, traveling figures and currencies. As the data processing associated with the study is rather resource demanding, the author is also open to share the processed datasets applied in the regression analyses. Nevertheless, the authors do not own all the raw data applied in the paper, so the they cannot be distributed freely.

Acknowledgement: I want to thank Per Botolf Maurseth and Rune Sorensen at BI Norwegian Business School, Oliver Falck at University of Munich and Steve Gibbons at London School of Economics and Political Science for good advice in relation to my work with this paper. In connection with two consultancy reports for Menon Economics, we conducted introductory empirical analyses, which inspired me to do research on this topic (i.e. Skogstrøm et al. 2013 and 2015). I want to thank Heidi Ulstein, Jens Fredrik Skogstrem, Leo Andreas Grünfeld and Peter Aalen at Menon Economics in this regard. Besides, I want to thank Gretar Ævarsson and Kjetil Haukás at the Institute of Transport Economics, Bjørn Ingeberg Fesche at Menon Economics and Tom Anders Engebakken at Statistics Norway, who contributed to estimating and preparing the raw traveling data applied in this paper. Note that the study was carried out in parallel to the work with Holmen (2022a).

Declaration of Interests: I declare no conflict of interests.

Supplementary Material: With the article follows on-line comprehensive appendixes with background material on the study region and a wide range of robustness checks https://zenodo.org/record/7276788#.Y2OtpeRBxPY. Otherwise, the article is closely related to the work in Holmen (2022a).

Rasmus Bøgh HOLMEN is senior researcher at Institute of Transport Economics. December 2020, he defended his doctoral dissertation in economics on the topic ‘Productivity and Mobility’ at BI Norwegian Business School. Previously, Holmen has several years’ experience from the consultancy industry as senior economist at Menon Economics. He holds a master’s degree in economics from University of Oslo from 2012.

ORCID ID: https://orcid.org/0000-0002-1964-1495
OPERATIONAL RISK FACTORS AND THE SUSTAINABILITY OF SMALL AND MEDIUM MANUFACTURING ENTERPRISES IN SOUTH AFRICA

Urim-Tummim Shipanga ¹, Suzaan Le Roux ², Jobo Dubihlela ³*

¹,²,³ Cape Peninsula University of Technology, District 6 Campus, Hanover and Tenant Street, Cape Town, South Africa

E-mails: ³* dubihlelaj@cput.ac.za (Corresponding author)

Received 15 July 2022; accepted 22 October 2022; published 30 December 2022

Abstract. In South Africa, the success of Small and Medium-sized Enterprises (SMEs) is essential in developing the economy as these businesses are known as drivers of economic growth. However, these SMEs have a high failure rate, mainly caused by operational risks. According to Scholarly literature, operational risk is any event that disrupts the normal flow of business processes. It generates financial loss or damage due to operational risk factors: people, process, system and external events. It was further established in South Africa that when these operational risk factors are not controlled, they can adversely impact the sustainability of SMEs. Most of these manufacturing enterprises do not implement operational risk management as they consider it a privilege of large enterprises. These Manufacturing companies must implement operational risk management to mitigate potential threats. This research study aims to investigate the operational risk factors affecting Manufacturing SMEs’ sustainability in the Cape Metropole. Quantitative data was acquired from 85 respondents (managers/or owners) of Manufacturing SMEs in the Cape Metropole. Based on the results, it was found that all four of these operational risk factors still affect the sustainability of these business entities. To deal with these operational risk factors effectively, SMEs need operational risk management to identify, minimise and mitigate risks that negatively affect their sustainability.

Keywords: Small Medium Enterprises; manufacturing SMEs; sustainability; operational risks; South Africa


JEL Classifications: M1, M13, D23

1. Introduction

South African Small and Medium Enterprises (SMEs) are known as the lifeblood of the national economy due to their potential to assist in attaining core socio-economic objectives (Naicker et al., 2017; Nanziri & Wamalwa, 2021). SMEs were officially introduced to the national economy by the South African government to create jobs, alleviate poverty, and stimulate the national economy (Smit & Watkins, 2012). Furthermore, SMEs are drivers of economic growth, are essential in reducing poverty (Morongwa, 2014), creating employment, contributing to the Gross Domestic Product (GDP) (Lekhanya & Mason, 2014), and play a crucial role in developing countries (Morongwa, 2014).
Despite the above-mentioned, Kalida, Magwentshu, and Rajagopaul (2020) aver that SMEs are struggling to fulfil their potential with great success as an estimate of 70 to 80 per cent of SMEs are reported to fail within their first five years of existence. Most SMEs face several obstacles hampering their sustainability, which, in turn, negatively affect their economic growth (Phaho & Pouris, 2008; Bvuma & Marnewick, 2020). Prior research has suggested that SMEs are exposed to risks, which influence daily processes, reduce proceeds, or increase overheads (Certified Practising Accountant (CPA) Australia, 2018). These risks adversely affect business sustainability, causing the failure of some of these businesses.

One of these major risks that business entities face is an operational risk (Engle, 2009; Bai, Gao & Sarkis, 2021), which is inherent in every human activity and ascends from activities such as acts of fraud, errors, negligence, violations, technological failure events, process deficiencies, system flaws, acts of terrorists and vandalism, as well as natural disasters like floods, earthquakes etc. (Hussain & Shafi, 2014). Operational risk is the potential loss resulting from inadequate or flawed internal processes, people and systems, or external events. However, operational risk is still not prevalent amongst manufacturing SMEs. The research study explores what operational risk factors affect manufacturing companies as there is less research done in this sphere of operational risk in SMEs. According to the survey by Renault Agumba & Ansary (2020), most SMEs do not develop a viable operational risk management program to identify internal risks.

Furthermore, there are no studies done on operational risk management within the manufacturing SMEs in South Africa. Therefore, this research will assist manufacturing companies in identifying their critical operational risks. Finally, it will fill the knowledge gap, add to the existing body of knowledge regarding operational risks and provide valuable guidelines to SME owners and/or managers on how to better manage operational risks more efficiently and effectively to improve the sustainability of their businesses.

2. Review of literature

2.1 SMEs in the manufacturing sector

The South African manufacturing sector is a primary contributor to the country's GDP and has significant potential to create employment opportunities and improve national economic growth (Signé, 2018). According to the DTI (2018), the manufacturing sector contributes directly to the country’s GDP, employment, exports and human capital development. The manufacturing industry in South Africa employs over 1.6 million people. It is rated among the top three multiplier sectors concerning value addition, employment creation, export incomes, and revenue generation for each rand invested (Rodseth, 2018). According to the Western Cape Economic Overview (2010) and Stats SA (2019), the Western Cape manufacturing industry is the third largest contributor to national manufacturing output and employment and the dominating private sector in the province. The manufacturing industry plays an essential role in the reduction of unemployment, alleviation of poverty and the increase in economic growth. The DTI (2008) aims to evaluate and improve the manufacturing capabilities of South Africa into a sustainable and globally competitive industry. However, the manufacturing sector is declining for its influence on employment and economic growth. According to Katanamrp (2022), some challenges affect manufacturing industries. One of the challenges that face manufacturing companies is operational risks. Operational risks are most prevalent in the manufacturing and mining industries (Qeke & Dubihlela, 2018; Bai, Gao & Sarkis, 2021). Manufacturing businesses encounter undesirable events and unwanted setbacks such as machine breakdowns, material shortages, accidents, and absenteeism that make the system unreliable and inconsistent (Islam & Tedford, 2012; Islam, Tedford & Haemmerle, 2008). Bai, Gao & Sarkis (2021) study found several operational risks such as fire, natural disasters and air pollution. The combined effect of different operational risk events could effectively cripple an SME's business performance, which may ultimately put it at risk of complete failure (Islam, Tedford & Haemmerle, 2008; Allen, 2016).
The failure rate of SMEs in South Africa is usually pinned on various economic factors such as management negligence, high-interest rates and ineffective internal control (Ahmad & Seet, 2009; Bushe, 2019; Saah & Musvoto, 2020). These factors ultimately lead to the weak sustainability rate within South African SMEs (Bruwer et al., 2013; Petersen, Bruwer & Le Roux, 2018; Bruwer et al., 2019; Mbomvu et al., 2021). The South African manufacturing industry is thus failing, considering its GDP contribution plummeting for over ten years from 18 per cent (1997) to 11 per cent (2017), resulting in devastating 105,000 job losses (Macpherson, 2018). The inference could be made that South African manufacturing SMEs are ineffective and inefficient as they suffer from various internal and external influences like people, internal processes, technological systems, and external events (Qeke & Dubihlela, 2018), which, in turn, generate operational risks (Bushe, 2019). These risks adversely affect business sustainability, causing the failure of some of these businesses. Hence, this study aims to determine the operational risk affecting manufacturing SMEs' sustainability in the Cape Metropole.

2.2 Sustainability of SMEs

SMEs contribute to a country’s national GDP and economic development in South Africa (Erdin, & Ozkaya, 2020; Saah, 2021). SMEs play numerous roles within an economy, such as generating employment and contributing to the GDP (Lekhanya, 2016; Bruwer, 2020). Stats SA, 2020 reports that small businesses generate just over one-fifth of total turnover. According to Kalidas, Magwentshu & Rajagopaul (2020), SMEs are recognised as an engine of growth, are imperative in reducing poverty, and play an essential role in mainly developing countries. Additionally, they are indispensable for ensuring a competitive and productive market. Business activities and functions should be instrumental in small businesses’ future growth and sustainability (Saah & Musvoto, 2020). SMEs' sustainability in a growing South African economy is of paramount importance as they represented 97.5 per cent of actively registered businesses in 2001 and 96.9 per cent between 2001 and 2007 while contributing approximately 30 per cent to the GDP and an estimated 80 per cent to all local employment opportunities (DTI, 2008). The sustainability of SMEs is a critical component in ensuring the future success of a business (Eccles, Ioannou & Serafeim, 2012; Hoffman, 2018; Saah & Musvoto, 2020).

Despite the importance of SMEs, most SMEs face several obstacles hampering their sustainability and, consequently, harming their economic development (Phaho & Pouris, 2008; Saah & Musvoto, 2020). According to Mbomvu, Hlongwane, Nxazonke, Qayi & Bruwer (2021), South African SMEs have been performing very poorly during the past few years, indicating that several factors are hindering their efforts to operate efficiently and sustainably. There are many reasons why SMEs fail; those mentioned above are attributed to some constraints SMEs face (Msomi & Olarewaju, 2021; Mbomvu et al., 2021). According to Rehman and Anwar (2019) and Yakob, Hafizuddin–Syah, Yakob, and Raziff (2019), risk management significantly affects business performance.

Business entities face several significant risks: strategic, reporting, compliance, and operational risks (Engle, 2009; Chowdhury, Lau, & Pittayachawan, 2019; Xing et al., 2022; Koeplin & Lele, 2022). Within the ambit of this study, the focus will be on operational risk.

2.3 Operational risk factors for SMEs

The Basel Committee on Banking Supervision (2001) defines operational risk as "the risk of loss resulting from inadequate or failed internal processes, people and systems or external events.

Operational risk refers to the possibility of error in information processing, insufficient documentation procedures, or delay in work completion. All manufacturing businesses face unwanted events and impediments (internal and external) in their daily operations (Islam & Tedford, 2012). Furtado, Kolaja, Mueller and Salguero (2020) are in support that manufacturing companies do face operational challenges. The internal causes of failure
include poor management, lack of risk management planning, and failure to adopt a risk limit threshold. The external causes included government policies, the vulnerability resulting from small size, competition from larger businesses, civil strife, natural disasters, and general economic downturns (Arif, Jan & Kulsoom, 2016).

Operational risk factors may be internal or external to the business and are usually generated by people, processes, and systems (Strzelczak, 2007). Internal factors involve people, products or services offered, and operational systems. External factors are the causes from which operational risk may arise (Global Association of Risk Professionals, 2011). These factors emphasised the need for reinforcing controls over operational risk, particularly in the financial area and utilising indicators to keep track of risk exposure tendencies (Ferretti & Birindelli, 2017). Internal and external events to its day-to-day operation put an SME at risk regarding production, safety, and the business itself (Islam, Tedford & Haemmerle, 2008; Bodnar et al., 2019). According to Bai, Gao & Sarkis (2021), small businesses are more likely to have operational risk events as large companies have more significant slack and capacity to absorb the risk. Operational risk can be found at various levels such as a 1) personal level (e.g., unintended errors, inexperience, fraudulent behaviour (recorded as destructions, fabricated or hidden information); 2) procedural level (inappropriate procedures and control regulations to generate detailed reports for national institutions responsible for operational risk, to observe and take decisions); 3) technique level (lack of insufficient tools used to measure operational risk); and 4) technological level (system errors) (Radu & Olteanu, 2008). Major accidents and emergencies rarely occur in SMEs, although small losses, near misses, unsafe acts and unsafe conditions are common occurrences. But, problems, failures and mistakes, as well as incorrect or ineffective actions, are very likely occurrences in the daily business of SMEs (Islam & Tedford, 2012; Ntshangase & Msosa, 2022).

Figure 1 graphically depicts the dimensions of operational risk – a loss that is caused by an operational event, which, in turn, is caused by four different factors: processes, people, systems, and external events (Van den Brink, 2002; BCBS, 2003; Van Grinsven, 2009; Islam & Tedford, 2012; South African Reserve Bank, 2020).

An operational risk loss arises only from an actual operational risk event (Coleman, 2010). According to Bai, Gao & Sarkis (2021), the effect of an operational risk event on a firm financial performance depends on the direct financial losses and the information timeliness of the risk event. Operational losses are mostly known to be caused by individuals but can be caused by all levels of staff, including the Boards of Directors, whether intentional or not (De Jongh et al., 2013). Further, operational losses result from weak management, subcontracting nontactical activities, or external factors (Coleman, 2010). Most substantial losses occur due to operational failures at the
senior level (Advisory & Perrin, 2010). According to Shahudin, Zulkeflee and Cipriano (2018), operational risk arises because of "failure in systems, processes, people and external events", and if this risk is not addressed appropriately, it can result in business failure.

Basel II provides a set of seven categories of operational risk event types having the potential to result in substantial losses, including the following:

- internal fraud (acts meant to deceive, steal property or avoid regulations, company policy or the law);
- external fraud (acts by a third party intended to defraud, avoid the law or misappropriate property);
- employment practices and workplace safety (unpredictable acts concerning health, safety agreements or employment, which result in the payment of personal wound claims or claims relating to multiplicity or discrimination issues);
- clients, products, and business practices (unintended or neglectful failure to meet a professional obligation to explicit clients or due to the nature or design of a product);
- damage to physical assets (loss or damage to physical assets from natural catastrophes or other events);
- business disruption and system failures; execution, delivery, and process management (fruitless transaction processing or process management, or relations with trade counterparties and vendors).

Operational risk events are growing radically, and organisations must develop ways to mitigate them. Operational risks can be mitigated efficiently if organisations learn the core operational vulnerabilities of their businesses and set the risk indicators accordingly.

