NOVA IMS
Information Management School

DOCTORAL PROGRAMME

Information Management

Using Network Analysis to Map Public Procurement Activities in Ceará Municipalities (Brazil)

Marcos da Silva Lyra

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor in Information Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

USING NETWORK ANALYSIS TO MAP PUBLIC PROCUREMENT ACTIVITIES IN CEARÁ MUNICIPALITIES (BRAZIL)

by

Marcos da Silva Lyra

Doctoral Thesis presented as a partial requirement for obtaining a Ph.D. in Information management

Co-Supervisor: *prof Doutor* Fernando Bação

Co-Supervisor*: prof Doutor* Flávio Luis Portas Pinheiro

December 2023

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledge the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisboa, December 20, 2023

DEDICATION

With profound gratitude, I dedicate this doctoral thesis to all those who stood by me unwaveringly throughout this unique journey. Especially, my dear wife, beloved daughter, dedicated parents, and caring sisters - my solid foundation and unwavering support in every phase of my life. Your words of encouragement, gestures of love, and infinite patience have been the anchor that kept me steady in the face of challenges.

To my loyal friends, whose constant presence brought joy, relief, and the sharing of unforgettable moments, my heartfelt acknowledgment. Your exchanges of ideas, stimulating discussions, and tireless support have immensely enriched this academic journey.

I cannot fail to mention the mentors and professors who guided me with wisdom and patience, shaping my understanding and honing my skills throughout this challenging academic path.

To those whose contributions might not be explicitly mentioned but have left an indelible mark on my journey in some way, my deepest thanks.

May this thesis be a modest tribute to the trust and support I have received. May my work resonate positively and contribute to the advancement of knowledge in our field.

ACKNOWLEDGEMENTS

First, I would like to thank God for providing me with the strength, guidance, and inspiration throughout this academic journey.

I extend my heartfelt gratitude to my esteemed supervisors, Fernando Bação and Flávio Pinheiro, for not only accepting the responsibility of overseeing my research but also for their unwavering guidance, insightful feedback, and valuable input that played an instrumental role in shaping and refining this work. Their dedication to academic excellence has been truly inspiring.

A profound acknowledgment goes to the distinguished faculty members of the information management course at Universidade Nova de Lisboa. Their commitment to imparting knowledge, fostering critical thinking, and nurturing a conducive learning environment has been invaluable. The technical expertise and wisdom they shared have significantly enriched the quality of this research.

I am profoundly appreciative of the Tribunal de Contas do Estado do Ceará for granting me the opportunity to conduct this Ph.D. research. Their support and cooperation have been pivotal in facilitating the data collection and analysis process, allowing for a comprehensive exploration of the subject matter.

In addition, I wish to express my sincere thanks to my peers and colleagues whose stimulating discussions, constructive debates, and camaraderie have provided a stimulating intellectual environment that has pushed the boundaries of my understanding and creativity.

I would also like to extend my gratitude to my family, whose unwavering love, encouragement, and understanding have been my bedrock throughout this academic endeavor.

Lastly, I am indebted to all those whose names might not appear on these pages but have nonetheless contributed significantly to my growth and learning experience.

This work stands as a testament to the collaborative efforts, support, and encouragement of all those mentioned and those unmentioned. May this endeavor contribute to the betterment of knowledge and the advancement of our field.

ABSTRACT

Illegal activities in public treasuries worldwide pose a significant threat to the financial stability of cities, states, and countries, resulting in monetary losses. Despite their widespread occurrence, empirical research has not furnished proper analytical tools for objective investigations. More practical insight remains scarce, limiting the effectiveness of public-private networks in efficient spending interactions. Public tenders, the primary mechanism for government procurement, are highly susceptible to fraud. Control institutions make efforts, albeit with limited success, to identify fraudulent activities between governments and firms during the procurement process.

This thesis aims to enhance the understanding of fraud in public procurement and explore the benefits of statistics and network analysis in detecting signs of fraud. Using procurement data from 184 municipalities of Ceará, Brazil, the research spans two empirical studies. The thesis begins with a literature review on data-driven methods for detecting fraud, corruption, and collusion in public procurement activities.

In practical application, it starts with a characterization of the entire procurement network to map the data's structure and complexity. Network science methods examined firms co-bidding relations, identifying communities susceptible to cartel formation. A composite corruption risk indicator (CRI) is adapted based on local empirical data. The CRI focused on binary relations between tenders and winners, using red flags at the contract level. Firms and municipalities were classified into four corruption risk bands based on their scores, allowing the identification of groups more susceptible to fraud.

Overall, this doctoral thesis advances our understanding of irregular activities in public procurement through statistics and network analysis methodologies, providing insights into fraud detection, risk assessment, and the identification of vulnerable groups within the public procurement process.

KEYWORDS

Fraud Detection; Corruption Risk; Network Science; Public Procurement; Co-bidding Network

Sustainable Development Goals (SGD):

INDEX

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ANN - Artificial Neural Networks

CRI – Corruption Risk Index

CRM - Customer Relationship Management

CURE - Clustering Using REpresentatives

EU – European Union

GDP - Gross Domestic Product

IEEE - Institute of Electrical and Electronics Engineers

OECD - Organization for Economic Cooperation and Development

PDF - Portable Document Format

PRISMA - Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Q1 - First Quartile Journal

Q2 - Second Quartile Journal

Q3 - Third Quartile Journal

Q4 - Forth Quartile Journal

RF - Random Forest

RQ: Research Question

ROC - Receiver Operating Characteristic Curve

SNA – Social Network Analysis

SVM - Support vector machine

TCE-CE – Tribunal de Contas do Estado do Ceará

USA – United States of America

VOSviewer - Visualization Of Similarities Viewer

W.O.S - Web Of Science

1. INTRODUCTION

Public Procurement is the process by which public administration entities purchase goods and services they need (e.g., food, consulting services, office supplies, etc.) and execute large infrastructure construction works (e.g., bridges, hospitals, etc.). Without surprise, Public Procurement account for a significant share of governments budgets. Indeed, among the OECD countries, public procurement represents between 12% to 20% of GDP (Diesing et al., 2019), in Jamaica it reached 30% of GDP in 2015 (Dawar & Oh, 2017), while in Angola and Eritrea it accounted, respectively, for 26% and 33% of GDP in 2018 (National Bureau of Statistics, 2019). These numbers highlight the substantial role of public procurement in the economy and emphasize the need for effective governance and oversight in a process that is often deemed too complex and often bureaucratic.

Although public procurement should be fair, transparent, and efficient, fraud and corruption constitute a severe problem for many countries, leading to significant monetary losses and undermining public trust in government institutions. For instance, in many developing countries, public procurement fraud is prevalent, resulting in substantial financial losses and eroding public trust in government institutions (Oberoi, 2014). Cases of bid rigging, bribery, and embezzlement undermine the integrity of the procurement process and divert resources away from vital public services (Hall, 2012). These illicit activities not only harm the economy but also hinder social development and exacerbate inequalities (Gasparikova, 2012). To make things worse, there is a lack of understanding of the procurement market dynamics and the potential traces left by irregular/fraudulent activities.

Auditing agencies, like the Court of Auditors, face many challenges in the auditing of public procurement contracts, from the selection of public agencies to limited access to analytical tools, but also a lack of appropriate analytical methods to narrow down the scope of the audit. In this thesis, we seek to mitigate some of the challenges auditors face when detecting irregularities and/or preventing fraud in public procurement contracts. With the support of the Tribunal de Contas do Estado do Ceará (TCE-CE), we use a rich empirical dataset to explore venues to extract valuable insights into fraud patterns and risk indicators in public procurement activity, enhancing our understanding of fraud's prevalence, characteristics, and underlying factors.

The TCE-CE faces many challenges in obtaining more concrete and fact-based guidance regarding the jurisdictional entities to be audited. These challenges are particularly prominent regarding the entities participating in public tenders and suspected of committing crimes against the treasury. In addition, the mandatory annual submission of the accounting information report does not contain sufficient elements to characterize possible irregularities in tenders. Also, they are not designed to provide a more straightforward approach to what should be audited. Furthermore, although the court's ability to find irregular activities and apply penalties, no track record of such activities is digitally available in the court records. As a result, exploratory and unsupervised data mining methods become necessary to identify potential fraudsters and narrow down the targets of field audits.

Examples of the consequences of these challenges are evident in the *Mão Dupla* (Carson & Prado, 2012) and *Fraternidade* (Miasato, 2017; O Povo, 2017) Operations in Ceará, which exposed cases of corruption and collusion in public procurement. The *Mão Dupla* Operation unveiled irregular bidding processes for public works, revealing schemes involving bribery, contract manipulation, and cartel formation among firms. Similarly, the *Fraternidade* Operation uncovered a corruption scheme related to the hiring of companies for school meals provision, involving collusive behavior, price-fixing, and bid manipulation. These illicit activities resulted in financial losses for the state in both cases and also compromised the quality of food for students in the second instance.

It is noteworthy that both operations started on individual allegations and without leveraging the data-driven intelligence derived from bidding data. As such, irregularities such as the ones underlying both those operations can be more prevalent, and relying solely on allegations to uncover them represents a reactive approach, which is ineffective and inefficient. Therefore, data-driven analytics is expected to play an increasingly central role in detecting and validating such irregularities in a timely manner, allowing for a proactive approach that would contribute to a more effective fight against corruption, ensuring transparency and integrity in the bidding processes.

To address these challenges, we conducted two studies focused on a historical dataset provided by the TCE-CE to develop empirical analyses. The first study examines a five-year period from 2015 to 2019, while the second one extends the analysis over an eight-year timeframe from 2013 to 2020, contemplating two different governments of the 184 municipalities. Currently, the choice of agents to render public administration accountability is primarily based on a systematic selection of the municipalities. It relies on criteria such as the materiality or representativeness of budgetary, financial, and equity values available to managers, as well as the volume of goods, public assets, and values under the management of administrators or those responsible for the jurisdictional unit. When it comes to the external audit, the emphasis is placed on the most economically relevant cases, regardless of a fraud risk assessment. In many cases, more targeted audits are usually based on reports and allegations of suspected fraud submitted by the public. That said, it is evident that numerous tenders may never undergo monitoring due to limitations in the Court's capacity for audits and even procedural instruction. Thus, many bidding processes remain unaudited or lack a comprehensive fraud risk assessment, and those chosen for inspection may not necessarily have the highest risk of fraud.

Despite the aforementioned challenges in fraud detection in public procurement tenders, there is a limited amount of literature on data-driven studies investigating this area (Smith, 2022). However, institutions and governments have made significant efforts to compile and provide access to relevant data. Initiatives like Open Data portals and the dedication of audit institutions have contributed to this cause (Dener et al., 2021). These efforts have led to the increasing availability of government procurement data through Open Data platforms, emphasizing the importance of transparency and accountability in public procurement (Duguay et al., 2022). Audit institutions have also utilized technology and data analytics to enhance their auditing processes and compile comprehensive databases of procurement-related information, shedding light on fraud and corruption issues (Aquino et al, 2021; De Catalunya, 2022). These efforts over the past 15 years have significantly improved data availability, creating opportunities for further research and analysis in fraud detection in public procurement (Sloane et al., 2021).

While advancements have been made, the lack of labeled data remains challenging for researchers. To overcome this limitation, alternative approaches are being explored. One such approach is network analysis, which utilizes the inherent structure of the network to uncover fraud-related patterns and anomalies. By applying unsupervised learning techniques within the network analysis framework, researchers can identify clusters, detect communities, and reveal irregularities indicative of fraudulent behavior. Furthermore, collaboration and data sharing among researchers, audit institutions, and government agencies further strengthen fraud detection efforts. This collaborative approach facilitates a deeper understanding of the complex network dynamics in public procurement and enables the identification of suspicious patterns and entities that require further investigation. Although the lack of labeled data remains a challenge, integrating network analysis with collaborative efforts provides a promising pathway for enhancing fraud detection and prevention, ultimately fostering transparency and accountability in public procurement processes.

From the perspective of the regulatory body, a poor understanding of fraud in public procurement leads to inefficiency in identifying and combating fraudulent activities, compromising control and oversight mechanisms, resulting in ineffective detection and punishment of illegal practices. Insufficient knowledge of fraudsters' tactics, collusion dynamics, and corrupt behavior indicators hinders the identification of suspicious activities and thorough investigations. The absence of expertise and specialized training among professionals exacerbates the inefficiency in detecting and preventing criminal practices. Consequently, fraudulent activities persist, undermining competition, inflating contract prices, and compromising service quality. A comprehensive understanding of these practices enables the adoption of effective prevention, detection, and punishment strategies. This fosters transparency, competitiveness, and integrity in public procurement, ensuring responsible resource utilization and enhancing public trust in oversight institutions.

1.1. Major crimes against the public administration treasury

Despite these being ancient practices, their emergence on a global scale is relatively recent (Martins, 2008), and the terminology associated with different types of crimes is often confused or misused. According to Spink (2019), it is essential to view the issue as a process resulting from coordinated steps between agents over time. Therefore, its study must cover the context in which fraudulent processes occurs and how corruptors develop and maintain the schemes.

Corruption in public contracting is the most recurring crime, given its significant impact on government budget. It manifests in two variants, passive and active corruption. The first occurs when someone offers or promises an undue advantage to public officials to induce them to perform, omit, or delay some acts (Andreucci, 2018). On the other hand, passive corruption is practiced by public officials who request or receive, for themselves or someone of their interest, either directly or indirectly, any undue advantage due to their position (Andreucci, 2018). This type of corruption relies extremely on relational behavior requiring interactions between two or more entities on different sides. It often involves the exchange of favors or, more commonly, through bribery.

Defining all situations classifiable as corruption presents a great challenge because the phenomenon is not limited to illegal acts. Furthermore, it encompasses different fields of knowledge, such as law, political science, and administration, among others, leading to semantic complexity and different meanings and approaches (Dewi, Mahmudi, & Maulana, 2021). The term “corruption” is commonly and erroneously linked to crimes involving theft of public funds, even when the administration is not involved (Prado & Cornelius, 2020). Nevertheless, corruption is just one of the types of these crimes, and others possess distinct characteristics (Bouman, 2021). In fact, we noticed that the term “corruption” is sometimes used as a synonym to refer to many crimes in public administration in the scientific articles we collected.

In contrast, the most prevalent crime identified against the public administration, without the involvement of a public official, was collusion. It represents an irregularity wherein companies engage in coordinated behavior regarding price, quantity, quality, or geographic presence, with the aim of inflating market values (Toth, Fazekas, Czibik, & Toth, 2014). It involves an arrangement among a group of participants intended to limit competition in a given process (Porter & Zona, 1993). In the active form, remuneration can be obtained by participating in the winning coalition, while in the passive form, subcontracting is used. The resulting market indicates that collusion outcomes in a monopolistic or market imitation structure, both non-competitive (Toth, Fazekas, Czibik, & Toth, 2014).

Although these typification of crimes against the public administration may occur together, or even one of them be contained in the other in procurement processes, we observed that corruption and collusion were the two most prevalent offenses leading to the diversion of public resources. Corruption was perceived as a relational activity between corrupting agents and corrupted entities, always involving the participation of a public body. On the other hand, collusion was identified as an agreement between third parties to defraud public administration operations without its participation.

1.2. Overall Goals and Objectives

One of the main objectives of this research was to provide a theoretical, methodological, and empirical basis to characterize the behavior of public administration bodies, companies, and service providers in line with potential fraudulent actions in procurement markets of the municipalities of the Ceará State. Within the realm of fraud detection in public procurement, two consecutive literature reviews were conducted to shed light on the most current techniques and methodologies employed in this domain. One review aimed to identify prevailing approaches specifically used for detecting fraud in public procurement, with a focus on empirical data from procurement processes. This comprehensive study revealed two prominent methodologies: machine learning and network science. However, it became evident during the review that machine learning techniques heavily relied on small datasets of labeled data that often focused on a small sector or specific contexts. Such data was, however, not available at the TCE-CE, making its application unviable. On the other hand, network science tools emerged as a promising avenue for exploring fraud detection using unlabeled data in an exploratory manner. As a result, network science was selected as the preferred methodology for further examination.

The following literature review shifted its exclusive focus towards exploring network science techniques, taking into consideration the nature of the available unlabeled data. Through this review, several relevant tools were identified, including community detection, modularity measure, degree centrality, and network projection. These techniques were deemed most suitable for achieving the study's objectives, providing valuable insights into the disclosure of fraud in public procurement. The broader review article addressed the research question: "*How have data-driven methods contributed to detecting, characterizing, and predicting fraud, collusion, and corruption in public procurement activities?*" This question aimed to investigate the broader landscape of data-driven approaches employed in fraud detection within the context of public procurement. In contrast, the more focused review article centered around the research question: "*How can network science help in the disclosure of fraud in public procurement?*" This question delved specifically into the application of network science techniques to uncover and expose instances of fraud in public procurement, considering the availability of unlabeled data. These two research questions provided a comprehensive framework for the literature reviews, guiding the exploration of relevant methodologies and tools while addressing the unique challenges associated with fraud detection in public procurement. As a result, the work focused on unsupervised methods that could bring insights on the properties of the procurement market and of the multiple actors involved.

Understanding and characterizing public procurement network structure is crucial for effective fraud detection. Network analysis tools provide insights into the complex connections among stakeholders, uncovering hidden relationships and identifying influential actors. By leveraging these tools, stakeholders can proactively detect fraud, enhance transparency, and strengthen accountability in public procurement processes. Network analysis plays a vital role in detecting fraud by revealing patterns of collusion, corruption, and fraudulent activities. By examining network metrics and algorithms, such as centrality measures and community detection, key players and groups can be identified. This understanding of the network structure enables stakeholders to implement targeted strategies, mitigate risks, and ensure the integrity of public procurement. The application of network analysis techniques offers a powerful toolset to combat fraud and promote responsible resource utilization in public procurement.

1.3.Types of agents and relationships in public procurement networks

Overall, public procurement markets can be represented as tripartite networks comprising issuers (public institutions or buyers), bidders (companies or sellers participating in a bid competition), and contract or tender types (products and services). Additionally, bipartite networks between issuers and bidders, tender types and bidders as well as issuers and tender types provide, in some instances, a sufficient projection for analytical studies. A step further consists in projecting these networks into co-activity patterns or co-bidding networks, connecting firms that co-bid significantly together in the same tender, competing to win the bid. These empirical networks exhibit characteristics that distinguish them from random networks since they have heterogeneous degree distributions and substantial local correlations (Fazekas, Skuhrovec, & Wachs, 2017). As such, network analysis of such structures is particularly suited for investigating fraud, especially in relational networks of procurement processes (Wachs & Kertész, 2019). Research on organized crime has shown that criminal actions tend to leave intricate fingerprints that networks are able to understand and capture (Wachs & Kertész, 2019).

The literature review revealed that most studies on public procurement fraud detection can be divided into two main groups: (1) research that explores the bipartite relationships between public tenders and firms (Fazekas & Tóth, 2016.2; Wachs & Kertész, 2019); and (2) studies that explore co-bidding relationships among firms derived from tender projections within the bipartite network (Toth, Fazekas, Czibik, & Toth, 2014; Reeves-Latour & Morselli, 2017; Morselli & Ouellet, 2018; Wachs, Fazekas, & Kertész, 2021). Both approaches have their merits, and each is suitable for identifying different mechanisms underlying the manipulation of the procurement process. Bipartite relationships are well-suited to identifying fraud arising from bribery and influence ties, while inter-firm relationships are better suited to identifying cartels and collusion. However, the network analysis used to study the relationship between companies or between tenders and winners is a relatively new undertaking.

Understanding and characterizing the network structure of relationships in public procurement is crucial for effective fraud detection. Network analysis tools provide insights into complex connections among stakeholders, uncovering hidden relationships and identifying influential actors. By leveraging these tools, stakeholders can proactively detect fraud, enhance transparency, and strengthen accountability in public procurement processes. Network analysis plays a vital role in detecting fraud by revealing patterns of collusion, corruption, and fraudulent activities. By examining network metrics and algorithms, such as centrality measures and community detection, key players and groups can be identified. This understanding of the network structure enables stakeholders to implement targeted strategies, mitigate risks, and ensure the integrity of public procurement. The application of network analysis techniques, including network projection, offers a powerful toolset to combat fraud and promote responsible resource utilization in public procurement. Figure 1 showcases the different network configurations, including tripartite, bipartite networks and monopartite networks derived from the projection of a bipartite net, providing visual representation of the relationships between procurement entities.

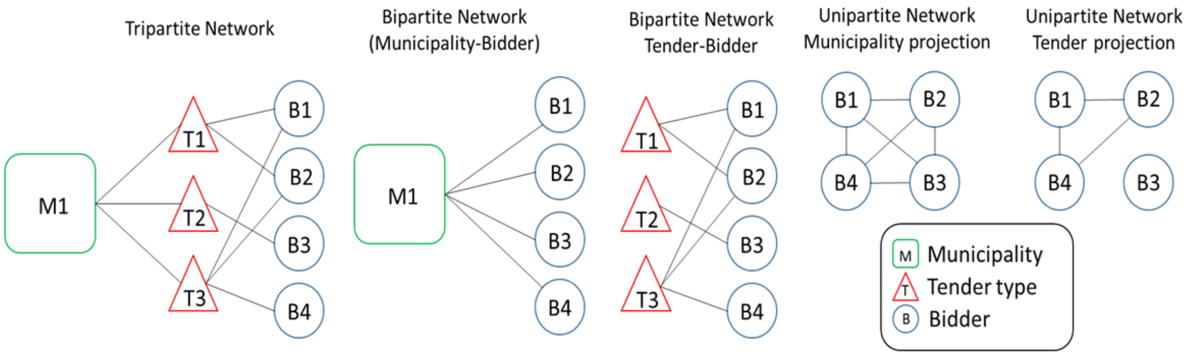

Figure 1: Scheme of different network types associated with public procurement data. Here, the underlying bid sequence for Municipality 1 is made by bidders B1, B2, B3, B1, B2, and B4, and these bids are associated with contract or tender types T1, T1, T2, T4, T4, and T4, respectively. The tripartite network connects a bidder to one or more tender types, and each tender type, in turn, can be connected to one or more municipalities (in this case, we exemplify with only one municipality). Tripartite networks provide the most comprehensive representation of the data and preserve all relevant information for network construction. The tripartite network can be projected into three different bipartite networks, each generated by omitting one type of node from the tripartite network; here, we show two bipartite networks, one connecting bidders and municipalities, and the other connecting tender types and bidders. Finally, any bipartite network can be projected into two different types of monopartite networks containing nodes of only one type. Here, we display a projection for each bipartite network, the projections in which the remaining nodes are bidders connected by shared municipalities or shared contract types.

1.4. Uncovering hidden connections through co-bidding networks

Even with high frequency in the public sector and significant economic impact, fraudulent interactions are still challenging to detect since individuals and organizations involved in these irregularities strive to keep them obscure (Hawken & Munck, 2009). Examples of convicted corruptors are challenging to access and generally scarce due to the lack of adequate methodology to capture them and the restriction of information. As a result, most frauds remain undetected or even go unpunished since various government entities may be captured by vested interests, defrauding contracts, or turning a blind eye to wrongdoing (Davis, 2019). The level of corruption in a government is closely linked to the effectiveness of its administration (Speck & Ferreira, 2012). Society perceives that the more persuasive the governance, the more a government is associated with low rates of corruption, often linked to stricter enforcement of the law (Faría, Montesinos-Yufa, Morales, Aviles B, & Brito-Bigott, 2013).

Based on this premise, it is essential to explore analytical frameworks to help governments identify agents more susceptible to irregular activities. Therefore, we conducted an empirical study titled "*Characterization of the firm-firm public procurement co-bidding network from the State of Ceará (Brazil) municipalities*" to uncover hidden relationships among firms participating in same tenders to investigate potential fraudulent behavior. The article uses standard network science methods to study the co-bidding relationships between competitors participating in public tenders undertaken by the 184 municipalities of the State of Ceará (Brazil) between 2015 and 2019. One objective of this research was to identify competitors' communities with standard and interconnected practices in procurement processes characterized by their geographic and activity routines. Also, examine whether the features of the communities allowed for spotlighting the organizations most exposed to market manipulation and uncommon actions. It would reinforce the potential application of network analysis in politics to reveal the complex nature of relationships between market agents in a scenario of short data. Another goal was to characterize the largest communities by the diversity of tenders and present their distribution within each community, region, and type of contract using the Simpson Diversity Index. This analysis aimed to compare contract diversity regarding the regional extent and the type of object contracted.

In addition, we also sought to investigate the degree of bidding manipulation by companies and issuers, analyzing the propensity of each community to participate in contracts as single bidders. The idea was to investigate the average number of times, per community, that a firm participates as a single bidder in different tenders since a high level of single bidders can be a strong indicator of spurious relationships between issuers and winners. Thus, we could evaluate the different levels of participation as single bidders to determine whether companies participate in tenders fairly with low levels of single bids, whether less competitive markets prevail at high levels, or whether it indicates fraud in the bidding process.

Finally, we were interested in measuring competitiveness in tenders by calculating the average number of bidders per bid to assess the existence of coordination between companies in a community with joint appeals. We took the average number of bidders per bid in each community, normalized by their size to estimate the tendency of the community to participate jointly in the same tenders predominantly.

Having identified a way to uncover hidden relationships between competitors in public tenders, our next objective was to assess the connections between tenders and winners by recognizing the second gap identified in the literature review, which pertains to the limited use of objective indicators to estimate corruption risk in tenders. The proposal was to evaluate interactions between sellers and buyers in bipartite tender networks and examine the integrity or corrupt nature of the relationships between them in two different governments.

1.5. Measuring the risk of fraud between tenders and winners in public procurements in different governments

The risks of fraudulent activities in public procurement can vary across different periods of governance depending on the actions and priorities of the incoming government and their interaction with other market players. Transitions of power can potentially impact its level, bringing reforms in policies, regulations, oversight mechanisms, and the overall institutional framework. It may enhance transparency, accountability, and integrity in public procurement by strengthening anti-corruption legislation, improving procurement regulations, establishing independent oversight bodies, implementing electronic procurement systems, promoting public participation and access to information, and consequently facilitating the detection of irregularities. It is important to recognize that the issue lies not only within the government itself but also in the preferential relationships that can develop between various actors involved in public procurement. By addressing these relationships and implementing measures that promote fairness, competition, and a level playing field for all participants, the detection of irregularities can be further facilitated.

One of the challenges in addressing fraud risks in public procurement lies in accurately measuring and assessing these risks. While subjective surveys have traditionally been used to gauge corruption in the public sector, objective data intrinsic to the tenders has emerged as a more tangible approach in recent studies (Nikolova, 2019). Surprisingly, there has been a scarcity of researchers focusing on the development of fraud risk indicators specifically tailored for public tenders (Amundsen, 2019). Nonetheless, this approach holds significant potential utility for audit courts, as it can provide clearer guidance in selecting jurisdictions to be audited. Thus, by employing such indicators, control bodies could conduct more precise and targeted inspection while implementing mechanisms to mitigate the financial losses incurred by governments.

Mihaly Fazekas, along with other authors and co-authors, has made significant contributions to the field of corruption research, particularly in relation to the development of the Corruption Risk Index (CRI) for public procurement (Fazekas, Toth, & King, 2016.1; Charron, Dahlström, Fazekas, & Lapuente, 2017; Dávid-Barrett & Fazekas, 2020; Fazekas & Kocsis, 2020; Czibik, Fazekas, Sanchez, 2021). The CRI is an indicator designed to assess the vulnerability of public procurement processes to corruption. It is developed through the analysis of large-scale datasets and sophisticated statistical methods, such as data preprocessing, indicator construction, weighting, benchmarking and ranking, and validation. The CRI serves as a robust tool for systematically assessing corruption risks across different countries and contexts. Fazekas's works (Fazekas, Toth & King, 2016.1; Fazekas & Toth, 2016.2; Fazekas & Kocsis, 2020) has not only focused on the conceptualization and construction of the CRI but has also explored its practical applications, identifying high-risk areas within public procurement systems and guiding targeted anti-corruption interventions. By combining the CRI with other qualitative and contextual information, a comprehensive understanding of corruption dynamics in public procurement can be achieved. The Corruption Risk Index serves as an essential tool in ongoing efforts to combat corruption and promote good governance in public procurement.

Building upon the need for reliable corruption indicators, our study sought to address this gap within the Brazilian context, specifically focusing on public tenders. In that sense, we aimed to introduce a high-level corruption risk indicator to the Brazilian landscape. Leveraging public procurement data provided by the TCE-CE, our study aimed to identify potential red flags within the process based on intrinsic information from tenders at the micro level, linking them to restricted competition and recurring awarding of contracts to the same firm on areas of concern that warrant further investigation. This examination of the data culminated in the article titled *"Fraud Risk Assessment in Public Procurement Activities - An Investigation in Two Different Governments of Ceará's Municipalities."*

The study examined the bipartite relationship in public procurement within the municipalities of the State of Ceará, utilizing empirical data spanning from 2013 to 2020 and encompassing two distinct government periods. The primary objective was to develop a corruption risk indicator, known as the Corruption Risk Index (CRI), which was based on intrinsic parameters of the bidding processes and adapted from Fazekas's previous work to suit the local context. This indicator incorporated nine red flags that were derived from audit reports and interviews conducted with auditors.

To determine the significance of each red flag, regression analysis was employed to estimate their respective weights. Furthermore, quartile categorization was utilized to group municipal tenders and winning firms based on their levels of corruption risk. By employing this approach, the study was able to organize and classify firms and municipalities in both governments into four corruption risk bands.

The findings of the study offer valuable insights into the interactions between issuers and bidders, the corruption risk magnitudes, and its underlying reasons. Ultimately, the methodology employed proved to be effective in identifying potential grouping of regions and contract types where corrupt practices may occur.

References

Amundsen, I. (2019). Extractive and power-preserving political corruption. In Political Corruption in Africa (pp. 1-28). Edward Elgar Publishing.

Andreucci, D. (2018). Populism, hegemony, and the politics of natural resource extraction in Evo Morales's Bolivia. *Antipode, 50*(4), pp. 825-845.

Aquino, A. C. B. D., Lino, A. F., & Azevedo, R. R. D. (2021). The embeddedness of digital infrastructures for data collection by the Courts of Accounts. Revista Contabilidade & Finanças, 33, 46-62.

Bouman, J. (2021). Preventing Corruption of Civil Servants in France, Romania and the Netherlands: A Comparative Study of Law and Policy. *Wolters Kluwer*, pp. 1-307.

Carson, L., & Prado, M. M. (2014). Brazilian anti-corruption legislation and its enforcement: Potential lessons for institutional design (p. 18). IRIBA Working Paper: 09. Manchester. UK, 2014.

Charron, N., Dahlström, C., Fazekas, M., & Lapuente, V. (2017). Careers, connections, and corruption risks: Investigating the impact of bureaucratic meritocracy on public procurement processes. The Journal of Politics, 79(1), 89-104.

Czibik, Á., Fazeka, s. M., Sanchez, A., & J., W. (2021). Networked Corruption Risks in European Defense Procurement. *Corruption Networks: Concepts and Applications*, 67-87.

Dávid-Barrett, E., & Fazekas, M. (2020). Grand corruption and government change: an analysis of partisan favoritism in public procurement. *European Journal on Criminal Policy and Research, 26*(4), 411-430.

Davis, K. E. (2019). Between Impunity and Imperialism: The Regulation of Transnational Bribery. *Oxford University Press*.

Dawar, K. and Oh, S.C. (2017), “The role of public procurement policy in driving industrial development, no. 8”, Inclusive and Sustainable Industrial Development Working Paper Series, Vienna.

De Catalunya, B. G. O. A. On Public Data Availability, Interoperability and Reusability.

Dener, C., Nii-Aponsah, H., Ghunney, L. E., & Johns, K. D. (2021). GovTech maturity index: The state of public sector digital transformation. World Bank Publications.

Dewi, H. R., Mahmudi, M., & Maulana, R. (2021). An analysis on fraud tendency of village government officials. *Jurnal Akuntansi dan Auditing Indonesia*, pp. 33-44.

Diesing, L., Unit, P. P., Magina, O. P., & Head, P. P. U. (2019). Innovation in the Evaluation of Public Procurement Systems. In *Joint public procurement and innovation: lessons accross borders* (pp. 349-372). Bruylant.

Duguay, R., Rauter, T., & Samuels, D. (2022). The impact of open data on public procurement. Available at SSRN 3483868.

Faría, H. J., Montesinos-Yufa, H. M., Morales, D. R., Aviles B, C. G., & Brito-Bigott, O. (2013). Does corruption cause encumbered business regulations? An IV approach. *Applied Economics, 45*(1), pp. 65-83.

Fazekas, M., & Kocsis, G. (2020). Uncovering High-Level Corruption: Cross-National Objective Corruption Risk Indicators Using Public Procurement Data. *British Journal of Political Science, 50*(1), pp. 155-164.

Fazekas, M., & Tóth, I. J. (2016.2). From corruption to state capture: A new analytical framework with empirical applications from Hungary. *Political Research Quarterly, 69 (2), 320–334.*

Fazekas, M., Skuhrovec, J., & Wachs, J. (2017). Corruption, government turnover, and public contracting market structure–insights using network analysis and objective corruption proxies. *Budapest: Government Transparency Institute*.

Fazekas, M., Toth, I. J., & King, L. P. (2016.1). An Objective Corruption Risk Index Using Public Procurement Data. European Journal on Criminal Policy and Research, 22(3), pp. 369-397.

Gasparikova, S. (2012). The Czech Republic's experience with competition policy and law. In Evolution of competition laws and their enforcement (pp. 26-42). Routledge.

Hall, D. (2012). Corruption and public services. Public Services International Research Unit (PSIRU), 1(1), 1-50.

Hawken, A., & Munck, G. L. (2009). Measuring corruption: A critical assessment and a proposal. *Perspectives on Corruption and Human Development, Macmillan*.

Martins, D. G. (2008). O estado da arte da capacidade institucional: uma revisão sistemática da literatura em língua portuguesa. *EBAPE.BR, 19*(1), pp. 165-189.

Morselli, C., & Ouellet, M. (2018). Network similarity and collusion. *Social Networks, 55*, pp. 21-30.

Miasato, C. S. (2017). Licitação: Da Perfeição Idealizada à Parcial Ineficácia do Sistema. Faculdade de Ensino Superior e Formação Integral – FAEF

National Bureau of Statistics (NBS) (2019), "National Accounts Statistics, Popular Version 2018", Dar es salaam, Tanzania

Nikolova, E. V. (2019). Using Data Analysis and Instrumental Variables to Study the Drivers of Corruption. SAGE Research Methods Cases.

O Povo (2017). Operação Fraternidade é deflagrada pela Polícia Federal no Ceará. Jornal O Povo. 13 de setembro de 2017. Consultado em 20 de novembro de 2022. http://blog.opovo.com.br/politica/operacao-fraternidade-deflagrada-pf/

Oberoi, R. (2014). Mapping the matrix of corruption: Tracking the empirical evidences and tailoring responses. Journal of Asian and African Studies, 49(2), 187-214.

Porter, R. H., & Zona, J. D. (1993). Detection of bid rigging in procurement auctions. *Journal of political economy, 101*(3), pp. 518-538.

Prado, M. M., & Cornelius, E. (2020). Institutional Multiplicity and the Fight Against Corruption: A Research Agenda for the Brazilian Accountability Network. *Revista Direito GV*, p. 16.

Reeves-Latour, M., & Morselli, C. (2017). Bid-rigging networks and state-corporate crime in the construction industry. *Social Networks, 51*, pp. 158-170.

Sloane, M., Chowdhury, R., Havens, J. C., Lazovich, T., & Rincon Alba, L. (2021). AI and Procurement-A Primer.

Smith, C. D. (2022). A Study of Public Procurement Fraud: Framing the Study. In White-Collar Crime and the Public Sector: An Interdisciplinary Approach to Public Procurement Fraud (pp. 47-64). Cham: Springer International Publishing.

Speck, B., & Ferreira, V. (2012). Sistemas de Integridade nos Estados Brasileiros. *Instituto Ethos de Empresas e Responsabilidade Social, 12*, pp. 30-82.

Spink, J. W. (2019). International Public and Private Response. *Food Fraud Prevention: Introduction, Implementation, and Management*, pp. 443-478.

Toth, B., Fazekas, M., Czibik, A., & Toth, I. (2014). Toolkit for detecting collusive bidding in public procurement with examples from Hungary. *Corruption Research Center Budapest (CRC-WP/2014:02)*.

Wachs, J., & Kertész, J. (2019). A network approach to cartel detection in public auction markets. *Sci Rep, 9*. doi:10818

Wachs, J., Fazekas, M., & Kertész, J. (2021). Corruption Risk in Contracting Markets: A Network Science Perspective. *International Journal of Data Science and Analytics*.

2. LITERATURE REVIEW - FRAUD, CORRUPTION, AND COLLUSION IN PUBLIC PROCUREMENT, A SYSTEMATIC LITERATURE REVIEW ON DATA-DRIVEN METHODS

Fraud, corruption, and collusion are the most common types of crime in public procurement processes; they produce significant monetary losses, inefficiency, and misuse of the public treasury. However, empirical research in this area to detect these crimes is still insufficient, requiring a literature review to identify studies that have been developed and the techniques that are being applied in detecting fraud in public tenders.

Here we exclusively present the second review article (Lyra et al., 2022) because it was more comprehensive and embraces more topics than just those referring to network science. However, in appendix C, at the end of this document, we present the other review article for consultation.

2.1. Introduction

Public procurement is the process by which the public sector purchases services, goods, and construction works from third-party companies. It represents a substantial share of the government's public expenditure (OECD, 2021), amounting to 12% of the global GDP (Bosio, Djankov, Glaeser, & Shleifer, 2022). According to the Organization for Economic Cooperation and Development (OECD), in 2017, public procurement represented 29.1% of the general government expenditures among OECD countries.

The reliability of these processes is key for an effective act of governance (Silva Filho, 2017). However, while governments seek the best cost-benefit regarding the price and quality of their procured services (Costa, Machado, & Martins, 2020), public procurement is also well known to be quite permeable to irregularities.

Indeed, Transparency International Organization estimates that losses in public contracts due to corruption account for 20% to 25% of the value of the contract and up to 50% in some cases (Volosin, 2015). Among public administration, public procurement is the process that contains the highest risks of illegality that can arise virtually at any stage of the process of acquiring goods, services, or construction works (Rustiarini, Sutrisno, Nurkholis, & Andayani, 2019). Crimes in this context are elaborated and refined, involving schemes such as abuse of power, corruption, peculate, bribery, favoritism, stunt, nepotism, and misappropriation (Padhi & Mohapatra, 2011). In that sense, these crimes are difficult to detect and measure (Mufutau &

Mojisola, 2016), particularly after awarded procurement contracts (Whiteman, 2019). As such, well-defined legislation is not enough to circumvent criminality in public procurement, and mechanisms to prevent it are frequently inadequate (Wensink & Vet, 2013).

In that sense, it is crucial to understand the phenomena and establish which methods and practices have been explored by the academic community to help identify actors and objects linked to fraudulent activities in the public procurement process. In particular, we focus on actors who act through the practice of Corruption - the relational activity between corrupting agents and corrupted entities, always with the participation of the public body - (Paganetto & Scandizzo, 2015); Collusion - an agreement between third parties to defraud operations of the public administration, without its participation - (Mvelase, 2015); and Fraud - crimes of forgery of documents, fraudulent proposals, use of shell companies, the introduction of false assessment of competitor, identity fraud, and data theft - (Mlondo, 2013) to impair the public administration from acting efficiently.

Here, we use the Preferred Reporting Items for Systematic Reviews and Meta-Analyzes (PRISMA) methodology to identify and review the use of data-driven methods in academic literature to study fraud, collusion, and corruption in public procurement processes. We look to identify the problems and contexts in which mainly network science and machine learning methods have been applied. Overall, we identify in our work 48 relevant manuscripts and discuss the methods used, the problems they explored, and their framing in the academic literature.

Recent works have leveraged the increasing availability of public procurement data, resulting from adopting electronic procurement (e-procurement) bidding processes/modalities. These have been introduced to increase transparency and competitiveness and decrease the bureaucracy associated with the procurement processes (Borras, Cousins, Liu, McKay, & Milhorance De Castro, 2015) (Weigel & Ruecker, 2017). E-procurement is, thus, a modality that offers a speed-up process at reduced costs (Motta, 2003). Although e-procurement has been around for a long time and has been widely adopted, there is limited research on its impact on reducing/mitigating crime in the public sector (Kartika, Kebijakan, Barang, & Pemerintah, 2020). Therefore, we extend our analysis to procurement activity in the private sector. This dimension has seen more research and explored more advanced techniques to detect and mitigate crimes in the procurement process (Wu Chebili, et al., 2022). It is the case, in our opinion, that the public sector can learn from the best practices of the private sector, in particular, given the similarities in the procurement processes in both cases.

This manuscript is structured as follows: Section 2 announces the method used for identifying relevant research and conducting a bibliometric analysis. In Section 3, we raise the results, introducing the keywords, citations, occurrences, methods, and research objectives. In Section 4, we discuss the outcomes and show the research limitations. Finally, Section 5 concludes the research, identifies some gaps, and gives directions for further work.

2.2. Methodology

We performed a systematic literature review using the PRISMA methodology. PRISMA is an auditable methodology based on a checklist of items often depicted in a flow diagram. The checklist includes 27 topics distributed throughout seven sections: title, abstract, introduction, methods, results, discussion, and other information. PRISMA enhances the meta-analysis of systematic review reports and promotes their reproducibility, Figure 2.

The PRISMA 2020 template for systematic reviews can be summarized by a three-step flowchart that includes identifying, screening, and selecting manuscripts.

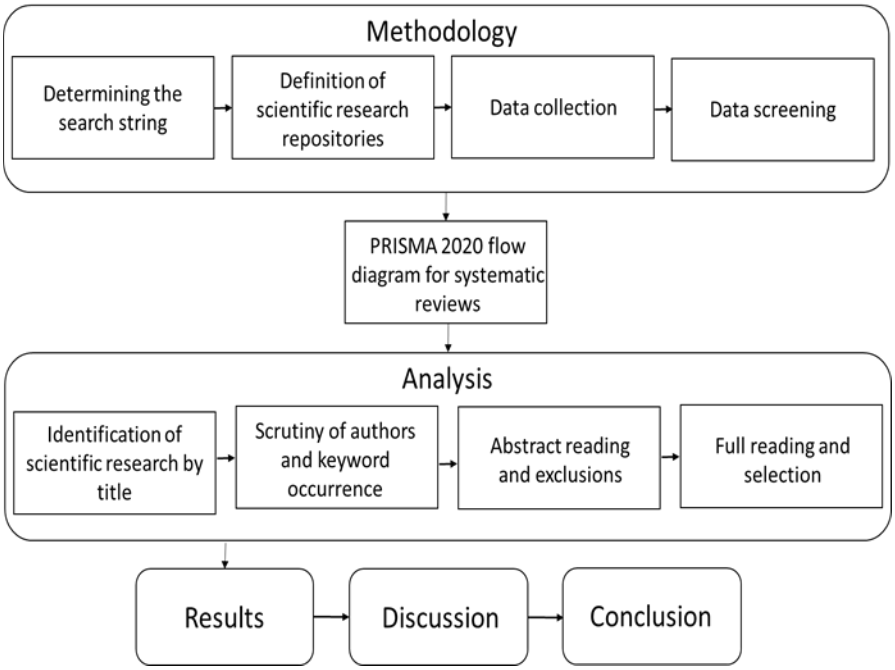

Figure 2: Research Workflow

Here, we proceed with the PRISMA methodology to answer the following question: “How have data-driven methods contributed to detecting, characterizing, and predicting fraud, collusion, and corruption in public procurement activities?”

We begin by identifying relevant scientific manuscripts. In that sense, we followed a systematic process of iterative query in the research repository Scopus to identify publications with the scope defined by the search string. The Scopus repository was chosen due to its comprehensive coverage of scientific venues, more immediate indexing methods, and the availability of more articles than other repositories.

a) Keyword tracking strategy and Bibliometric analysis

Bibliometric analysis is part of an interdisciplinary science concerning a quantitative study. It uses mathematical and statistical methods to capture specific knowledge and expertise from various authors (Merigó, 2016). It is a methodology usually used to identify the development of a particular scientific field (Merigó, Mas-Tur, Roig-Tierno, & Ribeiro-Soriano, 2015) (Železnik, Vošner, & Kokol, 2017). Based on this, we conducted an iterative tracking process with an assembled search string, shown below, encompassing words and phrases logically connected by Boolean operators in the titles, abstracts, and keywords. We chose the words aiming to answer the research question, serve the research objectives, and whose frequency, in most various articles, on the topic was high in the abstracts, keywords, and texts.

Search String: (fraud* OR corrupt* OR bribe OR rigg* OR cartel* OR collusi*) AND (predict* OR detect* OR SNA OR network* OR statistic*) AND (“public sector“ OR “public administration“ OR procurement OR bidd* OR auction OR tender OR government)

To sort, extract, organize, and store metadata (publications, citations, authors, and references), we used the Software Mendeley, which provides a free reference manager that assists in academic work, manage electronic files (PDF format), helps in the normalization of citations, and generates references automatically. To assemble the network that describes the relations between authors and topics, we used the open-source software VOSviewer (Van_Eck & Waltman, 2014). It is an engine for constructing and visualizing bibliometric networks, developed by Nees Jan van Eck and Ludo Waltman at the Center for Science and Technology Studies at Leiden University. These networks are built based on co-citation, bibliographic coupling, or co-authorship ratios. VOSviewer also offers text mining capabilities that allow building and viewing simultaneous occurrence networks of important terms extracted from the scientific literature. Furthermore, we investigate the citation’s framework to capture the conjectural basis of the research (Ding & Yang, 2020).

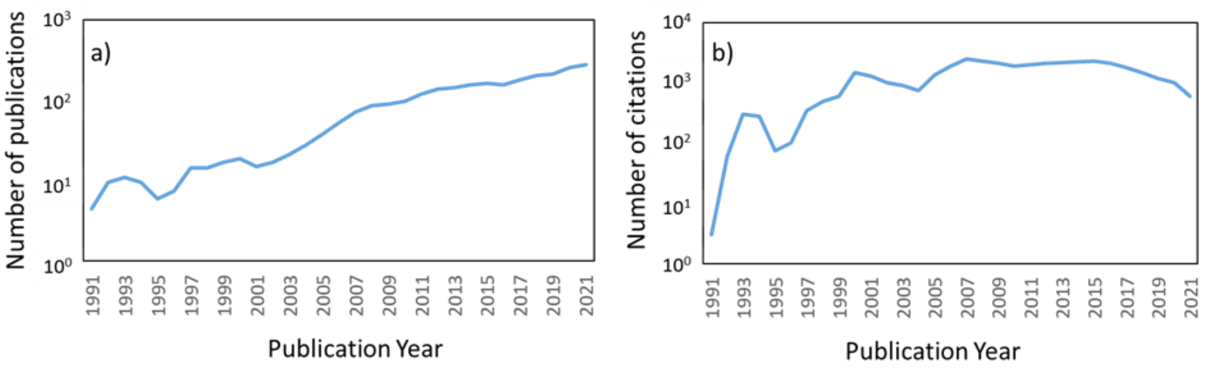

Figure 3: Number of article publications on corruption detection in the public sector – 2-year moving average $Log_{10.}$ a) Evolution of the number of publications. b) Evolution of the number of citations.

Based on measures of degree centrality, which shows how many times the writings are cited (illustrated by the size of the nodes), and the connection weight, which symbolizes how many times two articles are mentioned together in different articles (characterized by the link thickness), we created graphs based on data collected from the Scopus repository through VOSviewer and analyzed them using network science methods. Furthermore, we examined the data gathered from the simultaneous occurrence of keywords and authors, which indicates that the greater the node size, the more frequently the keyword emerges in writings. On the other hand, the thicker the connections between nodes, the grander the number of times these keywords appear in diverse papers. The same idea was established for article citations, i.e., the bigger the node size, the more cited the issued articles in the specified database, and the wider the linkage between them, the more frequently they operate together in publications.

2.3. Results

The data collection began in April 2021 and was later updated in December 2021 and January 2022. At first, we applied a broader search string without setting the specific area of interest included in the last "AND" to capture the dimension of the study being researched. As a result, we identified 39,445 documents. Thus, we applied the entire search string assembled, obtaining 2,914 articles, representing 7.4% of the scientific production in the sector. In Figure 3, we see the distribution of these articles by annual publication and annual citation amounts presented on a 2-year moving average. In Figure 3a, we noticed that one of the first articles presented by Scopus in the sector dates back to 1991, with a more significant and constant increase from the year 2000 on. From 2016 onwards, we see a steeper curve slope, indicating a greater interest in the subject over time. On the other hand, in 3b, we note that the peak of citations occurs in articles from the 2007/2008 biennium since those articles are available for longer for research and, therefore, are better able to receive more citations.

Table 1: Main Institutions

Institution	Publicat.	Citations
The Hong Kong Polytechnic University	5	68
University of Maryland	4	255
London School of Economics and Political Science	4	133
Illinois State University	4	87
University of Massachusetts Dartmouth	4	60
Australian National University	4	53
Universiti Teknologi Mara - Malaysia	4	32
Princeton University	4	32
Islamic University of Indonesia	4	31
University of Manchester	4	21

Table 2: Main Countries

Country	Publicat.	Citations
USA	732	19015
UK	270	4741
China	193	1310
India	160	976
Australia	136	1900
Canada	97	1273
Germany	82	996
Brazil	74	518
Italy	72	900
Nigeria	71	135

Among the 2914 studies identified, Table 1 presents the ten institutions that most contributed to the research using the specified string. Table 2 shows how these articles were distributed in the ten countries that have published the most.

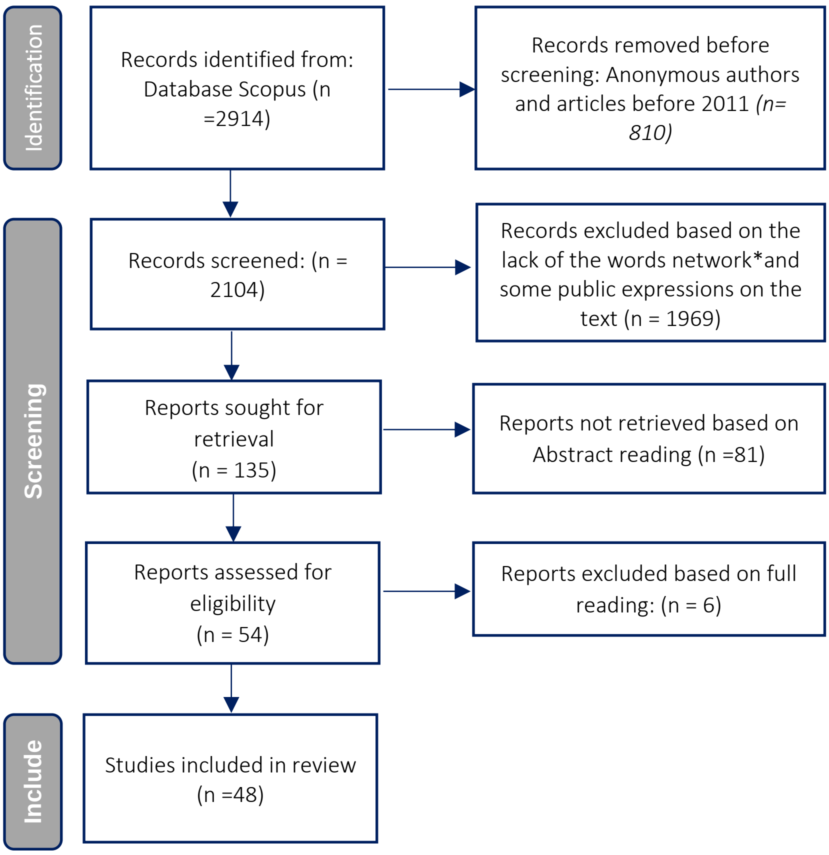

Figure 4: PRISMA 2020 flowchart three phases, n denotes the number of publications retrieved in each level.

Henceforward, we followed with the extraction of relevant publications through the PRISMA stages. We took all available publications with the author's identification, discarding 98 without a definition of authorship and 712 publications prior to 2011 to avoid picking up outdated articles, so we entered the screening phase with 2104 documents.

During this phase, we carried out suppressions based on the three sets of words: network*, some "public expressions" (public administration, public sector, tender*, auction, bid*, procurement), which were not presented in the title, abstract or keyword of the articles. Then we excluded 1969 publications and maintained 135 of them. Next, we analyzed the abstracts of the 135 documents and abandoned 81 that were beyond the scope, keeping 54. After making a complete reading of the remaining articles, we excluded six more publications, keeping 48 papers (Figure 4). Of the 48 articles, 36 were issued by scientific journals, while 13 were presented in conference procedures.

a) Scientific articles published in journals and conferences

The 48 selected publications are distributed over a decade in quantities ranging from two publications in 2011 to a maximum of 10 in 2021. It indicates that the topic has been gaining notoriety and being more investigated over time. The documents gathered from the Scopus from diversified journals and conference annals encompass a vast variety of scientific work, from business & economics to computer science, passing by law & government and science administration. The survey of scientific writings is fundamental for identifying pertinent material for our area of investigation. We selected papers mentioned up to 47 times, deriving from the list of publications in Table 3.

Table 3: **Articles selected from Scopus repository (ordered by number of citations)**

No	Title/Author/Year	Citation	Journal/Conference
1	An Objective Corruption Risk Index Using Public Procurement Data (Fazekas, Toth, & King, 2016.1)	47	European Journal on Criminal Policy and Research
2	From Corruption to State Capture: A New Analytical Framework with Empirical Applications from Hungary (Fazekas & Tóth, 2016.2)	41	Political Research Quarterly
3	Uncovering High-Level Corruption: Cross-National Objective Corruption Risk Indicators Using Public Procurement Data (Fazekas & Kocsis, 2020)	35	British Journal of Political Science
4	Internet auction fraud detection using social network analysis and classification tree approaches (Chiu, Ku, Lie, & Chen, 2011)	34	International Journal of Electronic Commerce
5	Bid-rigging networks and state-corporate crime in the construction industry (Reeves-Latour & Morselli, 2017)	32	Social Networks

No	Title/Author/Year	Citation	Journal/Conference
6	The impact of inter-organizational relationships on contractors' success in winning public procurement projects: The case of the construction industry in the Veneto region (Sedita & Apa, 2015)	28	International Journal of Project Management
7	Combining ranking concept and social network analysis to detect collusive groups in online auctions ((Lin, Jheng, & Yu, 2012)	23	Expert Systems With Applications
8	Finding the needle: A risk-based ranking of product listings at online auction sites for non-delivery fraud prediction (Almendra, 2013)	19	Expert Systems With Applications
9	Corruption and complexity: a scientific framework for the analysis of corruption networks (Luna-Pla & Carlock, 2020)	18	Applied Network Science
10	Fuzzy rule optimization for online auction frauds detection based on genetic algorithm (Yu & Lin, 2013)	15	Electronic Commerce Research
11	A network approach to cartel detection in public auction markets (Wachs & Kertesz, 2019)	14	Scientific Reports
12	Social capital predicts corruption risk in towns (Wachs, Yasseri, Lengyel, & Kertesz, 2019)	13	Royal Society Open Science
13	Leveraging social networks to combat collusion in reputation systems for peer-to-peer networks (Li, Shen, & Sapra, 2013)	13	IEEE Transactions on Computers
14	Distinguishing Characteristics of Corruption Risks in Iranian Construction Projects: A Weighted Correlation Network Analysis (Hosseini, Martek, & Banihashemi, 2019)	12	Science and Engineering Ethics
15	Network similarity and collusion (Morselli & Ouellet, 2018)	12	Social Networks
16	Machine learning with screens for detecting bid-rigging cartels (Huber & Imhof, 2019)	9	International Journal of Industrial Organization
17	Online Detection of Shill Bidding Fraud Based on Machine Learning Techniques (Ganguly & Sadaoui, 2018)	8	31st International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems (IEA/AIE)
18	A decision support system for fraud detection in public procurement (Velasco, Carpanese, Interian, & et, 2021)	7	International Transactions in Operational Research
19	Prediction of public procurement corruption indices using machine learning methods (Rabuzin & Modrusan, 2019))	7	11th International Conference on Knowledge Management and Information Systems (KMIS)
20	Comparison of ANN Classifier to the Neuro-Fuzzy System for Collusion Detection in the Tender Procedures of Road Construction Sector (Anysz, Foremny, & Kulejewski, 2018)	7	3rd World Multidisciplinary Civil Engineering, Architecture, Urban Planning Symposium (WMCAUS)
21	Using Self-Organizing Maps for fraud prediction at online auction sites (Almendra & Enachescu, 2014)	7	15th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC
22	Detection of feedback reputation fraud in Taobao using social network theory (Zhu, Zhang, & Yu, 2011)	7	Proceedings - 2011 International Joint Conference on Service Sciences, IJCSS 2011

No	Title/Author/Year	Citation	Journal/Conference
23	Corruption risk in contracting markets: a network science perspective (Wachs, Fazekas, & Kertész, 2021)	6	International Journal of Data Science and Analytics
24	Corruption and the network structure of public contracting markets across government change (Fazekas & Wachs, 2020)	5	Politics and Governance
25	Mapping Corruption Risks in Public Procurement: Uncovering Improvement Opportunities and Strengthening Controls (Sharma, Sengupta, & Panja, 2019)	5	Public Performance & Management Review
26	Big data system for analyzing risky procurement entities (Dhurandhar, Graves, Ravi, & Maniachari, 2015)	5	Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining
27	Online hybrid model for online fraud prevention and detection (Mundra & Rakesh, 2014)	5	Advances in Intelligent Systems and Computing
28	Uncovering the structure of public procurement transactions (Popa, 2019)	4	Business and Politics
29	Preventing rather than punishing: An early warning model of malfeasance in public procurement (Gallego, Rivero, & Martinez, 2021)	3	International Journal of Forecasting
30	Clustering and Labeling Auction Fraud Data (Alzahrani & Sadaoui, 2020)	3	International Conference on Data Management, Analytics and Innovation (ICDMAI)
31	Trust model based on Islamic business ethics and social network analysis (Kolan, Jailani, Abu Bakar, & Latih, 2018)	3	International Journal on Advanced Science, Engineering and Information Technology
32	How do strategic networks influence awarding contract? Evidence from French public procurement (Mamavi, Meier, & Zerbib, 2017)	3	International Journal of Public Sector Management
33	Detecting fraudsters in online auction using variations of neighbor diversity (Khomnotai & Lin, 2015)	3	International Journal of Engineering and Technology Innovation
34	Incremental collusive fraud detection in large-scale online auction networks (Dadfarnia, Adibnia, Abadi, & Dorri, 2020)	2	Journal of Supercomputing
35	Bidder Network Community Division and Collusion Suspicion Analysis in Chinese Construction Projects (Zhu, Wang, Li, & et, 2020)	2	Advances in Civil Engineering
36	Detecting the collusive bidding behavior in below average bid auction (Lei, Yin, Li, & Li, 2017)	2	13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)
37	Analysis on bidding behaviours for detecting shill bidders in online auctions (Majadi, Trevathan, & Bergmann, 2017)	2	Proceedings - 2016 16th IEEE International Conference on Computer and Information Technology, CIT 2016, 2016 6th International
38	Network Analysis for Fraud Detection in Portuguese Public Procurement (Carneiro, Veloso, Ventura, Palumbo, & Costa, 2020)	1	International Conference on Intelligent Data Engineering and Automated Learning
39	Improving fraudster detection in online auctions by using neighbor-driven attributes (Lin & Khomnotai, 2016)	1	Entropy

No	Title/Author/Year	Citation	Journal/Conference
40	Characterization of the firm–firm public procurement co-bidding network from the State of Ceará (Brazil) municipalities (Lyra, Curado, Damásio, Bacao, & Pinheiro, 2021.2)	0	Applied Network Science
41	Tackling corruption in urban infrastructure procurement: Dynamic evaluation of the critical constructs and the anti-corruption measures (Owusu, Chan, & Wang, 2021)	0	Cities
42	A social network-based examination on bid riggers' relationships in the construction industry: A case study of China (Xiao, Ye, Zhou, & Ye, 2021)	0	Buildings
43	Conspiracy of Corporate Networks in Corruption Scandals (Nicolás-Carlock & Luna-Pla, 2021)	0	Frontiers in Physics
44	The extra-legal governance of corruption: Tracing the organization of corruption in public procurement (Fazekas, Sberna, & Vannucci, 2021)	0	Governance
45	Networked Corruption Risks in European Defense Procurement (Czibik, Fazeka, Sanchez, & J., 2021)	0	Understanding Complex Systems
46	Corruption in the Implementation of Public Procurement from Small and Medium Businesses (Perevezentceva, Osipova, & Skrynnikova, 2021)	0	Lecture Notes in Networks and Systems
47	Real-Time Shill Bidding Fraud Detection Empowered with Fussed Machine Learning (Abidi, Daoudm, Ihnaini, & Khan, 2021)	0	IEEE Access
48	Adaptation of cluster analysis methods in respect to vector space of social network analysis indicators for revealing suspicious government contracts (Davydenko, Morozov, & Burmistrov, 2017)	0	IEEE 5th International Conference on Future Internet of Things and Cloud (FiCloud)

The material acquired from the repositories represents the most recent and updated studies currently used to detect fraud in the public sector. The more cited articles and authors are shown in Table 3. We verify that article no. 1 has the largest volume with 47 citations, followed by article no. 2 with 41, and article no. 3 with 35 citations, all related to the Scopus platform. The citations were examined in depth in the Author and Co-Authorship Citation Occurrences section, guided individually.

b) Publishing Journals and Conferences

Regarding the publication of articles in Journals, 36 studies were distributed in 32 distinct Journals and published by 26 editors. Based on Scimago Journal & Country Rank, 21 Journals (65%) are classified in the first quartile Q1, five journals (16%) in the second quartile Q2, four journals (13%) in quartile Q3, and two journals (6%) are classified in quartile Q4. The leading areas of action of the 32 journals are computer science, business, Social Science, and some multidisciplinary areas, see Table 4.