2.4 Operational Risk Management

Successful SMEs use operational risk management (ORM) to assist them with contemplating risk when making decisions, as it has been proven to be the critical success component of any SME organisation. Mohamud and Salad (2013) define ORM as a process of managing or reducing risks appearing from technical- and human errors. ORM is very effective in preventing the risk that causes operational losses by managing those risks. Therefore, organisations need to implement ORM to succeed.

2.4.1 Stages of operational risk management

According to Pearson (2020), there are different stages in regards to implementing ORM:

**Risk identification**: Understanding the risks specific to a business is critical, but many potential risks affect any business, and you need to identify all of them, both those that are recurring and those that can be once-off events. The identification process needs to involve staff from all levels of the business, bringing various backgrounds and experiences to make a cohesive result. Risks identified by work floor staff will be very different and no less critical than those identified from the boardroom.

**Risk Assessment**: Once the risks have been identified, they need to be assessed. This needs to be done from a quantitative and qualitative perspective, and factors like the frequency and severity of occurrence need to be considered. The assessment needs to prioritise managing these risks concerning those factors.

**Measurement and Mitigation**: Mitigating these risks (if not eliminating them) is the next stage, with controls that should limit the business's exposure to the risks and the potential damage caused by them.

**Mentoring and reporting**: Any ORM plan must have something in place for the ongoing monitoring and reporting of these risks, if only to demonstrate how effective the plan has been. Most of all, it is to ensure that the
solutions put in place continue to be effective and do their job in managing the risks.

The above stages are crucial in implementing ORM as they ensure a successful strategy.

ORM aids management in understanding the cause of the risk to determine what factors affect earnings and manage the risk effectively. This improves a business's overall operation as losses are reduced, and a financial crisis can be avoided.

3. Research Methodology

The research study was empirical, making use of a survey research questionnaire and quantitative research, and fell within the positivistic research paradigm as data were collected from the owners and/or managers of the manufacturing SMEs within the Cape Metropole. The owners and/or managers of manufacturing SMEs were invited to participate in the study, of which a total of 85 participated in this research study. Since all the owners and/or managers mentioned above were invited to participate in the study, the sampling method used was twofold and included both convenience and purposive sampling.

The owners and/or managers participated in the study out of free will (voluntary participation) and could withdraw from the study at any time and without any consequences. To justify a valid response, each respondent had to adhere to the following delineation criteria:

- Respondents must be owners and/or managers of their businesses.
- Respondents must be actively involved in the daily operations of the company.
- Respondents' SMEs must adhere to the South African definition of SMEs stipulated in the National Small Enterprise Act of 2019.
- Respondents' SMEs must employ between five and two hundred and fifty employees.
- Respondents' SMEs must be non-franchised.
- Respondents' SMEs must operate in the manufacturing industry in the Cape Metropole.
- Respondents' SMEs must be in existence for at least three years.

Descriptive statistics were used to present and interpret the basic features of the collected data and derive meaningful, informative results.

4. Results and Discussions

SMEs adhered to strict delineation criteria whereby respondents were asked questions about demographical matters. 85 questionnaires were completed in 2021 (January – June) by the owners and/or managers of manufacturing SMEs operating in the Cape Metropole. Table 1 below summarises the socio-demographic information of the participants from Section A.
Table 1. Demographic Summary of Sampled SMEs

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision-Making Power</td>
<td>95.3% had decision-making power</td>
</tr>
<tr>
<td></td>
<td>4.7% did not have</td>
</tr>
<tr>
<td>Position in SMEs</td>
<td>16.5% were owners</td>
</tr>
<tr>
<td></td>
<td>54.1% were managers</td>
</tr>
<tr>
<td></td>
<td>29.4% were managers and owners</td>
</tr>
<tr>
<td>Type of Business</td>
<td>22.4% operated in Clothing and textiles</td>
</tr>
<tr>
<td></td>
<td>14.1% were wood</td>
</tr>
<tr>
<td></td>
<td>11.8% were metal</td>
</tr>
<tr>
<td></td>
<td>4.7% were unknown</td>
</tr>
<tr>
<td></td>
<td>2.4% operated in Chemical</td>
</tr>
<tr>
<td></td>
<td>1.2% operated in Petroleum</td>
</tr>
<tr>
<td></td>
<td>3.5% operated in Paper</td>
</tr>
<tr>
<td>Education Level</td>
<td>1.2% had a Master's degree</td>
</tr>
<tr>
<td></td>
<td>9.4% had an Honors degree</td>
</tr>
<tr>
<td></td>
<td>28.2 had Higher Diploma</td>
</tr>
<tr>
<td></td>
<td>10.6% Higher Certificate</td>
</tr>
<tr>
<td></td>
<td>14.1% had below grade 12</td>
</tr>
<tr>
<td></td>
<td>23.5% had grade 12</td>
</tr>
<tr>
<td>Type of Sales used</td>
<td>31.8% use cash sales</td>
</tr>
<tr>
<td></td>
<td>5.9% use credit sales</td>
</tr>
<tr>
<td></td>
<td>62.4% use cash and credit sales</td>
</tr>
</tbody>
</table>

Decision-Making Power: The average respondent was a manager with decision-making power and between 11 and 50 full-time employees. Most respondents (95.3%) had decision-making power within their respective manufacturing businesses. The goodness of fit test shows that there are statistically significant more respondents with decision-making power within their business compared to those who do not have decision-making power. Position within the company shows that 16.5% of the respondents acted as the owner of their manufacturing business, 54.1% worked as the business manager, and 29.4% served as both the owner and manager. Statistically, significantly more respondents acted as managers compared to owners or both the owner and manager of the company.

Regarding the type of business operation for respondents, 22.4% of the respondents who completed the survey acted as the owner and/or manager of a clothing and textile business, followed by 14.1% in wood businesses and 11.8% in metal businesses. The smallest number of respondents came from paper businesses (3.5%), chemical businesses (2.4%), and petroleum businesses (1.2%). 4.7% of the respondents did not respond to the question and were indicated as "unknown". Three respondents who selected "other" types of businesses described their businesses as packaging, paint, and pottery. The goodness of fit test shows that the various types of businesses were not equally distributed. For the level of education, 14.1% of the respondent's highest level of education was lower than Grade 12, 23.5% of the respondents had Grade 12/Senior Certificate/Matric, 10.6% had a National Higher Certificate/Higher Certificate/National Certificate, 28.2% had a Higher Diploma/Diploma/National Diploma, 12.9% had a Bachelor's Degree/Advanced Degree, 9.4% had an Honours Degree/Postgraduate Diploma, and only 1.2% had a Master's Degree. Furthermore, respondents were not equally distributed over the highest level of education groups.

Type of Sales used: In total, 31.8% of the respondents’ businesses made use of cash sales only, 5.9% made exclusive use of credit sales, and 62.4% made use of cash and credit sales. The respondents were not equally distributed over the types of sales groups. Section B's design consisted of two Likert scale questions (see Table 2 overleaf). These questions were centred on achieving business objectives from which a conclusion can be made about the overall sustainability of manufacturing SMEs. Respondents were provided with a list of possible answers and had to rate their level of agreement or disagreement with the statement.
Table 2. Summary of perceived sustainability of sampled SMEs

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neither agree nor disagree</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income is greater than expenses (profitability)</td>
<td>63.5%</td>
<td>18.8%</td>
<td>8.2%</td>
<td>3.5%</td>
<td>3.5%</td>
</tr>
<tr>
<td>There is sufficient cash on hand (liquidity)</td>
<td>64.7%</td>
<td>16.5%</td>
<td>7.1%</td>
<td>4.7%</td>
<td>3.5%</td>
</tr>
<tr>
<td>Assets are greater than liabilities (solvency)</td>
<td>68.2%</td>
<td>21.2%</td>
<td>4.7%</td>
<td>1.2%</td>
<td>1.2%</td>
</tr>
</tbody>
</table>

From the results in Table 2, the inference can be made that the sustainability of the sampled SMEs by respondents is perceived as very good. These business entities agreed to have income higher than expenses (63.5% of the time). However, some of these businesses did experience unprofitability as they strongly disagreed that their income is greater than their expenses (3.5% of the time). Most respondents did not perceive their businesses as having weak business liquidity (64.7%) and weak business solvency (68.2%). However, these entities show weak business liquidity (3.5% of the time) and weak business solvency (1.2% of the time).

Table 3. Types of operational risks encountered in SMEs

<table>
<thead>
<tr>
<th>Risk</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neither agree nor disagree</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff errors</td>
<td>1.2%</td>
<td>7.1%</td>
<td>8.2%</td>
<td>9.4%</td>
<td>65.9%</td>
</tr>
<tr>
<td>System processing errors</td>
<td>1.2%</td>
<td>3.5%</td>
<td>9.4%</td>
<td>12.9%</td>
<td>64.7%</td>
</tr>
<tr>
<td>Inadequate auditing procedure</td>
<td>2.4%</td>
<td>4.7%</td>
<td>2.4%</td>
<td>9.4%</td>
<td>69.4%</td>
</tr>
<tr>
<td>Inadequate segregation of duties</td>
<td>2.4%</td>
<td>1.8%</td>
<td>3.5%</td>
<td>12.9%</td>
<td>69.4%</td>
</tr>
<tr>
<td>Failed systems and transactions</td>
<td>0%</td>
<td>3.5%</td>
<td>8.2</td>
<td>5.9%</td>
<td>71.8%</td>
</tr>
<tr>
<td>Poor systems design</td>
<td>1.2%</td>
<td>3.5%</td>
<td>3.5%</td>
<td>7.1%</td>
<td>74.1%</td>
</tr>
<tr>
<td>Poor human resources policies</td>
<td>1.2%</td>
<td>2.4%</td>
<td>7.1%</td>
<td>8.2%</td>
<td>63.5%</td>
</tr>
<tr>
<td>Inadequate security measures</td>
<td>1.2%</td>
<td>2.4%</td>
<td>5.9%</td>
<td>7.1%</td>
<td>68.2%</td>
</tr>
<tr>
<td>Internal and external frauds</td>
<td>1.2%</td>
<td>2.4%</td>
<td>3.5%</td>
<td>9.4%</td>
<td>70.6%</td>
</tr>
<tr>
<td>Lack of management supervision</td>
<td>2.4%</td>
<td>1.2%</td>
<td>3.5%</td>
<td>9.4%</td>
<td>68.2%</td>
</tr>
<tr>
<td>Lack of internal control</td>
<td>1.2%</td>
<td>3.5%</td>
<td>3.5%</td>
<td>7.1%</td>
<td>65.9%</td>
</tr>
<tr>
<td>Inadequate staff training</td>
<td>1.2%</td>
<td>1.2%</td>
<td>1.2%</td>
<td>7.1%</td>
<td>67.1%</td>
</tr>
<tr>
<td>Non-compliance issues</td>
<td>1.2%</td>
<td>1.2%</td>
<td>2.4%</td>
<td>11.8%</td>
<td>75.3%</td>
</tr>
<tr>
<td>Insufficient training</td>
<td>1.2%</td>
<td>2.4%</td>
<td>2.4%</td>
<td>5.9%</td>
<td>69.4%</td>
</tr>
<tr>
<td>Customer attrition</td>
<td>1.2%</td>
<td>1.2%</td>
<td>0%</td>
<td>7.1%</td>
<td>67.1%</td>
</tr>
</tbody>
</table>

Although respondents were mainly unaware of the term operational risk, Table 3 shows that respondents had to indicate whether they experienced certain types of operational risks or not. A high percentage of those who strongly disagreed is expected to be the ones that stated that they do not know the term operation risk, and also, some respondents did encounter these risks. However, Islam & Tedford's (2012) study found that the internal events were found to be the most significant once than the external events.

A conclusion can be drawn that most of these entities do not encounter these risks as they strongly disagree (more than 50% of the time). However, a minority implementing operational risk management has indicated that they have encountered these risks as they strongly disagreed (less than 5% of the time).
Table 4. Operational risk factors

<table>
<thead>
<tr>
<th>Operational risk factors</th>
<th>Very much</th>
<th>Much</th>
<th>Average</th>
<th>Little</th>
<th>Very little</th>
</tr>
</thead>
<tbody>
<tr>
<td>External risk</td>
<td>1.2%</td>
<td>3.5%</td>
<td>11.8%</td>
<td>8.2%</td>
<td>69.4%</td>
</tr>
<tr>
<td>System risk</td>
<td>3.5%</td>
<td>7.1%</td>
<td>15.3%</td>
<td>3.5%</td>
<td>65.9%</td>
</tr>
<tr>
<td>Process</td>
<td>3.5%</td>
<td>7.1%</td>
<td>11.8%</td>
<td>8.2%</td>
<td>62.4%</td>
</tr>
<tr>
<td>People</td>
<td>9.4%</td>
<td>9.4%</td>
<td>11.8%</td>
<td>8.2%</td>
<td>57.7%</td>
</tr>
</tbody>
</table>

Table 4 shows results about the respondents’ identification of their level of agreement with the aspects of operational risks and losses in their organisations. Most SMEs indicated very little to none of all the operational risks. These findings contradict the findings by Islam & Tedford (2012) as SMEs face operational risk within their businesses, especially with 49.4% of the vetted SMEs that were not implementing operational risk management.

5. Limitations and recommendations

One of the main limitations of this research is the small sample size, which could give different results with a larger sample or if clustered across industrial regions. Therefore, it is proposed that further research be conducted with larger sample size. The second limitation is that most of the sampled SMEs have been in existence for 15.4 years on average, which is a long time considering that most SMEs fail within the first two years of existence. It could be that many of them have attained some level of sustainability due to other reasons that may have been important other than operational risk management. Hence, it is recommended that future research look at SMEs with a higher age of existence and find out how they can attain sustainability without implementing operational risk management. Triangulation of methodology with a qualitative flair could also provide rigour to the research results.

6. Conclusions

Despite SMEs being regarded as the drivers of economic growth, academic literature shows that South African SMEs have the worst sustainability rates globally, as 70-80% fail within three years. It is believed that this stems from the harsh environment South African SMEs have to operate in, which affects their sustainability and various factors, including numerous internal and external challenges. If left unmanaged, these internal and external challenges may lead to operational risks, resulting in losses. To prevent this occurrence of operational risks translating to loss events, it is highly recommended that these businesses implement operational risk management. This is one of the most effective ways of detecting operational risks, which can help prevent and mitigate the risk. This study focused on operational risk factors that negatively influence the sustainability of the sampled manufacturing SMEs in South Africa, specifically in the Cape Metropole.

From the research, it is clear that the sustainability of the sampled SMEs is above average, even though 49.4% of them do not implement operational risk management. This questions the significance of operational risk management as these businesses have survived, on average, for 15.4 years without it. Despite the preceding, it was discovered that the people are the internal factor that results in the most risk, which in turn causes loss events and affects the SMEs' sustainability. However, the overall findings do provide reasoning for further research.