Table 4: Published Journals

Journal	Qt	Rank	Publisher	Country	Field
Social Networks	2	Q1	Elsevier	Netherlands	Anthropology; Sociology
Applied Network Science	2	Q1	Springer Open	Switzerland	Computer Science
Expert Systems With Applications	2	Q1	Elsevier Science LTD	England	Computer Science
Lecture Notes in Networks and Systems	1	Q4	Springer International Publishing AG	Switzerland	Computer Science
British Journal of Political Science	1	Q1	Cambridge University Press	UK	Social Science
Business and Politics	1	Q1	Cambridge University Press	UK	Social Science
Cities	1	Q1	Elsevier Ltd.	UK	Social Science
Electronic Commerce Research	1	Q1	Springer	Netherlands	Economics
European Journal on Criminal Policy and Research	1	Q1	Springer Netherlands	Netherlands	Social Sciences
Governance	1	Q1	Wiley-Blackwell Publishing Ltd	UK	Social Sciences
IEEE Access	1	Q1	Engineers Inc.	USA	Computer Science
IEEE Transactions on Computers	1	Q1	IEEE Computer Society	USA	Computer Science
International Journal of Electronic Commerce	1	Q1	M.E. Sharpe Inc.	USA	Business, Management & Accounting
International Journal of Forecasting	1	Q1	Elsevier	Netherlands	Business, Management & Accounting
International Journal of Industrial Organization	1	Q1	Elsevier Inc.	USA	Business, Management & Accounting
International Journal of Project Management	1	Q1	Elsevier Science LTD	England	Business & Economics
International Transactions in Operational Research	1	Q1	Wiley	USA	Business & Economics
Policing & Society	1	Q1	Routledge Journals, Taylor & Francis LTD	England	Criminology & Penology
Political Research Quarterly	1	Q1	Sage Publications Inc	USA	Government & Law
Public Performance & Management Review	1	Q1	SAGE Publications Inc.	USA	Business, Management & Accounting
Royal Society Open Science	1	Q1	Royal Soc	England	Science & Technology
Science and Engineering Ethics	1	Q1	Springer	Netherlands	Multidisciplinary Sciences
Scientific Reports	1	Q1	Nature Publishing Group	England	Science & Technology
International Journal of Public Sector Management	1	Q2	Emerald Group Publishing LTD	England	Business & Economics; Public Administration
Politics and Governance	1	Q2	Cogitatio Press	Portugal	Government & Law

Journal	Qt	Rank	Publisher	Country	Field
Buildings	1	Q2	MDPI Multidisciplinary Digital Publishing Institute	Switzerland	Building and Construction
Entropy	1	Q2	MDPI	Switzerland	Physics, Multidisciplinary
Frontiers in Physics	1	Q2	Frontiers Media S.A.	Switzerland	Materials Science
International Journal of Data Science and Analytics	1	Q2	Springer	Switzerland	Computer Science
Advances in Civil Engineering	1	Q3	Hindawi LTD	England	Construction & Building Technology
Advances in Intelligent Systems and Computing	1	Q3	Springer Verlag	Germany	Computer Science
International Journal on Advanced Science, Engineering and Information Technology	1	Q3	-	Indonesia	Agricultural and Biological Sciences
Journal of Supercomputing	1	Q3	Springer Netherlands	Netherlands	Computer Science
International Journal of Engineering and Technology Innovation	1	Q4	Taiwan Association of Engineering and Technology Innovation	Taiwan	Civil and Structural Engineering
Understanding Complex Systems	1	Q4	Springer Verlag	Germany	Computer Science

The selected papers were issued by publishers belonging to eleven nations. The Netherlands, Switzerland, and the USA were the most recurring, publishing six works each. In turn, the most commonly used editors were Springer and Springer Netherlands.

c) Conference Proceedings

The conferences were responsible for editing 12 documents published by seven editors belonging to five countries of the selected articles. The main field of investigation once again was computer science, Table 5.

Table 5: Conference Annals

Conference	Qt.	Publisher	Country	Field
International Conference on Intelligent Data Engineering and Automated Learning	1	Springer International Publishing AG	Switzerland	Computer Science
International Conference on Data Management, Analytics and Innovation (ICDMAI)	1	Springer Verlag Singapore PTE LTD	Singapore	Computer Science
11th International Conference on Knowledge Management and Information Systems (KMIS)	1	SCITEPRESS	Portugal	Computer Science

Conference	Qt.	Publisher	Country	Field
13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)	1	IEEE	USA	Computer Science
31st International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems (IEA/AIE)	1	Springer Nature Switzerland AG	Switzerland	Computer Science
3rd World Multidisciplinary Civil Engineering, Architecture, Urban Planning Symposium (WMCAUS)	1	IOP Publishing LTD	England	Architecture; Engineering; Urban Studies
Proceedings - 2016 16th IEEE International Conference on Computer and Information Technology, CIT 2016, 2016 6th International	1	IEEE	USA	Computer Science
12th Iberian Conference on Information Systems and Technologies (CISTI)	1	IEEE	USA	Computer Science
Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining	1	ASSOC COMPUTING MACHINERY	USA	Computer Science
15th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC	1	IEEE	USA	Computer Science; Engineering
Proceedings - 2011 International Joint Conference on Service Sciences, IJCSS 2011	1	IEEE	USA	Computer Science
IEEE 5th International Conference on Future Internet of Things and Cloud (FiCloud)	1	IEEE	USA	Computer Science; Telecommunications

Of the 12 conference proceedings, 11 were classified as computer science. The publishers were also varied; the IEEE published papers, particularly in Conferences on Fuzzy Systems, Artificial Intelligence, the Internet of Things, Information Systems and Technologies, and Internet Computing. Springer Nature issued articles on Intelligent Systems and Technologies, CISTI. The other five editors published articles on Intelligent Data Engineering, Data Management, Multidisciplinary Civil Engineering, and Data Mining.

d) Articles classification

To classify the articles, we took the specific themes addressed in each one. It is worth remembering that the public procurement articles referred to various types of contracts; however, not all papers offered a more comprehensive detailing of the contracted object. Furthermore, E-procurement* was presented as a resource for the Public E-procurement surveys. They are broader research and have more outstanding fraud detection capabilities, which may assist the public sector with solutions.

Table 6: Article classification

Article Classification	Number of articles	Corruption cases	Fraud cases	Collusion cases
General Public Procurement	18	17	3	5
Construction industry	9	5	0	8
Public online auction	1	1	1	0
Shell companies contracting	1	1	0	0
Bidding markets	1	0	0	1
Defense	1	1	0	0
*Private online auction	17	0	16	4
TOTAL	**48**	**25**	**20**	**18**

Table 6 shows that general public procurement dominates the researched occurrences due to the greater availability of open data in this segment, followed by the construction industry, which represents a large part of the expenditure associated with procurement contracts. On the other hand, the three types of crime studied perceived the same magnitude order in terms of monetary values. The most common crimes identified were corruption in public procurement and collusion in construction. Fraud appeared more in online auctions or e-procurement; however, this topic was better explored in the private sector.

e) Keyword Occurrences

To examine in detail sensitive themes, we performed a co-occurrence assessment by employing keywords occurrence based on the count of the individual or combined words in portions of scientific research (title, keyword, and abstract). Besides, the investigation is extended to various documents concomitantly, helping researchers review and exploit the writings statistically (Horvat, Havaš, & Logozar, 2015).

The analysis of keywords occurrence has been performed employing VOSviewer to assemble and view bibliometric networks. The examination was carried out through the complete counting method, containing ten keywords, in which the lowest occurrence boundary was fixed as six. The ten keywords classified in descending order of occurrence are displayed in Table 7.

Table 7: Major keywords occurrence

Keyword	Link strength	Occurrences	Avg. pub. Year
crime	276	19	2017
public procurement	146	15	2018
corruption	81	14	2019
fraud detection	144	11	2017
social network analysis	63	8	2013
construction industry	82	7	2018
social networking (online)	93	6	2014
learning systems	92	6	2018

Keyword	Link strength	Occurrences	Avg. pub. Year
data mining	76	6	2017
machine learning	71	6	2017

We can also see the average year of keyword usage in the selected documents. We perceive that the most relevant keywords, crime, public procurement, and corruption, have their average launch years between 2017 and 2019, and social network analysis (SNA) and social networking more between 2013 and 2014, when this expression was still widely used for network science tools.

The identified keywords appearing the most had the following link strength and occurrence, respectively: crime (276, 19), public procurement (146, 15), corruption (81, 14), fraud detection (144, 11), and social network analysis. In Figure 5, we see the co-occurrence connections between keywords in various scientific papers collected. Here we perceive the formation of two communities in the keyword co-occurrence network with ten nodes and 30 links.

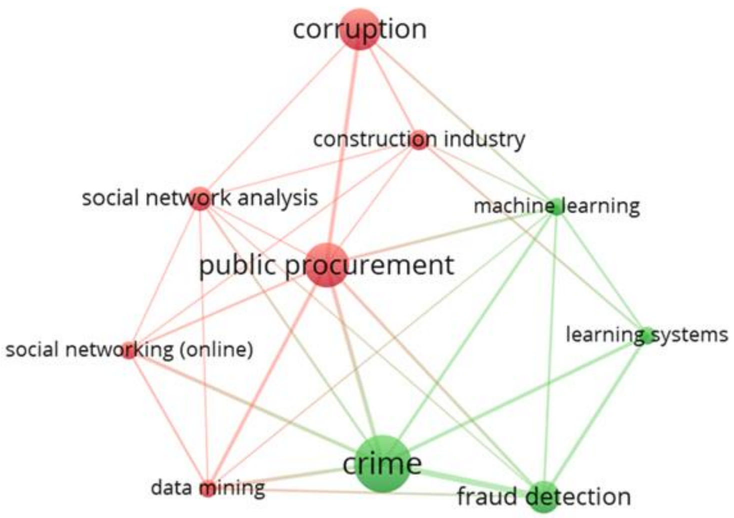

Figure 5: Keyword Occurrence Network - Community visualization

Keyword nodes are linked by their concurrent existence in separate documents, where keyword communities are formed based on the similarity between their connections. The node size denotes the occurrence of keywords, the bigger the node, the greater the keyword occurrence. Figure 5 shows the previously demonstrated statistics, where we perceive that the nodes representing public procurement, corruption, and crime keywords are the largest, meaning that they are the ones that appear the most. We named the two communities based on their most representative and centralized nodes, Public Procurement (red) and Crime (green).

The network emphasizes "public procurement" as the most studied area. On the other hand, the keyword "corruption", "social network analysis", and "data mining" appears in the same community as the problem being investigated and the techniques used for its identification. In turn, the community Crime presents machine learning and learning systems as techniques for fraud detection.

f) Author and Co-Authorship Citation Occurrences

Once again, we used the VOSviewer tool to investigate the author's and co-author's citation occurrence and the link strength. We employed the full counting method and established the minimum link strength as six. The documentation brought in 11 authors and is presented in Table 8 by link strength order.

Table 8: Noteworthy citations of authors and co-authors

Author	Strength	Documents	Citations	Avg. pub. Year
Fazekas, Mihály	12	7	134	2019
Wachs, Johannes	10	5	38	2020
Chan, Albert.P.C.	7	2	12	2020
Kertész, János	6	3	33	2019
Abidi, Wajhe Ul Husnian	6	1	0	2021
Alzahrani, Ahmad M.	6	1	0	2021
Alyas, Tahir	6	1	0	2021
Daoud, Mohammad S.	6	1	0	2021
Fatima, Areej	6	1	0	2021
Ihnaini, Baha	6	1	0	2021
Khan, Muhammad.Adnan	6	1	0	2021

The authors and co-authors that stand out the most for their Link strength and citations are Fazekas M., with seven articles and 134 citations, and Wachs J., with five documents and 38 citations. We observed that these two authors are very interconnected. Of the 12 articles they wrote, three they did together; Fazekas was the author in two and Wachs in one. The average publication period of the articles from the 11 authors is very recent, varying from 2019 to 2021.

As it is an emerging and incipient community, the network representation assembled by VOSviewer is still very disconnected and dispersed, but it is a community with high potential for future collaboration and should coordinate efforts to systematize techniques and approaches. Figure 6 presents the giant component of the citation network. Here we see the formation of 5 clusters with 21 nodes (articles) and 41 links, where Fazekas and Wachs appear in two communities. Fazekas has the most articles written, mainly in the blue community, and Wachs is the central link between the different communities and probably the most influential author in this subnet.

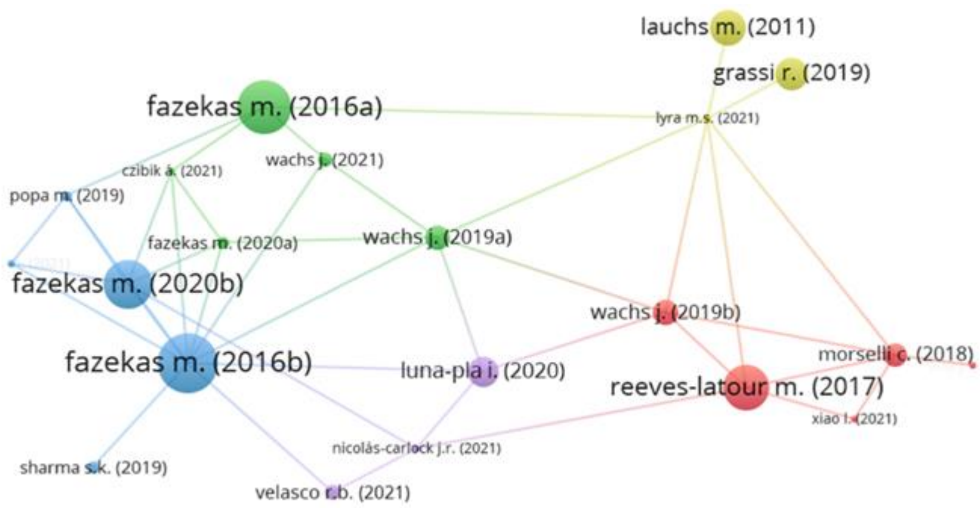

Figure 6: Citation network – giant component, community visualization

g) Methods, connectivity, and investigation type of the articles

In the list of scientific articles collected, we identified two most popular types of techniques, together or separately, to detect crimes against public administration: machine learning and network science. Furthermore, we verified the connectivity of the data presented in the papers, classifying them as connected networks if they had these characteristics. The status of connected networks was determined by a route between any random pair of nodes, even if intermediate nodes are needed to establish the course (Dong, Wang, Mostafizi, & Song , 2020). The high density of a network considerably improves the potential for collaborative action and cooperation between actors (Olsson, Folke, & Berkes, 2004), as well-connected networks foster communication, promote common confidence, and help contain or control conflicts (Bodin & Norberg, 2005). Furthermore, the high density of links promotes joint action, particularly when there are many connections between different agents (Sandström & Carlsson, 2008). In our study, 34 or 71% of articles were related to connected networks. It allowed researchers to explore the relationships between agents and uncover hidden connections.

Finally, we examined the type of investigation carried out, if exploratory or using labeled data. Labeled data research was used to confirm predictions and train models based on patterns of the information used. If the data are representative of the population, these will confirm or not the results of a finding, or a machine learning model will make predictions when using new data (Xiao, Chang, Geng, & Liu, 2018). On the other hand, exploratory data analysis sought to research little-known problems with no confirmatory data and investigated information to develop hypotheses (Robson, 2002). It aimed to develop knowledge that is still new (Kuncoro, 2013).

Our research has 52% reports using labeled data and 48% exploratory research. It allowed researchers to explore the data and confirm the results when using labeled data, and in the absence of confirmatory data, the exploratory study could draw significant conclusions. Notably, 85% of machine learning research used pre-labeled data in their analyses, while network science used it in 44% of its research. The information is presented in more detail for each article in Appendix A.

h) Objectives of the works

This section presents the primary purposes of each study grouped into higher-order objectives**,** Table 9.

Table 9: Central purpose of the articles

RESEARCH OBJECTIVE	ARTICLES NUMBER
Identify the risks of fraud, corruption, and collusion in public procurement	1, 2, 3, 19, 23, 24, 25, 26, 28, 29, 32, 38, 41, 43, 44, 46
Detect the risks of fraud, corruption, and collusion in the construction industry	5, 6, 14, 15, 16, 20, 35, 42
Uncover corruption in co-bidding relationships	11, 40
Recognize fraud in product listings, and fraudulent dealers on online public auctions	36, 48
Identify potential crimes in suspected shell companies	9, 18
Detect the risks of corruption and collusion in public defense contracts	45
Recognize collusion in the public administration	12
*Identify fraud in product listings, fraudulent dealers, or bidder reputation on private online auctions	4, 7, 8, 10, 13, 21, 22, 17, 31, 33, 34, 39
*Detect shill bidding fraud in private online auctions	17, 30, 37, 47

We divided the articles into ten macro-objective groups to better understand the proposals presented. A significant part of the articles falls under the detection of fraud, corruption, and collusion in public procurement, with 11 articles. Ranking second to detect identical crimes, we have the construction industry with eight articles. We also present two macro-objectives studied from the private* segment referring to online auctions, which are widely used in the public domain, but little explored in scientific research to detect crimes in this segment. To have a more specific glimpse of the objectives of each scientific article, a synopsis of the purposes of the 48 topics is presented in Appendix B.

2.4. Discussion

The present study uses the PRISMA methodology and network analysis to spot the most recent research and techniques for crime identification in public procurement processes. This section introduces a straightforward assessment comparing the outcome of the Scopus repository. It analyzes how we answer the research question and addresses the limitations of the research as follows:

RQ: How have data-driven methods contributed to detecting, characterizing, and predicting fraud and corruption in public procurement activities?

a) Most commonly used techniques using labeled data

Concerning the main techniques applied, 85% of the approaches used in machine learning papers employed labeled data alone or in conjunction with another method for detecting crimes in public contracts. Different techniques were applied: Artificial Neural Network - ANN (Articles no. 16, 20, 47); Support Vector (Articles no. 17, 19, 34, 47); Random Forest - RF (Articles no. 16, 28, 44); Logistic Regression (Articles no. 19, 44); Lasso Regression (Articles no. 16, 29); Gradient Boosting Classification Model (Article no. 29); Naive Bayes (Article no. 19); Boosted trees (Article no. 8); binary classification (Article no. 21).

Random Forest was employed to find bid-rigging cartels, identify close connections between government and private firms in repeated interaction and geographic dispersion, and detect single bidding and government spending concentration. It reached an adequate performance in finding competitive tenders and classifying collusive ones. It performed well in both collusive and non-collusive periods, appropriate for detecting bid-rigging cartels. RF also identified the emergence of geographically close and repeated connections between private entities and public authorities, suggesting corrupt or socially undesirable behavior and favoring illicit schemes. When used to detect single bidders or government spending concentration, it showed encouraging results with internal data but only moderate when externally validated. Nonetheless, it achieved higher explanatory power than Logistic Regression due to a more flexible parameterization.

Logistic Regression helped find suspicious tenders with text-mining techniques and detect single bidding and government spending concentration. Applied jointly with text mining to recognize contract corruption, it showed intermediate precision, recall, and Receiver Operating Characteristic (ROC) results. It had better results than SVM but lower than Naive Bayes. Similar to RF, Logistic Regression for detecting single bidders and concentration of government spending showed significant results when using internal data; however, only reasonable outcomes when validated with external information. On the other hand, the Logistic

Regression obtained less explanatory strength than the RF due to a less flexible parameterization.

The Lasso Regression was intended to detect bid-rigging cartels and predict inefficiency and corruption in public procurement. The lasso performed slightly better in collusive than non-collusive periods, but the results were quite satisfactory in detecting bid-rigging. When used to predict inefficiency and corruption in tenders, it was outperformed by Gradient Boosting Machine in all scenarios.

Regarding network science techniques using labeled data, the most employed were cluster analysis (community detection/clustering) used by eight studies (Articles No: 24, 31, 35, 36, 39, 40, 43, 48), where clustering was formed when the similarity between nodes is high, and nodes are separated into different communities when similarity is below the threshold defined by the algorithm. Centrality measures were used by four studies (Articles No: 33, 38, 42, 48) which usually reflect the prominence, status, prestige, or visibility of a node and often explain behaviors on the network (Marsden, 2005). One (Article No: 40) applied bipartite network projection to generate a co-participation network and capture the hidden relationships between players.

They examined the endurance of fraudsters in a weak competitive community over the years after a government turnover to identify corrupt connections between agents. The centrality measures were used to help identify the most central, active, influential players and those presumed of fraudulent connections. The degree centrality, for example, excels in revealing the most linked players of the net. Nevertheless, while it is able to assist in indicating network hubs, criminal interfaces are not always intermediated by them (Morselli & Ouellet, 2018); therefore, other measures and analyses were necessary to identify suspicious connections. On the other hand, regardless of employing closeness centrality, some criminals were conscious of the existence of risk indicators and were persevering to bypass them by decreasing the mean closeness values through data tampering. An analysis of clusters demonstrated that the most effective indicators to expose suspect features were based on contract expenditures, contract title, closeness weights, and eigenvector mean values, demonstrating, via these measures, that a considerable portion of government deals held infractions.

Some other methods that deserve further comments used combinations of models. The labeled data based on hierarchical Clustering Using REpresentatives (CURE) is a clustering model that identifies bidding strategies. This technique delivers an adequate number of clusters, and k-mean clustering supplies elements concerning the players' methods in each cluster. Furthermore, a combination of approaches via Machine Learning techniques (hybrid Decision Support System, decision trees, Support Vector Machine, and Neural Networks) and network science algorithms (PageRank, density, and centralities) was used to reveal collusion and corruption associated with defrauding conduct on networks, and together with the concept of neighbor diversity to differentiate fraudsters from standard actors.

b) Most applied techniques using exploratory research

Through network science techniques, community detection was the most applied mechanism to verify the formation of clusters, particularly cartels. It was chosen to examine collusive trades in significant markets in which market shares with appropriate prerequisites for cartel appearance were uncovered in previously hidden connections. Nevertheless, the approach to detecting cartels was sometimes merely representative and could not always convincingly indicate crimes without employing labeled data.

Community detection also allowed witnessing organizations making extensive agreements with each other rather than making alliances beyond their communities. It pointed to communities practicing transgressions in their contracts, spotting groups without agreements providing the awarded objects. The diversification of co-bidding communities formed in different regions was also evaluated, in which alleged cartel formation was identified based on the communities' low regional and contractual diversification, suggesting collusion of companies in some regions to supply specific products.

In turn, it was analyzed to what extent the volume of single bids differed in diverse communities within a network. The diagnosis was that fraud is not erratically disseminated in procurement processes; the higher the clustering of single bidders, the more likely the analysis of their neighbors would bring up more fraud cases. Besides, communities with extra bonding social capital stood at a greater risk of fraud, and communities with more assorted extrinsic connectivity were less susceptible to It. Furthermore, some networks were split into different clusters, categorizing intersections as new communities, and classifying them as conspiracy cases when there was a high degree of nodes overlaying.

Over time, an interchangeable and nonlinear association between a corruption risk index and the closeness centrality was pointed out. It indicated the closeness measure effectively revealing fraud risk after government changes.

Betweenness was used to estimate the likelihood of agents having weak links with each other but still being a necessary mediator in trades and partnership management. They were able to underline the relational effects of organizations that could control the award of contracts.

It should be noted that, regardless of using some measures like network diameter, mean path length, and degree measures, these metrics stayed unchanged in certain studies, not supplying sufficient facts to predict the risk of crimes. Besides, in networks of bipartite nature, these measures alone were not sufficient to demonstrate evidence of a crime.

Based on the facts exposed above, we can realize that some mechanisms of network science are not always capable of providing accurate sole outcomes. Therefore, they are usually applied with different frameworks, such as neighbor diversity, single bidder, diversity of winners, victory rate of a bidder, or even creating indexes from intrinsic process parameters.

c) Research Gaps

We found three crucial gaps during this research; the first is related to the online auction in the public sector context. Although the process is very similar in the public and private domains, it has been more research in the private sector, contributing to developing fraud detection techniques in this segment. This type of bidding has a substantial share of public spending and should be considered when studying crimes in public contracts. A good example is electronic trading accounts which represent more than 90% of public contracting expenditures in Brazil's federal government (Governo_Federal_do_Brasil, 2022).

The second gap is related to the limited use of objective indicators to measure corruption risk in tenders (Lyra, Bacao, & Pinheiro, 2021.1). Corruption indices are usually obtained based on qualitative research and tend to reflect individual researchers' perceptions and biases. Objective microdata such as the number of bidders, market share of winners, bidding modality, time-lapse for awarding a contract, amendment of contractual clauses during its regular term, duration of the contract, and extension of the contract period after the end of the regular term are intrinsic to biddings operations. They promote greater transparency and accuracy in developing a fraud indicator, especially when well-calibrated and weighted. Although We collected five articles from the same group of authors and co-authors, using data related to EU countries. Currently, corruption indicators are mostly subjective and somewhat unreliable (Fazekas & Kocsis, 2020). Different red flags in public procurement processes could be better explored and taken as elements of a composite indicator, using regression coefficients to estimate the weights.

Third, it is essential to continue looking for solutions that allow the use of unlabeled data more accurately to detect crimes against the public administration. Despite the increase in the amount of data available from the public sector, the vast majority of those with access are not classified. In this sense, a methodology that proved to be more feasible in the absence of pre-classified data is using network science tools.

d) Research Limitations

The methods applied in this systematic literature review contain limitations for identifying papers beyond the keywords defined in the search string. The chosen essays may not comprise extensive studies in crime detection in public procurement since some keywords or expressions, such as Red Flag, Hoax, Trick, Blow, and Fake, may have been neglected. Besides, Boolean connectors might have restricted the search query. Furthermore, only the repository Scopus were investigated, and maybe vaster results could have been reached using different platforms.

However, concerning the research period between 2011 and 2021, we believe it is not demanding to search periods prior to 2011 since the search query for the period before it returned 745 documents, representing 25% of the total. In addition, the fraud detection methodology in the public sector, mainly regarding procurement, has begun to attain visibility more lately. Taking a more specific search string regarding the area of interest, for example, such as ("network analysis" OR SNA OR "Network science") AND ("public procurement") AND (corruption OR fraud), we found only eight articles in Scopus, where the publications start to appear only in 2016 with one paper, 2017 with another one, 2020 with three and 2021 with another three articles. It suggests that the specific subject is recent and very little investigated.

2.5. Conclusions

Many existing methods for detecting fraud in the public sector rely on rules established by an expert. However, access to an increasing number of data sources and the ability to analyze large volumes of data through machine learning, network science, and other methods offers the possibility to understand and predict fraudulent activities.

Here, we used the PRISMA methodology to identify the most relevant contributions to the topic. We characterize the research efforts in the field, identifying the most active authors, venues, methodologies, and application contexts. Finally, we emphasized the increasing importance of network science and machine learning in this domain.

We highlighted 48 contributions chosen based on their relevance to respond to the research question regarding the methods used to detect crimes in public procurement. The survey allowed us to investigate deeper into machine learning methods, most of them using supervised techniques and network science approaches to detect crimes in public contracts.

Machine learning techniques have proved to be very efficient in identifying instances of inadequate public spending use. However, there is no specific best tool to detect all types of crime. We found that the most effective and most used approaches rely on Random Forest algorithms. At the same time, ANN was particularly adequate in detecting shill bidding in real-time, SVM in finding fraudsters through neighbor diversity features, Naive Bayes detecting inefficiency and corruption in tenders, and Random Forest in uncovering bid-rigging cartels.

Network science methods have been used to estimate centrality measures and relevant communities among the participating actors, particularly to verify the existence of cartels. The use of network science is particularly relevant when studying unlabeled data. That is when there is no clear identification of which contracts have been flagged as subject to illegal procedures (e.g., fraudulent activity). The absence of confirmatory information was addressed through exploratory investigation and identifying hidden relationships between network agents. That said, we perceive that network science is a central framework for developing more sophisticated methods for studying fraud, collusion, and corruption activities lacking confirmatory data. Future work should direct efforts to investigate the issues addressed in the uncovered gaps, especially identifying fraudulent actions in electronic public auctions, since this is becoming the most used means by world governments in public acquisition.

Appendix A: Articles classification

Table 10 presents the methods, data connectivity, and type of investigation carried out in each article.

Table 10: Method, connectivity, and investigation type

Article No.	Method	Connected network	Investigation type
1	Statistics	No	Exploratory
2	Network science	Yes	Exploratory
3	Statistics	No	Exploratory
4	Network science and machine learning	Yes	Labeled data
5	Network science	Yes	Exploratory
6	Network science	Yes	Exploratory
7	Network science	Yes	Labeled data
9	Machine learning	No	Labeled data
9	Network science	Yes	Exploratory
10	Network science and machine learning	Yes	Labeled data
11	Network science	Yes	Labeled data
12	Network science	Yes	Labeled data
13	Network science	Yes	Labeled data
14	Network science	Yes	Exploratory
15	Statistics	Yes	Exploratory
16	Machine learning	Yes	Labeled data
17	Machine Learning	No	Labeled data
18	Network science and statistics	Yes	Exploratory
19	Machine learning	No	Labeled data
20	Machine Learning	Yes	Labeled data
21	Machine learning	No	Labeled data
22	Network science	Yes	Labeled data
23	Network science	Yes	Labeled data
24	Network science	Yes	Labeled data
25	Machine learning and statistics	No	Exploratory
26	Network science and machine learning	Yes	Exploratory
27	Network science and statistics	Yes	Exploratory
28	Machine learning	Yes	Labeled data
29	Machine learning	No	Labeled data
30	Clustering Using REpresentatives (CURE)	No	Labeled data
31	Network science	Yes	Exploratory
32	Network science and statistics	Yes	Exploratory
33	Network science	Yes	Labeled data
34	Machine learning	Yes	Labeled data
35	Network science	Yes	Exploratory
36	Network science	Yes	Exploratory
37	Machine learning and statistics	No	Exploratory

Article No.	Method	Connected network	Investigation type
38	Network science and machine learning	No	Labeled data
39	Network science	Yes	Exploratory
40	Network science	Yes	Exploratory
41	Network science	No	Exploratory
42	Network science	Yes	Labeled data
43	Network science	Yes	Labeled data
44	Machine learning	Yes	Labeled data
45	Statistics	Yes	Labeled data
46	Statistics	No	Exploratory
47	Machine learning	No	Labeled data
48	Network science	Yes	Labeled data

From Table 10, we can see that the most frequent approaches when using label data are generally related to machine learning, representing 100% of articles exclusively concerning it. They represent 50% of these studies regarding network science using label data. On the other hand, studies with label data using other methods represent 53% of articles. That said, it emerges that in studies where label data is available, the use of machine learning as a preferred method prevails. In contrast, approaches such as network science are evenly distributed with the use of label data and exploratory research.