Furthermore, the data gleaned from the research results were not in congruence with past literature as the majority of the manufacturing SMEs have been in existence on average for 15.4 years without implementing operational risk management. Therefore, it is astonishing that these SMEs have achieved this level of sustainability, as literature has shown that they fail within their first three years of existence. It is not the intention of this research study to discuss possible reasons for this success as the questionnaire did not have questions about this; however, it is worth discussing a few points.
The suspected reason for this sustainability level is that most manufacturing SMEs might have been family enterprises. Employees in these kinds of entities always feel like they are part of the family, which makes it very personal. Therefore, a lot of trust and teamwork is exhibited in family enterprises, and there is a willingness to follow instructions as leadership is perceived to be mentoring. These kinds of organisations can manage knowledge well through sharing it and often create a culture of learning and creating knowledge. Research studies have shown that correctly capturing and sharing knowledge can assist the business in mitigating operational risks. Lastly, it seems these entities have managed to be sustainable through knowledge management and not operational risk management.

The main question is what type of operational risk factors affect manufacturing SMEs. We find that most manufacturing SMEs do not encounter these risks, which is worrisome as these were the same responses that stated that they do not know the term operational risk. Thus, the government should offer a course on operational risks to encourage SME owners and/or managers to improve their operational risk knowledge and skills. These entities should ensure that managers/owners have the skills to identify, assess and measure the risks. Such a process could assist them in identifying, addressing, and understanding operational risk factors, improving their businesses’ sustainability and helping to create favourable conditions in which small businesses can grow and flourish.

References


Data Availability Statement: Data Availability Statement: More primary data can be obtained from the authors on a reasonable request.

Author Contributions: Conceptualisation: Urim-Tummin Shipanga, Suzaan Le Roux; methodology: Suzaan Le Roux, Urim-Tummin Shipanga; data analysis: Urim-Tummin Shipanga, Jobo Dubihela, writing—original draft preparation: Urim-Tummin Shipanga, writing; review and editing: Suzaan Le Roux, Jobo Dubihela; visualisation: Urim-Tummin Shipanga. All authors have read and agreed to the published version of the manuscript.
Urim-Tummim SHIPANGA graduated with Mtech in Internal Auditing from the Department of Internal Auditing and Financial Information Systems, Cape Peninsula University of Technology (CPUT), Cape Town, South Africa. Research interests: Operational risk; innovation and small business finance; sustainability and risk management.
ORCID ID: https://orcid.org/0000-0002-7989-868X

Dr. Suzaan LE ROUX is a senior lecturer in the Department of Internal Auditing and Financial Information Systems, Cape Peninsula University of Technology (CPUT), Cape Town, South Africa. She has presented papers in national and international conferences in Africa, Europe and the USA. She is currently the research coordinator in the Faculty of Management and Business studies at CPUT, South Africa. Research interests: Financial information systems; technology adoption and innovation.
ORCID ID: https://orcid.org/0000-0002-3511-429X

Prof. Jobo DUBIHLELA is the Associate Professor in the Department of Internal Auditing and Financial Information Systems, Cape Peninsula University of Technology (CPUT), Cape Town, South Africa. His work has attracted best research paper awards in both the African and the USA conferences. He is currently Head of Department in the Faculty of Management and Business studies at CPUT, South Africa. Research interests: Social entrepreneurship; innovation and small business finance; sustainability and corporate governance.
ORCID ID: https://orcid.org/0000-0001-6228-6524
LOCATING CRIME IN THE FIELD OF SUSTAINABILITY: A BIBLIOMETRIC MAPPING ON "SUSTAINABILITY AND CRIME" FROM 1995 – 2022

K.G.N.U Ranaweera

Department of Criminology and Criminal Justice, University of Sri Jayewardenepura, Sri Lanka

E-mail: ranaweera@sjp.ac.lk

Received 18 August 2022; accepted 20 October 2022; published 30 December 2022

Abstract. Emerging as a field of study and paradigm shifts parallel to the societal changes are inherited consequences of social sciences. Criminology, as a field of study, originated from sociology and developed as a particular field of study, later initiating sub-disciplines within the subject. Moreover, moving back and forth in the criminological evolution, multidisciplinary research connected the crimes with indirectly related fields and generated more total outputs to emphasize the ultimate goal of criminology, "crime control and prevention". Identifying crimes within the area of sustainability is vairy. Discussions of crime prevention as improving quality of life, locating different criminal activities within Sustainable Development Goals (SDGs), and using SDGs for crime prevention are among the ample dialogues. This study reviews the literature on sustainability and crime with VOSviewer, a bibliometric visualizing software.

Keywords: bibliometric mapping; crime; Scopus; sustainability; VOSviewer


JEL Classifications: Q56

Additional disciplines: law; sociology; ecology and environment

1. Introduction

With the delicate balance between constructive ambiguity and blatant disagreement, sustainable growth could not resist the weight of different interpretations. It has been a decade since the term "sustainable development" was first coined. Despite its seeming good fortune for environmentalists, this decision reveals several severe issues. In addition, economists have taken over the sustainable development narrative, and wealthy nations have become less concerned about developing their economies. Many people agree that sustainability is a beneficial move, but there are many different interpretations of what it means to be "sustainable," and these views are becoming increasingly polarized. It is possible to trace these misinterpretations back to the first English usage of the term "sustain." Taking care of one's own needs is an apparent extension of the notion of sufficiency. Keeping what is already in mind naturally leads to a preservation focus in a second family of meanings. First, the focus is on
"meeting the needs of today," whereas, in the second, it's all about looking to the future and planning for the future (Jamieson, 1998).

The Brundtland Report of 1987 established sustainability as a policy term concerned with the conflict between humankind's ambitions for a better existence on the one hand and the restrictions imposed by nature on the other. Over time, the notion has been re-interpreted to include three dimensions: social, economic, and environmental. A change, change in meaning (a) obscures the fundamental contradiction between the goals of welfare for all and environmental conservation; (b) risks diminishing the importance of the environmental dimension; and (c) separates social and economic aspects, which are, in reality, the same. It is advocated instead to revert to the original definition, where sustainability is concerned with the wellbeing of future generations, particularly with irreplaceable natural resources as opposed to the satisfaction of current demands, which we term wellbeing. A balance must be struck between the two, but not by claiming three sides of the same coin. Although depleting natural resources at the cost of future generations, also developing capital (including knowledge) that improves future wellbeing. A crucial issue is how much one compensates for the other (Kuhlman & Farrington, 2010).

Since the 1970s, the concept of sustainability has increasingly been linked to the concept of human sustainability, and as a result, the most often quoted definition of sustainability and sustainable development is "human sustainability on planet earth." Sustainable development, as defined by the United Nations World Commission on Environment and Development, is "development that meets the demands of the present without compromising the ability of future generations to meet their own needs" (United Nations). The word "sustainable development" refers to the simultaneous pursuit of three goals that overlap: environmental, economic, and social, and since 1987, there has been a substantial increase in the number of disciplines in which the idea of sustainable development has been implemented (Goni et al., 2015). This trend is shown by the increase in the number of sustainability publications combined with different fields.

The Sustainable Development Goals (SDGs), often known as the Global Goals, are 17 interconnected goals intended to serve as a "blueprint to build a better and more sustainable future for everyone." The United Nations General Assembly established the SDGs in 2015, intending to complete them by 2030 (United Nations, n.d.). As SDGs addressed a broad spectrum of human life, social issues were also combined with the concept of sustainability sooner. The combination of sustainability and crime addresses the "quality of life" directly and is attached to all SDGs, publications related to the theme of "sustainability and crime" has become more interesting with time.

Literature reveals definitions, themes, and descriptions of sustainability in many complex shapes and sizes, ranging from strategy, framework, phrases, concepts, indexes, indicators, weak, strong, externality, internal, and criterion. As a result, there is a vast diversity of opinions, which leads to confusion regarding the literal implementation of sustainability. It does this by exposing the concept of "sustainable" as a function of transdisciplinary factors, highlighted in four common themes: social, economic, ecological, and technological (Hasna, 2007).

Researchers hypothesize that there is a connection between being a victim of crime and the victim's reported level of life satisfaction and wellbeing, related to the social theme of sustainability. Studies that investigate linked concepts with crime and the quality of life often focus on indicators like fears of crime and worries for personal safety, happiness, and overall levels of satisfaction with the quality of life. It would suggest that having a history of being a victim of crime does not have a significant impact on these factors, on average. Crime-related difficulties have a relatively minor influence on people's happiness levels, life satisfaction, and contentment with their quality of life. The overall relationships were relatively low, leading the authors to conclude that victimization did not significantly impact these factors. Even though victimization was associated with some concerns about neighbourhood crime and problems, a tendency to engage in more defensive behaviours, less
satisfaction with their own and their family's safety in neighbourhoods, and lower levels of satisfaction with the quality of life, the authors concluded that victimization did not have a significant overall impact on these factors (Hanson et al., 2010).

The most prevalent source of information about the crime is police reports. Many crimes go unreported for a variety of reasons. As a result, police reports tend to understate crime and cover various locations and kinds of crime inconsistently. Victimization surveys are a better source of crime statistics since they are less likely to be skewed. Surveys question a representative sample of people if they have been victims of crime, how much they were harmed, and under what conditions. Many of these polls are only accessible in a few countries and typically for a short period. Victimization surveys, on the other hand, aren't flawless. Crime and violence committed by criminals and other fringe groups, such as gangs, are often overlooked by the media. Police reports are the primary data source for international and inter-period comparisons (Bourguignon, 2000). Sustainability reports indicate the absolute crime rates as it reveals the factors affecting the quality of life, measuring the levels of crime fear of individuals and the community.

The study's main objective is to review the existing literature on sustainability and crime with VOSViewer, a bibliometric visualizing software, answering 05 questions: What is the volume and document citation by the time of distribution from 1995 to 2022? Which countries contributed most to the knowledge of sustainability and crime between 1995-2022? What top journals have published the 100 most cited articles on sustainability and crime? What is the intellectual structure of sustainability and crime? What are the key concepts that have been explored in the field of sustainability and crime? and how they are related? What is the nature of collaboration that is evident in the publications on sustainability and crime?

2. Methodology

Although the literature review is based on bibliometric analysis, different methods were used to answer each question (see Table 1).

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Method Used</th>
<th>Used software</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the volume and document citation by the time of distribution from 1995 to 2022?</td>
<td>Bibliometric Analysis</td>
<td>Microsoft Excel</td>
</tr>
<tr>
<td>Which countries contributed most to the knowledge of sustainability and crime between 1995-2022?</td>
<td>Bibliometric Distribution</td>
<td>Microsoft Excel</td>
</tr>
<tr>
<td>What top journals have published the 100 most cited articles on sustainability and crime?</td>
<td>Citation counting</td>
<td>Vosviewer and Microsoft Excel</td>
</tr>
<tr>
<td>What is the intellectual structure of sustainability and crime?</td>
<td>Citation Network Analysis</td>
<td>Vosviewer</td>
</tr>
<tr>
<td>What are the key concepts that have been explored in the field of sustainability and crime? And how they are related?</td>
<td>Keyword co-occurrence Analysis</td>
<td>Vosviewer</td>
</tr>
<tr>
<td>What is the nature of collaboration evident in the publications on sustainability and crime?</td>
<td>Co-authorship Network Analysis</td>
<td>Vosviewer</td>
</tr>
</tbody>
</table>
Table 2. Data collection (Data filtering)

<table>
<thead>
<tr>
<th>Phase</th>
<th>Search Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Search &quot;((&quot;Sustainability&quot;) AND (&quot;Crim*&quot;)) in Scopus database</td>
</tr>
<tr>
<td>02</td>
<td>Search Result - 572</td>
</tr>
<tr>
<td></td>
<td>Limitation of subject areas: Sociology, Environment, Psychology, Arts and Humanities and Multi-disciplinary</td>
</tr>
<tr>
<td>03</td>
<td>Search Result - 405</td>
</tr>
<tr>
<td></td>
<td>Fixing article type - Article(Final)</td>
</tr>
<tr>
<td>04</td>
<td>Search Result - 393</td>
</tr>
<tr>
<td></td>
<td>Limitation of the Language: English Only</td>
</tr>
<tr>
<td>05</td>
<td>Search Result - 393</td>
</tr>
<tr>
<td></td>
<td>Final Manual screening</td>
</tr>
</tbody>
</table>

The original list derives from the Scopus database (.csv format). The keywords "Crime" and "Sustainability" examined in the linked literature were utilized. In the Scopus database, the following Boolean operation and truncation were used as "(("Sustainability") AND ("Crim*"))". Total publications were 868. As the second limitation, data was limited to 5 subject areas: Sociology, Environment, Psychology, Arts and Humanities, and Multi-disciplinary, the total number of publications was limited to 572. Publication type was selected as "Article" and limited results to 414; Publication Stage was fixed as "Final" and resulted from 405. As the final limitation, the Language was selected as "English," and the total units were finalized as 393 (see Table 2).

The search string was TITLE-ABS-KEY (("Sustainability") AND ("Crim*")) AND (LIMIT-TO (SUBJAREA, "SOCI") OR LIMIT-TO (SUBJAREA, "ENVI") OR LIMIT-TO (SUBJAREA, "PSYC") OR LIMIT-TO (SUBJAREA, "ARTS") OR LIMIT-TO (SUBJAREA, "MULT") AND (LIMIT-TO (DOCTYPE, "ar") AND (LIMIT-TO (PUBSTAGE, "final") AND (LIMIT-TO (LANGUAGE, "English")) available data was limited to the years from 1995 to 2022.

3. Results and discussion

3.1 Identification of the volume and distribution of the field of sustainability and crime

Tabulation of the bibliometric data extracted from the Scopus database indicates a gradual increment in the field of sustainability and crime. Although the publications in the field are less than ten each year between 1995 – 2010, from 2011, a gradual rise is visible. Except for 2015, from 2013, the number of publications is higher than 20. The era of 2020 to 2021 indicates a higher number of publications, exceeding the number 45 (see Fig. 1).
Fig. 1. Number of publications on sustainability and crime

The increment year point, 2015, is the year of the officially formatted year of the Sustainable Development Goals. The 2030 Agenda for Sustainable Development provides a shared road map for achieving global peace and prosperity for people and the planet in the here and now and in the years to come. In the year 2015, every membered state of the United Nations gave their stamp of approval to this agenda. The 17 Sustainable Development Goals (SDGs) are at the centre of this movement. These goals represent an urgent call for action by all developed and developing countries within the framework of a global partnership. They understand that ending poverty and other forms of deprivation must go hand in hand with strategies that improve health and education, reduce inequality, and spur economic growth, all while tackling climate change and working to preserve our oceans and forests. To put it another way, eradicating poverty and several other types of deprivation won't be achievable until these tactics are implemented.

Discussion of SDG and Crime arose with SDG 16. The goal of Sustainable Development Goal 16 is to create societies that are peaceful and inclusive, based on the protection of those who are the most vulnerable, the rule of law, and effective governance at all levels. In addition, it foresees institutions that are open to the public, efficient in their operations, and answerable to those they serve. These institutions would promote non-discriminatory laws and policies, fight organized crime, bribery, and corruption, and prevent acts of violence, terrorism, and criminality. Goal 16 of the Sustainable Development Goals asks for decision-making processes that are responsive, inclusive, participatory, and representational, along with an increased involvement for developing nations in institutions of global governance. But, later, researchers identified the combination of all 17 SDGs and crimes.
Fig. 2. Number of publications and total citations

Articles published in 2007 were the most cited (see Figure 2). Although a significant drop can be identified in 2008, a gradual increase and stability could be recognized with available data. The minor citation within the period of 2021-2022 due to the period of time, and according to the normalizing data, a significant change cannot be addressed. According to Figure 3, the field of sustainability and crime is an uncommon but slightly popular topic among researchers.

3.2 The most contributed countries to the knowledge of sustainability and crime between 1995-2022

Most of the articles published on sustainability and crime were in the United States and the United Kingdom to measure the distribution of the publications. Australia, Italy, and Canada also bear the higher number (Map 1, Figure 3).

Map 1. Distribution of publications according to the country
Fig. 3. Distribution of publications
3.3 The top journals that have published the 100 most cited articles on sustainability and crime

<table>
<thead>
<tr>
<th>Most cited Journals (More than 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proceedings of the National Academy of Sciences (PNAS) of the United States of America</td>
</tr>
<tr>
<td>Journal of Cleaner Production</td>
</tr>
<tr>
<td>Cities</td>
</tr>
<tr>
<td>Sustainability (Switzerland)</td>
</tr>
<tr>
<td>American Psychologist</td>
</tr>
<tr>
<td>Journal of Planning Literature</td>
</tr>
<tr>
<td>Resources, Conservation and Recycling</td>
</tr>
<tr>
<td>Global Environmental Change</td>
</tr>
<tr>
<td>Journal of Interpersonal Violence</td>
</tr>
<tr>
<td>Computers, Environment and Urban Systems</td>
</tr>
<tr>
<td>Recourse conservation and recycling</td>
</tr>
</tbody>
</table>

**Fig. 4. Most cited journals**

With the highest number of publications on sustainability and crime, the Proceedings of the National Academy of Sciences (PNAS) of the United States of America is the most-cited journal. PNAS is a multidisciplinary scientific publication that experts from various fields evaluate. It is often abbreviated as PNAS or PNAS USA. Original research, scientific reviews, comments, and letters are all published in this magazine, which has been in publication since 1915 and serves as the official publication of the National Academy of Sciences. The journal is said to have an impact factor for the year 2020 of 11.205, as reported by Journal Citation Reports. The Proceedings of the National Academy of Sciences (PNAS) is the scientific journal that has received the second-highest number of citations between 2008 and 2018. The Public Library of Art and Science (PNAS) has been characterized in numerous ways by the mainstream media, such as "prestigious," "sedate," "famous," and "high impact" (PNAS, n.d.).

The second most cited Journal, The Journal of Cleaner Production (JCP) is a scholarly publication that is evaluated by experts in the relevant fields and publishes research that crosses disciplinary boundaries. Elsevier is the issuing company, and the Journal of Cleaner Production acts as a platform for sharing information and research ideas, strategies, and technologies that are meant to assist in ensuring progress is made toward making societies and areas more sustainable. This forum is worldwide and cross-disciplinary. Its purpose is to foster inventiveness and creativity, new and better products, and develop unique structures, systems, processes, goods, and environmentally friendlier services. In addition, it is intended to encourage the formulation and execution of preventatively oriented educational programs and policies by government agencies (JCP, n.d).

Moreover, Plos One, Cities, Sustainability (Switzerland), American psychologists, Journal of Planning Literature, Recourse conservation, and recycling, Global environmental change, journal of interpersonal violence, computers, environment, and urban system are the journals included in more than 100 cited journal category on sustainability and crime. The fact that the crime and related components are directly and indirectly rooted in the sustainability field is evident due to the differently distributed areas of journals, including the journal on interpersonal violence and computers, environment, and urban systems, for instance (see Figure 4).
3.4 The intellectual structure of sustainability and crime

After the completion of a manual reading phase, the intellectual structure of knowledge on sustainability and crime is composed of twelve different schools of thought: Sustainability and environmental justice (Cluster 01), Sustainability and psychology of crime (Cluster 02), Sustainability and economy (Cluster 03), Sustainability and built environment (Cluster 04), Sustainability and urban crimes (Cluster 05), Sustainability and evidence-based practice (Cluster 06), Sustainability and regional development (Cluster 07), Sustainability and deviant behaviour (Cluster 08), Sustainability and urban planning (Cluster 09), Sustainability and cultural criminology (Cluster 10), Sustainability, technology and security (Cluster 11), Sustainability and investment (Cluster 12), and Sustainability and exploitation (Cluster 13). Figure 6 indicates 13 clusters in different colours while visualizing the citation network (figure 5).

![Fig. 5. The citation network](image)

*Note: The minimum number of citations of an author is fixed as 05, and of the 29051, only 755 meet the threshold. In the network of 755, only 731 were connected.*
3.5 Key concepts that have been explored in the field of sustainability and crime

In order to identify the key concepts explored in the field of sustainability and crime, keyword analysis was done. Vosviewer software was set up as the minimum number of occurrences of a keyword as two (2), 326 keywords were found, and 722 met the threshold. Country names, repeated words, hyphenated words, and singular plural, for instance, were manually removed, and 12 clusters (themes) were identified according to the keyword occurrence. The generated CSV data of keyword occurrence (Vosviewer) were sorted according to the weight of occurrence in each cluster (12 clusters in total) and identified a common theme (Table 3).

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Theme</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sustainability and Social Strata</td>
<td>male, female, adults, human, economics</td>
</tr>
<tr>
<td>2</td>
<td>Sustainability and Governance</td>
<td>sustainability, governance approach, safety, urban development, social sustainability</td>
</tr>
<tr>
<td>3</td>
<td>SDG and Environmental Crimes</td>
<td>crime, sustainable development, environment protection, conservation, forestry</td>
</tr>
<tr>
<td>4</td>
<td>Sustainability and Urban Planning</td>
<td>urban planning, neighborhood, urbanization, perception, urban designing</td>
</tr>
<tr>
<td>5</td>
<td>Regional crime prevention</td>
<td>crime prevention, strategic approach, criminology, developing countries, illegal fishing</td>
</tr>
<tr>
<td>6</td>
<td>Sustainability and human behaviour</td>
<td>urban population, public policy, human experiments, quality of life, cities</td>
</tr>
<tr>
<td>7</td>
<td>Environmental sustainability</td>
<td>urban area, ecological sustainability, environmental policy, water management, environmental monitoring</td>
</tr>
<tr>
<td>8</td>
<td>Statistical explanations of sustainability</td>
<td>economic development, correlation, regression analysis, questionnaire survey, united nations</td>
</tr>
<tr>
<td>9</td>
<td>Sustainability and social problems</td>
<td>decision making, poverty, waste management, public health, unemployment</td>
</tr>
<tr>
<td>10</td>
<td>Sustainability and quality of life</td>
<td>housing, community, mental health, wellbeing, air quality</td>
</tr>
<tr>
<td>11</td>
<td>Sustainability and law</td>
<td>law enforcement, corruption, legislation and jurisprudence, civil society, environmental legislation</td>
</tr>
<tr>
<td>12</td>
<td>Sustainability and social impact assessment</td>
<td>land use, fear of crime, social impact, designing out crime</td>
</tr>
</tbody>
</table>

The network visualization of the identified 12 clusters related to each other is as follows (Figure 6, Figure 7).
3.6 The nature of collaboration that is evident in the publications on sustainability and crime
Fixing the minimum number of documents of an author as 02, among 1154 authors, only 46 met the threshold (Figure 8). The authorship analysis shows that a total of 46 authors contributed to the 100 most-cited articles. Among these 46 authors, 11 had linked once, 4 clusters of authors linked with two lines, while the other 14 authors remained alone. Figure 8 represents the co-occurrence network of the collaboration of these distinctive authors. It can be seen that there are 28 distinct clusters, including single-author clusters.

Conclusions

From 1995 to 2022, there has been a steady rise in the number of publications on sustainability and crime, with minor oscillations, notably after 2016. There has been a growing interest in "sustainability and crime" which can be observed by looking at where the research comes from and the location.

A recent bibliometric study on differentiated instruction's impact on crime reports and sustainability yielded some surprising results. The subject is fascinating to many academics in light of the rise in publications and citations. According to the study, the most significant number of publications occurred between 2020 and 2021. The analysis also found that in journal citations, Proceedings of the National Academy of Sciences of the United States of America has garnered the most acclaim.

The USA and The UK authors have contributed the most to the most referenced articles on differentiated teaching. A list of notable authors, keywords, and study locations for future academics interested in the issue of sustainability and crime. These results enrich the foundation of knowledge used to provide differentiated education by providing an overview of the current body of research.

The findings of the bibliometric study, in particular the co-authorship network, point to the following as the primary causes of the uncertainty over the scientific quality of sustainability and crime research: Because there is a dearth of intellectual interaction among individuals who are doing study in the topic of sustainability and crime, writers are more likely to be subjective and pursue their own particular research pathways in isolation from one another.

The scientific community members investigating sustainability and crime need to improve the amount of intellectual communication they have with one another to confront the problem that this scenario provides. The improvement of regional research helps construct a collaborative and interdisciplinary environment, which is required to generate a possible consensus regarding the approach to conceptualizing and enabling sustainability through the reduction of crime and the enhancement of the quality of life.

References


152


K.G.N.U. RANAWEERA is a Senior Lecturer attached to the Department of Criminology and Criminal Justice, University of Sri Jayewardenepura, Sri Lanka. After completing her BA with the highest GPA, she completed a Master of Philosophy Degree with a research component titled "Victimization factors of human trafficking in Sri Lanka" from the University of Sri Jayewardenepura. To combine crime study with quality of life, she is reading for a Ph.D. in Sustainability and Social Innovation at Thammasat University, Thailand. She also deepened her knowledge in several proficiency fields of her interests, fear of crime, gender studies, and juvenile delinquency.

ORCID ID: https://orcid.org/0000-0002-7682-5733
TOWARD QUERYING THE NATIONAL PERIL OF KIDNAPPING IN NIGERIA

Funmilayo Idowu Agbaje

Department of Peace, Security and Humanitarian Studies, Faculty of Multidisciplinary Studies, University of Ibadan, No 9 Parry Road, Nigeria

E-mail: funmibolajiagbaje@gmail.com

Received 10 August 2022; accepted 23 October 2022; published 30 December 2022

Abstract. In Nigeria, the dynamics of kidnapping is framed in complexities. There are different types of kidnapping going on within the country ranging from resource-based kidnapping to political kidnapping and terrorist-based kidnapping (Agbaje, 2020; Onduku, 2011). While some are based on agitations for political and economic recognition, others are based on the need for compulsory recruitment by insurgents and terrorists. Existing gaps in the literature are demanding urgent attention from researchers and policymakers. This paper focuses on the need to look into the recent sporadic kidnappings in contemporary Nigeria to project sustainable solutions for them. This paper relies solely on secondary sources of data for its analysis. The paper concludes that a viable solution to the kidnapping problem in Nigeria depends mainly on full recognition of all forms of kidnapping in the country while strategically addressing them.

Keywords: kidnapping; peace; security; Nigeria


JEL Classifications: F52

Additional disciplines: political sciences

Introduction

Kidnapping is a vice that is perpetuated worldwide. It is described as a global phenomenon that permits the abduction and captivity of a person to obtain a ransom (Uzochuckwu, 2013). It is a wicked act committed by criminals against human beings for personal gains. It can also be described as an act of terror against humanity because it creates perpetual fear and a sense of insecurity in people's minds (Guerette and Headley, 2019). Scholars have identified different types of kidnappings thus; child/parent kidnapping, corporate kidnapping, terrorist-based kidnapping, love kidnapping, economic-based kidnapping, and political kidnapping, among several other typological and identification of kidnappings (Auerbach, 1998; Johnson, Doug and Greg, 2003). Specifically, Okoli and Agada (2014) categorized the different concepts driving kidnappings in Nigeria into four thus; abduction, hostage-taking, captivity and ransom. Even though works of literature abound on the types of kidnapping, the factors fuelling the act, the responses by government, media and individuals, the diverse challenges of kidnapping on human security as well as ways of addressing the scourge in the society, there is still a gap in the areas of an empirical study into the patterns and trends of kidnapping in Nigeria as a major step in proffering a long-lasting solution to the problem (Essien and Ema, 2013; Briggs, 2009).
In Nigeria, the dynamics of kidnapping is framed in complexities. There are different types of kidnapping going on within the country ranging from resource-based kidnapping to political kidnapping and terrorist-based kidnapping (Agbaje, 2020; Onduku, 2011). While some are based on agitations for political and economic recognition, others are based on the need for compulsory recruitment by insurgents and terrorists. In essence, there had always been a struggle behind kidnappings in the past. Also, the people connected directly or indirectly with the demands or grievances of the kidnappers are usually the victims. For example, the Niger-Delta youths would instead kidnap the officials of oil companies such as Shells and Chevron or target their anger on the community chiefs who are suspected of perfidy or avarice to the detriment of the community (Osumarh and Ekpenyong, 2006). However, contemporary Nigeria is grappling with a composite form of kidnappings webbed in heavy mystery. This is a difficult situation whereby kidnappings are based on several reasons (economic, political, ritual, terrorism) with no particular susceptible class of the populace targeted. The vulnerabilities cut across ages, gender, class and race making it difficult for concrete analysis to be made by experts and security analysts. This random but sporadic form of kidnapping ravaging the country demands urgent attention from the government, scholars and policymakers, hence the reason for this paper that delves into the complexities of kidnapping insecurity in Nigeria while making recommendations for addressing it.

Statement of the Problem

Since the beginning of the 21st century, kidnapping has been a significant organized crime in Nigeria (Ibaba, 2008). In 2000, the Niger Delta agitated youths kidnapped several oil companies and top government officials to register their grievances while demanding a huge ransom. Likewise, in 2009, the Boko Haram insurgents started their kidnapping spree by abducting young men and women while conscripting them into becoming members of the sect (Abdulbasit and Ankpa, 2018). Kidnapping by bandits and suspected herders gained recognition in 2019. Subsequently, the act of kidnappings for ransom became an almost daily occurrence in all states in the country. The situation escalates in northern Nigeria, where bandit abounds with hundreds of primary, secondary and tertiary school students being sporadically removed with ransoms put on them (Okoli and Agada, 2014).

Across the thirty-six states in Nigeria, kidnapping is carried out. This has posed a huge national challenge and threat to the country's security. It has led to national economic and developmental sabotage. It has created fear in the people's minds, increased poverty, unemployment and health-related issues among the citizens. A lot of lives have been lost in the process. It has adversely affected Nigeria’s reputation in the international community while hampering tourism and foreign investment. The situation is worsened because the security officers are not adequately equipped to tackle the menace (Human Rights Watch, 2022).

The perpetrators are usually equipped with much more sophisticated weapons and are vast in the act of surfing the complex geographical terrain of the forests and several porous borders in the country. It has also been postulated (Alemika, 2012) that corruption among government officials, especially the security officers, is also a significant setback in the fight against kidnapping in Nigeria.