Appendix B: Articles' purpose

Table 11: Objective of the articles (ordered by number of citations)

No	Author/Year	Objective
1	(Fazekas, Toth, & King, 2016)	Intend to measure grand corruption and state capture at the micro-level of individual contracts and tenders based solely on "objective" behavioral data.
2	(Fazekas & Tóth, 2016)	Identify clusters at high-corruption risk in public procurement in a contractual network of procuring authorities and suppliers.
3	(Fazekas & Kocsis, 2017)	Contemplate the absence of feasible corruption indicators to estimate the level of fraud in public procurement.
4	(Chiu et al., 2011)	Proposes a model to discover fraudsters based on Internet auction transaction records using data collected from Yahoo! Auctions
5	(Reeves-Latour & Morselli, 2017)	Outlines the appearance of bidding processes and corporate criminal networks in the building industry. It describes the development of corrupt networks around dubious and continued actions of criminality, collusion, and bribery and analyzes relations among players.
6	(Sedita & Apa, 2015)	Perform an empirical investigation into a network with relationships between companies involved in public procurement schemes in the construction sector in the Veneto region between 2008 and 2012.
7	(Lin et al., 2012)	Detect irregular conduct and supply a classification method to evaluate the level of risk of defrauding online auctions by collusive groups.
9	(Almendra, 2013)	Recognize fraudulent dealers in auctions and combat non-delivery of non-existent goods before the fraudsters can act.
9	(Luna-Pla & Carlock, 2020)	Conduct an empirical process for modeling corruption practices concerning a network of hundreds of shell firms.

No	Author/Year	Objective
10	(Yu & Lin, 2013)	Uncover fraud cases on online auction sites and detect the accounts of the most active fraudsters on the network, in addition to helping buyers identify them.
11	(Wachs & Kertesz, 2019)	Intend to detect the potential cartels involved in bidding processes
12	(Wachs et al., 2019)	Investigate the social factors of corruption and evaluates the risk of fraud due to suppressed competition and the absence of clearness in the public contracts awarded to the deal. It compares the risk of fraud in agreements with fragmented social networks and excess compulsory social capital.
13	(Li et al., 2013)	Seeks to measure and improve the predictive capacity of reputation systems in the fight against collusion in peer-to-peer networks.
14	(Hosseini et al., 2019)	Determine the character of the endemic corruption threats in the Iranian construction initiative, its grade of occurrence, and the strength of its effect.
15	(Morselli & Ouellet, 2018)	Investigate whether the companies' market share can be identified through the co-bidding similarity in bidding patterns and whether this measure forecasts their market share in processes with recognized conspiracy.
16	(Huber & Imhof, 2019)	Predict collusion through bid-rigging cartels in tenders within the Swiss construction sector
17	(Ganguly & Sadaoui, 2018)	Identify and characterize Shill Bidding in e-auctions employing real data from eBay.
18	(Velasco et al., 2021)	Identify the principal patterns of corruption risk in public expenses, such as stunts, conflicts of interest, and the participation of shell companies in bidding processes.
19	(Rabuzin & Modrusan, 2019)	Compare prediction models of suspicious bids and develop a model to detect suspicious one-bid in public procurement processes.
20	(Anysz et al., 2018)	Compare two machine learning techniques for detecting collusion between bidders in road construction tenders.
21	(Almendra & Enachescu, 2014)	Identify fraud in product listings on online auction sites for fair operators, customers, and vendors.
22	(Zhu et al., 2011)	Seeks to extract characteristic traits in the behaviors of fraudsters from the feedback reputation system, who manipulate their feedback scores by engaging in various legitimate sales.
23	(Wachs et al., 2021)	Aim to investigate the risk of fraud in public procurement agreements of the Member States of the European Union.
24	(Fazekas & Wachs, 2020)	Explore the association between corruption and market anatomy in public procurement approaches.
25	(Sharma et al., 2019)	Identify the risks of corruption in public procurement and discover the contribution of suspected actors and their relationships in the network.
26	(Dhurandhar et al., 2015)	It analyzes corporation loss due to fraud and proposes tools and processes to quickly and cheaply identify fraud/risks related to public and private procurement and potential wrongdoers.
27	(Mundra & Rakesh, 2014)	Investigates how a contractor's network position affects its success in winning public procurement projects through its ability to partner and influence the network.
28	(Popa, 2019)	It analyzes a set of public procurement data from European countries in order to identify the emergence of close links between bidders and public administration, defined in terms of repeated interaction and geographical dispersion.
29	(Gallego et al., 2021)	Determine trades that can become troublesome and forecast inefficiency and fraud in public procurement by employing a dataset with about two million public procurement processes in Colombia.
30	(Alzahrani & Sadaoui, 2020)	Identify shill bidding fraud in online auctions.
31	(Kolan et al., 2018	It intends to measure the user reputation (trust score) by considering the feedback reliability status for all transactions in the e-commerce market, such as e-auction.
32	(Mamavi et al., 2017)	Explore conniving networks deriving from collaborative associations among firms. It suggests a protocol to investigate networks within the public sector

Nₒ	Author/Year	Objective
		based on a quantitative method. It also emphasizes the interactive effects of companies that might impact the granting of contracts in decision-making, besides the influence of weak and strong bonds between firms on the contracts.
33	(Khomnotai & Lin, 2015)	Recognize irregularities on online auctions by exploring the concept of neighborhood diversity as an effective resource to identify competitiveness and joint participation between companies that may be acting together to defraud bids regularly.
34	(Dadfarnia et al., 2020)	Detect fraudulent users in collusive acts in online auctions.
35	(Zhu et al., 2020)	Design a social networking standard to detect possible bidder conspiracy in the construction sector.
36	(Lei et al., 2017)	Detect the formation of scheming groups in auctions through collusive activities or manipulation of the average price of the bids towards a higher or even unrealistic winning price.
37	(Majadi et al., 2017)	Provides a brief overview of leading research on bidding patterns to detect shill bidders in online auctions, illustrates the characteristics of such bidding patterns, and presents case studies identifying shill bidding behaviors in eBay's datasets.
38	(Carneiro et al., 2020)	Detect fraud in public tenders of Portugal, investigate the clearness of the deals between network actors and promote access to relevant data.
39	(Lin & Khomnotai, 2016)	It seeks to determine the change in the reliability perception of competitors within the bidder's reputation system on online auction sites when creating false transactions.
40	(Lyra et al., 2021.2)	Explore analytical structures that can help public controls recognize agents more sensitive to irregular activities in Public Procurement in Ceará (Brazil).
41	(Owusu et al., 2021)	It examines the effects of corruption at the stages of procurement processes and develops a framework for modeling the impacts of corruption and improving anti-corruption measures.
42	(Xiao et al., 2021)	Examine the features of conniving bidding networks and the types of collusive bidding groups in the construction industry.
43	(Nicolás-Carlock & Luna-Pla, 2021)	Understand the criminal collusion of companies implicated in corruption scandals in public tenders through exclusive data from reported corruption cases in Mexico, where several companies were manipulated to misappropriate billions of dollars.
44	(Fazekas et al., 2021)	Explores through empirical data the role of governance played by criminal groups organized in corruption networks, facilitating corrupt transactions by reducing the costs of searching, negotiating, and enforcing public purchases.
45	(Czibik et al., 2021)	It studies the corruption risks in EU defense procurement that present a significant potential for corruption and state capture. It uses a large set of contract data covering ten years of investigation.
46	(Perevezentceva et al., 2021)	It studies fraud in public procurement and the development of the control role in the bidding system and suggests introducing an anti-corruption investigation of the records of candidates participating in competitive processes.
47	(Abidi et al., 2021)	Detect Shill Bidding (SB), which is when the seller presents false bidders to increase the final price of a bid.
48	(Davydenko et al., 2017)	Detect likely risks of corruption in public tenders competition using empirical data from procurement processes and labeled data encompassing the registration of suppliers with lawsuits in fraudulent cases of public acquisitions.

Table 11 presents the objectives of the selected writings to disclose supposed fraud, corruption, or collusion in public contracting. Here we can glance at the scope of each survey, its instructions, and the outcomes presented. We verified that the most adopted methods to capture crimes in public acquisition vary widely, ranging from network science tools (social network analysis and complex network techniques) to statistical analysis and machine learning approaches.

Out of 48 articles, 22 (45%) are related to Network Science (Articles no. 2, 5, 6, 7, 9, 11, 12, 13, 114, 22, 23, 24, 31, 33, 35, 36, 39, 40, 41, 42, 43, 48); 11 (23%) to Machine Learning (Articles no. 8, 16, 17, 19, 20, 21, 28, 29, 34, 4447); five (10%) statistical methods (Articles no. 1, 3, 15, 45, 46); and 10 (21%) used combined methods like “network science & machine learning” (Articles no. 4, 10, 26, 38), “network science & statistics” (Articles no. 18, 27, 32), “machine learning & statistics” (Articles no. 25, 37) and “Clustering Using REpresentatives – CURE, (Article no. 30).

From the selected articles 34 (71%) used connected Network in their analysis (Articles no. 2, 4, 5, 6, 7, 9, 10, 11 ,12, 13 ,14 ,15 ,16, 18, 20, 22, 23, 24, 26, 27, 28, 31, 32, 33, 34, 35, 36, 39, 40, 42, 43, 44 ,45, 48), and 29 used labeled data (Articles no. 4, 7, 8, 10, 11, 12, 13, 16, 17, 19, 20, 21, 22, 23, 24, 25, 29, 31, 33, 34, 38, 42, 43, 44, 45, 47, 48).

References

Abidi, W., Daoudm, M., Ihnaini, B., & Khan, M. (2021). Real-Time Shill Bidding Fraud Detection Empowered with Fussed Machine Learning. *IEEE Access*.

Almendra, V. (2013). Finding the needle: A risk-based ranking of product listings at online auction sites for non-delivery fraud prediction. *Expert Systems With Applications*.

Almendra, V., & Enachescu, D. (2014). Using Self-Organizing Maps for fraud prediction at online auction sites. *15th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*.

Alzahrani, A., & Sadaoui, S. (2020). Clustering and Labeling Auction Fraud Data. *International Conference on Data Management, Analytics and Innovation (ICDMAI)*.

Anysz, H., Foremny, A., & Kulejewski, J. (2018). Comparison of ANN Classifier to the Neuro-Fuzzy System for Collusion Detection in the Tender Procedures of Road Construction Sector. *3rd World Multidisciplinary Civil Engineering, Architecture, Urban Planning Symposium (WMCAUS)*.

Bodin, O., & Norberg, J. (2005). Information network topologies for enhanced local adaptive management. *Environ Manag, 35*(2), pp. 175-193.

Borras, J., Cousins, B., Liu, J., McKay, B., & Milhorance De Castro, C. (2015). Emerging trends in global commodities markets: the role of Brazil and China in contemporary agrarian transformations. *International Institute of Social Studies, The Hague*.

Bosio, E., Djankov, S., Glaeser, E., & Shleifer, A. (2022). Public procurement in law and practice. *Am Econom Rev, 112*(4), pp. 1091-1117.

Carneiro, D., Veloso, P., Ventura, A., Palumbo, G., & Costa, J. (2020). Network Analysis for Fraud Detection in Portuguese Public Procurement. *Conference on Intelligent Data Engineering and Automated Learning.*

Chiu, C., Ku, Y., Lie, T., & Chen, Y. (2011). Internet auction fraud detection using social network analysis and classification tree approaches. *International Journal of Electronic Commerce*.

Costa, G. A., Machado, D., & Martins, V. (2020). The Efficiency of Social Control in Municipal Bidding: A Study in Social Observatories. *Sociedade Contabilidade e Gestão 14(4):112.*

Czibik, Á., Fazeka, s. M., Sanchez, A., & J., W. (2021). Networked Corruption Risks in European Defense Procurement. *Corruption Networks: Concepts and Applications*, 67-87.

Dadfarnia, M., Adibnia, F., Abadi, M., & Dorri, A. (2020). Incremental collusive fraud detection in large-scale online auction networks. *Journal of Supercomputing*.

Davydenko, V. I., Morozov, N. V., & Burmistrov, M. I. (2017). Adaptation of cluster analysis methods in respect to vector space of social network analysis indicators for revealing suspicious government contracts. *IEEE 5th International Conference on Future Internet of Things and Cloud (FiCloud)*.

Dhurandhar, A., Graves, B., Ravi, R., & Maniachari, G. (2015). Big data system for analyzing risky procurement entities. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Ding, X., & Yang, Z. (2020). Knowledge mapping of platform research: a visual analysis using VOSviewer and CiteSpace. *Electronic Commerce Research*.

Dong, S., Wang, H., Mostafizi, A., & Song , X. (2020). A network-of-networks percolation analysis of cascading failures in spatially co-located road-sewer infrastructure networks. *Physica A: Statistical Mechanics and its Applications, 538, 122971*.

Fazekas, M., & Kocsis, G. (2017). Uncovering High-Level Corruption: Cross-National Objective Corruption Risk Indicators Using Public Procurement Data. *British Journal of Political Science, 50*(1), pp. 155-164.

Fazekas, M., & Tóth, I. J. (2016.2). From corruption to state capture: A new analytical framework with empirical applications from Hungary. *Political Research Quarterly, 69 (2), 320–334.*

Fazekas, M., & Wachs, J. (2020). Corruption and the network structure of public contracting markets across government change. *Politics and Governance, 8*(2), pp. 153-166.

Fazekas, M., Sberna, S., & Vannucci, A. (2021). The extra-legal governance of corruption: Tracing the organization of corruption in public procurement. *Governance*.

Fazekas, M., Toth, I. J., & King, L. P. (2016.1). An Objective Corruption Risk Index Using Public Procurement Data. *European Journal on Criminal Policy and Research, 22*(3), pp. 369-397.

Gallego, J., Rivero, G., & Martinez, J. (2021). Preventing rather than punishing: an early warning model of malfeasance in public procurement. *International Journal of Forecasting, 37*(1), 360-377.

Ganguly, S., & Sadaoui, S. (2018). Online Detection of Shill Bidding Fraud Based on Machine Learning Techniques. *31st International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems (IEA/AIE)*.

Governo_Federal_do_Brasil. (2022). Decreto aprimora regras do pregão eletrônico. Fonte: https://www.gov.br/compras/pt-br/acesso-a-informacao/noticias/decreto-aprimoraregras-do-pregao-eletronico. Accessed on 08/04/2022

Grassi, R., Calderoni, F., Bianchi, M., & Torriero, A. (2019). Betweenness to assess leaders in criminal networks: New evidence using the dual projection approach. *Social Networks, 56 , 23–32.*

Horvat, T., Havaš, L., & Logozar, R. (2015). The Analysis of Keyword Occurrences Within Specific Parts of Multiple Articles — the Concept and the First Implementation. *SSN 1846-6168 UDK 001.8.*

Hosseini, M., Martek, I., & Banihashemi, S. (2019). Distinguishing Characteristics of Corruption Risks in Iranian Construction Projects: A Weighted Correlation Network Analysis. *Science and Engineering Ethics.*

Huber, M., & Imhof, D. (2019). Machine learning with screens for detecting bid-rigging cartels. *International Journal of Industrial Organization*.

Kartika, D., Kebijakan, I., Barang, P., & Pemerintah, J. (2020). The impact of e-procurement implementation on public procurement's corruption cases – evidences from Indonesia and India. *J Kajian Wilayah, 11*, pp. 193-212.

Khomnotai, L., & Lin, J. (2015). Detecting fraudsters in online auction using variations of neighbor diversity. *International Journal of Engineering and Technology Innovation.*

Kolan, N., Jailani, N., Abu Bakar, M., & Latih, R. (2018). Trust model based on Islamic business ethics and social network analysis. *International Journal on Advanced Science, Engineering and Information Technology, 8*(6), pp. 2323-2331.

Kuncoro, M. (2013). Research methods for business and economics. *Indonesia Erlangga Publisher*(4).

Lauchs, M., Keast, R., & Yousefpour, N. (2011). Corrupt police networks: uncovering hidden relationship patterns, functions and roles. *Policing & Society 21 (1), pp.110-12.*

Lei, M., Yin, Z., Li, S., & Li, H. (2018). Detecting the collusive bidding behavior in below average bid auction. *13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD).*

Li, Z., Shen, H., & Sapra, K. (2013). Leveraging social networks to combat collusion in reputation systems for peer-to-peer networks. *IEEE Transactions on Computers*.

Lin, J., & Khomnotai, L. (2016). Improving Fraudster Detection in Online Auctions by Using Neighbor-Driven Attributes. *Entropy, Vol. 18, Ed:1, N:e18010011*.

Lin, S., Jheng, Y.-Y., & Yu, C. (2012). Combining ranking concept and social network analysis to detect collusive groups in online auctions. *Expert Systems With Applications.*

Luna-Pla, I., & Carlock, N. (2020). Corruption and complexity: a scientific framework for the analysis of corruption networks. *Applied Network Science*.

Lyra, M., Bacao, F., & Pinheiro, F. (2021.1). Public Procurement Fraud Detection – A Review using Network Analysis. *Complex Network 2021*.

Lyra, M., Curado, A., Damásio, B., Bacao, F., & Pinheiro, F. (2021.2). Characterization of the Firm-Firm Public Procurement Co-bidding Network from the State of Ceara (Brazil) Municipalities. *Applied Network Science, 6 (1)*.

Majadi, N., Trevathan, J., & Bergmann, N. (2017). Analysis on bidding behaviours for detecting shill bidders in online auctions. *Proceedings - 2016 16th IEEE International Conference on Computer and Information Technology, CIT 2016, 2016 6th International*.

Mamavi, O., Meier, O., & Zerbib, R. (2017). How do strategic networks influence awarding contract? Evidence from French public procurement. *International Journal of Public Sector Management*.

Marsden, P. (2005). Network Analysis. *In Encyclopedia of Social Measurement.*

Merigó, J. (2016). Academic research in innovation: A country analysis. *Scientometrics 108, 559–593*.

Merigó, J., Mas-Tur, A., Roig-Tierno, N., & Ribeiro-Soriano, D. (2015). A bibliometric overview of the Journal of Business Research between 1973 and 2014. *J. Bus. Res., 68, 2645–2653.*

Mlondo, N. (2013). Effectiveness of knowing your customer policy in combating money laundering in commercial banks in Tanzania: a case of bank of Africa. *(T) Limited (Doctoral dissertation, Mzumbe University)*.

Morselli, C. (2008). Inside Criminal Networks. Studies of Organized Crime. *Springer New York*.

Morselli, C., & Ouellet, M. (2018). Network similarity and collusion. *Social Networks, 55*, pp. 21-30.

Motta, F. (2003). *Teoria das Organizações: evolução e crítica* (2 ed.). São Paulo: Pioneira Thomson Learning.

Mufutau, G., & Mojisola, O. (2016). Detection and prevention of contract and procurement, fraud Catalyst to organization profitability. *Journal of Business and Management*, pp. 9-14.

Mundra, A., & Rakesh, N. (2014). Online hybrid model for online fraud prevention and detection. *Advances in Intelligent Systems and Computing*.

Mvelase, T. (2015). The impact of non-compliance with Eskom procurement policies. *Nelson Mandela Metropolitan University - Doctoral dissertation*.

Nicolás-Carlock, J., & Luna-Pla, I. (2021). Conspiracy of Corporate Networks in Corruption Scandals. *Frontiers in Physics*.

OECD. (2021). Government at a Glance 2021. *OECD Publishing, Paris*.

Olsson, P., Folke, C., & Berkes, F. (2004). Adaptive comanagement for building resilience in social-ecological systems. Environ Manag. *Environ Manag, 34*(1), pp. 75-90. Fonte: https://doi.org/10.1007/s00267-003-0101-7

Owusu, E., Chan, A., & Wang, T. (2021). Tackling corruption in urban infrastructure procurement: Dynamic evaluation of the critical constructs and the anti-corruption measures. *Cities. Hong Kong Polytechnic University, 119. N. 103379*.

Padhi, S., & Mohapatra, P. (2011). Detection of collusion in government procurement auctions. *J. Purch. Supply Manag. 17, 207–221.*

Paganetto, L., & Scandizzo, P. (2015). Governance, moral and economic values in achieving dynamism in an anaemic Europe. *Springer, Cham*.

Perevezentceva, E., Osipova, K., & Skrynnikova, K. (2021). Corruption in the Implementation of Public Procurement from Small and Medium Businesses. *Lecture Notes in Networks and Systems*.

Popa, M. (2019). Uncovering the structure of public procurement transactions. *Business and Politics*.

Rabuzin, K., & Modrusan, N. (2019). Prediction of public procurement corruption indices using machine learning methods. *11th International Conference on Knowledge Management and Information Systems (KMIS)*.

Reeves-Latour, M., & Morselli, C. (2017). Bid-rigging networks and state-corporate crime in the construction industry. *Social Networks*, *51*, pp. 158-170.

Robson, C. (2002). Real world research: a resource for social scientists and practitioner-researchers. *Wiley-Blackwell*.

Rustiarini, N., Sutrisno, T., Nurkholis, N., & Andayani, W. (2019). Why people commit public procurement fraud? the fraud diamond view. *Journal of Public Procurement 19(4), 345–362*.

Sandström, A., & Carlsson, L. (2008). The performance of policy networks: the relation between network structure and network performance. *Policy Stud J*, *36*(4), pp. 497-524.

Sedita, S., & Apa, R. (2015). The impact of inter-organizational relationships on contractors' success in winning public procurement projects: The case of the construction industry in the Veneto region. *International Journal of Project Management*.

Shaffril, H. A., Samsuddin, S. F., & Samah, A. A. (2020). The ABC of systematic literature review: The basic methodological guidance for beginners. *Quality & Quantity, 1–28*.

Sharma, S., Sengupta, A., & Panja, S. (2019). Mapping Corruption Risks in Public Procurement: Uncovering Improvement Opportunities and Strengthening Controls. *Public Performance & Management Review*.

Silva Filho, J. B. (2017). A eficiência do controle social nas licitações e contratos administrativos. *Master's thesis - Universidade Nove de Julho, São Paulo*.

Van_Eck, N., & Waltman, L. (2014). Visualizing bibliometric networks. *In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), Measuring scholarly impact: Methods and practice (pp. 285-320). Springer.*

Velasco, R., Carpanese, I., Interian, R., & et, a. (2021). A decision support system for fraud detection in public procurement. *International Transactions in Operational Research*.

Volosin, N. (2015). Datos abiertos, corrupción y compras públicas. *RIGA - ILDA*.

Wachs, J., & Kertesz, J. (2019). A network approach to cartel detection in public auction markets. *Scientific Reports*.

Weigel, U., & Ruecker, M. (2017). e-procurement. *In: the strategic procurement practice guide*, 179-209.

Wensink, W., & Vet, M. (2013). Identifying and Reducing Corruption in Public Procurement in the EU. *European Commission. Bruxelles*.

Whiteman, R. (2019). Fraud and corruption tracker. *The Chartered Institute of Public Finance and Accountancy – CIPFA*.

Wu Chebili, B., La Cascia, H., Collineau, F., Salomon, A., Calvet, B., & Moreau, Y. (2022). Electronic government procurement implementation types.

Xiao, L., Ye, K., Zhou, J., & Ye, X. (2021). A social network-based examination on bid riggers' relationships in the construction industry: A case study of China. *Buildings*.

Xiao, Q., Chang, H., Geng, G., & Liu, Y. (2018). An ensemble machine-learning model to predict historical PM2. 5 concentrations in China from satellite data. *Environ Sci Technol, 52*(22), pp. 13260-13269.

Yu, C., & Lin, S. (2013). Fuzzy rule optimization for online auction frauds detection based on genetic algorithm. *Electronic Commerce Research*.

Železnik, D., Vošner, H., & Kokol, P. (2017). A bibliometric analysis of the Journal of Advanced Nursing, 1976–2015. *J. Adv. Nurs, 73, 2407–2419.*

Zhu, J., Wang, B., Li, L., & et, a. (2020). Bidder Network Community Division and Collusion Suspicion Analysis in Chinese Construction Projects. *Advances in Civil Engineering*.

Zhu, Y., Zhang, W., & Yu, C. (2011). Detection of feedback reputation fraud in Taobao using social network theory. *Proceedings - 2011 International Joint Conference on Service Sciences, IJCSS*.

3. CHARACTERIZATION OF THE FIRM–FIRM PUBLIC PROCUREMENT CO-BIDDING NETWORK FROM THE STATE OF CEARÁ (BRAZIL) MUNICIPALITIES

Fraud in public funding can have deleterious consequences for societies' economic, social, and political well-being. Fraudulent activity associated with public procurement contracts accounts for losses of billions of euros every year. Thus, exploring analytical frameworks that can help public authorities identify agents more susceptible to irregular activities is of utmost relevance.

This study aimed to characterize the public procurement network of the municipalities of Ceará by first evaluating fraudulent activities associated with procurement contracts in a co-bidding relationship. The lack of explicit connection information between bidders is one of the biggest challenges for government control bodies to identify collusion or the formation of cartels between bidders in public tenders in Brazil. This information is not directly contained in the data obtained from those under the jurisdiction of the Court of Auditors. However, we are able to capture it through co-bidding relationships using Network Science tools to convert a Firm-Tender network into a Firm-Firm network. Second, we used two diversity indicators, regional and contract type, as potential arrows of communities that establish some influence over the public tender. Furthermore, we analyzed the dimensions of the bids to verify the volume of single bidders in each community, besides scanning the mean competition in each tender through the average number of bidders per tender.

Here, we used standard network science methods to study the co-bidding relationships between firms participating in public tenders issued by the 184 municipalities of the State of Ceará (Brazil) between 2015 and 2019. We identified 22 groups/communities of firms with similar patterns of procurement activity, defined by their geographic and activity scopes. The profiling of the communities allowed us to highlight organizations that were more susceptible to market manipulation and irregular activities. Our work reinforced the potential application of network analysis in public policy to unfold the complex nature of relationships between market agents in a scenario of scarce data.

3.1. Introduction

Despite the weight of public procurement in governmental budgets (OECD, 2017), procurement activity is still one of the most vulnerable vehicles open to corruption (Murray, 2014; OECD, 2015). In that context, corruption can have many forms (Angulo Garzaro 2018) and occur at any point in the procurement cycle —from the pre-tendering to the tendering and the past-award phases— making it difficult to detect and measure (Mufutau & Mojisola, 2016; Rustiarini, Sutrisno, Nurkholis, & Andayani, 2019; Whiteman, 2019). In Brazil alone, corruption in procurement contracts can represent an additional cost of 20% to 30% over the expected price, which represents losses of around 200 billion Reais annually (Zeferino, 2020). Likewise, in Europe, it is estimated that losses figure around 5 billion Euros annually (Hafner, et al., 2016). Naturally, these losses undermine the ability of governments and public authorities to forge ahead with essential investments in health, education, infrastructure, security, housing, and social services (Søreide, 2002; Beittel, Meyer, Seelk, Taft-Morales, & Gracia, 2019). Unsurprisingly, there is a considerable effort to develop analytical solutions to understand and mitigate the effects of corruption in the public procurement process (Fazekas, Cingolani, & Tóth, 2018). The increasing availability of open data concerning public administration activities (Curado, Damásio, Encarnação, Candia, & Pinheiro, 2020) has recently renewed the scientific community's efforts to uncover hidden connections between participating agents and how their relationships can link to fraudulent activities (Herrera, 2019; Kertész & Wachs, 2021).

One of the most challenging aspects of identifying corruption in public procurement contracts is the lack of labeled data. Indeed, it is largely impossible to know which instances stem from corruption (Wachs & Kertesz, 2019). However, a fundamental principle in public procurement is transparency in bidding (Adjei-Bamfo P., 2019; Angulo Garzaro, 2018; Nowrousian, 2019; Spagnolo, 2012), and efficiency can be obtained through independent and open competition between firms. In that sense, past works have approached this problem from an unsupervised learning perspective, meaning that they look to extract information about the relationships between the involved parties and, thus, flag groups of agents with patterns that might be linked to a high risk of corrupting activities. Indeed, firms can achieve leverage to manipulate the tender process by establishing the right relationships among themselves and coordinating their activity (Hanák & Serrat, 2018).

An open issue remains, can communities of firms obtained from co-bidding patterns allow us to highlight groups that are more susceptible to collusion and market manipulation? The use of network analysis for the study of corruption is not new (Lauchs, Keast, & Yousefpour, 2011; Chang, 2018; Grassi, Calderoni, Bianchi, & Torriero, 2019). In the context of public procurement, past studies can be divided into two main groups: (1) works that explore bipartite relationships between public bodies and firms (Fazekas & Tóth, 2016.2; Wachs & Kertesz, 2019); and (2) studies that explore firm–firm co-bidding relationships in public tenders (Toth, Fazekas, Czibik, & Toth, 2014; Reeves-Latour & Morselli, 2017; Morselli & Ouellet, 2018; Wachs, Fazekas, & Kertész, 2021). Both approaches have their merits, and each is suitable for identifying different mechanics underlying the manipulation of the procurement process. For instance, bipartite relationships are suitable for identifying fraud stemming from bribes and influence ties, while firm–firm relationships are more suited to identifying cartels and collusion. Despite these, the use of network analysis to study the relationship between firms in procurement bids is a relatively new venture (Reeves-Latour & Morselli, 2017). More evidence is required to understand the universality of the existing patterns and mechanics across cultural and socio-economic contexts.

Here, we use network science and complexity sciences methods to map and characterize the co-bidding network (Wachs & Kertész, 2019; Piccolo, Lehmann, & Maier, 2018; Ramalho, Almeida, & Fraga, 2020; Reeves-Latour & Morselli, 2017) between firms that participated in public tenders issued by the 184 municipalities of the state of Ceará (Brazil). In that sense, we characterize the relationships between competing firms and identify the major communities of firms that often compete for tenders with a similar scope. Moreover, we argue that some such communities have characteristics that place them at a higher risk of market manipulation and irregular activities often associated with corruption.