A critical problem to be addressed in this study is breaking down the 'epidemic' of kidnapping in Nigeria, which is a difficult task. Due to its diverse complex dynamics, understanding the riddle of kidnapping in Nigeria is an uphill adventure. Noteworthy is the fact that scholars (Raheed, 2008) have elucidated two significant factors behind the kidnapping in Nigeria political vendetta and economics. It has also been mentioned that officials of international oil companies, top politicians, and their close relatives are kidnapping targets (Detotto, McCannon and Vannini, 2012). In addition, scholars have also identified the diverse reasons insurgents would abduct beyond political and economic reasons to be forceful recruitment, reproduction and objectification (Agbaje, 2020; Unam, Arua and Out, 2018; Ezeibe and Eze, 2012). However, observing the current kidnapping trends in Nigeria has revealed the inconsistency in the modus-operandi of the kidnappers.
Different types of kidnappings go on concurrently with diverse interests and targets. It is more worrisome that people are being kidnapped not because of their wealth, gender, class, age or status. Instead, the present kidnappers just abduct their victims based on the fact that they are human beings. Therefore, it is not uncommon for people to be taken away from churches, hospitals, schools, public vehicles, and homes; in fact, newborns have been carted away from hospitals (Sunday Trust, 2021). This is contrary to Okoli and Aguda's (2014) notion that the crime of kidnapping in Nigeria is primarily class sensitive. They posit that organized kidnap syndicates in Nigeria have often targeted the rich, the powerful and the influential. The current situation's root cause has been poverty, avarice, revenge and terrorism. Adebajo (2021) reveals that the figure for kidnapped persons in Nigeria has risen from 110 in 2015 to 437 in 2020. A significant figure of 315 has been recorded for the first half of 2021, projecting the likelihood that the highest number of kidnappings will have occurred since 2015. Although the Nigerian government is working relentlessly to end kidnapping, its efforts have not yielded much success. Perhaps, the reason for limited success may not be far-fetched from the fact that the root causes of kidnapping have not been appropriately identified and addressed. This study focuses on the emerging trends in kidnapping in Nigeria to proffer methods of ending it.

**Kidnapping in Nigeria**

Kidnapping, according to Guerette and Headley (2019), is the abduction of an individual against his or her will, usually followed by some duration of captivity. Akwash (2016) describes kidnapping as one of the most psychologically damaging crimes against humanity. He further describes kidnapping as a global pandemic with no real solution. He mentions how kidnappers may abduct for sexual, revenge or financial reasons and the gory fact that the kidnapped may not be released alive by the kidnappers for several reasons. He highlights the negative impacts of kidnapping on the societal psyche. The shift in focus has been tremendous, from the kidnapping of expatriate oil workers to relatives of politicians to relatives of those in the Diaspora and now anybody on sight. Kidnapping takes place anytime and anywhere, in churches, streets, hospitals, and even at home (Ugabi, 2014).

This random abduction makes kidnap anxiety-inducing to the general populace because anyone could be a victim (Uzochukwu, 2015). Kidnapped persons are treated according to the motives of the kidnappers (Alexander, 2009). In the past year, kidnapping has become commonplace in Nigeria, with bandits and insurgents raiding and killing for ransom. Specifically, between January and August 2021, one hundred and one (111) kidnapping incidents occurred, particularly in schools and children as victims (Ojiego, 2021). The total figure is derived from confirmed cases reported by the police and the media. It has been noted that an average of 13 persons were abducted daily in Nigeria. This further underscores the rampant criminal activity in the country. Kidnapping, therefore, is a fast-growing industry and has become so virulent and largely uncontrolled.

Scholars have highlighted some reasons behind increased kidnapping in Africa (Ndegwa, 2021) as a deadly mix of money and power. Moreover, some of the physical and psychological effects of kidnapping on victims have also been identified to be withdrawal from society, hyper-vigilance, feelings of fear and hopelessness, depression, irritability and even suicide (Symonds, Eichelman, Soskis and Reid, 1983; Alexander, Brown and Campbell, 2021). In Nigeria, it has instilled tremendous fear into the people's minds that people are now afraid of travelling, going to school and worship centres. Abdulkabir (2017) stressed the increase in kidnapping rates in Nigeria by mentioning that between the years 2014 to 2017, over 2000 people, including young Chibok girls, government officials, politicians and kings, were kidnapped. Furthermore, Ezemenaka (2018) posits that kidnapping is not a mere capital security issue but also a part of the terrorist attack, which attributes to threats on individuals and the state. It, therefore, should not be dealt with frivolously. He further explains that Kidnapping in Nigeria goes beyond the conventional typology of kidnapping in terms of global trends. In this regard, this paper focuses on examining the extant body of literature on the various forms of kidnapping ongoing in Nigeria to identify the factors responsible for the consequences for present and future security while suggesting the way forward.
Economic Kidnapping

The economic motivation behind the kidnapping was mentioned by Ngwama (2014). He notes that kidnapping in Nigeria became prominent when it was used as a weapon to fight for economic and environmental justice in the Niger Delta. According to him, the financial motivation was intermittently used to fund and sustain the fight. This was a collective mission by the agitated minority. Similarly, Townsend (2008) refers to the beginning of kidnapping in Nigeria as natural resources nationalism. According to him, this is a situation where an ethnic nationality seeks fair shares of the returns from natural resources, primarily when it is sourced from their locality. Such was the case in the Niger Delta (Essien and Ema, 2013). A related form of kidnapping is carried out due to poverty caused by unemployment that has pervaded the country (Nseabasi, 2008, Umoh, 2010). The perpetrators of this form of kidnapping often engage in the act as a matter of survival. It is thus common for them to target the wealthy and prominent individuals in society. In addition, they can also target persons with prospects of high and lucrative ransom, including teenagers, children and adults alike (Ngwama, 2014). The demand for ransom by kidnappers has made the crime to be categorized as "a big and sophisticated business" (Ryssdal, Hollenhorst and Palacios, 2019).

Scholars have examined the diverse phenomenon of kidnapping in Nigeria with its attendant socio-economic implications (Benjamin, Ajah, Nwokeoma and Okpan, 2018). One major challenge is the ubiquitous nature of kidnapping that has resulted in a considerable security challenge and hence a reversed economy of the nation. Inyang and Ubong (2013) also note that Nigeria loses a lot of revenue when expatriates working in multinational oil companies are kidnapped. Therefore, one major factor responsible for kidnapping in Nigeria has been established to be motivated and sustained by the criminal quest for material accumulation. This is a national problem that has become a cankerworm eating deep into the nation's fabric and thus, requires the attention of scholars, individuals, corporate organizations and policymakers for a sustainable solution. Furthermore, Afun, Ogboye and Eze (2021) note a report by SB Morgen Intelligence, that at least, Eighteen Million and Thirty-Four Thousand United States Dollars (US$18,034,000) had been paid as ransom to kidnappers in Nigeria between June 2011 and March 2020 (Afun, Ogboye and Eze, 2021).

This shows that the majority of kidnaps were economically motivated. Consequently, the Nigerian Senate is working on the Terrorism Prevention Amendment Bill. This seeks to be an amendment to the Terrorism Prevention Act of 2011. The bill says:

Anyone who transfers funds makes payment or colludes with an abductor, kidnapper or terrorist to receive any ransom for the release of any person who has been wrongfully confined, imprisoned or kidnapped is guilty of a felony and is liable on conviction to a term of imprisonment of not less than 15 years.

(Afun, Ogboye and Eze, 2021)

From the above, it is evident that the Nigerian government is taking steps to address the increased insecurity in the country. However, there seems to be a problem with the bill. First, the punishment slated for a ransom payer does not correspond with the crime's gravity. This is because a kidnapped victim stands the risk of being molested and/or killed; hence the family members are always agitated and often would want to source funds to set the victim free. As much as the Nigerian Senate is working to curb kidnapping in the society, there is a need for a strategic approach that will not further aggravate the insecurity of the citizens.

Ritual Kidnapping

One prominent form of unconventional kidnapping in Nigeria is what is known as ritual kidnapping (Oyewole, 2016). There is a lack of data on this theme. The insufficient data may not be far-fetched because, in most cases, scientific evidence to back up claims may not be precise enough. Notwithstanding, the crime is being committed
regularly, with countless casualties attributed to it over the years (Nwolise, 2013; Ebohon and Ifeadi, 2012). This phenomenon has been linked with the high level of religiosity among Nigerians. It is a common practice for many Nigerians to believe in one or two supernatural beings. There is also a constant quest for supernatural power to acquire wealth, position, control, and protection, among other reasons (Adefila and Opeola, 1998; Hund, 2000). Many Nigerians dread kidnapping for rituals because it is believed that victims of rituals are misfortune and can never live an everyday life again. According to Gbinije (2014), victims can run mad, die, be poor, deformed or afflicted by ailments and diseases. Jordan (2001) situates ritual killings within the quests for wealth and power.

In Nigeria, the use of human beings for ritual has existed since the pre-colonial era (Ayegbonyin, 2009; Oyewole, 2016). However, there has never been a consensus on the subject. In the words of La Fontaine, "human sacrifice is a blood sacrifice that involves the killing of a living creature as a ritual offering to a god or spirit, usually in expectation of a return in the form of good fortune, whether generalized or as the granting of a particular prayer" (La Fontaine 2011:4). Susceptible persons are identified as slaves, hunchbacks, dwarfs, pregnant women, female virgins and albinos (Igwe, 2004; Badiora, 2015). The scourge of ritual kidnapping is increasing at an alarming rate in contemporary Nigeria. This criminal act is being perpetrated by different classes of people (religious leaders, politicians, fraudsters) for various purposes, some of which have already been highlighted in this paper. Nevertheless, kidnapping for ritual constitutes a considerable threat to the nation's peace and security, and its socio-religious and traditional underpinnings have not received enough action from the government.

**Kidnapping for Trafficking Purposes**

Kidnapping for human and organ trafficking is a phenomenon submerged under modern-day slavery (Adesina, 2014). It is very common to report human trafficking from the perspective of victims' knowledge and consent rather than from a conscripted or abductive point of view (Makinde, 2015; Adepoju, 2005). Often, studies on human trafficking are carried out with a focus on the economic and security implications (Ofuoku, 2010; Kazeem, 2009). More so, extant studies abound on how victims are cajoled into trafficking for financial reasons, with little attention paid to how many are kidnapped for slavery or body parts harvesting (Adio-Moses, Agbaje and Eselebor, 2019; Okojie, 2005). Similarly, much scholarly attention has been given to women and children as victims of human trafficking for sexual and labour exploitation (Benn, 2007; Mbakogu, 2004). There is a need for a conscious awakening to the possibility of kidnapped victims being trafficked in person or body parts. Noteworthy is the fact that the medical field is fast becoming advanced with technology, and organ transplants are being carried out on patients more than ever before. This has resulted in increased requests for human body parts. Even though the quest for human body parts has increased, there is a meagre supply in the market. Consequently, it is a commonplace for desperate indisposed persons to seek alternatives from the black market (Rennie, 2018). This trend has reinforced a dangerous and criminal act of forceful organ harvesting from kidnapped victims. In 2017, The National Agency for the Prohibition of Trafficking in Persons (NAPTIP) commenced an investigation into Nigeria's illegal human organ harvesting. The Agency further linked human organ trafficking with some of the country's supposedly kidnappings and ritual killings (Ayodele, Punch Newspaper, 2017). The outcome of this investigation has not transformed into any policy implementation so far. There is a need to torchlight the numerous kidnappings in Nigeria to the possibilities of being done for human and body organs' trafficking purposes.

**Terrorism Kidnapping**

Hostage-taking as a terrorists' tactic has been a long-term discourse in the literature (Jenkins, 1974). Terrorism has been defined as a campaign and acts of violence waged by a small group of people to create fear and also attract broad media coverage (Jenkins, 1974). The motives for kidnapping continue to emerge as terrorists evolve in their tactics (Cook, 2011; Agbaje, 2020). In essence, kidnapping by terrorist groups is deadly and increasing globally (Ragazzi, 2017). While kidnapping for terrorism is not new, it is usually categorized and exhibited within
domestic contexts. For instance, while terrorist groups elsewhere in the world would kidnap expatriates and aid workers above all other victims, in the case of Nigeria, the attacks are difficult to categorize. However, frequent attacks are meted out on schools and local populations.

Scholars (Nwokedi, 2010; Adelakun, 2010) captured Nigeria's entire drama kidnapping trade and linked it to terrorism. It has been noted that the Boko Haram terrorist group in Nigeria have abducted several people intending to instrument them for various terrorist activities such as militants, intelligence gate-keeping, suicide bombing, and reproduction, among several other reasons (Agbaje, 2020; Markovic, 2019; Comolli, 2015). This is not to rule out that they also kidnap for financial gains in occasional instances (Perri, Lichtenwald and MacKenzie, 2009). Kidnapping by terrorists in Nigeria is rampant and hence is one of the organized crimes within the country. It can happen anywhere and anytime. People have been kidnapped on the highway, churches, homes, schools, markets and offices. Even though the Nigerian government has taken several measures to tackle the crime of kidnapping, the measures have yielded little to no success due to high poverty rates, religious bigotry, sabotage and grave socio-political factors. Another reason kidnapping thrives is the willingness of the government to negotiate with kidnappers and the lack of effective deterrence mechanisms against perpetrators. This paper thus advocates a need for urgent anti-kidnapping mechanisms and techniques by the Nigerian government to create a more effective system of combating crime.

**Consequences of Kidnapping in Nigeria**

There are significant implications of kidnapping on the victims, relatives and the entire country. Some of the identified consequences are highlighted thus:

**Fear and Insecurity**

Insecurity and fear are one major challenge Nigerians face due to the increase in kidnapping cases across the nation. An average Nigerian is now afraid of being kidnapped. The types of brutality reported by victims have caused excessive fear, significant distress and heightened anxiety in the minds of the citizens as they worry that the same plight might befall them (Ngwama, 2014). This fear is not limited to Nigerians living in the country alone but also, affects Nigerians in the diaspora and foreigners have been discouraged from visiting the country either for business or leisure.

**Economic Loss**

Kidnapping has resulted in a massive loss of money, manpower and investments in Nigeria. The actual estimated value of money lost to the kidnappers is yet unknown. However, there was evidence of ransoms paid in different parts of the country concurrently by government officials (though mostly denied), friends and families of kidnapped persons and religious sect members (Campbell, 2021). Kidnapping in Nigeria and across the Sahel has been described as an extraordinarily lucrative enterprise by the insurgents and some unruly citizens who engage in the crime for financial gains. The consequence of this is the lack of trust by potential investors, loss of money by investors and outright loss of reputation and goodwill as a result of increased kidnapping in the country.