3.2. Data

We used data from the State of Ceará Audit court authority—*Tribunal de Contas do Estado Ceará* (Brazil), covering public tenders issued by the 184 municipalities of the State of Ceará between 2015 to 2019. Each observation informs about a firm's bid to a tender and whether the bid was one of the winning bids. It also includes information about the municipality that issued the tender and whether a firm won a contract. Hence, the data is naturally represented through a bipartite nature (Fierăscu, 2017), which connects firms to tenders (see Figure 7a). The data set contains 196,608 observations that account for the bids of 45,502 firms to 84,835 tenders.

Information about the firms and tenders is anonymized, and bidding values are not available. Moreover, the data set does not contain information about which contracts/ firms have been investigated for irregularities in the past.

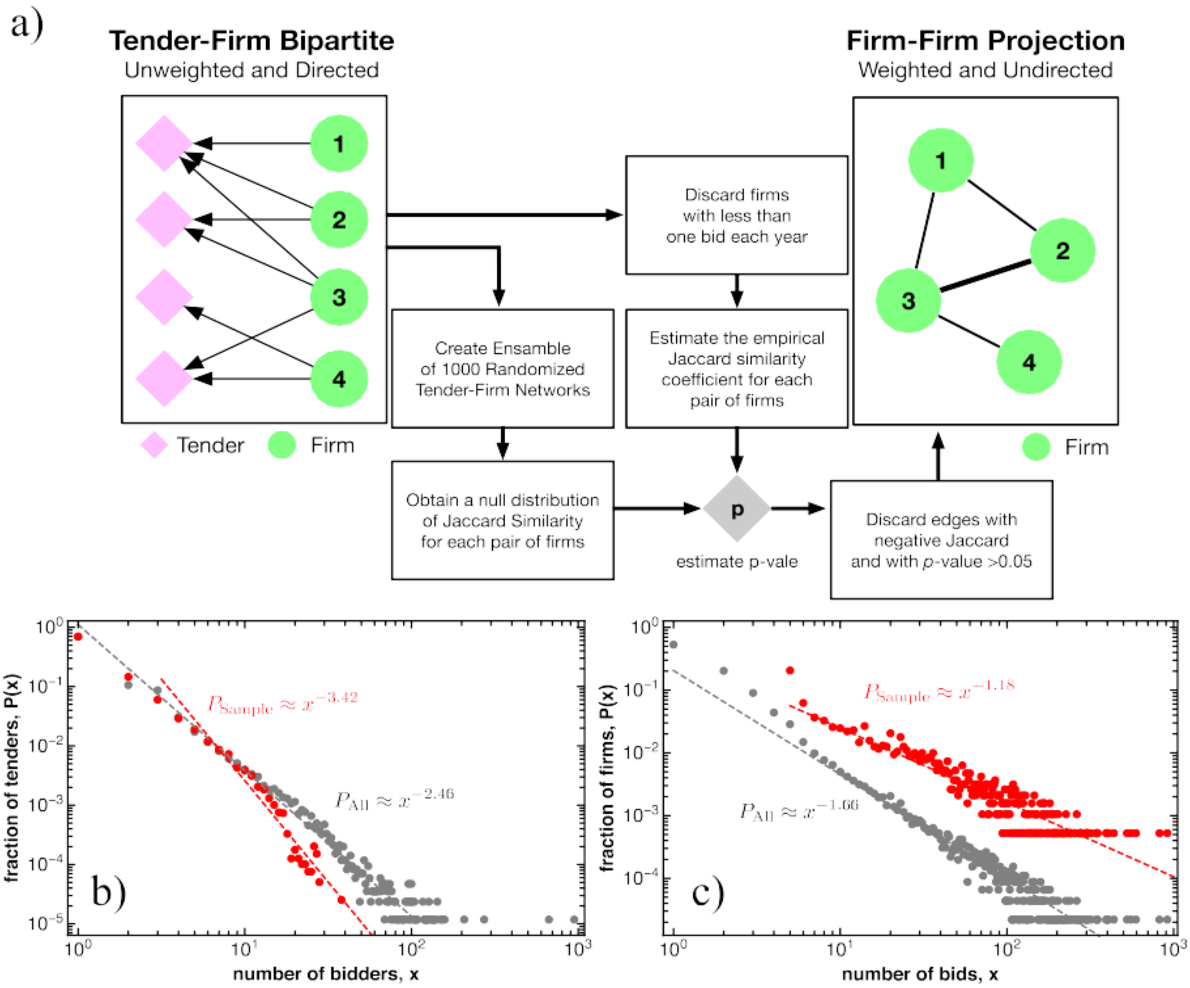

Figure 7: Panel a, graphical representation of the process employed to infer the Firm–Firm co-bidding network. Panel b, comparison between the frequency of bidders per tender in the original data set (gray) and in the working data set (red) after filters have been applied. Panel **c**, comparison between the frequency of bids per firm in the original data set (gray) and in the working data set (red) after filters have been applied. In panels (**b**) and (**c**) dashed line represents the OLS regression lines, the domain of the line indicates the domain used for fitting the curve.

3.3. Network inference

Since we are interested in studying the relationships between firms, we focus on the Firm–Firm projection. We estimated the projection from the co-bidding patterns of firms (Piccolo, Lehmann, & Maier, 2018) using the Jaccard similarity coefficient (Veech, 2013; Mainali, et al., 2017; Chung, Miasojedow, Startek, & Gambin, 2019; Wachs & Kertész, 2019). Figure 7a shows a graphical illustration of the data structure and depicts the steps conducted to infer the Firm–Firm network from the original Tender-Firm bipartite structure.

In order to infer the Firm–Firm co-bidding network, we started by discarding all firms that did not bid at least once during each year under analysis. By doing so, we were able to extract the core of active firms, while removing firms with sporadic activity. Figure 7b and

Figure 7c compare the original (PALL) with the filtered data set (PSample). In particular, it shows that filtering removes excess participants from tenders while not affecting the distribution of the number of bids done by each firm. Likewise, we refer to firms present in the firm–firm co-bidding network as Established firms. The final working data set includes 1906 firms, which account for 72,078 bids to 39,523 tenders. Hence, next, we compute the centered Jaccard/Tanimoto coefficient (Chung et al. 2019) between each pair of firms. The centered Jaccard coefficient measures the similarity between the bidding of two firms, accounting for the occurrence probabilities of each firm. Formally, it can be computed as:

$$J_{ij}^{c} = \frac{\sum_k b_{it} b_{jt}}{\sum_k (b_{it} + b_{jt} - b_{it} b_{jt})} - \frac{p_i p_j}{p_i + p_j - p_i p_j}, \tag{1}$$

Where k is the union set of the number of tenders that firm i and firm j participate, b_{it} is one if firm i made a bid to tender t, being zero otherwise, b_{jt} is one if firm j made a bid to tender t, being zero otherwise; and p_i is the fraction of tenders in which firm i participated ($p_i = \sum_k b_{it}$). The second term in Eq. (1) provides the expected number of observations when the bids from both firms are independent and identically distributed through a Bernoulli process (Chung, Miasojedow, Startek, & Gambin, 2019). Hence, the centered Jaccard coefficient allowed us to distinguish between positive and negative associations between firms, accounting for their individual level of activity.

Finally, we estimate the significance of the observed J^c_{ij}, to test the hypothesis that $J^c_{ij} > 0$. To that end, we bootstrapped a null distribution (J_{ij}) of centered Jaccard coefficient for each pair of firms by generating an ensemble of 1000 randomizations of the initial bipartite network. Data were randomized in order to ensure that the number of bids observed per firm and per year remained constant while preserving the number of firms bidding to each tender. Then, we estimate the one-tailed *p*-value associated with J^c_{ij} by calculating the upper tail probability of obtaining a value equal or greater than J^c_{ij} from the cumulative frequency of the null-distribution J_{ij} (Gotelli, 2000). Links with a *p*-value greater than 0.05 were discarded.

The resulting firm–firm co-bidding network contains 1529 nodes and 12,892 edges. Relationships are treated as undirected and unweighted. The network exhibits an average degree of 16.86, with a cluster coefficient of 0.52 (Newman, Cantwell, & Young, 2020), and 56 connected components. Figure 8 shows that the Degree Distribution (Figure 8a) decays exponentially with the degree, which suggests that the underlying mechanics of co-bidding can be approximated by a random attachment process (Albert & Barabás, 2002). Alternative distribution (e.g., power-law) has been tested but shown the worst likelihood given data. However, the average clustering coefficient shows a power-law inverse relationship with the degree (Figure 8b), suggesting the existence of some level of hierarchy in the structure of the network. It is noteworthy to mention that the largest connected component contains 1141 nodes, 10,630 edges, and a clustering coefficient of 0.43.

3.4. Results and discussion

Figure 9 presents the giant component of the firm–firm co-bidding network. Using the Louvain algorithm (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008) we identified 22 communities with modularity of 0.66. We refer to these communities as C1, C2, . . . , C22, and they are indexed in descending order with respect to their size. For readability, we have colored the eight largest communities in Figure 8. However, the high modularity of the network is unsurprisingly and can be explained by the fact that the network primarily represents competing firms that are specialized in supplying different services–works, goods, services, etc – in different regions. Hence, the most prominent communities divide the network into two major groups of firms that operate mainly in the northern (Red, Blue, and Green) and southern (Purple and Yellow) regions of the state of Ceará but also on contracts that supply Food services (Blue and Purple) or construction works (Red and Yellow).

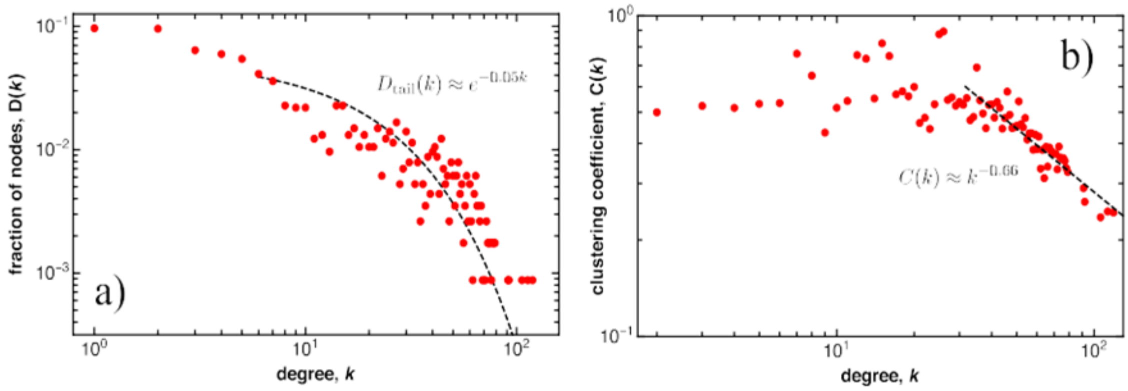

Figure 8: Panel **a** shows the degree distribution, $D(k)$. The dashed black line represents the best exponential distribution fitted to the empirical distribution's tail ($k > 5$). Panel **b** shows the average clustering coefficient average per degree, $C(k)$. The dashed black line shows the best linear fit. The results have been estimated from the entire graph.

Interestingly, the remaining communities operate at a state-wide level (Pink and light Blue), and one particular community (Violet) operates exclusively in the mesoregion of Jaguaribe and only supplies Food services.

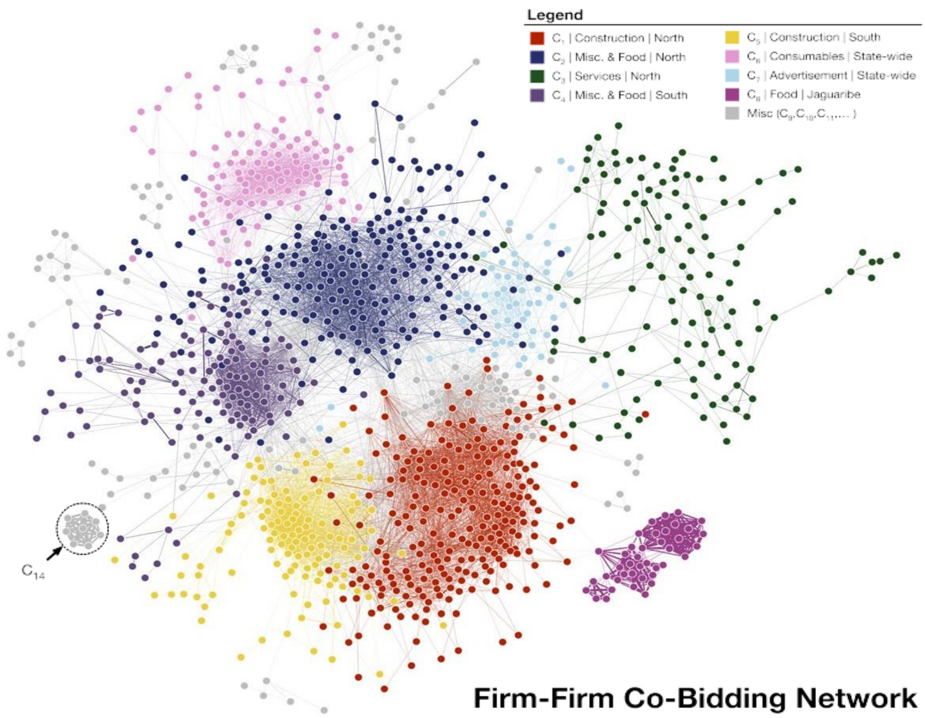

Figure 9: Graphical representation of the Firm–Firm network, which relates firms with similar bidding patterns. In order to build the network, we only considered the most active firms and edges with a significant Jaccard similarity index. Represented is the giant component with some relevant disconnected components. Nodes in the giant component are colored according to one of the eight major communities (out of 22) identified using the Louvain algorithm (Blondel et al. 2008). The presented partition of the network achieved a modularity of 0.61. North and South categorization of the firms location was used here in a broader sense and should not be confused with the Mesoregions, which we identified with their Portuguese names. For reference, the North includes the Mesoregions of Norte and Noroeste; while the South includes Sul and Centro Sul

In some cases, firms form densely connected subgraphs (e.g., C14). These structures can be a first indicator to flag groups of firms that present a high risk of collusion and procurement manipulation. As such, next, we explore possible additional metrics to classify each community of firms through their activity in order to narrow down further which groups of firms might deserve a more profound investigation by audit officials.

3.5. Activities diversity

We started by looking at the regional diversity on which firms performed their activities (e.g., bid on tenders in order to supply services) and the diversity of the type of contracts for which they bid. While a firm with low diversity in both regional reach and contract type can simply indicate a firm that is narrow in both scope and domain, the existence of groups of connected firms (i.e., a community of firms) that share a low diversity in both dimensions can highlight a more troublesome scenario. In particular, it can indicate the conditions for firms to coordinate and cooperate to control a specific market and regional context and should be investigated with further discernment.

To that end, we estimate the Simpson's diversity index1 for each community. The Simpson's diversity index (λ) measures the probability that two randomly sampled elements from a set share a given characteristic in common. In that sense, $\lambda = 0$ is associated with the highest diversity possible, and $\lambda = 1$ with the lowest diversity. Formally we estimate the Simpson's index for each community as

$$\lambda_{\text{cat}}^{C_i} = \sum_{t \in \gamma} (p_{C_i}^t)^2 \qquad (2)$$

where p_{Ci}^{t} corresponds to the fraction of bids done in a procurement contract type γ :{Consumables Health, Services, Construction, Events, Food, Fuel,…} or mesoregion γ :{Metropolinana, Norte, Sul, Noroeste,…} by the firms in community Ci, $\forall i \in \{1, 2, \ldots, 21, 22\}$. The quantity ptci is normalized per community so that $\sum_t p_{Ci}^{t} = 1.0$. We estimated $\boldsymbol{\lambda}_{cat}^{Ci}$ independently for each community (Ci), and for contracts according to the region that issue the tender and the tender contract type (e.g., services, food, tenancy, construction, etc). Our choice of Simpson's index over other alternatives (e.g., entropy) is due to its straightforward interpretation in our context: the probability that two bids made by firms within the same community share the same characteristic (e.g., region or contract type).

Figure 10a illustrates the empirical distributions (p_{Ci}^{t}) of procurement activity for the ten most prominent communities. We show the results for both the Regional distribution of activities and by Contract Type. Blue colors denote a low relative frequency of bids, while red identifies a high frequency. These indicators allowed us to infer the degree of specialization and agglomeration of a community. In particular, we found that Community 8 (C8) activities are agglomerated in a single region (Jaguaribe), and firms specialize in one type of contract (Food). The same conclusion can also be inferred from the high levels of $\boldsymbol{\lambda}_{cat}^{C_8}$, which means that Community 8 has low diversity of activity distribution. Figure 10b compares all 22 communities in terms of the two diversity indicators defined above. We find a clustering of communities in the bottom left quadrant—a low level of agglomeration and specialization—that we associate with healthy markets composed of firms that, on average, have a diversified portfolio of activities and regional distribution. In contrast, in the top right quadrant, we found communities that relied on procurement contracts of a single type and agglomerated in a small number of regions.

The combination of these two diversity indicators, at the community level, provides a powerful feature to identify groups of firms that can dominate over a niche market or, in the worst case, develop undesirable leverage, as a group, in negotiating procurement contracts. Hence, lowering the desirable efficiency that public procurement aims at achieving in the tendering process. However, it is important to stress that these metrics are just indicative of potential problems, and thus the true nature of the activities of the firms in each community should be carefully investigated by the corresponding local authorities.

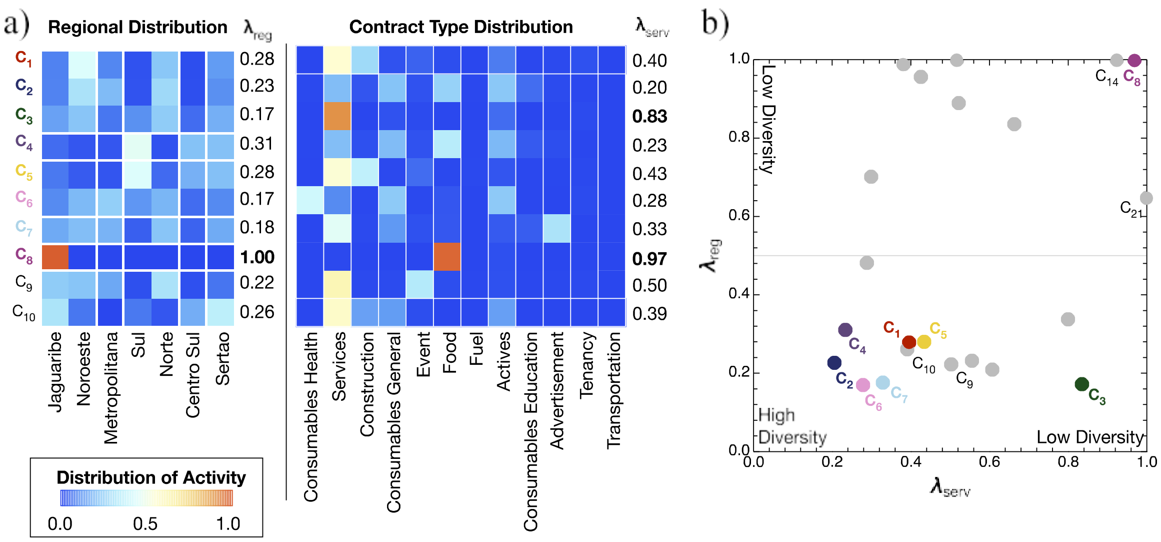

Figure 10: Characterization of the ten largest communities by the diversity of bids done by region and type. Panel **a** shows the distribution of bids within each community by mesoregion (left) and contract/tender type (right). For each community we compute the Simpson's diversity index (λ_{reg} and λ_{serv}). The full and official names of the mesoregions are: Jaguaribe; Noroeste Cearense (Noroeste); Metropolitana de Fortaleza (Metropolitana); Sul Cearense (Sul); Norte Cearense (Norte); Centro-Sul Cearense (Centrol Sul); and Sertoes Cearenses (Sertão). We use simplified references to these names for visualization purposes. Panel b compares communities by their diversity of contracts in terms of regional span and type. Note that in panel a we only show results for the ten largest communities, which are representative of the results. Communities not identified by a color code in Figure 9 are shown in gray in panel **b**, in the particular community C14 corresponds to the gray clique easily identifiable in the bottom left of the network in Figure 9

3.6. Bidding coordination

To further investigate the risk/susceptibility of market manipulation by firms, we next looked at each community's propensity to participate in "single bidder" contracts. Another pattern often associated with corruption and loss of efficiency. Hence, what is the susceptibility of each community to such practice? To answer this question, we started by investigating the average number of times per community that a firm is the single bidder of a tender. Figure 11a shows the results for all 22 communities in the most significant component of the Firm–Firm network. Traditionally, a high level of single bids can be an indicator of firms acting with some level of informal advantage in the tendering process or due to a lack of competition in a specific market. At the community level, such an indicator can be indicative of unusual activity from a group of firms. Hence, low levels of single bidding indicate the risk of coordination (e.g., firms participating coherently in the same contracts), while high levels can sign the prevalence of less competitive markets or informal advantage in the tendering process. Overall, of the largest ten communities, only Community 8 exhibits low levels of single bidders, a pattern that also extends to Communities 14 and 21. In contrast, we saw that community 12 strongly deviates from the baseline with an average value of single bidding that is roughly four times that of a typical firm.

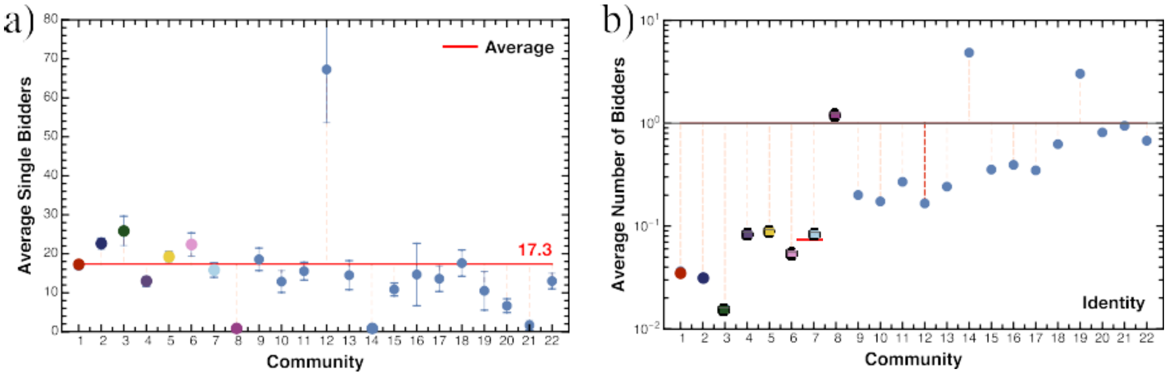

Figure 11: Characterization by bidding activity. Panel **a** shows the average number of single bids per community (i.e., the average number of times a firm in a community participated in a tender as the single bidder). We compared the values of each firm with the average of the entire population of firms (horizontal red line). Panel **b** shows the average number of bidders in tenders in which firms within a community typically participate. We normalized the value obtained for each community by the number of firms in that community. The horizontal red line shows the threshold that marks the size of the community

In addition, we looked at the average number of bidders per tender in order to assess the potential existence of coherent behavior, that is, coordination between the firms in a community. To that end, we estimated the average number of bidders per tender for each community, which we normalized by the community size (i.e., the number of firms in a community). Interestingly, Figure 11b shows that in Community 8, firms tend to participate in tenders with several firms that match almost exactly the community's size. While, in some cases—Communities 14 and 19—firms tend to bid to tenders that are several times larger than their communities. Noteworthy to mention that this analysis is biased by the size of the communities, so the expectation would be to see a smoothly increasing relationship, with the largest community achieving the smallest value, and in the limiting case of a community with a single firm, we would obtain the maximum. However, it is clear that, in some cases there are apparent deviations — Communities 8, 14, and 19.

3.7. Conclusions

In this manuscript, we explored the potential of mining a large data set of public tenders collected from firms' activity to compete for procurement contracts issued by the municipalities of the State of Ceará (Brazil). By matching firms with similar bidding patterns, we have inferred a firm–firm network comprising a total of 1141 nodes and 10,630 edges.

We showed that we were able to identify communities of firms with similar bidding patterns. The network exhibits a high modular structure partitioned in 22 communities. These communities cluster firms that have a similar scope in procurement activities both in the nature of the contracts they celebrated and in the regional reach of their activities. Moreover, we looked at two diversity indicators—regional diversity and procurement contract nature diversity—as a sign of the potential of certain communities to develop leverage over the procurement process. In other words, in affecting the expected efficiency of the market. Finally, we looked at the sizes of the tenders, first by looking at the abundance of single bidders in communities, and secondly by looking at the average number of bidders in each tender. Overall we identified a particular community (Community 8) that combines several undesirable properties. Community 8 involves a group of firms that offers Food services in the region of Jaguaribe. They have an unusually low number of single bids; the average number of participating firms per tender matches the number of firms in the community, and they exhibit a high specialization and agglomeration in their activities. Nevertheless, having said that, such an odd combination of characteristics can be useful to narrow down the activities of audit officials, but it does not allow us to conclude much about the true nature of the activities of firms involved in Community 8.

Finally, it is essential to highlight some shortcomings in our analysis and future working directions. The lack of pre-labeled data on past corruption cases significantly limits our ability to make any causal link between the network structure, its motifs, and the location of firms in the network with irregular procurement behavior. In that sense, our results are merely exploratory and show the potential of combining network science methods with descriptive statistics to highlight relevant groups of firms according to their activity pattern in a data-scarce environment. Future works should look at the evolution of the network, that is, if a larger temporal window is available, to capture the evolution and segregation of communities of interest but also of their parametric path in terms of the diversity of their activities.

References

Adjei-Bamfo P., M.-N. T. (2019). The role of e-government in sustainable public procurement in developing countries: a systematic literature review. *Resour Conserv Recycl*, p. 142:189. doi:203

Albert, R., & Barabás, A.-L. (2002). Statistical mechanics of complex networks. *Rev Mod Phys, 30*(15). doi:47

Angulo Garzaro, N. (2018). Eu competition law and public procurement: competition-driven limits imposed to public bodies when they source goods, works and services. *Works and Services*.

Beittel, J., Meyer, P., Seelk, C., Taft-Morales, M., & Gracia, E. (2019). Combating corruption in Latin America: congressional considerations. *Congressional Research Service, Report number R45733*.

Blondel, V., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J Stat Mech: Theory Exp, 2008*(10). doi:10008

Chang, Z. (2018). Understanding the corruption networks revealed in the current Chinese anti-corruption campaign: a social network approach. *J Contemp China, 27*(113), pp. 735-747.

Chung, N., Miasojedow, B., Startek, M., & Gambin, A. (2019). Jaccard/Tanimoto similarity test and estimation methods for biological presence-absence data. *BMC Bioinform, 20*(15), pp. 1-11.

Curado, A., Damásio, B., Encarnação, S., Candia, C., & Pinheiro, F. (2020). Scaling behavior of public procurement activity. *arXiv preprint arXiv:2007.15276*.

Dong, S., Wang, H., Mostafizi, A., & Song , X. (2020). A network-of-networks percolation analysis of cascading failures in spatially co-located road-sewer infrastructure networks. *Physica A: Statistical Mechanics and its Applications, 538, 122971*.

Fazekas, M., & Tóth, I. J. (2016.2). From corruption to state capture: A new analytical framework with empirical applications from Hungary. *Political Research Quarterly, 69 (2), 320–334.*

Fazekas, M., Cingolani, I., & Tóth, B. (2018). Innovations in objectively measuring corruption in public procurement. *Governance indicators. Approaches, progress, promise*, pp. 154-185.

Fierăscu, S. (2017). The networked phenomenon of state capture: network dynamics, unintended consequences, and business-political relations in Hungary. *Central European University, Budapest*.

Gotelli, N. (2000). Null model analysis of species co-occurrence patterns. *Ecology, 81*(9), pp. 2606-2621.

Grassi, R., Calderoni, F., Bianchi, M., & Torriero, A. (2019). Betweenness to assess leaders in criminal networks: New evidence using the dual projection approach. *Social Networks, 56 , 23–32.*

Hafner, M., Taylor, J., Disley, E., Thebes, S., Barberi, M., Stepanek, M., & Levi, M. (2016). The cost of non-Europe in the area of organised crime and corruption: annex II-corruption. *RAND Corporation, Santa Monica, CA.*

Hanák, T., & Serrat, C. (2018). Analysis of construction auctions data in Slovak public procurement. *Adv Civ Eng*, pp. 1-13.

Lauchs, M., Keast, R., & Yousefpour, N. (2011). Corrupt police networks: uncovering hidden relationship patterns, functions and roles. *Policing & Society 21 (1), pp.110-12.*

Mainali, K., Bewick, S., Thielen, P., Mehoke, T., Breitwieser, F., Paudel, S., . . . Karig, D. (2017). Statistical analysis of co-occurrence patterns in microbial presence-absence datasets. *PLoS ONE, 12*(11). doi:0187132

Morselli, C., & Ouellet, M. (2018). Network similarity and collusion. *Social Networks, 55*, pp. 21-30.

Mufutau, G., & Mojisola, O. (2016). Detection and prevention of contract and procurement, fraud Catalyst to organization profitability. *Journal of Business and Management*, pp. 9-14.

Nowrousian, B. (2019). Combatting public procurement criminality or simple rules for complex cases. *J Financ Crime, 26*, pp. 203-210.

OECD. (2015). *Preventing corruption in public procurement.* Fonte: OECD.org: https://www.oecd.org/gov/public-procurement/integrity/

Piccolo, S., Lehmann, S., & Maier, A. (2018). Design process robustness: a bipartite network analysis reveals the central importance of people. *Des Sci, 4:e1*.

Ramalho, H., Almeida, A., & Fraga, A. (2020). Detecção de Casos Suspeitos de Conluio em Licitações Públicas: uma Aplicação do Algoritmo a Priori de Aprendizado de Máquina para o Estado da Paraíba. *Teoria e Prática em Administração, 10*(2), pp. 5-22.

Reeves-Latour, M., & Morselli, C. (2017). Bid-rigging networks and state-corporate crime in the construction industry. *Social Networks, 51*, pp. 158-170.

Rustiarini, N., Sutrisno, T., Nurkholis, N., & Andayani, W. (2019). Why people commit public procurement fraud? the fraud diamond view. *Journal of Public Procurement 19(4), 345–362*.

Søreide, T. (2002). Corruption in public procurement. Causes, consequences and cures. *Chr. Michelsen Institute*.