**Physical and Psychological Trauma**

Victims of Kidnap tend to experience stress reactions including shock, anger, sadness, anxiety, guilt, insomnia and depression. It is usual to experience one or more of the mentioned psychological setbacks in captivity and upon release (Speckhard, Tabrina, Krasnov, and Mufel, 2005). Retaining to the real world after being held hostage can be just as difficult as abruptly leaving it (Wessely, 2005). More so, victims are not uncommon to be subjected to physical and psychological brutality resulting in long-term trauma. Beatings, rape, drugging and locking up are some of the vices committed by kidnappers to intimidate the captive into conforming to their demands.
Consequently, are health hazards associated with these actions, such as sexually transmitted infections, viral infections, drug overdose, injury, high blood pressure, insomnia, memory loss, disorientation and unwanted pregnancy (Akwash, 2018). While some of these health issues are treatable within a period, others can have lifetime effects on the victim. For example, the mothers and children of unwanted pregnancies from kidnappers can suffer a long-lasting stigma from society (Moore, 2014). Another physiological and psychological effect of kidnapping is what McKenzie (1987) refers to as Stockholm Syndrome. This, according to him, is the body's response to anxiety and stress. He relates this psychological issue with a kind of positive bonding that hostages often develop with their captors. This bonding may arise due to their inability to deal with the shock caused by being taken captive. It is a very dangerous feeling that may have lifelong effects on victims.

Loss of Lives
Deaths have occurred during kidnaps, be it due to victims' resistance or health-related issues experienced in captivity. On July 15, 2021, a man was killed in the Bauch-Jos highway from sporadic shootings from kidnappers in the area (Davies Ngere, 2021). Similarly, on 30 June 2021, over 100 students of an Islamic school were kidnapped, and a life was lost during the raid (Ahmad Ibrahim and Stephanie Busari, CNN, June 2021). Also, on 21 April 2021, three persons were killed out of the twenty six people kidnapped in Greenfield University, Kaduna. These are few examples compared to the high number of killings of kidnappers in Nigeria. The height of it all was the attack on the Nigerian Defence Academy, where two officers were killed and one carted away by bandits on 24 August 2021. One major reason for quick ransom payment by relatives of the kidnapped is the fear of being killed by the assailants. It is not uncommon for kidnappers to intentionally desecrate human lives through brutal killings (Nimi Princewill, 2021). Kidnappers will kill anyone who resists them, and they would kill to create fear and thus, facilitate quick ransom payment.

Conclusion and Recommendations
This paper has examined the plague of kidnapping ravaging Nigeria while identifying some of the factors responsible for kidnapping and the consequences of increased kidnapping on human security in Nigeria. It has also been established that efforts by the Nigerian government and citizens have not yielded many positive results. The reason for the low-rated success may not be farfetched from the complex and random nature of kidnapping in Nigeria. To combat kidnappings in Nigeria, this paper will recommend the following. There is a potential for kidnapping cases to decrease if the Nigerian government can carry out sincere poverty alleviation programmes (Government Corruption - The Borgen Project)

Many random kidnappings have been traced to family members, workers and friends of the kidnapped who want to extort money from them. Also, corrupt security officers should be exposed and removed from the system. There is a need for synergy among the citizens and government officials to expose the hideouts of the kidnappers in society. There is a need for collective security measures across all boards, including devised technologies designed to combat the menace of kidnapping in Nigeria. These measures include counter-insurgency and strong anti-criminal security units across the nation.

In addition, victims of kidnapping should be adequately counselled alongside their family members and associates. This will prepare them for all physical and emotional reactions. Similarly, efforts should be made to enrol victims in trauma healing programmes that translate to long-term physical and psychological stability. This paper further recommends a form of hybrid security measures to encompass all forms of kidnaps going on in Nigeria. For instance, ritual kidnapping and killing have not been fully explored. However, there is no denying that it is a frequent occurrence in Nigeria. Therefore, stakeholders in the forms of religious leaders, traditional leaders, security operatives and community leaders need to be oriented toward reducing or outright eradicating kidnapping in Nigeria.
References


---

**Funmilayo Idowu AGBAJE, Ph.D.**

ORCID ID: https://orcid.org/0000-0002-4665-366X

---

Make your research more visible, join the Twitter account of INSIGHTS INTO REGIONAL DEVELOPMENT: @IntoInsights

Copyright © 2022 by author(s) and VsI Entrepreneurship and Sustainability Center

This work is licensed under the Creative Commons Attribution International License (CC BY).

http://creativecommons.org/licenses/by/4.0/

Open Access
GEOPOLITICS AND GEOECONOMICS IN FUNDING HUMANITARIAN MINE ACTION – THE CASE OF ANGOLA

Éva Hegedűs

Doctoral School for Safety and Security Sciences / Africa Research Centre, Óbuda University, Népszínház str. 8, 1081 Budapest, Hungary

E-mail: hegedus.eva@phd.uni-obuda.hu

Received 25 August 2022; accepted 24 October 2022; published 30 December 2022

Abstract. As a result of forty years of armed conflict, Angola is today among the eight countries in the world with the highest contamination of landmines and other Explosive Remnants of War (ERWs). Since 1994, a number of international donors supported humanitarian mine clearance with varying amounts and for varying periods as part of their foreign policy agenda. The study finds that, as it is generally true in the case of foreign policy and aid, international mine action support is closely linked to underlying geopolitical considerations. Given Angola’s natural resource wealth and market potentials, geoeconomics also plays an important role in funding decisions, particularly for influential powers like the United States and Japan, but not (yet?) for China, for instance.

Keywords: humanitarian mine action; geoeconomics; geopolitics; Angola; China; The European Commission; Japan; United States

Reference to this paper should be made as follows: Hegedűs, É. 2022. Geopolitics and geoeconomics in funding humanitarian mine action – the case of Angola. Insights into Regional Development, 4(4), 166-183. http://doi.org/10.9770/IRD.2022.4.4(10)

JEL Classifications: F35, F50, H56, N40

Additional disciplines political sciences

1. Introduction

As a result of thirteen years of war for independence (1961-1974) and twenty-seven years of civil war (1975-2002), Angola is today among the eight countries in the world with the largest contamination of landmines and other Explosive Remnants of War (ERWs)\(^1\). This deadly legacy poses a daily risk to lives and limbs and hinders inclusive and lasting socio-economic development at all levels: local (household and community), regional and national.

Angola is also among those countries which have an enormous natural resource wealth and thus foreign investment and market potential (Besenyő, 2019, 63) While the country’s economic growth is held back by a range of factors, the country has always been of interest for foreign actors for various geopolitical and

\(^1\) Beside landmines, Angola has a significant problem of other ERWs, which include Unexploded Ordnance (UXO), and certain, though limited, presence of cluster munition remnants. Humanitarian mine action deals with the clearance of all these. For the sake of word economy, this study will use the term ‘landmines’ or ‘mines’ in most places, but it also implies ERW and cluster munition contamination.
geoeconomic reasons. These include political and ideological alliances; promotion of international stability, security, peace, and conflict resolution; acceleration of socio-economic development; and scaling up of bilateral commercial and trade relations.

There is already extensive research into the interrelations of foreign policy, geopolitics, geoeconomics and the provision of aid, and the way aid may be allocated to serve donors’ (but also recipients’) political, economic, military, and other goals. This study looks at the specific case of international support to humanitarian mine action and the case of Angola exploring the web of apparent or likely motives behind financial allocations from geopolitical and geoeconomic perspectives. It gives an overview of Angola’s geopolitical and geoeconomic endowments, and its existing landmine and ERW contamination including the consequences thereof. Then it analyzes the current main international funding trends and focus in mine action; and explores in depth the intersections of humanitarian mine action, geopolitics and geoeconomics. It also examines the specific case of the top three funders in the past decade: the United States (US), the Commission of the European Union (EC) and Japan. Finally, it shares an outlook to future trends, among them those relating to new emerging powers and donors, some of which would constitute interesting future areas of research.

2. A land of potentials: Angola’s geopolitical and geoeconomic endowments

As Hodges puts: ‘If human progress depended on natural resources alone, Angola’s people would be among the most fortunate in Africa’ (Hodges, 2004, 101). Indeed, from a geoeconomic perspective, Angola offers a range of opportunities by possessing, according to estimates, 35 of the 45 most important world trade commodities (Governo de Angola, n.d.). Oil is by far the most attractive for external investors. With around 7.2 billion barrels of proven crude oil reserves and an average production of 1.4 million barrels a day, Angola is currently the second largest producer in Africa (after Nigeria), and the 17th in the world (Organization of the Petroleum Exporting Countries (OPEC) 2021). Most of the export goes to China (56 percent on average in the period 2012-2020), European countries (18 percent combined), India (9 percent), Asian countries (7 percent) and the United States (5 percent). The reserves are in northern offshore fields off the coast of the Cabinda enclave and the Lower Congo and Kwanza basins. Associated to the oil fields, Angola also holds proven natural gas reserves of some 301 million m$^3$ (OPEC 2021). While in world comparison, this is not significant, it provides additional investment and trading potential.

Diamonds are Angola’s other strategic trading commodities, generating an export value of around USD1.1 billion per year from 8.5 million carats produced (Kimberley Process 2021; Statista 2021). Angola has some of the most valuable surface and underground deposits of the world (World Trade Organization 2015, 59), with the greatest concentrations in Lunda Sul and Lunda Norte provinces in the north-east.

Angola also abounds in renewable natural resources. Its varied tropical and moderate climate, and fertile soil enable the production of numerous crops and the raising of cattle (Hodges 2004, 101-102) on its agricultural lands making up around 45 percent of its 1.24 million km$^2$ of territory. The country’s rivers provide large hydroelectric potentials with a total estimated capacity of around 18 gigawatts (Búr 2020, 176). Its coastline of 1,650 kilometers provides ideal conditions for maritime trade and industrial fishing, with some of the richest fishing grounds along Africa’s southern coast. Its natural sights and parks, flora and fauna and ethnic and cultural

---

2 Due to limitations in length and the relatively low level and varying frequency of fund provision, the rest of the donors are not analyzed by the study. The study also does not look at the pre-1994 period as humanitarian mine action activities in Angola with bilateral international donor support started in 1994 only. Due to limitations in data availability, and to focus on the most recent and impactful trends, the period of analysis is from 2010 to 2020.

3 Average between 2017-2021. Data prior to 2017 is not available in public OPEC statistics.

4 Average between 2004-2020.
heritage also make Angola as a potential eco-, agro- and coastal tourism destination (Global Tourism Forum 2019).

Some analysts refer to Angola as one of the new investor hotspots in Africa (Mahajan 2009, 41). Business environment remains rather difficult though with the country ranking only 177th out of 190 on the World Bank’s Doing Business rating (World Bank 2020). Since the end of the civil war, foreign direct investment (FDI) into Angola has been low and volatile (World Bank 2021d), and predominantly concentrated in the extractives (oil and gas) sector, with smaller investments in energy, power, and construction.

While the oil and other energy sources may provide the country with a stronger perceived position than its actual economic weight (Szilágyi 2018, 206; Tordoff 2002, 14), Angola struggles with several challenges, which, in turn, affect its geopolitical and geoeconomic potential. Angola still needs to diversify its oil-based economy, improve its political and economic governance, curb corruption and money laundering, and rebuild and upgrade infrastructure, much of which was destroyed during the war.

Angola only ranks 148th in the world in terms of human development, with 51 percent of its population of approx. 33 million living in multidimensional poverty. Unemployment rates are 7 percent, with 16 percent among youth. Agriculture employs half of the national work force, but mainly for subsistence production only, with missing knowledge, equipment, and market infrastructure to drive the sector’s growth. Some 34 per cent of the adult population is still illiterate. The combined impact of oil price fluctuations and the Covid-19 pandemic has kept Angola in economic recession in the past five years with a continuously contracting GDP, from USD122 billion in 2017 down to USD62 billion by 2020 (World Bank 2021b; World Bank 2021c).

Amidst rising economic difficulties and social discontent, Angola’s political environment is at least relatively stable and predictable (World Bank 2021a). Internationally, it is among the promoters of peace and stability across the continent, with friendly relations with all its neighbors.

It is an active member of several global and regional entities, such as the Organization of African Unity, the Southern African Development Community, the United Nations, the World Trade Organization and others.

In sum, based on the systemization of Mendes Dias (as cited by Szilágyi 2018, 203-209), Angola demonstrates a promising and diverse profile in terms of geopolitical and geoeconomic factors. For now, the country lacks sufficient force and power in structural (political, economic, and military system), technological, demographic, and human, as well as transport and communication aspects. Its physical factors and natural resources, however, render it an important factor in world economy and politics, and, as the third largest economy in sub-Saharan Africa, a potential future regional power (Búr 2020, 179).

3. Landmine and ERW contamination: significance and consequences

The extensive presence of landmines and other ERW poses a critical challenge for Angola on its path to economic and social progress, infrastructural development, and attraction of further domestic and foreign investments, particularly in the most remote and contaminated areas.

---

5 Unless otherwise referenced, all statistics based on United Nations Development Program n.d.
In its latest, 2020 Mine Ban Treaty Article 7 Report, Angola reported a total land of 85.42km² yet to be cleared in 17 of its 18 provinces (Government of Angola, 2021). Of this, 84.41km² is reported to be contaminated by antipersonnel mines (the rest by antivehicle mines). Some 82.51km² are classified as a confirmed hazardous area, and 2.91km² as a suspected hazardous area. In addition, some 3,749 kilometers of roads are estimated to be contaminated (Mine Action Review 2020, 32). International survey and clearance operators argue that based on available evidence, the total size of contaminated land is less than the above, ‘only’ 30km² (Mine Action Review 2020, 30). Regardless, Angola still falls in the group of countries with large/heavy contamination. As people gradually return to abandoned areas, new, previously unrecorded mined areas are also likely to be discovered yet (Mine Action Review 2020, 32).

Angola joined the 1997 Mine Ban Treaty on 1 January 2003, committing to clear all its contamination within the ten-year deadline set by the treaty. Following two extension requests, it is now bound to remove all mines and ERW by at latest 31 December 2025 – a target which, based on its current progress, it is very unlikely to meet. In fact, Angola already flagged that based on the current contamination and level of funding, clearance will only finish by 2028 the earliest (“Desminagem vai ser concluída apenas em 2028”, 2021).

Mines were laid both by government and rebel UNITA (União Nacional para a Independência Total de Angola) forces, and their allies (mainly Portuguese, Cuban and South African troops), often in a sporadic and unrecorded manner (Búr et al., 2013, 158) to prevent the enemy’s movement and access to critical infrastructure. While all provinces are affected to a certain extent, the most contaminated ones are Kuando Kubango, Mexico, Kwanza Sul, Zaire, Lunda Sul, Kwanza Norte and Bié (in this order) with a contaminated land of between 5.5km² and 17.9km² in each (Mine Action Review 2020, 32).

The following publicly available map in Angola’s latest Article 7 report (Government of Angola 2021, 13) shows the extent of current known contamination. The red dots indicate open minefields (where clearance not yet started); the orange dots indicate fields with ongoing clearance work; and the green dots mark closed/cleared mine fields.