Toth, B., Fazekas, M., Czibik, A., & Toth, I. (2014). Toolkit for detecting collusive bidding in public procurement with examples from Hungary. *Corruption Research Center Budapest (CRC-WP/2014:02)*.

Veech, J. (2013). A probabilistic model for analysing species co-occurrence. *Glob Ecol Biogeogr, 22*(2), pp. 252-260.

Wachs, J., Fazekas, M., & Kertész, J. (2021). Corruption Risk in Contracting Markets: A Network Science Perspective. *International Journal of Data Science and Analytics*.

Wachs, J., Yasseri, T., Lengyel, B., & Kertesz, J. (2019). Social capital predicts corruption risk in towns. *Royal Society Open Science*.

Whiteman, R. (2019). Fraud and corruption tracker. *The Chartered Institute of Public Finance and Accountancy – CIPFA*.

World_Bank_Group. (2017). A fair adjustment: efficiency and equity of public spending in Brazil. *Brazil public expenditure review.[S. l.]: World Bank Group, 1*.

Wu Chebili, B., La Cascia, H., Collineau, F., Salomon, A., Calvet, B., & Moreau, Y. (2022). Electronic government procurement implementation types.

Xiao, L., Ye, K., Zhou, J., & Ye, X. (2021). A social network-based examination on bid riggers' relationships in the construction industry: A case study of China. *Buildings*.

Xiao, Q., Chang, H., Geng, G., & Liu, Y. (2018). An ensemble machine-learning model to predict historical PM2. 5 concentrations in China from satellite data. *Environ Sci Technol, 52*(22), pp. 13260-13269.

Yu, C., & Lin, S. (2013). Fuzzy rule optimization for online auction frauds detection based on genetic algorithm. *Electronic Commerce Research*.

Zeferino, L. (2020). A corrupção na construção de edifícios públicos no brasil: análise de instrumentos inibidores e facilitadores na etapa de projeto arquitetônico. *Universidade de De Brasília – UnB Master's thesis*.

4. FRAUD RISK ASSESSMENT IN PUBLIC PROCUREMENT ACTIVITIES - AN INVESTIGATION IN TWO DIFFERENT GOVERNMENTS OF CEARA'S MUNICIPALITIES

In chapter 4, we investigate possible fraudulent relationships between public and private entities participating in procurement processes and analyze the evolution of the risk of corruption within these processes in two different governments. Our contribution proposes an application of the Corruption Risk Indicator (CRI) to the Brazilian procurement landscape, targeting binary relationships between organizations and identifying red flags at the contract level. Furthermore, we identified groups of firms and municipalities with varying levels of corruption risk, categorizing them into four risk bands. Additionally, we have conducted a thorough analysis by cross-referencing the 7 mesoregions of Ceará with the main auctioned contract type to determine their respective CRIs, variations, and vulnerability to fraud within each government. Lastly, we cross-plotted the CRIs of each municipality on the map of Ceará, allowing us to visualize and compare their magnitudes through color differentiation in the context of two sequential governments.

4.1. Introduction

Public procurement is an administrative process used by public administration to acquire services and goods and develop construction works. It holds significant importance in government budgets, accounting for over 16% of the European Gross Domestic Product (GDP) (Ratcliff, Martinello, & Kaiser, 2022), about 14% of the Brazilian GDP (Rauen & Paiva, s.d.), and 4.18% of the Portuguese GDP in 2019 (IMPIC, 2018). One major challenge to the execution and management of public procurement contracts is fraud and corruption. In Brazil irregularities in public procurement contracts is estimated to result in an additional cost of up to 30% and in a waste of approximately 200 billion reals[1] per year (Zeferino, 2020). Despite considerable effort in developing analytical solutions to understand the phenomena better and develop solutions to mitigate its impact, there is still a lack of studies focusing on fraud detection in this sector, especially concerning the use of objective fraud indicators (Lyra, Curado, Damásio, Bacao, & Pinheiro, 2021.2).

Estimating corruption risk in public procurement has been a widely debated topic since the introduction of the *Corruption Perceptions Index* (CPI) in 1995 (Transparency

[1] 40 billion US dollars

International, 2022), which relies on the analysis of surveys, legal procedures, and audits (Clark, 2017). However, such indicators often present a low level of confidence as they are primarily based on individual perceptions, making their use for fraud identification somewhat unreliable (Roca, Orme, & Brown, 2010). Leveraging the increasingly available open datasets on public procurement contracts across EU member states, Fazekas et al (2016.1) proposed the *Corruption Risk Index* (CRI) to estimate the risk of corruption at the contract level.

Here, we build on the same principles to develop a CRI indicator tailored to the case of Brazil (Bauhr, Czibik, Licht, & Fazekas, 2019). Using intrinsic parameters of the bidding processes in Brazil and drawing a comprehensive understanding of the Brazilian bidding rules and the perception (from an auditor's point of view) of how fraudulent activities can occur in public tenders. To that end, we use objective micro-level data on bidding processes and contrast temporal information across the various relationships between public and private entities on contracts issued by the municipalities of the state of Ceará. Investigating transactions and grouping among allegedly fraudulent entities allows for estimating the level of corruption within a relational network (Wachs & Kertesz, 2019). As such, our analysis focuses on measuring the risk of fraud in public procurement networks at the micro-level of interaction, that is, between two entities. Moreover, we investigate the evolution of the corruption risk over an eight-year period in two separate networks, exploring the change of government in the municipalities of Ceará from 2016 to 2017.

The remaining manuscript is organized as follows: Section 2 details the dataset used in this study, providing a detailed description of the steps used to clean and pre-process the data, Section 3 elucidates the methodologies utilized, while Section 4 delves into examining of red flags. Subsequently, Section 5 describes the regression specification. Furthermore, Section 6 presents the findings, whereas Section 7 identifies quartiles indicative of varying levels of corruption risk. Finally, Section 8 offers conclusive remarks, discussing the limitations inherent in this study and outlining potential avenues for future research.

4.2. Data

We used data shared by the *Tribunal de Contas do Estado do Ceará* [2] (Brazil) that informs on tenders carried out by 184 municipalities from Ceará over eight years between 2013 and 2020. The data was segmented into two periods of consecutive governments, the first one referring to the years between 2013 and 2016, and the second one covering 2017 to 2020. The

[2] Court of Auditors of the State of Ceará

Court receives information on all public procurement contracts carried out in the municipalities and maintains a permanent database. The data details participating agents (municipal tenders and bidding firms), tender references, dates, quantities, and values of goods and services that a public administration body intends to contract. However, the dataset does not contain labeled data identifying which agents have previously faced sanctions for irregularities. Overall, the dataset contains 355,401 connections (links between a firm and a tender or the bids made by firms in tenders) involving 68,812 companies bidding on 154,972 tenders or tender items.

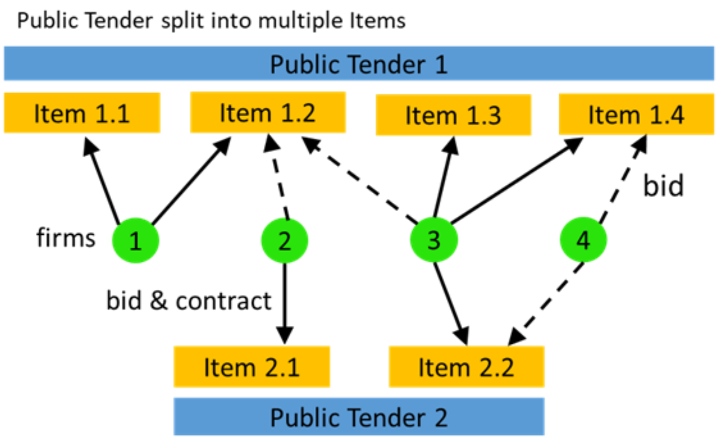

Figure 12. Illustration of the structure of a Public Tender (blue). A proposal can consist of several items (orange). Companies (green circles) bid on specific items, which can win and generate contracts (full lines) or lose (dashed lines). The dataset tells us whether a firm has bid for a public tender. Information about the items won is only available to the companies that submitted the winning bid.

It is noteworthy to mention that the public tenders under analysis in this study adhere to a set of premises, Figure 12:

1. Tenders are issued by a municipality (Public Institution);
2. Each tender may correspond to multiple items;
3. Each item may have various bidders (firms);
4. A contract identifies the winning bid and bidder of each item;
5. Each item will have only one winning firm;
6. A firm might bid on different items in the same tender;
7. A firm can win multiple items, including winning all items in a tender.

In that sense, we treated each tender item as an independent tender, as each item has its own set of competitors and a single winner. Each item within a tender is evaluated separately, as we cannot assume that the winner of one item automatically wins the entire tender unless the tender consists of only one item. This approach enabled us to establish a clear association between each item or tender and its corresponding winner.

4.3. Methods

Following past works (Fazekas, Toth, & King, 2016.1; Charron N. , Dahlström, Fazekas, & Lapuente, 2017), we use the methods proposed by Fazekas et al. (2016) to estimate *Corruption Risk Index* (CRI) at institutional-level — winning firms and municipal tenders — based on performance variables intrinsic to bidding processes applied to Brazil. The CRI is measured by weighting the regression coefficients derived from multiple models. These weights reflect the proportional influence of each red flag's risk on the composite index.

To that end, we leverage the knowledge obtained from field audits, interviews with external control auditors, and the examination of audit reports spanning from 2013 to 2020. These investigations yielded a list of nine potential *Red Flags* divided into two groups. The first focused on the bidding competition or lack thereof, as indicated by cases with single bidders during the bidding stage. The second group involved examining the winner's market shares in terms of volume and value, capturing instances of recurring contracts. Table 12 presents the parameters recognized as red flags that will be used to estimate the CRI.

To assemble the CRI, we apply weights based on the regression coefficients derived from one logistic regression and two linear regressions. These weights reflect the proportional influence of each red flag's risk on the composite index. The determination of weights for primary red flags, such as the dependent variables Competitiveness and Market Share, is based on their high degrees of influence on corrupt activities as revealed in audits conducted by the Court of Auditors. Therefore, for the primary red flags, Single Bidder, Market Share Volume, and Market Share Value are assigned the highest regression weighted coefficient found during the regressions, Table 13.

The independent variables or secondary red flags and their items are assigned values based on their specific regression weighted coefficient. Once the weights are determined, they are normalized across the nine red flags and their subitems ensuring that the maximal sum of weights equals 100%. To calculate the normalized weighted coefficient, we considered only the highest weighted regression coefficient among all items of each red flag. This means that for each red flag group, we use only the regression coefficient of the item with the greatest magnitude in the normalizing denominator. This normalization process aims to ensure that each red flag's contribution to the overall CRI estimate is appropriately scaled based on the significance of each red flag group. Subsequently, the values of the red flags are summed according to their respective normalized weights to calculate the *Corruption Risk Index* (Fazekas, Toth, & King, 2016.1).

The CRI, estimated from a set of red flags and ranging from 0 to 1, represents the anticipated likelihood of fraudulent contract awards in public procurement and reflects the level of corruption in each transaction between a municipal tender (tender) and a winning firm (winner), estimated from a set of red flags. It provides insights into the contribution of each red flag in determining a corrupt contract. The CRI is defined as follows:

$$\mathrm{CRI_i} = \sum_j W_{ji}\beta_{ji} \tag{3}$$

where $\mathrm{CRI_i}$ represents the corruption risk index of bid *i* between the transaction of a tender and a winner, β_{ji} denotes the j^{th} Red Flag observed in bid i, and W_{ji} indicates the weight of the Red Flag j. Moreover, and by definition, $\sum_j W_{ji} = 1$. The CRI of the entities is obtained by calculating the average of their CRI transactions over a given period.

The CRI values for tenders and winners are obtained by averaging their respective CRI transactions over a period. This comprehensive measure of corruption risk considers multiple red flags and their weighted contributions, providing a reliable indicator for assessing the potential risk of corruption in public procurement processes.

4.4. Red Flags

The research incorporates nine red flags that serve as parameters for detecting irregularities by identifying fraudulent attempts to eliminate competition. These red flags were identified through assessments of municipal accountability reports and interviews with internal control auditors (Tribunal de Contas do Estado do Ceará, 2022). Furthermore, in studies by Fazekas et al. (2016.1), the use of such red flags have proven to be significant predictors in the regression analyses, suggesting logical and linear correlations with fraud in bidding processes. Table 12 summarizes the nine red flags used as inputs for forming the Corruption Risk Index (CRI).

Table 12: List of red flags to assemble the CRI

RED FLAG	DEFINITION	IMPACT
Competitiveness	**Number of competitors on each issuer's tender** a) Single bidder	Single bidder contracts facilitate issuers continually awarding contracts with the same well-connected company. The greater the competitiveness, the higher the probability of achieving fair and unbiased competition. However, It must be analyzed together with the company's market share since a high market share of a firm with an exclusive issuer can lead to corruption regardless of the high competitiveness in the tender.
Tender Modality	**Modality used in the bidding process** a) Invitation letter b) Price taking c) Waived d) Auction e) Competition f) Other	Less demanding, less competitive, and less transparent modalities generate better chances for customers to frequently award contracts to the same colluding firm.
Market share value	**Winner's share of the total monetary value of the issuer's contracts.** Continuous values	It represents the percentage of participation of a given company in the issuer's bids, both in terms of value and quantity. It also indicates how much the issuer is diversifying the winning companies.
Market share volume	**Winner's share of the total number of the issuer's contracts.** Continuous values	
Threshold value for bidding exemption	**The limit value for waiving a bidding process** a) Below or equal waiving limit b) Below or equal 10x waiving limit c) Below or equal 100x waiving limit d) Below or equal 1000x waiving limit e) above 1000x waiving limit	Repeated awards from the same issuer with values below the waived modality threshold increase the possibility of waived tenders and fractioning of expense, weakening competition and allowing the issuer to frequently award contracts to a well-connected company.
Contract award time	**Time-lapse for awarding a contract** a) Award <= 30 days b) 30 days < Award <= 90 days c) 90 days < Award d) Award Missing	Reduced or extended judgment periods may indicate a legal challenge to the bidding processes, suggesting that the issuer may wish to restrict competition by awarding the contract to an allied player.
Contract amendment	**Addendum or change of clauses in the contract during its regular term** a) No contract addendum b) Contract addendum	Not being able to eliminate competition, a well-connected company can still win a competitive bid at unattractive values, but later modifications during the contract term still allow it to earn an extra profit. These modifications suggest that the issuer may be corruptly favoring a well-connected company repeatedly.
Contract extension	**Extension of the contract period after the end of the regular term without a new bidding process.** a) No contract extension b) Contract extension	An extension of the contract without rebidding after the end of the term could suggest that the issuer may be corruptly favoring a well-connected company because it could not eliminate competition at the primary bidding event.
Contract duration	**Initial contract duration in days** a) Contract duration = 0 days or Immediate delivery b) days < Contract duration <= 180 days	Contracts with a life longer than one year represent less than 1% of cases and can be seen as an opportunity to ensure well-connected companies a longer duration for a contract with characteristics of corruption.

	c) 180 days < Contract duration <= 366 days d) Contract duration > 366 days or missing	

One primary red flag used to identify corruption in public procurement processes is competitiveness, with a focus on the presence of single bidders. This pattern is often associated with corruption and loss of efficiency in bidding processes (Lyra, Bacao, & Pinheiro, 2021.1). Observing the occurrence of a unique connection between a municipal tender and the winning firm reveals the absence of competition, since the proposal receives only one offer, indicating a lack of competitiveness in a public contract (Amaral, Saussier, & Yvandre-Billon, 2013). However, it is worth noting that this red flag may wrongly indicate high corruption risk in tenders where the participation of a single bidder is necessary and allowed, such as in exclusive bidding situations or cases of emergencies or public calamities. Although a single bidder may indicate a high risk of corruption in these specific scenarios, it does not necessarily imply any fraudulent activity (Goryunova, Baklanov, & Ianovski, 2022).

Another primary red flag and dependent variable focus on the market share of winning companies, indicating the firm's success rate in winning tenders from the same issuer or municipality (Heggstad & Heggstad, 2011). A high market share suggests a higher likelihood of collusion between public and private entities. This red flag was further divided into two components: Market Share Value, representing the ratio between the contract value received by the winner and the total value of contracts awarded by the issuer, and Market Share Volume, which captures the ratio between the number of contracts won by a bidder and the total number of contracts awarded by the issuer. While market share volume helps identify smaller bidders with a high win rate and potentially engaging in fraudulent activities, it may underestimate corruption when a corrupt network employs multiple companies to extract bribes.

The data collected reveals significant variations in the market share of the companies. The firm holding the largest market share in the municipalities of Ceará over eight years presented a market share of 18.3% in value and 2.9% in volume of tenders. On the other hand, the firm that won the largest volume of bids, representing 6.7% of the total in a given municipality, had a market share value of 1.6%. A significant difference between the market share volume and market share value suggests the possibility of fractioning expenses, a practice where expenses are divided among multiple contracts to distribute the financial impact. This approach allows for the dilution of high values typically associated with more demanding tender modalities and displays a more balanced distribution of financial resources

In addition to primary red flags, we consider secondary red flags used as independent variables in the regression analysis. One such red flag is the tender modality adopted in bidding processes in Brazil, which can either facilitate or hinder the occurrence of irregularity (Speck & Ferreira, 2012). Different corruption risk levels are assigned to each modality, with stricter procedures, greater transparency, and more competition posing lower risks. On the other hand, more flexible, less transparent, and less competitive modalities carry higher fraud risks. The study considers Brazil's five most-used bidding modalities: Invitation Letter, Price Taking, Auction, Competition, and Waiver. It is worth noting that, based on reports and interviews with auditors from the court of auditors, we have identified that the invitation letter is the modality with the highest recorded irregularity rate. This observation led to its exclusion from Brazil’s allowed modalities as of 2023, following new regulations. Additionally, this modality item required at least 3 bidders chosen at the issuer's discretion. Therefore, it could not be included in the single bidder regression and had to be weighed separately from the other items related to single bidder analysis.

Another secondary red flag is the threshold value for bidding exemption. Contracts with values below the bid exemption threshold are likelier to employ bidding waivers, eliminating competition. Moreover, they may employ the fractioning of expenses to dilute larger values in many small tenders (Gomes, 2016). Furthermore, contracts with values up to ten times higher than the bid exemption limit can also benefit from less demanding bidding modalities.

The contract award itself serves as another secondary red flag. The time taken to award a contract, if too short or too long, can indicate spurious corruption intentions. Immediate decisions, for example, may provide evidence of predetermined decisions, while delayed decisions may suggest implementing legal remedies to circumvent a legitimate initial outcome (Fazekas & Tóth, 2016.2). The contract amendment red flag can also act as camouflage for targeted tenders. Winning bids may initially appear less attractive to the bidders at first glance but undergo subsequent price and longevity changes that turn the contract into a profitable agreement (Owusu, Chan, & Wang, 2021).

Moreover, contract extensions, occurring after the end of the term without a new bidding process, can increase the supplier's profitability. By extending the deadline or even increasing the contract value, delivery costs can be reduced (Fortini & Motta, 2016). Conversely, contract duration plays a significant role in determining the risk of fraud. Generally, the longer the duration of administrative contracts, the greater the risk of fraud since it may undergo price corrections above the market reference and inflation corrections throughout its term.

4.5. Regression Specification

Regression techniques are widely employed in the analysis of various datasets, and in this study, we use regression models to investigate binary logistic data pertaining to single bidders, as well as linear data related to the winner's market share within issuer contracts. The primary objective is to understand the factors influencing the probability of having a single bidder in tenders and to predict the volume and value of the winner's market share. Two regression models were implemented to quantify corruption risks: a simple binary logistic regression for single bidding and a linear regression for market share.

The first regression model aimed to identify corruption risks related to single bidding contracts, where only one bid is submitted in a tender on an otherwise competitive market. This situation allows the awarding of contracts above market prices and may indicate the presence of corruption. The probability of a contract being a single bidding contract was modeled using binary logistic regression. The regression techniques employed to analyze the binary logistic data involving single bidders in tenders are represented by the following equations:

$$P(\text{condition}_i = 1\,) = \frac{1}{1+e^{Z_{1_i}}} \tag{4}$$

$$Z_i = \alpha + \beta_{1j} S_{ij} + \beta_{2m} C_{im} + \epsilon_i \tag{5}$$

where

- $P(condition_i = 1)$ equals one if the i^{th} contract met the criteria set by the condition — single bidder — and otherwise they take value zero;
- Z_i represents the logit of a contract i being a single bidder contract;
- α is the constant of the regression, representing the expected value of the dependent variable when all the independent variables are zero;
- β_{1j} is a vector of coefficients associated with S_{ij}, a matrix that captures a set of j independent features associated with each contract i;
- β_{2m} is a coefficient vector associated with a matrix of m control variables C_{im} for contract i;
- ϵ is the error term.

In addition to the logistic regression model, we conducted two linear regression analyses to examine the winner's market share, which is divided into two parameters. Specifically, we investigated both the volume and value of the winner's market share within the issuer contracts.

The linear regression equation used for the market share model is as follows:

$$Y_i^{\text{condition}} = \alpha + \beta_{1j}S_{ij} + \beta_{2m}C_{im} + \epsilon \quad (6)$$

where $Y_i^{\text{condition}}$ represent the two winner's market share variables within the issuer's contracts. We regress the model for the two conditions described above. The regression constant α, the regression coefficients β_1 and β_2, and the variables "S_{ij}," and "C_{im}" have the same interpretations as in the previous equation.

In these equations, the regression coefficients α, β_1, and β_2 represent the impact of the respective corruption risk inputs and control variables on the probability of occurrence of a single winning bidder, high volume of market share, and high value of market share. The regression models aim to test and validate the impact of corruption inputs on corruption outcomes in public procurement, based on the theoretical expectations of institutionalized corruption. The primary purpose of the regressions is to evaluate whether corruption inputs can contribute to outcomes in line with theoretical expectations reflecting corruption risk in the procurement market. They provide the primary source of internal validity of the composite indicator.

In regressions, a significant and large positive coefficient is interpreted as indicating that a given corruption risk input is typically used to reach a corruption risk output, taking into account alternative explanations such as contract extension or duration and others corruption risk entries. Component weights of the composite indicator are derived from regression coefficients; whereby, the larger coefficient means higher component weight. This reflects the view that the more often a corruption input is used in combination with corruption outcomes, the more confident we can be that corruption lies behind its use. Regression models were built based on the above theoretical expectations by entering each corruption risk input and controls step-by-step. Here, only final regression results are reported for the sake of brevity.

4.6. Regression Results

The regression models for fraudulent activity in public procurement aim at producing the component weights for assembling the Corruption Risk Indicator (CRI). This involved assigning weights to each red flag based on the regression variables of competitiveness, and market share. The magnitude of the regression coefficients played a crucial role in establishing the weight of each component, ensuring that the most significant irregularity predictors were given greater importance. As mentioned, we used binary logistic regression for competitiveness-based regression and linear regression for market share-based regression. This approach allowed us to assign red flag weights determined by empirical patterns of the network under study, rather than relying on subjective judgments (Kenny & Musatova, 2010).

The regression coefficients provide insights into the degree of correspondence between each red flag and the risk of fraudulent activities in procurement bids, see Table 13. Hence, based on the regressions coefficients, we estimate the weights of the components that will compose the Corruption Risk Index (CRI). As mentioned, to the three primary red flags we assigned a weight equal to the highest regression coefficient found, reflecting their reference status, as they are considered the key factor for irregularities. Besides, the invitation letter was treated differently in relation to the single bidder regression, as it requires a minimum of 3 competitors. Therefore, given its classification as a high-risk modality for irregularities, the highest regression coefficient obtained in the single bidder regressions was also used as the regression coefficient for this modality.

Finally, the remaining red flag group received values proportional to their regression coefficients in the three regression parameters, with weights varying between 0 and 1 depending on the coefficient magnitude. Based on the Regression-Based Weight obtained for each red flag, we then normalized the weight of the components so that the final composite indicator (CRI) ranged between 0 and 1, Table 13. The component weights were distributed so that when reaching the maximum score in each red flag group, the CRI value would equal 1, the highest risk of irregularity possible.

Table 13: Regression models – Presentation of the resulting coefficient and weight

Red Flags - Variables		Models: **Single winner**	**MKT volume**	**MKT value**	**Regression weighted coefficient**	**Normalized weighted coefficient**
Mkt Share Volume					0,319	0,168
Mkt Share Value					0,319	0,168
Competitiveness	Single Bidder				0,319	0,168
Modality	a - Invitation letter	0,796	0,014	-0,018	0,264	0,139
	b - Price taking	0,425	0,025	0,021	0,157	0,083
	c - Waived	0,416	-0,267	0,000	0,050	0,026
	d - Auction	0,113	0,000	0,000	0,038	0,020
	e - Competition	0,119	-0,043	0,046	0,041	0,021
	f - Other	0,114	0,034	-0,680	-0,177	0,000
Contract amendment		-0,334	-0,002	-0,331	-0,222	0,000
Contract extension		0,183	-0,001	0,149	0,110	0,058
Threshold value for bidding exemption	a - Below or equal	0,796	0,115	0,045	0,319	0,168
	b - Below or equal 10x	0,360	0,011	0,057	0,143	0,075
	c - Below or equal 100x	-0,080	0,038	0,091	0,016	0,009
	d - Below or equal 1000x	-0,479	-0,026	-0,101	-0,202	0,000
	e - above 1000x	-0,706	-0,004	-0,033	-0,248	0,000

Contract award time	a - Award <= 30	0,226	-0,017	-0,021	0,063	0,033
	b - 30 < Award <= 90	-0,301	-0,021	0,147	-0,058	0,000
	c - 90 < Award	-0,144	-0,028	-0,008	-0,060	0,000
	d - Award Missing	0,052	0,001	-0,015	0,013	0,007
Contract duration	a - Duration = 0	-0,710	-0,017	-0,281	-0,336	0,000
	b - 0 < Duration <= 180	0,028	-0,031	0,562	0,186	0,098
	c - 180 < Duration <= 366	0,220	0,014	0,010	0,081	0,043
	d - Duration >366 or missing	0,310	0,012	0,042	0,121	0,064

4.7. CRI segmentation into quartiles of corruption risk levels

This section aims to identify suspicious binary interactions characterized by corporate arrangements supporting corrupt coordinated activities. Here we start using the CRI values to capture corruption risk in each tender-winner relationship and establish different corruption ranges for these interactions. To calculate the CRIs of winning firms and municipalities separately, we applied the average of all CRIs associated with their transactions in a given period.

Considering firms that achieved an average of at least four wins per year in each government, referred to as the winning elite, allowed us to capture the most successful relationships and, consequently, those most susceptible to extended corruption. In this way, we focused on the most prosperous relationships and monitored the number of firms that did not transition to the second government and those that joined it. Categorizing all the CRIs of the 8-year study into quartiles for both municipalities and firms, we established four corruption risk bands: corruption-free, low corruption, medium corruption, and high corruption. The analysis of both governments' four bands of winning firms and municipalities revealed a low CRI standard deviation and an increasing mean magnitude as the corruption risk bands transitioned from corruption-free to high CRI, as shown in Tables 14 and 15.

Table 14: Corruption risk level bands of firms in two government periods

Firms - Corruption Risk Bands	**Government 1**				**Government 2**			
	CRI mean	CRI std dev	Number of Firms	% Firms	CRI mean	CRI std dev	Number of Firms	% Firms
Corruption free	0,177	0,032	483	26,1%	0,176	0,030	488	25,6%
Low corruption	0,250	0,017	501	27,0%	0,248	0,017	418	21,9%
Medium corruption	0,303	0,014	468	25,2%	0,304	0,014	349	18,3%
High corruption	0,376	0,046	402	21,7%	0,383	0,044	652	34,2%
Total corruption	0,272	0,077	1854	100,0%	0,286	0,088	1907	100,0%

Table 15: Corruption risk level bands of municipalities in two government periods

Municipality - Corruption Risk Bands	Government 1				Government 2			
	CRI mean	CRI std dev	Number of Munic.	% Munic.	CRI mean	CRI std dev	Number of Munic.	% Munic
Corruption free	0,228	0,018	64	34,78%	0,222	0,016	43	23,4%
Low corruption	0,257	0,006	35	19,02%	0,258	0,006	43	23,4%
Medium corruption	0,281	0,008	43	23,37%	0,280	0,006	39	21,2%
High corruption	0,319	0,018	42	22,83%	0,325	0,021	59	32,1%
Total corruption	0,267	0,039	184	100,00%	0,276	0,043	184	100,0%

Based on the CRI bands obtained, we observed structural differences between the municipal groupings and firm groupings of the two governments. In the second period, there was an increase of 17 municipalities in the high corruption band, representing a 40% rise of the municipalities exhibiting the highest risk of corruption, Table 15. On the other hand, the number of firms with the highest risk of corruption surged by 150, a substantial 62% increase, compared to the first period, Table 14. In addition, the number of connections in the second administration, increased by 11,6%, representing a 20% rise in network density.

Among the 652 companies with the highest corruption risk in the second period, 38.3% had already exhibited a high risk of corruption in the first administration, while 49.5% of firms were previous suppliers in the first government but with a lower CRI. Furthermore, 12.1% of firms with high corruption risk were new additions to the second government and had not been part of the network during the previous government. Among 152 companies that were initially classified in the highest corruption risk band exclusively during the first government, 32 firms completely exited the subsequent network, while the remaining 120 moved to a lower corruption risk band.

The number of winners with an average of four or more victories per year, the winning elite, increased by 2,9% from the first to the second government. Among the 1907 firms in the second government network, 85.6% had also been part of the first government. Within this group, 34.2% experienced a decrease in their CRI band, 19.4% witnessed an increase, and 46,5% maintained the same band level.

Of the 221 suppliers that were no longer part of the winning elite in the second government, 31.7% were corruption free during the first government, 31.2% had low corruption risk, 22.6% had a medium corruption risk, and 14.5% had a high corruption risk. In other words, 62.9% of the companies that left the winning elite belonged to the two lowest corruption risk bands in the first government. Moreover, 62,2% of the firms in the high corruption band remained in this band, while an additional portion representing 80% joined it, which significantly contributed to the increase in the CRI during the second government.

On the other hand, out of the 274 firms that joined the winning elite in the second period, 29.9% became corruption-free, 25.5% had low corruption risk, 15.7% had medium corruption risk, and 28.8% had high corruption risk. It indicates a substantial influx of corruption-free companies, but also a significant rise in the number of firms joining the high corruption risk band. As a result, there was a significant increase of 37.3% in the number of companies classified in the high corruption risk band. This overall surge significantly contributed to the 5.2% increase in the magnitude of the Corruption Risk Index in the second government.