---

6 In full name: Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on their Destruction.
7 Submitted annually in compliance with Mine Ban Treaty obligations.
8 Landmine clearance and risk education operations in Angola are carried out by national operators (National Institute for Demining - INAD and Association of Mine Professionals – APACOMinas) and international operators (APOPO, The Halo Trust, Mines Advisory Group and Norwegian People’s Aid). Each operator is tasked to a specific province in the national demining work plan.
9 The provinces of Huambo, Malange, and Namibe are approaching completion.
10 More detailed maps are held by the coordinating government agency and national and international clearance operators.
Landmines pose a risk to people’s lives and physical security. In the absence of a comprehensive and up-to-date national casualty surveillance system, the number of victims (killed or injured but survived) in Angola is estimated to be between 60,000 and 80,000 (Landmine & Cluster Munition Monitor 2021a). As many accidents go unreported, this number may be higher. Generally, 80 percent of victims are civilians and at least 30 percent are children, most of them boys (Landmine & Cluster Munition Monitor 2021c, 34). All victims, including those directly affected by an accident (whether killed or survived) and their close family and/or household members require continuous support, such as emergency medical care, long-term rehabilitation, psycho-social assistance,
and socio-economic integration. New casualties still occur, at least one every other week, with (at least) 98 people killed or injured in 2019-2020 (Landmine & Cluster Munition Monitor 2021a, 5; Landmine & Cluster Munition Monitor 2021c, 45).

Landmines also prevent Angolans to access basic services, such as healthcare, education, jobs, and markets. People cannot move around safely and freely; use agricultural lands, roads, rivers, bridges, and railways; (re)build homes; and develop infrastructure. Large areas remain uninhabited and uncultivated due to landmine and ERW contamination, which is a critical issue in a country, where more than half of the population is engaged in subsistence farming. Landmines and ERW hinder post-conflict reconstruction, economic and social progress, and the achievement of long-term and stable peace, social cohesion, and human security. Mine contamination is an obstacle to meet the 2030 Sustainable Development Goals (SDGs) since development projects cannot proceed without prior clearance of land (Downs, 2009, 82-84).

4. Funding mine action activities – focus and trends

State parties to the 1997 Mine Ban Treaty are legally bound to stop the use, production, and transfer of mines and ERW; to destroy all stockpiles and to clear contamination in areas under their jurisdiction. Those in the position to do so, are also required to assist landmine survivors, their families, and communities; and support or conduct risk education to help prevent new accidents. Assistance to victims is also included in a range of other binding legal instruments on human rights, children’s rights, women’s rights, disability rights and political, economic, social, cultural, and civil rights.

Angola, just as most countries, is unable to fulfil all these obligations without international support. Over the past ten years (2010-2020), it has been the ninth biggest recipient of international funding for humanitarian mine action, with a total of USD159.93 million (some 2.8 per cent of all international mine action support worth USD5.7 billion globally) (Landmine & Cluster Munition Monitor 2021g, 1).

Fig. 2. Angola’s share of all international support between 2010-2020.
Table 1 shows the international institutional donors that have allocated dedicated funding for either of the five pillars of mine action (stockpile destruction, clearance, mine risk education, victim assistance and advocacy) in Angola between 2010-2020. It also includes national contributions provided by the government of Angola to its own mine action sector in the same period.

### Table 1. International and national institutional contributions to humanitarian mine action activities in Angola between 2010-2020.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>9.85</td>
<td>8.35</td>
<td>8.68</td>
<td>6.00</td>
<td>6.00</td>
<td>5.30</td>
<td>3.75</td>
<td>2.50</td>
<td>4.00</td>
<td>4.10</td>
<td>7.00</td>
<td>65.63</td>
<td>1.00</td>
<td>69.56%</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.06</td>
<td>2.84</td>
<td>4.70</td>
<td>4.11</td>
<td>10.97</td>
<td>4.00</td>
<td>6.86%</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>2.75</td>
<td>1.12</td>
<td>1.68</td>
<td>2.10</td>
<td>0.80</td>
<td>0.20</td>
<td>0.55</td>
<td>0.84</td>
<td>0.79</td>
<td>1.57</td>
<td>2.22</td>
<td>15.76</td>
<td>3.00</td>
<td>8.97%</td>
</tr>
<tr>
<td>Norway</td>
<td>0.66</td>
<td>-</td>
<td>0.08</td>
<td>-</td>
<td>-</td>
<td>0.30</td>
<td>0.18</td>
<td>0.25</td>
<td>0.31</td>
<td>1.00</td>
<td>2.98</td>
<td>8.00</td>
<td>1.58%</td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.37</td>
<td>0.38</td>
<td>0.75</td>
<td>9.00</td>
<td>0.47%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.19</td>
<td>0.05</td>
<td>-</td>
<td>-</td>
<td>0.24</td>
<td>11.00</td>
<td>0.15%</td>
<td></td>
</tr>
<tr>
<td>EC</td>
<td>26.52</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>51.62</td>
<td>2.00</td>
<td>32.20%</td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>1.07</td>
<td>1.05</td>
<td>1.16</td>
<td>1.30</td>
<td>1.40</td>
<td>1.15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.15</td>
<td>3.00</td>
<td>4.40%</td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.64</td>
<td>1.93</td>
<td>0.80</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.37</td>
<td>6.00</td>
<td>2.12%</td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>0.48</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.48</td>
<td>16.99</td>
<td>0.20%</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>0.20</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.20</td>
<td>12.00</td>
<td>0.15%</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>0.07</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.07</td>
<td>13.00</td>
<td>0.04%</td>
<td></td>
</tr>
<tr>
<td>Czech Republic</td>
<td>0.05</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.05</td>
<td>14.00</td>
<td>0.03%</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>0.01</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.01</td>
<td>15.00</td>
<td>0.01%</td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>0.84</td>
<td>1.39</td>
<td>0.60</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.30</td>
<td>7.00</td>
<td>2.08%</td>
<td></td>
</tr>
<tr>
<td>ALL INTERNATIONAL</td>
<td>42.30</td>
<td>13.29</td>
<td>13.71</td>
<td>10.08</td>
<td>32.58</td>
<td>6.65</td>
<td>4.79</td>
<td>3.13</td>
<td>7.08</td>
<td>11.11</td>
<td>16.21</td>
<td>159.93</td>
<td>n/a</td>
<td>100%</td>
</tr>
<tr>
<td>Government of Angola</td>
<td>29.18</td>
<td>59.61</td>
<td>76.71</td>
<td>115.46</td>
<td>121.30</td>
<td>59.17</td>
<td>24.50</td>
<td>45.68</td>
<td>26.46</td>
<td>15.70</td>
<td>6.11</td>
<td>579.02</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>GRAND TOTAL</td>
<td>71.48</td>
<td>72.39</td>
<td>90.22</td>
<td>125.68</td>
<td>153.68</td>
<td>68.82</td>
<td>48.21</td>
<td>43.34</td>
<td>32.61</td>
<td>26.81</td>
<td>21.32</td>
<td>738.95</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

*Note: Contributions are rounded to the nearest two decimals
Source: own compilation from data in Landmine & Cluster Munition Monitor 2021h.

As Table 1 shows, mine action activities in Angola were supported by a total fifteen international institutional donors in the period of 2010-2020 in a total value of USD159.93 million. International funding has significantly dropped after 2010 and has generally seen a trend of shrinking volume, especially between 2011-2013 and 2015-2018.

By far the biggest mine action donor in Angola is the United States, having contributed some 40.6 percent of all international support (USD65.03 million) in the period of analysis. It is followed by the EC, with some 32.3 percent of all foreign funds (USD51.62 million). Other two bigger donors have been (in this order) Japan and the United Kingdom, although their total shares are significantly lower compared to the United States and the EC (some 8.6 percent and 6.9 percent, respectively). This is in line with global trends, with the top six donor states consistently including the United States, the EC and Japan (in this order) and the United Kingdom (in place six) (Landmine & Cluster Munition Monitor 2021b).

---

11 Although this study only focuses on the last ten-year period (2010-2020), it must be added that Angola also received a total of USD144.65 million of international contribution and a total of USD80.58 million of national government contribution in the period 2005-2009. International contributions per year were as follows: 2005: USD33.57 million; 2006: USD48.11 million; 2007: USD19.79 million; 2008: USD22.14 million; 2009: USD18.84 million. Detailed donor breakdown of international contributions is not publicly available. Therefore, this five-year period was not included in the present analysis. Source: Landmine & Cluster Munition Monitor 2021h.

12 The significant decrease from 2010 to 2011 and from 2014 to 2015 is partially due to that statistics list EC contributions under one single year, when they were awarded. However, these allocations cut across several years, thus if broken down per year, the annual total international contributions would be higher than those in the table 1.
In 2019, Angola also reported receipt of funds from oil and gas companies such as British Petroleum and ENI, worth USD39.19 million, which is considerably higher than the USD26.81 received from institutional donors only (Landmine & Cluster Munition Monitor 2021a, 7).

In terms of sectors, the predominant part of international contributions has been (and continue to be) earmarked for clearance, risk education and capacity-building (Landmine & Cluster Munition Monitor 2021h). Although Angola has one of the highest casualty rates (and number of survivors to support) globally, since 2012 no funds have been dedicated by any donor for victim assistance. Prior to that, only three donors – Japan (in 2009), Germany (in 2010) and the Netherlands (in 2010 and 2011) – contributed some funds to this pillar.

Angolan government contributions must also be briefly highlighted when analyzing funding trends in mine action, since these have consistently exceeded the volume of international contributions, in most cases to a significant extent. Between 2010-2020, Angola contributed some USD579.02 million to its own mine action efforts, which is 3.6 times more than all international contributions in the same period, and just over 78 percent of the entire mine action budget. Government funds support national entities in support of national development plan priorities, such as the Intersectoral Commission on Demining and Humanitarian Assistance and the Executive Commission for Demining. The latter funds development projects and associated demining operations carried out by the Angolan armed forces, the National Reconstruction Office, and the largest clearance operator in country, the National Demining Institute.

5. Intersections of humanitarian mine action, geopolitics and geoeconomics

Following nineteenth century colonization and twentieth century cold war rivalry, the twenty-first century sees a new round of ‘scramble for Africa’. The main motives of the third round are as before: resources, energy, and raw materials (Marshall 2016, 170). The scramble’s exploitative nature has not changed much over time: foreign actors still focus on maximizing the benefits and opportunities offered by African lands and economies (Pásztor 2020, 57). The market potential of sub-Saharan Africa also attracts a growing number of new actors such as China and mid-size power emerging economies, such as Brazil, India, Russia, and Turkey.

Generally, the economic development of a country is in the interest of both parties: the host country and its foreign partners. By definition, official development aid is provided by foreign donors to developing countries to further economic development and welfare (Organisation for Economic Co-operation and Development, n.d.). It is long-known that, beyond altruistic reasons, the provision of aid also serves to promote donors’ own political, economic, and often military-security priorities (Holdar 1993, 453). Luttwak and Lorot have already emphasized the new era and increasing weight of geoeconomics (vis-à-vis geopolitics) (Szilágyi 2018, 199). Blackwill and Harris argue that aid remains one of today’s main geoeconomic instruments, together with trade and investment policy, economic and financial sanctions, financial and monetary policy, energy and commodities, and cyber space (Schneider-Petsinger, 2021).

The promotion of economic development and livelihoods features high on donor states’ agenda when it comes to funding humanitarian mine action activities. Donor focus has gradually shifted from the early 2000s from simply removing mines and supporting victims to maximizing the impact of clearance for socio-economic development.

---

13 Exact amounts for 2009 and 2010 are not broken down in public statistics. In 2011, the amount was only USD1.92 million provided by the Netherlands. It must be noted, however, that victim assistance is also and often provided as part of other, general social sector support budget allocations (e.g., for emergency health care system development, disability inclusion, economic integration of vulnerable groups etc.). Thus, other international development and humanitarian funding may potentially cover and benefit mine and ERW victims, even though this may not always be possible to ascertain, including the exact amount.

14 Angola’s national budget also has some lines that may also include mine and ERW victims as part of other envelopes for financing socio-economic development.
including in Angola (Devlin 2010, 5). This includes enabling economic and human development plans through clearing critical and non-critical infrastructure at future or ongoing development project and investment sites. It also includes victims’ socio-economic reintegration and helping them rebuild their livelihoods. In addition, it aims at reducing the risk of new accidents to minimize the number of new victims and, ultimately, to strengthen human security.15

The most common factors that generally influence funding allocations (by bilateral donors) are as follows16 (in this order of ranking): the magnitude of identified needs (that is, results and impact that can be gained with return on investment); integration of mine action needs in consolidated funding appeals (that is, easiness of providing coherent support to validated requests); links to development opportunities, national development plans and donor development programming (that is, socio-economic impact and economic development gains); links to peace-building (that is, indirect security and conflict prevention considerations); other external reasons (e.g. ‘competing’ crises, foreign policy aims, etc.). Socio-economic impact and the ability to report back on results are also named by donors as two key factors that are necessary for continued funding (alongside better survey results on minefields, enhanced local capacities and national ownership).

In terms of which countries are selected for support, donors typically follow fourteen main criteria17, namely (in order of donors’ ranking): the target country’s commitment to the Mine Ban Treaty; the needs and the humanitarian imperative; the donor’s geographic strategy; link to the donor’s overall development and humanitarian policy; level of national ownership; constraints to development; concentration of other donor funds, the donor’s thematic priorities; and lastly, impact, sustainability, capacity-building and peace-building. In most cases, the countries which receive support for mine action are also recipients of other forms of development and/or humanitarian aid from the same donor. Donors rather focus on a smaller number of countries (in line with the Aid Effectiveness agenda) and select those with which they already have a ‘special relationship’.

All this generally reinforces that ‘humanitarian mine action’ is no longer considered purely humanitarian. In contrast, it is part of the broader foreign policy agenda; and falls under donors’ broader development and humanitarian support to further, among others, socio-economic development. In many cases, mine action is also linked to national and international security considerations in support of peacebuilding, disarmament, conflict prevention and post-conflict recovery.

5.1 The case of the United States

Interestingly, the United States is the ‘number one’ mine action donor (in terms of volume of funding) both in Angola and worldwide even though it has not acceded to the Mine Ban Treaty18. The majority of US mine action funding is channelled through the Office of Weapons Removal and Abatement, which sits within the Bureau of Political-Military Affairs (BPMA) of the Department of State (DoS). BPMA is the direct link of the DoS to the Department of Defense, and it is main mission is to build ‘enduring security partnerships to advance US national security objectives’ and provide ‘policy direction in the areas of international security, security assistance,

---

15 The increased importance of human security is also reflected in the fact that donor policies now cover all types of explosive devices that may or do pose a threat to human security, not only landmines but also ERWs, improvised explosive devices (IEDs), and cluster munitions.
16 Own analysis in italic, based on Devlin 2010, 22.
17 Paragraph based on Devlin 2010, 11.
18 The United States has not signed the 2008 Convention on Cluster Munitions either, but it is party to the 1980 Convention on Certain Conventional Weapons (including its protocols). The country has a particular stance on the use of anti-personnel landmines by itself. As of 31 January 2020, the Department of Defense has adopted a new policy that allows the use of landmines in major combat in future conflicts if those are equipped with a self-destruct/self-deactivation function, are detectable with commonly available detection equipment and can self-destruct within thirty days. This decision represented a major step back in previous US efforts to eradicate landmines and received a strong wave of criticism from domestic and foreign human rights, humanitarian and even military circles. Source: United States Department of State 2021e.
military operations, defense strategy and plans, and defense trade’ (United States Department of State 2021b). This demonstrates that humanitarian mine action is embedded in the security field, which, in turn, is normally influenced by geopolitical considerations. The United States is among the four donors (alongside Austria, Belgium, and Germany) that have stated that the biggest part of their mine action funding is ‘not linked to the broader development cooperation agenda’ (Devlin 2010, 9).

Regardless, US funding does further economic development objectives (alongside security ones), at least in its official communication. The two mine action related aims of the US’s Conventional Weapons Destruction (CWD) program (part of which is mine action) are (United States Department of State 2021c): a) ‘to remediate explosive hazards contamination, returning land to safe and productive use’; and b) ‘to promote US foreign policy interests by broadening international support for CWD efforts’ (the latter also extends to the destruction and securing of small arms and light weapons at risk of proliferation to terrorists, insurgents etc.).

The latest annual report on mine action achievements cites similar objectives, with only slight rephrasing: a) to ‘improve stability and prosperity by clearing ERW and returning land to productive use’; and b) to ‘build trust and deepen relationships with key partners to accelerate achievement of broader US foreign policy objectives’ (United States Department of State 2021d). The report makes clear: ‘The measurable, tangible results that flow from the US government’s commitment to CWD programs strongly support US foreign policy priorities. In addition, these programs help to protect the lives and livelihoods of civilians so they can more safely remain in their own countries.’ CWD programs aim at mitigating both transnational threats to the security of the United States, and national and regional threats to stability and human security. Analysts such as Kennedy have already pointed earlier to the importance of stability in developing states from the perspective of ensuring US national interests (Szilágyi 2018, 139).

Globally, Angola ranks seventh on the list of countries supported by US mine action funding from 1993; and it ranks number one among the 26 African countries that have received US support ‘to promote peace-building, economic growth and prosperity’ (United States Department of State 2021d, United States Department of State 2021a). The funds of USD145.7 million received by Angola represent 28.6 percent of the total funds (of USD509 million) spent by the United States on the continent, and 3.6 percent of the total funds (of nearly USD4 billion) globally.

**Fig. 3. Top ten countries funded by US CWD programs between 1993-2020 (in thousand USD)**

![Image of Fig. 3](image-url)
Despite the low share (of 3.6 percent) in global US CWD funding, including for mine action, Angola has long had a firm place on the priority list of the United States. During the Cold War, the country was important from both geoeconomic perspectives (as source of oil) and geopolitical ones: being one of the venues of the political-ideological ‘battle’ and expansion efforts vis-à-vis the Soviet Union. Backing the rebel movement UNITA, the United States in fact has indirectly contributed to today’s landmine problem in the country; the same issue that it works to tackle today. Angola also saw one of the ‘most bizarre constructions of the Cold War’ (Búr 2020, 177). The investments of the US multinational oil company Chevron were protected by Angolan government forces backed by Cuba and the Soviet Union against the rebel forces of UNITA, who were, in turn, supported by the United States and others. At the same time, the high extraction fees and taxes paid by US oil companies helped Angola to finance the military support received from Cuba and the Soviet Union.

Today, US companies have the largest market share in Angolan oil production (Chevron with 26 percent, Exxon Mobil with 19 percent) (International Trade Administration, 2021). The Angolan government intends to engage more US companies for multi-billion-dollar projects in the oil industry, including exploration, development, transportation and storage, refineries, and associated infrastructure (International Trade Administration, 2021). US government sources point to a range of investment opportunities as Angola works to diversify its oil-based (and dependent) economy, for instance, in public transportation, tourism, alternative energy, extractives, agriculture, fisheries, telecoms, and ports rehabilitation and management (United States Department of State 2021f). Latest US government policies aim at strengthening economic and trade relations across sub-Saharan Africa, promoting US investments, creating a favorable business environment for US firms, and maintaining peace and security, stabilizing resource supplies and other infrastructural developments (Pásztor 2020, 56). Under the African Growth and Opportunity Act, the US ensures preferential trade benefits for Angola, its third largest trading partner on the continent (thanks to petroleum), although only the 85th globally.

Other US support, mostly under official development assistance programs, includes health system strengthening, infectious disease prevention, professional military education, maritime security, technical assistance to the financial sector, strengthening of democratic institutions, and development of tourism through environmental conservation and mine clearance (United States Department of State 2021g). These objectives – as part of overall economic growth and investment promotion – require safe and accessible lands. Analyzing its mine action funding history, the United States to date has shown being considerate of the need to fund clearance resources on known or suspected hazardous areas to support such expansion plans.

5.2 The case of the European Commission

Despite its withdrawal after 2017, the EC has been the second biggest mine action donor in Angola (with USD51.62 million, that is, 32.28 percent of all foreign support) in the period of 2010-2020. The EC is also Angola’s main partner for import and third biggest commercial partner (Delegação da União Europeia em Angola, 2021).

In 2010, the EC allocated a total of USD26.52 to mine action in Angola in line with its 2008-2013 Angola country strategy from the 10th European Development Fund (EDF) (Landmine and Cluster Munition Monitor 2021h). Sub-grants were awarded to international NGOs and a French commercial company for clearance and for building national mine action coordination and operational monitoring capacities\(^\text{19}\) over multi-year project timeframes. In 2014, the EC disbursed another USD25.1 million (still from the 10th EDF) for the period of 2014-2016, albeit with eight months of delay.

\(^{19}\) Previously, capacity-building activities were funded as part of a comprehensive capacity-building initiative by the United Nations Development Program between 2003-2011. After co-funding certain activities, the EC took it over fully from 2011 until 2015.
At the 14th Meeting of States Parties in 2015, the EC stated that it would continue funding those countries ‘that need support in meeting their commitments under the Ottawa Convention’ (AP Mine Ban Convention 2015, 2). This was (and has still been) certainly the case in Angola, even though the country eventually fell off the EC’s priority list. Since 2017 onwards, the EC did not provide any funds for mine action activities in Angola, even though needs were flagged and discussed in in-country consultations (AP Mine Ban Convention 2018, 3). Mine action was also not mentioned anymore as a priority in the joint National Indicative Program (NIP) for 2014-2020 of the EC and Angola (European Commission 2015). Instead, technical, and vocational education and training and higher education, sustainable agriculture and water and sanitation were selected for support. Additional priorities noted in the document were: strengthening the democratic political culture; increasing institutional capacities; intensifying the fight against poverty and against corruption; improving transparency and accountability; establishment of a competitive and diversified economy and more favorable trade and investment climate; sustainable and inclusive growth, creation of decent jobs, extension of the national social protection floor and strengthening of civil society.

The withdrawal of EC stands in contrast with the fact that is among the top three donors in mine action globally. Between 2016-2020, it allocated more than USD415 million, which was 35 percent more than between 2011-2015 across at least nine countries (Landmine and Cluster Munition Monitor 2021d), even though not in Angola anymore. Based on official statements and other communication, supporting mine action is among the ‘high priorities of European Union foreign policy’ (Landmine and Cluster Munition Monitor 2021d; European External Action Service 2018, 8). Mine action is linked to development programming (notably, the achievement of the 17 SDGs), humanitarian response, conflict prevention and post-conflict rehabilitation. Funding is allocated from the EC’s Neighborhood Development and International Cooperation Instrument that finances sustainable development, democracy promotion, human rights, stability and peace-related initiatives, among them humanitarian mine action. Thus, regardless of its withdrawal from Angola, the EC also clearly demonstrates a close interconnection of geopolitical and geoeconomic considerations and the thematic areas it funds, including mine action, as it has been the case while it was still a main donor for the country.

5.3 The case of Japan

Japan has been the third largest and a consistent mine action donor in Angola between 2010-2020 (just as it is also among the top donors globally). Even though its share has been much lower than those of the United States and the EC, it has provided 8.5 percent of all international contributions in a total value of USD13.7 million in this period.

Japan’s mine action objectives are three-fold: a) support to seriously affected countries; b) promote south-south and regional cooperation; and c) support mine victims and survivors (United Nations Mine Action Service 2021, 12). Accordingly, Japan generally provides comprehensive support including clearance, risk education and victim assistance, and related support such as construction of schools and revitalization of economic activities to promote stability and development (Landmine & Cluster Munition Monitor 2021e). Although in the context of Angola, Japan has to date focused on clearance and risk education, only supported victim assistance once in 2009. Japan also emphasizes the need for efficiency and building on technological development including utilizing Japanese demining technology and engineering expertise for advancements and local capacity strengthening. Japan’s long-term support to mine action, including financial contributions from its official development assistance, has been continuously reinforced at the various pledging conferences (Delegation of Japan to the Conference on Disarmament 2021).

20 Japan was a former mine producer (in addition to importing mines from the US) until its accession to the Mine Ban Treaty.
Japanese foreign policy clearly links humanitarian mine action with the broader development, poverty reduction and social reintegration agenda through its support being channeled through the official government agency, Japan International Cooperation Agency (JICA) Infrastructure and Peacebuilding Department (Japan International Cooperation Agency Infrastructure and Peacebuilding Department 2016, 2). Further, the Tokyo Guidelines also state that landmines are not only of humanitarian concern but are a threat to peace and stability, and obstacles to reconstruction and development (Ministry of Foreign Affairs of Japan 1997). Japanese foreign policy and development aid has human security at its center in the effort to promote sustainable development (Carvalho 2011, 315). This explains why Japan has long supported Angola (among other war-torn African countries) to transition to long-term peace and development, an essential component of which is to help enable progress in humanitarian mine action. Carvalho argues that the human security and peace-building focused development approach of the Tokyo International Conference on African Development (TICAD) – originally established in 1993 to create trade and investment opportunities and promote sustainable development – has contributed to Japan’s soft power role in international politics. Human security has helped Japan to create a formal link between aid and security for development purposes, the latter being clearer as Japan’s direct national security interests were not at stake in case of the Angolan landmine problem.

Other analysts claim that by referring to the aims of promoting peace and human security, Japan furthers own national security interests, whereby its new official development assistance regulations now allow for providing aid to military forces. Goto argues that Japan’s support to African countries (through TICAD) ultimately aims at promoting own economic interests (Goto 2015), and only partially an altruistic image. As Japan’s demand for natural resources, in particular fossil fuel, increases, it focuses on the most important sub-Saharan counterparts for supply and markets. In addition, it is trying to counter-balance the expansion of emerging donors, most importantly its major competitor on the continent, China. These point to a combination of commercial (geoeconomic) as well as strategic-political (geopolitical/geostrategic) considerations behind the provision of the different forms of aid, including for security purposes.

The Angolan - Japanese economic and trade relations are not yet very significant (USD60 million of exports to Japan, mainly oil and raw materials, and USD35 million imports from Japan in 2019 (Ministry of Foreign Affairs of Japan 1997), but both parties seem to be keen to scale up cooperation in the future (Embassy of the Republic of Angola in Japan 2015).

Conclusion and future perspectives

Most donors link the issue of humanitarian mine action and provision of bilateral international support to traditional geopolitical and geoeconomic considerations and priorities. These may range from humanitarian efforts and development agendas to stability, peace, and security-focused motives (Bindseil and Mansfield 2020, 9). These allow donors to develop or strengthen their ‘soft power’ role and build bilateral linkages with the supported country with potential spill-over effect(s) in other areas of common interest, such as political and economic cooperation, trade, and a strong social and human capital.

Securing financial resources for the mine action sector has been a critical challenge over the past years for Angola (Mine Action Review 2020, 34). It also means the biggest challenge for the country to meet its current, 31 December 2025 clearance deadline, extended for the second time already (Landmine & Cluster Munition Monitor 2021a, 3). According to latest government and clearance operator calculations, an estimated amount of between USD180 and USD 265 million would still be required through 2021-2025 to achieve full clearance of all known contamination (Government of Angola 2021, 11; Mine Action Review 2020, 34)\(^{21}\).

\(^{21}\) The latter not including already registered 2020 contributions.
Donors’ aid budgets have also been greatly affected by the pandemic (as a consequence of global economic impacts) with subsequent cuts in mine action envelopes. The United Kingdom has for instance reduced its total mine action budget by 75 to 80 percent for the period of 2022-2024 (Landmine & Cluster Munition Monitor 2021f), with no funding planned for Angola from 2022 (Cormack 2021). This is despite that the country became a new mine action donor for Angola from 2017; and in terms of fund volume, quickly became the overall fourth largest supporter in the period of 2010-2020.

The question arises which countries will remain the key funders over time and for what reasons, as the global economy remains being hit by shocks influenced by oil prices, pandemics, and security and development challenges.

Will, for instance, China, that is already Angola’s biggest trade partner (through the oil exports) appear on the landscape of humanitarian mine action donors? Will it purse purely humanitarian or developmental or rather specific economic and commercial motives? China did fund mine action activities in the past in certain countries through its Ministry of Foreign Affairs and Defense based on needs, local conditions, and guarantees for capacity-building and sustainability, though with a relatively low annual budget. It also provided demining staff, equipment, and trainings, among them to Angola. Since 2008, it did not officially report however any financial contributions neither globally nor to Angola (Landmine & Cluster Munition Monitor 2010), even though at a recent United Nations Security Council debate, it noted examples of support in some countries from 2015 to date (United Nations Security Council 2021). Regarded as one of the biggest geo economic players of our days, China’s main interests in Africa are access to resources, markets and investments and political stability to ensure undisturbed transport (Marshall 2016, 166). While it pays little attention to issues such as human rights, governance, equal distribution of wealth and opportunities, corruption and so on.

Will the UK eventually return as a key mine action donor after recovering its economic situation and the earlier level of its aid budget? Will other emerging mid-size or larger powers, including from the BRICS group (Brazil, Russia, India, and South Africa) or Turkey decide to fund mine action activities alongside increasingly focusing on economic and investment opportunities on the continent? These provide intriguing themes for further research from both the perspectives of humanitarian mine action (funding) as well as the disciplines of geopolitics and geoeconomics.

References


Organization of the Petroleum Exporting Countries (OPEC). 2021. Table 1.2: OPEC Members' crude oil production allocations (1,000 b/d); Table 2.2: OPEC Members’ GDP at current market prices (m $); Table 5.1: OPEC Members’ crude oil exports by destination (1,000 b/d); Table 9.1: World proven natural gas reserves by country (bn standard cu m). Annual Statistical Bulletin 2021. https://asb.opec.org/data/ASB_Data.php.


Éva HEGEDŰS is a PhD candidate at Óbudai University, Doctoral School for Safety and Security Sciences and a junior member of the Africa Research Centre. She has been working in the international humanitarian and development sector since 2009, including in Humanitarian Mine Action between 2014-2015 as Programme Officer of Mines Advisory Group in Laos. She also has field experience from other contexts such as Bangladesh, Colombia, Laos, Mozambique, North-East Syria, the Occupied Palestinian Territory and Uganda, among others, in programme development and management roles with Save the Children, the Norwegian Refugee Council, CARE, GIZ, the IFRC and the UN Refugee Agency. She holds an M.A. in International Relations from Central European University (Hungary/Austria), and a double M.A in Portuguese Language and Literature as well as in Communications (Journalism) from Eötvös Loránd University (Hungary). Research interests: humanitarian mine action, development and humanitarian aid, armed violence reduction and conventional weapons destruction, Portuguese Africa, Latin-America

ORCID ID: https://orcid.org/0000-0002-9159-8903