We observed a notable trend where companies categorized in the high corruption risk band tend to secure more contracts in municipalities also part of the high corruption risk band — both periods displayed a high Pearson correlation, exceeding 0.6 between the CRI of each municipality and the average CRI of its suppliers (refer to Figure 13 for visual representation).

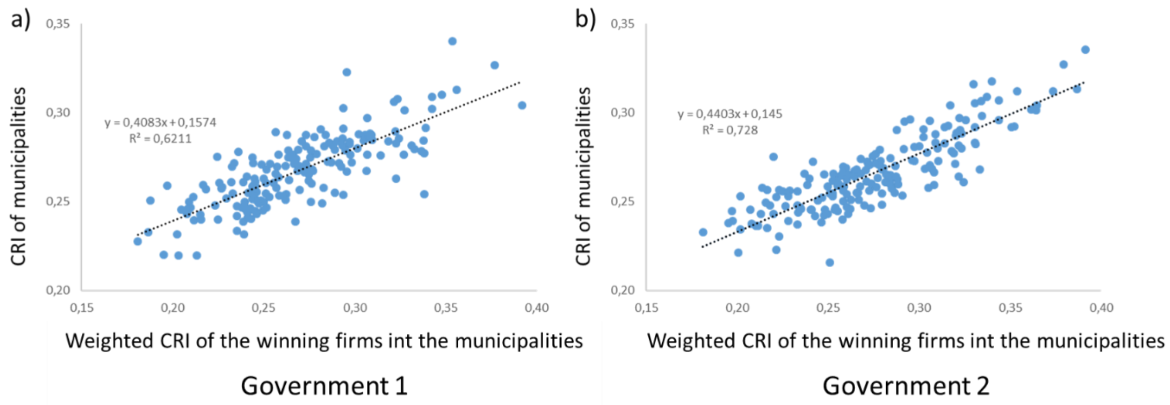

Figure 13: Comparison between the CRI of the municipalities and the average CRI of the winning firms in each municipality. The Y-axis - CRI of municipalities is calculated by the average of all CRIs between tenders and winners in the municipality. X-axis - CRI of all winners in a given municipality, calculated through the average of CRIs between tenders and winners, regardless of the municipality they participated in. a) First government period. b) Second government period

In Figure 14, we examined how municipalities are classified into the four corruption risk bands during government 1 (Figure 14a) and government 2 (Figure 14b). Additionally, Figure 14c depicts the migration patterns of each municipality after the change of government administration, illustrating the variations in their CRI bands. The CRIs of municipalities either moved up or down by up to three risk bands or remained within the same range.

From the data, it becomes evident that a significant part of municipalities, 37.5%, maintained their position within the same risk band. However, a considerable municipality migration towards higher risk bands was also observed, accounting for 39.6%, while a smaller migration towards lower risk bands amounted to 23.8%. Overall, these movements resulted in a 3.3% increase in the average municipal CRI level for the entire state during the second government's tenure.

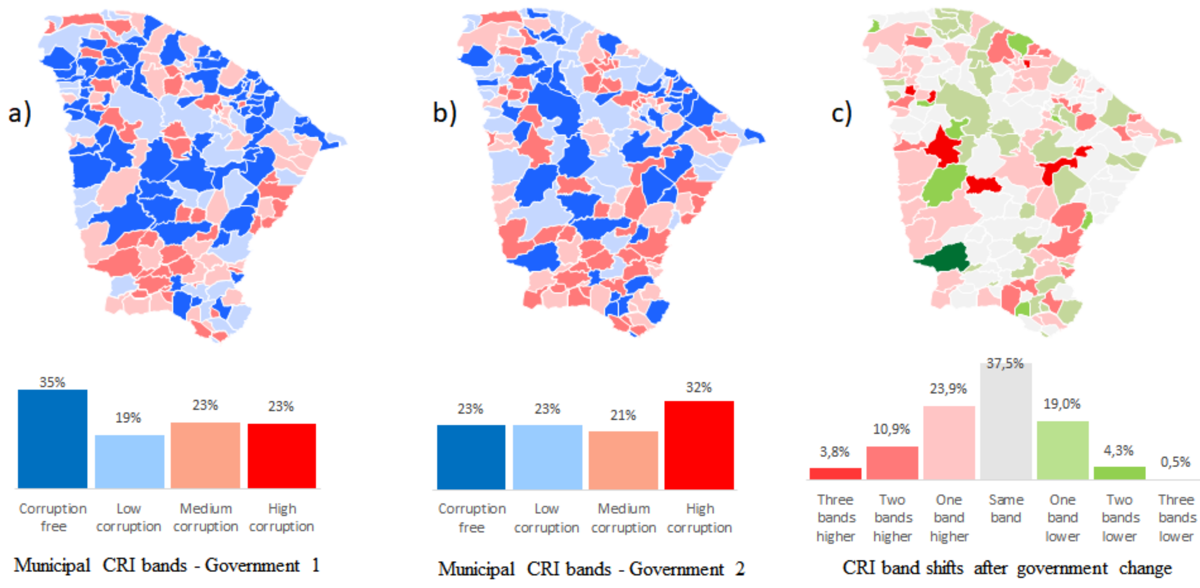

Figure 14: a) Classification of 184 municipalities in Ceará within 4 corruption risk bands during the first government. b) Classification of 184 municipalities in Ceará within 4 corruption risk bands during the second government. c) Corruption risk band changes after municipal government turnover.

According to the above findings, it is important to note that a municipality belonging to the high corruption risk band does not necessarily contain firms from the high corruption risk band. Theoretically, it may not contain any such firms; it is enough to have several corrupt associations with different companies that do not participate in many fraudulent associations in other municipal tenders, thus belonging to lower corruption risk bands. Likewise, a firm may belong to a high corruption risk band even without winning municipal contracts from the high corruption risk band.

Table 16: Change in CRI band level after government transition

Municipality change characteristics		Same Mayor and same party	Same Mayor and different party	Different Mayor and same party	Different Mayor and different party	Total municipalities
Number of municipalities		10	16	18	140	184
Municipality CRI Band moved to	Lower band	10.0%	25.0%	16.7%	25.7%	23.9%
	Same band	50.0%	43.8%	55.6%	33.6%	37.5%
	Upper band	40.0%	31.3%	27.8%	40.7%	38.6%

An alternative approach to examining the dynamics among network players in the two periods involves investigating the behavior of CRI levels following government turnover in the various municipalities. This analysis can provide valuable insights into the impact of mayors and political parties changes on the CRI bands. Table 16 provides a comprehensive overview of the government transition patterns observed across different cities of Ceará, highlighting the relationship between shifts in mayoral position, political party affiliations, and the corresponding CRI bands shown in Tables 14 and 15.

While the majority of municipalities, accounting for 76%, chose to change both the party and the mayor during the government transition, an intriguing pattern emerges when examining the correlation between these changes and shifts in the municipality's CRI band. Notably, we observed a significant migration of municipalities towards higher CRI bands in all scenarios, while a smaller number of municipalities migrated towards lower CRI bands.

A stronger association becomes evident when the same mayor is reelected under the same political party, accounting for 5.4 % of all cities, compared to different mayors being elected under other parties, which occurs in 76% of all cities. In the specific scenario where there is continuity in the mayor’s position and party affiliation, the CRI bands exhibit noteworthy variations. Specifically, 40% of the municipalities experienced an upward shift to high CRI bands, 10% transitioned to lower bands, and 50% remained in the same grouping. In the case of municipalities that changed both their mayor and party, we observed a more significant migration towards the higher CRI band, comprising a total of 40.7%. Additionally, 25.7% of the municipalities migrated to lower CRI bands, while 33.6% remained within the same band.

Conversely, when the same mayor and different party assume in municipalities, their CRI remain the same in 43.8% of towns, increases in 31,3% and decreases in 25%. Lastly, in 55.6% of the towns, where different mayor and same political party occur, they maintain the same corruption risk band, with 27.8% moving to upper bands and 16.7% transitioning to lower bands. This last trend can be partially attributed to changes in public administration behavior and shift in political policy at the same time, resulting in a managerial mismatch that may have contributed to more lenient decisions regarding irregularity control, thereby leading to an increase in the risk of corruption in public contracts.

To gain a more comprehensive understanding, we also assessed the average CRI of different contract types by region, Figure 15. This analysis aims to identify regions and specific contract types that exhibit the highest corruption risks. By examining this data, we gained valuable insights into the variations in corruption risks across different regions and contract types. This approach allowed us to better grasp the complex dynamics of corruption in public procurement and identify specific areas that require targeted interventions and preventive measures.

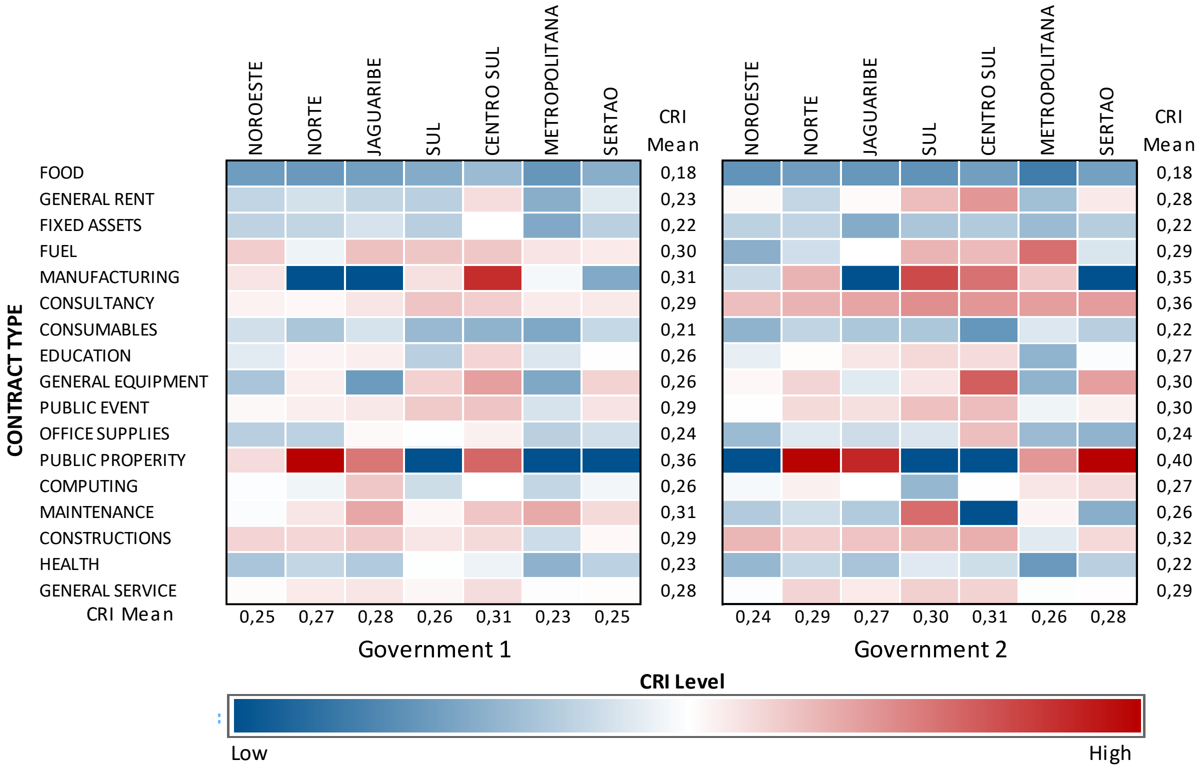

Figure 15: Representation of the cross-marketing interaction between the mesoregions of Ceará and the types of contracts carried out in the various tenders. a) market interaction during government 1 b) market interaction during government 2.

In Figure 15, distinct patterns emerge regarding corruption risks in different contract types and regions during the first and second governments. Throughout the first government, two contract types stood out with the highest corruption risks. The "manufacturing" was primarily concentrated in the centro-sul region, while the "public property", exhibited elevated corruption risks in the norte, Jaguaribe, and centro-sul regions.

In contrast, during the second government, the corruption risk in the "manufacturing" contract type remained high in the centro-sul region but also expanded to include the sul region. Similarly, the contract type "public property" maintained a high corruption risk in the norte region, experienced a significant increase in the Jaguaribe region, diminished substantially in the centro-sul region, and emerged with high magnitude in the Sertão region. The contract type Consultancy emerged with high CRI in all regions. The contract type that showed the most significant CRI reduction were "fuel" and "maintenance," despite "fuel" in the metropolitan area experiencing a considerable 40% increase in its CRI.

The elevated Corruption Risk Index in the "public property" contract type is largely attributed to its inherent characteristics. Public property leasing often involves single-bidder tenders with substantial values and may include construction works, contract extensions, and extended durations. Moreover, it's worth noting that public property lease tenders might not always be deemed necessary in certain governments and regions, as properties from past administrations are frequently utilized. This practice contributes to lower CRIs in specific regions, primarily due to the absence of tenders for this contract type.

4.8. Conclusions

In this article, we investigate a database of public procurements carried out in the municipalities of the State of Ceará between the years 2013 to 2020, representing a period of two different governments whose exchange occurred between the years 2016 and 2017. The research utilizes empirical data from public tenders representing the interactions between suppliers and public entities subdivided into municipalities, regions, periods, and contract types. Based on parameters intrinsic to the Brazilian bidding process, we set up a corruption risk indicator based on previous studies adapted to the Brazilian reality. The indicator was obtained using nine red flags associated with fraudulent activities captured from the analysis of audit reports and interviews with auditors from the court of auditors of Ceará (Brazil). These red flags include factors such as the presence of single bidders, market share of winning companies, bidding modalities, contract awards, contract extensions, and contract durations.

The corruption risk indicator is supposed to indicate the degree of corruption in a binary relationship between municipal tenders and winning firms in biddings to acquire products and services through public resources with the best cost-benefit. To foster the proper influence of each red flag in the composite indicator, we employed statistical techniques namely binary logistic regression and linear regression. These methods allowed us to analyze how corruption risk inputs influence various aspects of the procurement network. By conducting regressions based on the two key variables, competitiveness and market share, we estimated the weights of the components, ensuring their proper representation in the overall indicator. Next, we categorized tenders and winners into four bands using quartiles based on their relationships' average CRI, namely: corruption-free, low corruption, medium corruption, and high corruption. This categorization effectively measured the level of corruption risk present in their relations and provided valuable insights into the consistency of their behavior across different connections established in both government periods.

Organized categorizations helped to identify intersections with significant CRI level, prompting the necessity for precise interventions and preventive measures. It effectively highlighted contract types and regions linked to enduring high corruption risks, emphasizing the need for heightened external control to tackle corruption challenges in these domains.

The CRI provides valuable insights into the level of corruption in each transaction between a municipal tender and a winning firm, offering a reliable indicator for assessing potential corruption risks in the procurement processes. Furthermore, the research examines the impact of changes in government administrations on corruption dynamics. It reveals a notable trend where municipalities and firms belonging to high corruption risk bands tend to be more concentrated in regions characterized by higher corruption risks. In addition, government turnover is associated with shifts in corruption risk bands, indicating a possible influence of political changes on corruption risks in public procurement.

The study offers an extensive examination of corruption risks within public procurement across municipalities in Ceará. Through an innovative Composite Corruption Risk Indicator, the research offers valuable insights into the corruptive patterns. Furthermore, it pinpoints specific areas requiring focused attention to promote transparent and accountable public procurement practices. Implementing interventions derived from the study's findings not only holds the promise of potential mitigation in corruption risks, but also serves as a catalyst for nurturing a culture of integrity within procurement procedures.

Overall, the results revealed a dynamic scenario scaled in approaches to capture the degree of corruption risk at the organizational level. The method provides immediate applicability to public procurement markets when adapted to specific local regulations, not only to the Brazilian market. As this is exploratory research without labeled data on previous corruption cases, it was not possible to establish a degree of explanatory confidence in the model to prove its effectiveness. Besides, some red flags could not be used due to the lack of information in the database, such as the time between the public notice and the bidding process, the number of bidders excluded before the bidding, and the scores of bidder assessment. Despite that, we introduced other red flags not considered in previous studies, such as the Threshold value for bidding exemption.

The methodology proved to be of great value in identifying possible corrupt organizations, and the results can be used mainly in directing field audits based on findings with updated data. Further studies should look at replicating these techniques by employing them in a more targeted way in smaller groups, reducing the scope to focus on making a more regional and detailed assessment.

References

Amaral, M., Saussier, S., & Yvandre-Billon, A. (2013). Expected Number of Bidders and Winning Bids. *Journal of Transport Economics and Policy, 47:1*.

Atkinson, C. (2019). Full and open competition in public procurement: values and ethics in contracting opportunity. *International Journal of Public Administration, 43*(13), pp. 1169-1182.

Bauhr, M., Czibik,, A., Licht, J., & Fazekas, M. (2019). Lights on the shadows of public procurement: Transparency as an antidote to corruption. *Governance, 33*(3), pp. 495-523.

Blondel, V., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J Stat Mech: Theory Exp, 2008*(10). doi:10008

Charron, N., Dahlström, C., Fazekas, M., & Lapuente, V. (2017). Careers, connections, and corruption risks: Investigating the impact of bureaucratic meritocracy on public procurement processes. *J Polit, 79*(1), pp. 89-104.

Clark, A. (2017). Measuring corruption: transparency international's "corruption perceptions index". *In Corruption, Accountability and Discretion, 29*, 3-22. Publishing, Emerald

Czibik, Á., Fazeka, s. M., Sanchez, A., & J., W. (2021). Networked Corruption Risks in European Defense Procurement. *Corruption Networks: Concepts and Applications*, 67-87.

Dávid-Barrett, E., & Fazekas, M. (2020). Grand corruption and government change: an analysis of partisan favoritism in public procurement. *European Journal on Criminal Policy and Research*, *26*(4), 411-430.

Fazekas, M., & Kocsis, G. (2020). Uncovering High-Level Corruption: Cross-National Objective Corruption Risk Indicators Using Public Procurement Data. *British Journal of Political Science, 50*(1), pp. 155-164.

Fazekas, M., & Tóth, I. J. (2016.2). From corruption to state capture: A new analytical framework with empirical applications from Hungary. *Political Research Quarterly, 69 (2), 320–334.*

Fazekas, M., & Wachs, J. (2020). Corruption and the network structure of public contracting markets across government change. *Politics and Governance*, *8*(2), pp. 153-166.

Fazekas, M., Toth, I. J., & King, L. P. (2016.1). An Objective Corruption Risk Index Using Public Procurement Data. *European Journal on Criminal Policy and Research, 22*(3), pp. 369-397.

Fortini, C., & Motta, F. (2016). Corrupção nas licitações e contratações públicas: sinais de alerta segundo a transparência Internacional. *Revista de Direito Administrativo e Constitucional*.

Gomes, C. (2016). Dispensa Indevida de Licitação: O Dano ao Erário na Caracterização do Ato de Improbidade Administrativa. *Universidade Federal de Campina Grande – UFCG.*

Goryunova, N., Baklanov, A., & Ianovski, E. (2022). A Noisy-Labels Approach to Detecting Uncompetitive Auctions. *7th International Conference on Machine Learning, Optimization, and Data Science (LOD) / 1st Symposium on Artificial Intelligence and Neuroscience (ACAIN), 13163*, pp. 185-200.

Heggstad, K., & Heggstad, M. (2011). The basics of integrity in procurement. *Bergen, Norway: Anti-Corruption Resource Centre Chr. Michelsen Institute (CMI)*.

IMPIC. (2018). Contratação Pública em Portugal. *Relatório anual 2018. Lisboa: IMPIC.*

Kenny, C., & Musatova, M. (2010). Red Flags of Corruption in World Bank Projects: Na Analysis of Infrastructure Contracts. *Policy Research Working, Artc. 5243, World Bank, Washington DC.*

Lisciandra, M., Milani, R., & Millemaci, E. (2021). A corruption risk indicator for public procurement. *European Journal of Political Economy, 73, 102141*.

Locatelli, G., Mariani, G., Sainati, T., & Greco, M. (2017). Corruption in public projects and megaprojects: There is an elephant in the room! *Int. J. Proj. Manag, 35*, pp. 252-268.

Lyra, M., Bacao, F., & Pinheiro, F. (2021.1). Public Procurement Fraud Detection – A Review using Network Analysis. *Complex Network 2021*.

Lyra, M., Curado, A., Damásio, B., Bacao, F., & Pinheiro, F. (2021.2). Characterization of the Firm-Firm Public Procurement Co-bidding Network from the State of Ceara (Brazil) Municipalities. *Applied Network Science, 6 (1)*.

Owusu, E., Chan, A., & Wang, T. (2021). Tackling corruption in urban infrastructure procurement: Dynamic evaluation of the critical constructs and the anti-corruption

Ratcliff, C., Martinello, B., & Kaiser, K. (2022). *Factsheets on the European Union: Public Procurement Contracts*. Fonte: https://www.europarl.europa.eu/factsheets/en/sheet/34/public-procurement-contracts

Rauen, A., & Paiva, B. (s.d.). *Impacts of Public Procurement on Business R&D Efforts: The Brasilian Case*. Fonte: Instituto de Pesquisa Econômica Aplicada – IPEA. Report

Roca, T., Orme, W., & Brown, J. (2010). Fear and Loathing of the corruption Perception Index: Does Transparency International Penalize Press freedom? *Available at SSRN 1694211*.

Søreide, T. (2006). Is It Wrong to Rank? A Critical Assessment of Corruption Indices. *Chr. Michelsen Institute*.

Speck, B., & Ferreira, V. (2012). Sistemas de Integridade nos Estados Brasileiros. *Instituto Ethos de Empresas e Responsabilidade Social, 12*, pp. 30-82.

Tribunal de Contas do Estado do Ceará (2022). https://www.tce.ce.gov.br/cidadao/consulta-de-processos. Accessed on 07/07/2022.

Wachs, J., & Kertész, J. (2019). A network approach to cartel detection in public auction markets. *Sci Rep, 9*. doi:10818

Wachs, J., Yasseri, T., Lengyel, B., & Kertesz, J. (2019). Social capital predicts corruption risk in towns. *Royal Society Open Science*.

Zeferino, L. (2020). A corrupção na construção de edifícios públicos no brasil: análise de instrumentos inibidores e facilitadores na etapa de projeto arquitetônico. *Universidade de De Brasília – UnB Master's thesis*.

5. CONCLUSIONS

Throughout this dissertation we explored the possible application of Network Analysis methods to investigate irregularities in the Public Procurement Process by levering data at contract-level from municipal governments of Ceará state in Brazil. Our goal was to identify fraud detection methodologies by addressing common limitations in available data, such as the lack of labels on past audited contracts and firms.

Despite previous efforts in developing analytical solutions to understand corruption in public procurement and mitigate its effects (Fazekas et al., 2018), empirical studies focusing on fraud detection in this field, especially those employing objective fraud indicators, remain scarce (Lyra et al., 2021.2). However, the growing availability of open data on public procurement activities has recently revitalized the efforts of the scientific community to uncover hidden connections among agents and bring light to the phenomena (Curado et al., 2020).

In that sense, this thesis starts by accounting for most current studies and methodologies developed for detecting fraud in public administration, with a specific focus on tender processes and network analysis. We achieve that with two systematic literature reviews focusing on data-driven methods for detecting fraud, collusion, and corruption in public procurement processes. The first review, titled *"Public Procurement Fraud Detection: A review using Network Analysis"*, Appendix C, specifically explored the application of Network Analysis methods for fraud detection in public procurement processes. A second review, titled *"Fraud, corruption, and collusion in public sector expenditures*, *a systematic literature review on data-driven methods"*, Chapter 2, focused on the most contemporary data-driven techniques for crime detection in this domain.

Both review efforts revealed that researchers are increasingly adopting advanced analytical techniques such as data mining, machine learning, and predictive modeling to enhance fraud detection in public procurement. These methodologies leverage large datasets, including historical procurement data, financial records, supplier information, and other relevant variables, to identify and characterize procurement patterns, anomalies, and potential indicators of fraudulent/irregular activities. Moreover, the integration of network analysis and network-based algorithms to uncover hidden connections and collusive behaviors among suppliers, contractors, and public officials involved in procurement processes (Toth et al., 2014; Fazekas at al., 2016.1; Wach & Kertz, 2019) has also been pursued. By examining relationships and interactions within the procurement network, these methods aim to identify suspicious patterns indicative of fraudulent activities, such as bid rigging, collusion, cartel, and corruption.

Our work highlights the growing importance of incorporating domain knowledge and expert judgment into fraud detection, even when relying on big data-driven techniques. Researchers have emphasized the need for a multidisciplinary approach that combines quantitative analysis with qualitative assessments based on expert opinions and domain-specific insights. This fusion of methodologies enables a more comprehensive evaluation of the risks and vulnerabilities associated with public procurement, enhancing the accuracy of fraud detection models (De Maria & Howai, 2021). Indeed, our analysis underscores the importance of considering public procurement processes' specific context and characteristics when designing fraud detection methodologies. Factors such as the complexity of procurement procedures, multiple stakeholders' involvement, and regulatory frameworks necessitate tailored approaches to identify and prevent fraud effectively. Additionally, data quality and accessibility in fraud detection are key. Obtaining reliable and comprehensive data on procurement transactions, supplier profiles, and contract details is crucial for accurately assessing fraud risks (Lemke, 2020). Therefore, efforts should be made to enhance data governance practices, promote transparency, and establish mechanisms for data sharing and collaboration among relevant authorities (Mahdillou & Akbary, 2014).

The need for continuous monitoring and proactive fraud prevention measures in public procurement. While fraud detection techniques play a vital role, early intervention, and preventive actions can significantly reduce the occurrence and impact of fraudulent activities (Ahmed et al., 2016). Implementing robust internal controls, conducting regular audits, and fostering a culture of ethics and accountability are essential components of a comprehensive fraud prevention strategy (McAlister & Ferrell, 2016). Emerging trends in fraud detection in public procurement highlight the progress made in developing methodologies and strategies to mitigate its impact. Significant improvements have been made in identifying and preventing fraud in public procurement by incorporating advanced analytics, network analysis, and expert judgment while leveraging state-of-the-art technology (Sharkawy, 2022). However, continuous research, collaboration, and adaptation to evolving fraud schemes and technological advancements are necessary to avoid fraudulent activities and safeguard public resources.

As part of our empirical analysis, we conducted two studies in the context of public tenders across 184 municipalities of the Ceará state. In that sense, the article titled "*Characterization of the firm-firm public procurement co-bidding network from the State of*

Ceará (Brazil) municipalities", Chapter 3, looked to characterize the public market by identifying which firms regularly compete in public tenders from a network perspective. Suh a view of the market can offer insights into possible groups of firms that have the potential to engage in undesirable bidding patterns. The lack of explicit connection information between bidders presented a challenge for government control bodies in detecting collusion or cartel formation in public tenders (Alencar, 2021). In the article, we overcome this challenge using network analysis methods to convert a Firm-Tender network into a Firm-Firm network that captures co-bidding relationships (i.e., the competition procurement market). We identified 22 groups (*i.e.*, network communities) of firms with similar procurement activity patterns, categorized based on their geographic and activity scopes. By profiling these communities, our study shed light on organizations more susceptible to market manipulation and irregular activities, providing valuable insights for public authorities to combat fraud in public procurement. The findings emphasized the potential application of network analysis in public policy to unravel the complex relationships between market agents, particularly in scenarios with limited data availability. Corruption in public procurement contracts was a significant issue with substantial economic implications.

It is important to note that the use of network analysis in studying corruption in public procurement is not new, with previous studies exploring bipartite relationships between public bodies and firms, as well as firm-firm co-bidding relationships (Lauchs et al., 2011). Despite the strengths of the study, there were some challenges and limitations that need to be acknowledged. One limitation was the scarcity of labeled data, while this approach allowed for identifying patterns associated with a high risk of corrupt activities, it also indicates the need for additional research and data to validate and generalize the findings. Furthermore, the study primarily focused on the firm-firm co-bidding network, neglecting the potential influence of public bodies and other stakeholders involved in the procurement process. To overcome these limitations and further advance the understanding of corruption in public procurement, it is essential to incorporate bipartite relationship analysis between public bodies and firms.

In the context of fraud risk assessment in public procurement through bipartite networks, we investigated fraudulent relationships between public and private entities involved in a winner-tender relationship (the tender represents the issuing municipality), which resulted in the article titled *"Fraud risk assessment in public procurement activities - an investigation in two different governments of Ceara's municipalities"*, Chapter 4. The research examines the evolution of corruption risk within these processes under two different governments. To achieve this objective, we adapted the Corruption Risk Indicator (CRI) introduced recently by Fazekas et al (2016) to obtain an objective measure of fraud and corruption risk in public tenders conducted in the municipalities of the state of Ceará. The development of the CRI required the identification of primary and secondary red flags at the contract level. Furthermore, the study's focus on the evolution of corruption risk over time under different governments and provides valuable insights into the impact of political changes on the prevalence of fraudulent activities. By comparing the corruption risk levels between administrations, we can assess the effectiveness of governance measures and anti-corruption policies implemented by different governments. This knowledge can guide future policy-making decisions to strengthen accountability and transparency in public procurement processes.

However, it is important to acknowledge some limitations of this study. Firstly, the analysis is based on empirical data from public tenders in the state of Ceará, which may not fully represent the entire country. Therefore, the findings should be interpreted within the context of this specific region. Secondly, the effectiveness of the CRI relies on the accuracy and reliability of the data used for its calculation. Any limitations or biases in the data may impact the accuracy of the risk assessment. Lastly, the study focuses on detecting and analyzing fraud risk but does not delve into investigating specific fraudulent activities or legal implications.

It is important to note that this study contributes to the existing literature by providing an objective corruption risk indicator built specifically for Brazilian public procurement that relies on empirical data rather than subjective perceptions. By utilizing a composite indicator and incorporating various local red flags at the contract level, the Corruption Risk Indicator enhances the accuracy and reliability of fraud detection in public tenders. This objective approach is particularly valuable in addressing the limitations of traditional corruption risk indexes that heavily rely on subjective data.

Future research will focus on integrating additional data sources including financial records, supplier information (such as partner identification), details about companies with common ownership and familial relationships, and social contact information. By incorporating these diverse data sets, we aim to gain a more comprehensive understanding of procurement patterns, identify anomalies, and uncover potential indicators of fraudulent activity. Concurrently, efforts are underway to enhance data governance practices, promote transparency, and establish effective mechanisms for data sharing and collaboration among relevant government authorities. These initiatives aim to strengthen the overall integrity of public purchasing processes and facilitate a more coordinated approach to combating crimes related to public procurement.

6. APPENDIX C

Public Procurement Fraud Detection: A review using Network Analysis

This section presents an article containing the first literature review version that only considered network science as a methodology for identifying fraud in public tenders.

Public procurement is a key step in providing public services to the population in the health, education, and infrastructure sectors (World_Bank_Group, 2017). It is an instrument by which the public administration seeks to obtain the highest volume of goods, services, and construction at the lowest cost (Costa, Machado, & Martins, 2020). Its sustainability is one of the elements that lead to the efficiency of the public sector (Silva Filho, 2017). This process is considered one of the areas with the highest risk of fraud (Whiteman, 2019). It can occur at any point in the supply chain, making it difficult to detect and measure, especially after awarding a contract (Whiteman, 2019). Fraud in this sector are complex and sophisticated; they might occur at any stage of the purchase process (Rustiarini, Sutrisno, Nurkholis, & Andayani, 2019) and can be one of the most problematic crimes to investigate (Mufutau, 2016). These stratagems occur widely, from corruption, bribery, embezzlement, collusion, abuse of power, favoritism, misappropriation, and nepotism (Padhi & Mohapatra, 2011). However, even suitably defined public procurement legislation is insufficient to control fraud, and efficient mechanisms for its prevention are inaccurate (Wensink & Vet, 2013).

That said, it is essential to carry out scientific investigations to provide theoretical and methodological bases to identify the behavior of companies and service providers in line with potential fraudulent actions in the bidding markets. In that sense, this systematic review aims to survey the available literature regarding methodologies used to identify and combat fraud in public procurement through network analysis, being a valuable instrument to detect solutions conceived and unveil possible existing gaps.

The challenge of this article in helping the fight against corruption is to conduct a review of the state-of-the-art to identify the latest scientific contributions based on Network Analysis techniques. According to the overview of the current subject and with the help of some literature on the issue (Shaffril, Samsuddin, & Samah, 2020), we came up with a research question: How can network science help in the disclosure of fraud in public procurement? To answer this question, we performed a cross-sectional consultation and a qualitative analysis following the adoption of 18 scientific articles.

Method

This systematic literature survey intends to capture the contributions of the scientific community to the topic of corruption in public procurement, where we sought to identify the most recurrent authors, their interactions, the number of citations, identification of keywords, and their repetitions in the various articles. We used an auditable methodology based on the set of PRISMA items (Preferred Report Items for Systematic Reviews and Meta-Analyzes). This method comprises a checklist with 27 topics distributed in 7 sections, which is used to improve meta-analyses of periodic review reports. Its framework guides the setup of each section based on information from scientific articles: Title, Abstract, Introduction, Methods, Results, Discussion, and other information. The PRISMA 2020 approach was designed to carry out systematic reviews, including synthesis (meta-analyzes or other concepts of statistical synthesis) consisting of five phases (Page, McKenzie, & Bossuyt, 2021): 1. Identification of related articles in the field of interest; 2. Sorting titles and abstracts; 3. Acceptance analysis; 4. Exclusion analysis; 5. Final selection.

Afterwards, we started to identify and collect research articles to assist in the assembly of the current study. For this purpose, we used a systematic computerized query in two search repositories, Web of Science (W.O.S) and Scopus, chosen as primary search sources due to their extensive coverage of scientific writings, faster indexing methods, and based on a curated database of published and peer-reviewed content, often not considered in more comprehensive repositories.

Bibliometric Analysis and Keyword Search Strategy

We carried out an iterative search process with an interval period limitation from 2011 to 2021. The query identified publications covering terms and expressions logically combined by Boolean connectors in their titles, keywords, and abstracts. The resulting search string used in the repositories mentioned above is shown below:

Search String: (corruption OR fraud OR collusion OR bribe OR bribery OR rigging OR cartel OR collusive) AND (Procurement OR tender OR bid OR bidding OR auction) AND (network OR SNA OR detection OR detecting OR detect OR prediction OR predict).

To perform network analysis on the collected data set and represent the interactions between the actors or nodes of the net, we used the open-source engine VOSviewer (Van_Eck & Waltman, 2014). It allows us to identify the publications, authors, journals, institutions, keywords, and countries with the most significant impact on the research in repositories of scientific articles. From parameters such as degree centrality (indicates the number of times a paper is mentioned, represented by the size of the nodes) and edge weight (designates the number of times that two articles are cited together in different documents - depicted by the thickness of the link), we used VOSviewer to generate graphs based on the data collected from the repositories and interpreted them applying network analysis. We also analyzed the information collected about the concomitant occurrence of keywords and authors, which establishes that the larger the node size, the more often the keyword appears in publications. On the other hand, the thicker the links between two nodes, the greater the number of times these pairs of keywords occur in different articles. In turn, the same concept was applied to the authors, i.e., the larger the node size, the more published articles these two authors have in the selected database, and the thicker the connection between them, the more often they act in joint publications.

Results

Data collection took place in August 2021 and was complemented with more articles in September 2021. Subsequently, we proceeded to the selection of relevant publications following the three PRISMA stages. In the first identification phase, we employed the search string for the two electronic repositories. Filtering the years 2011 to 2021, we found 351 articles in Scopus and 311 in the Web of Science. Next, we adopted as an initial basis all writings found in both recipients and identified 279 identical publications. Counting the non-intersected publications, we reached a quantity of 383 different articles and entered the screening phase. At this stage, we made exclusions based on titles, eliminating 301 articles and keeping 82. Afterwards, we read the abstracts of the 82 articles and rejected 64 that were out of scope, remaining with 18. After reading the 18 articles in full, we did not generate exclusions and kept them all. Of these, 15 were published in scientific journals, while four were published in conferences.

Publications by Author, Number of Citations & Journal / Conference

The number of selected articles related to the words used in the search query totals 18 publications. It started with one publication in 2011 and grew to four in 2017, peaking in 2020 with six selected articles. It indicates that the issue is being researched more over time. The articles collected in Scopus and Web of Science repositories, both from journals and conference annals, embrace an extensive variety of scientific research, from Computer Science to

Government and Law, passing through Business, Economics, and Criminology. The scrutiny of scientific articles was essential for the identification of works relevant to our research area. We raised articles cited up to 35 times, resulting in the publications shown in Table 17.

Table 17: Articles by Year of Publication

Title	Year	Author	Citations	
			Scopus	W.O.S
Corruption risk in contracting markets: a network science perspective	2021	Johannes Wachs; Mihály Fazekas; János Kertész	3	4
Corruption and complexity: a scientific framework for the analysis of corruption networks	2020	Issa Luna-Pla; José R. Nicolás-Carlock	10	8
Distinguishing Characteristics of Corruption Risks in Iranian Construction Projects: A Weighted Correlation Network Analysis	2020	Hosseini, M. Reza; Martek, Igor; Banihashemi, Saeed; et. al	10	8
Bidder Network Community Division and Collusion Suspicion Analysis in Chinese Construction Projects	2020	Zhu, Jiwei; Wang, Bing; Li, Liang; et al.	2	1
Corruption and the network structure of public contracting markets across government change	2020	Fazekas, Mihaly; Wachs, Johannes	1	1
Network Analysis for Fraud Detection in Portuguese Public Procurement	2020	Carneiro, D., Veloso, P., Ventura, A., et al.	1	N/A
Betweenness to assess leaders in criminal networks: New evidence using the dual projection approach	2019	Grassi, R., Calderoni, F., Bianchi, M., & Torriero, A	16	14
Social capital predicts corruption risk in towns	2019	Wachs, Johannes; et al.	9	5
A network approach to cartel detection in public auction markets	2019	Wachs, Johannes; Kertesz, Janos	9	5
The Analysis of Water Project Bid Rigging Behavior Based on Complex Network	2017	Cheng, Tiexin; Liu, Ting; Meng, Lingzhi; et al	N/A	0
Bid-rigging networks and state-corporate crime in the construction industry	2017	Reeves-Latour, Maxime; Morselli, Carlo	28	26
Detecting the collusive bidding behavior in below average bid auction	2017	Lei, Ming; Yin, Zihan; Li, Shalang; Li, Hao	2	0
Adaptation of cluster analysis methods in respect to vector space of social network analysis indicators for revealing suspicious government contracts	2017	Davydenko, V. I.; Morozov, N. V.; Burmistrov, M. I.	0	0
From Corruption to State Capture: A New Analytical Framework with Empirical Applications from Hungary	2016	Fazekas, Mihaly; Toth, Istvan Janos	37	29

Improving Fraudster Detection in Online Auctions by Using Neighbor-Driven Attributes	2016	Lin, Jun-Lin; Khomnotai, Laksamee	1	1
The impact of inter-organizational relationships on contractors' success in winning public procurement projects: The case of the construction industry in the Veneto region	2015	Sedita, Silvia Rita; Apa, Roberta	27	25
Combining ranking concept and social network analysis to detect collusive groups in online auctions	2012	Lin, Shi-Jen; Jheng, Yi-Ying; Yu, Cheng-Hsien	22	19
Corrupt police networks: uncovering hidden relationship patterns, functions and roles	2011	Lauchs, Mark; Keast, Robyn; Yousefpour, et al	20	18

The articles collected comprise the most up-to-date concepts currently applied to identify fraud in public contracts using network analysis. The most cited publications in Table 17 were: (Fazekas & Tóth, 2016.2), with 37 citations; (Sedita & Apa, 2015), with 27 and (Lin, Jheng, & Yu, 2012), with 22 mentions, all related to Scopus. Besides, citations are also analyzed in section 3.4.2, referring individually to the authors and co-authors.

Major Journals

Concerning journals, a total of 14 articles were collected from 13 different journals and published by ten publishers. 38% of the journals are ranked by Scimago Journal & Country Rank (SJR) in the first quartile Q1, 48% in the second quartile Q2, and 14% in the Q3 quartile. The main fields of investigation are Computer Science, Government & Law and Business & Economics, Table 18.

Table 18: Major Journals

Journal	Qt.	Rank	Publisher	Country	Field
Expert Systems With Applications	1	Q1	Elsevier Science LTD	England	Computer Science;
Social Networks	2	Q1	Elsevier	Netherlands	Anthropology; Sociology
Politics and Governance	1	Q2	Cogitatio Press	Portugal	Government & Law
International Journal of Project Management	1	Q1	Elsevier Science LTD	England	Business & Economics
Science and Engineering Ethics	1	Q1	Springer	Netherlands	Multidisciplinary Sciences
Scientific Reports	1	Q1	Nature Publishing Group	England	Science & Technology

Applied Network Science	1	Q2	Springer	Switzerland	Computer Science
Entropy	1	Q2	MDPI	Switzerland	Multidisciplinary
International Journal of Data Science and Analit.	1	Q2	Springer	Switzerland	Computer Science
Royal Society Open Science	1	Q2	Royal Soc	England	Science & Technology
Policing & Society	1	Q2	Routledge Journals	England	Criminology & Penology
Advances in Civil Eng.	1	Q3	Hindawi LTD	England	Construction & Building
Political Research Quarterly	1	Q3	Sage Publications Inc	United States	Government & Law

Conference Proceedings

Regarding conferences, a total of four articles were collected from four different proceedings and published by three publishers belonging to three countries. The main field of investigation is computer science as shown in Table 19.

Table 19: Main Conferences

Conference	Qt.	Publisher	Country	Field
13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)	1	IEEE	USA	Computer Science
IEEE 5th International Conference on Future Internet of Things and Cloud (FiCloud)	1	IEEE	USA	Computer Science; Telecommunications
International Conference on Applied Mathematics, Modeling and Simulation (AMMS)	1	Atlantis Press	France	Computer Science; Mathematics
International Conference on Intelligent Data Engineering and Automated Learning	1	Springer International	Switzerland	Computer Science

The four conference proceedings identified in this study were listed in the Computer Science field of investigation. Regarding article publishers, the IEEE published two papers, specifically in Conferences on Fuzzy Systems and the Future Internet of Things and Cloud. Springer International Publishing AG has published one paper in the proceedings on Intelligent Data Engineering and Automated Learning. Atlantis Press published one paper on Mathematics, Modeling, and Simulation procedures.

Analysis of Word Incidence in Sections of Varied Articles

To further explore the most sensitive topics, we conducted a co-occurrence analysis. Here we use the concept of keyword occurrence, which is based on the count of words in parts of the articles (Keyword, Title, and Abstract). Besides, the analysis is applied to many articles together, serving researchers for review and statistical exploration.

a) Analysis of Keyword Occurrences

The study of keywords occurrence was developed using the VOSviewer, a software tool for constructing and visualizing bibliometric networks. We performed the search using the complete counting method, comprising 12 keywords in the Scopus and Web of Science repositories, whose minimum occurrence limit was established as two. In Table 20 we present the 10 Keywords with the most occurrence and link strength.

Table 20: Keywords with the highest occurrences

SCOPUS			Web of Science		
Keyword	Link Strength	Occurrence	Keyword	Link Strength	Occurrence
public procurement	13	7	corruption	12	7
corruption	9	7	government	7	4
crime	12	4	networks	6	3
social network analysis	8	4	social network analysis	4	3
social networking	8	3	markets	2	3
construction industry	7	3	public procurement	5	2
risk management	8	2	construction industry	4	2
anti-corruption	7	2	management	4	2
networks	5	2	centrality	3	2
human	4	2	collusion	2	2

The most identified keywords in the two repositories, Scopus and Web of Science, had the following occurrences and link strengths, respectively: public procurement (Scopus: 7, 13 and W.O.S: 2, 5), government (Scopus: 0, 0 and W.O.S: 4, 7), crime (Scopus: 4, 12 and W.O.S: 0, 0), corruption (Scopus: 7, 9and W.O.S: 7, 12), and social network analysis (Scopus: 4, 8 and W.O.S: 3, 4), networks (Scopus: 2, 5 and W.O.S: 3, 6).

The two keyword networks together (Scopus and Web of Science) emphasize the research areas of public procurement and corruption. They also indicate social network analysis as the science for detecting fraud in tenders. The two networks complement each other fully. Despite having some identical keywords, which reinforce their importance, they add, separately, others essential for the research, such as anti-corruption, risk management, collusion, and government.

b) Author Co-Authorship - Documents and Citations Within Multiple Articles

The study of authors' and co-authors' occurrence was carried out with the VOSviewer tool using a complete counting and limiting the number of author documents to a minimum of one. The search performed included 49 authors / co-authors in the Scopus and 48 in the Web of Science repositories. In Table 21, we present the 10 most-cited authors / co-authors in each repository.

Table 21: Highest citations from authors and co-authors

SCOPUS				Web of Science			
Author/ Co-author	Link strength	Documents	Citation	Author/ Co-author	Link strength	Documents	Citation
fazekas m.	4	3	40	fazekas, mihaly	4	3	34
tóth i.j.	1	1	36	toth, istvan janos	1	1	29
morselli c.	1	1	28	morselli, carlo	1	1	26
reeves-latour m.	1	1	28	reeves-latour, m.	1	1	26
apa r.	1	1	27	apa, roberta	1	1	25
sedita s.r.	1	1	27	sedita, silvia rita	1	1	25
jheng y.-y.	2	1	22	jheng, yi-ying	2	1	19
lin s.-j.	2	1	22	lin, shi-jen	2	1	19
yu c.-h.	2	1	22	yu, cheng-hsien	2	1	19
wachs j.	7	4	22	wachs, johannes	7	4	18

Combining the two repositories without intercessions, the authors/co-authors that stand out the most for their Link strength and the number of citations are Fazekas with three documents and Wachs with three articles.

Looking at the author and co-author network in Figure 16, filtered by authors with at least four citations, we present for the Web of Science and Scopus repositories the five biggest communities formed, with 21 nodes and 42 links each. Here we observe that the largest and most interconnected nodes belong to authors Mihaly Fazekas and Johannes Wachs. They have two and three articles as authors, respectively, and have written one paper together.

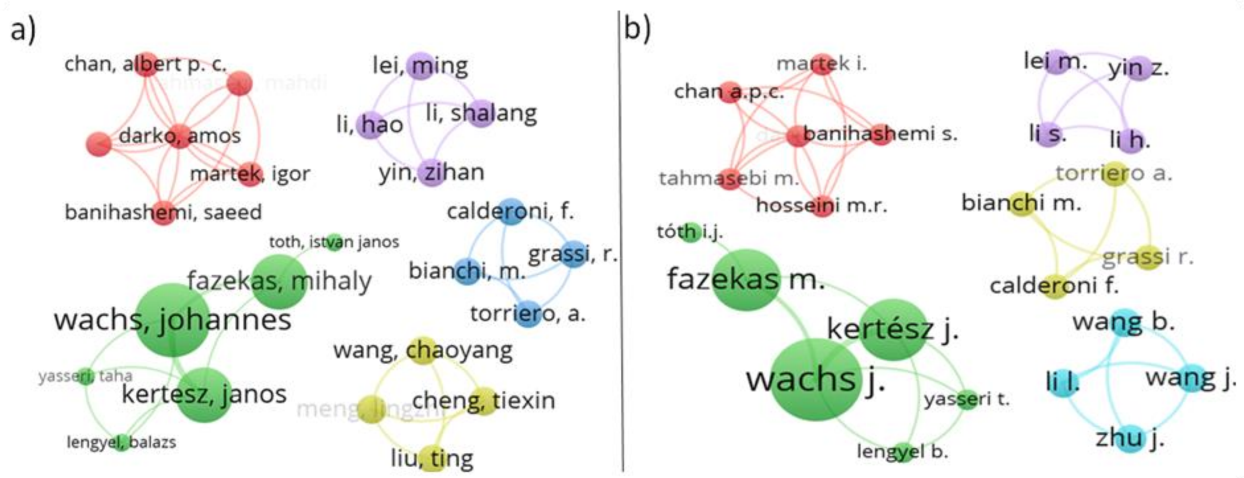

Figure 16: Author and co-authorship network view. a) Web of Science. b) Scopus.

c) Purpose and Methodology

All articles used connected networks for their evaluations. Concerning studies using labeled data or doing exploratory research, we have seven articles applying labeled data (Lin, Jheng, & Yu, 2012; Grassi, Calderoni, Bianchi, & Torriero, 2019; Hosseini, Martek, & Banihashemi, 2019; Lei, Yin, Li, & Li, 2017; Fazekas & Wachs, 2020; Wachs & Kertész, 2019; Davydenko, Morozov, & Burmistrov, 2017). The remaining 11 papers used exploratory analysis (Fazekas & Tóth, 2016.2; Lauchs, Keast, & Yousefpour, 2011; Sedita & Apa, 2015; Reeves-Latour & Morselli, 2017; Wachs, Yasseri, Lengyel, & Kertesz, 2019; Luna-Pla & Carlock, 2020; Cheng, Liu, & Meng, 2017; Lin & Khomnotai, 2016; Wachs, Fazekas, & Kertész, 2021) (Zhu, Wang, Li, & et, 2020; Carneiro, Veloso, Ventura, Palumbo, & Costa, 2020).

Out of the 18 articles, 13 (Cheng, Liu, & Meng, 2017; Davydenko, Morozov, & Burmistrov, 2017; Fazekas, Toth, & King, 2016.1; Fazekas & Wachs, 2020; Fazekas & Kocsis, 2020; Hosseini, Martek, & Banihashemi, 2019; Lei, Yin, Li, & Li, 2017) (Lin & Khomnotai, 2016; Lin, Jheng, & Yu, 2012; Wachs & Kertész, 2019; Wachs, Yasseri, Lengyel, & Kertesz, 2019; Wachs, Fazekas, & Kertész, 2021; Zhu, Wang, Li, & et, 2020) used cluster analysis (community detection or clustering) to classify the nodes and make analysis. Eight (Davydenko, Morozov, & Burmistrov, 2017; Fazekas, Toth, & King, 2016.1; Grassi, Calderoni, Bianchi, & Torriero, 2019; Hosseini, Martek, & Banihashemi, 2019; Luna-Pla & Carlock, 2020; Lauchs, Keast, & Yousefpour, 2011; Reeves-Latour & Morselli, 2017; Sedita & Apa, 2015) applied centralities and measures tools, which usually reflect the prominence, status, prestige, or visibility of a node and often explain behaviors on the network (Marsden, 2005). Two (Grassi, Calderoni, Bianchi, & Torriero, 2019; Wachs, Yasseri, Lengyel, & Kertesz, 2019) used bipartite network projection to create a firm-firm net and capture the hidden relationships between bidders. One (Lin & Khomnotai, 2016) used the concept of neighbor diversity, which is an effective resource to distinguish fraudsters from regular nodes.

The application of Network Analysis techniques varies, not only concerning the tools used but also regarding the objectives. The most cited article in the area (Fazekas, Toth, & King, 2016.1), for example, besides using community detection, centralities and measures, created a fraud risk index comprising intrinsic parameters of tenders and used it to identify a supposed state capture and centralization of the corrupt network after government turnover. On the other hand, Grassi et al. (Grassi, Calderoni, Bianchi, & Torriero, 2019) pointed out the betweenness centrality as the most standard measure to identify criminal agents and used the projection approach to capture strong brokerage positions in a mono-mode network and spot criminal leaders.

Discussion

The current research seeks to identify the most recent studies and their contributions regarding detecting fraud in public procurement using tools from network science. This section presents a brief analysis comparing the response of the repositories Scopus and Web of Science, analyze how we answer our research question, as follows:

a) Scopus and Web of Science Repositories

Concerning the two repositories, we found that they present an intersection of approximately 90% concerning the selected articles, with both adding the same amount of writings. Of the intersecting articles, Scopus has more citations in 75% of cases, Web of Science in 6%, and the same number of references in 19% of writings. Scopus also generates more keywords and related words in titles and abstracts, besides having greater link strength and word occurrence.

b) Research Question RQ1

RQ1: How can network science help in the disclosure of fraud in public procurement?

The review of the articles allowed the identification of Network Science techniques used to detect fraud in public procurement. We found that cluster analysis (community detection and clustering) and centrality measures are the most used tools in the 18 documents analyzed, with 13 and 9 occurrences, respectively. Community detection is one of the most used tools in studies to identify the formation of clusters, especially cartels. It was used by Wachs et al. (Wachs, Yasseri, Lengyel, & Kertesz, 2019) to investigate a cohesive and exclusive interaction in a large market, in which they found a niche with favorable conditions for the emergence of cartel collusion. However, the approach to detecting cartels was only suggestive and could not conclusively prove fraud without using labeled data.

Regarding clustering analysis, Cheng et al. (2017) identified that closely linked nodes with relatively large weights indicate companies in the communities with cooperative behavior and likely bidding fraudulently. Fazekas et al. (2016.2) detected companies that make extensive contracts among themselves rather than with other organizations outside their communities. Hosseini et al. (2019) analyzed communities in a contract network that refers to cases of violation of procedural principles in their concession, where they identified clusters in which none of the contracts had the capacity to deliver the awarded projects.

In turn, Wachs et al. (2021) quantified the extent to which a single bid varies across different communities within a network and interpreted it as significant evidence that corruption is not randomly distributed in the public sector. They concluded that the greater the grouping of single bidders in a market, the more likely it is that investigation of their neighbors would raise cases of fraud. In other research, Wachs & Kertesz (2019) identified that communities with an excess of bonding social capital are at higher risk of corruption, and communities with more diverse external connectivity are less exposed to fraud.

Zhu et al. (2020) divided a network into separate communities classifying overlaps as new communities and identifying collusion cases from the high degree of nodes overlapping. Fazekas & Kocsis (2020) analyzed the persistence of corrupt buyers in a low-competitive cluster over the years after government change to determine the corrupt link between public and private entities.

The centrality measures were used in the articles to help identify the most active, most influential, most central actors and those suspected of spurious associations. The degree centrality stood out in some articles (Fazekas & Tóth, 2016.2; Sedita & Apa, 2015; Lauchs, Keast, & Yousefpour, 2011; Luna-Pla & Carlock, 2020; Hosseini, Martek, & Banihashemi, 2019; Davydenko, Morozov, & Burmistrov, 2017) exposing the most connected actors in the network. However, while it helps to identify network hubs, criminal network interfaces are not necessarily hub-mediated as it becomes very evident and obvious on the network (Morselli, 2008). Therefore, the use of other measures and analyses is necessary to identify suspicious connections.

Fazekas et al. (2016.2), for instance, identified a non-linear and changeable relationship over time between the closeness centrality and a corruption risk index, making the closeness a strong indicator of the risk of corruption after a government turnover. Davydenko et al. (2017), on the other hand, despite using closeness centrality, identified that offenders were aware of the main indicators of doubt and were striving to circumvent them by lowering their average values through manipulation of information. In a cluster analysis, he showed that the most significant indicators to reveal suspicious elements were based on the contract price, contract name, closeness values, and average values of eigenvectors, revealing, through these measures, that 13% of government contracts had violations.

Testing the efficiency of the betweenness centrality in identifying brokers, Grassi et al. (2019) investigated a set of known leaders in a criminal network, which revealed a core group of about 20 heads with high betweenness values. On the other hand, they also found a small group of outlier leaders whose betweenness was low, which did not indicate them as possible brokers.

Luna-Pla & Carlock (2020), despite using network diameter, average path length, and average degree metrics, notably these remained essentially unaltered, not providing enough information to indicate the risk of corporate fraud, and the clustering coefficient was equal to zero due to the bipartite nature of the network. These network metrics were non-conclusive to establish clear corporate corruption risk indicators, except for the number of multi-edge node pairs, that provided a good measure of the degeneracy of both companies and people.

Since Network Science tools are not always able to promote precise results alone, it is often used with other parameters such as single bidder, neighbor diversity, winner diversity, winning rate, or even creating indicators based on tenders' parameters, as done in some studies (Fazekas, Toth, & King, 2016.1; Fazekas & Wachs, 2020; Luna-Pla & Carlock, 2020).

Conclusions

This research used the PRISMA methodology, an evidence-based collection of items that intend to help authors report on a broad spectrum of systematic reviews and meta-analyses. The method allowed for an organized presentation of studies and techniques used to detect fraud in public procurement. To select scientific articles, we used two repositories, Scopus and Web of Science. We compared their results separately to identify the contribution and determine the significance of each one in the research. The 18 selected articles were chosen based on the attempt to answer the research question referring to the techniques applied to detect fraud in public procurement using network science.

The most addressed technique was cluster analysis, which encompasses clustering and community detection. It was mainly used to explore locations of actors grouped with similar properties, helping to reveal hidden relations among them in the network. Another widely used technique was centrality, mainly degree centrality, closeness centrality, and betweenness centrality, helping to detect fraud. However, some articles reported that the use of these measures was not enough to detect suspicious episodes, but could assist the characterization and analysis of the network. Nevertheless, other articles created fraud indicators based on intrinsic bidding parameters and used cluster analysis to effectively identify fraud.

Regarding gaps, the literature review showed that the vast majority of fraud detection analyzes in public procurement occur in bipartite networks. Besides, the studies were not able to reconcile an analysis of projected co-bidding networks with the use of fraud indicators in a cluster investigation using network science.

Taking into account the gap presented, we suggest, as a guide for future studies, within the extensive subject of fraud detection in public contracts, investigating a projected co-bidding network using specific fraud risk indicators based on procurement process parameters.

References

Carneiro, D., Veloso, P., Ventura, A., Palumbo, G., & Costa, J. (2020). Network Analysis for Fraud Detection in Portuguese Public Procurement. *Conference on Intelligent Data Engineering and Automated Learning.*

Cheng, T., Liu, T., & Meng, L. (2017). The Analysis of Water Project Bid Rigging Behavior Based on Complex Network. *International Conference on Applied Mathematics, Modeling and Simulation (AMMS)*.

Costa, G. A., Machado, D., & Martins, V. (2020). The Efficiency of Social Control in Municipal Bidding: A Study in Social Observatories. *Sociedade Contabilidade e Gestão 14(4):112.*

Davydenko, V. I., Morozov, N. V., & Burmistrov, M. I. (2017). Adaptation of cluster analysis methods in respect to vector space of social network analysis indicators for revealing suspicious government contracts. *IEEE 5th International Conference on Future Internet of Things and Cloud (FiCloud)*.

Fazekas, M., & Tóth, I. J. (2016.2). From corruption to state capture: A new analytical framework with empirical applications from Hungary. *Political Research Quarterly, 69 (2), 320–334.*

Fazekas, M., & Wachs, J. (2020). Corruption and the network structure of public contracting markets across government change. *Politics and Governance, 8*(2), pp. 153-166.

Grassi, R., Calderoni, F., Bianchi, M., & Torriero, A. (2019). Betweenness to assess leaders in criminal networks: New evidence using the dual projection approach. *Social Networks, 56 , 23–32.*

Hosseini, M., Martek, I., & Banihashemi, S. (2019). Distinguishing Characteristics of Corruption Risks in Iranian Construction Projects: A Weighted Correlation Network Analysis.

Lei, M., Yin, Z., Li, S., & Li, H. (2017). Detecting the collusive bidding behavior in below average bid auction. *13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD).*

Lin, J., & Khomnotai, L. (2016). Improving Fraudster Detection in Online Auctions by Using Neighbor-Driven Attributes. *Entropy, Vol. 18, Ed:1, N:e18010011*.

Lin, S., Jheng, Y.-Y., & Yu, C. (2012). Combining ranking concept and social network analysis to

Luna-Pla, I., & Carlock, N. (2020). Corruption and complexity: a scientific framework for the

Marsden, P. (2005). Network Analysis. *In Encyclopedia of Social Measurement.*

Morselli, C. (2008). Inside Criminal Networks. Studies of Organized Crime. *Springer New York*.

Mufutau, G., & Mojisola, O. (2016). Detection and prevention of contract and procurement, fraud Catalyst to organization profitability. *Journal of Business and Management*, pp. 9-14.

Padhi, S., & Mohapatra, P. (2011). Detection of collusion in government procurement auctions. *J Purch Supply Manag, 17*, pp. 207-221.

Page, M., McKenzie, J., & Bossuyt, P. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International Journal of Surgery, 88, N. 105906*.

Reeves-Latour, M., & Morselli, C. (2017). Bid-rigging networks and state-corporate crime in the

Rustiarini, N., Sutrisno, T., Nurkholis, N., & Andayani, W. (2019). Why people commit public procurement fraud? the fraud diamond view. *Journal of Public Procurement 19(4), 345–362*.

Sedita, S., & Apa, R. (2015). The impact of inter-organizational relationships on contractors' success in winning public procurement projects: The case of the construction industry in the Veneto region. *International Journal of Project Management*.

Shaffril, H. A., Samsuddin, S. F., & Samah, A. A. (2020). The ABC of systematic literature review: The basic methodological guidance for beginners. *Quality & Quantity, 1–28*.

Silva Filho, J. B. (2017). A eficiência do controle social nas licitações e contratos administrativos. *Master's thesis - Universidade Nove de Julho, São Paulo*.

Wachs, J., & Kertesz, J. (2019). A network approach to cartel detection in public auction markets. *Scientific Reports, 9*. doi:10818

Wachs, J., Fazekas, M., & Kertész, J. (2021). Corruption Risk in Contracting Markets: A Network Science Perspective. *International Journal of Data Science and Analytics*.

Wachs, J., Yasseri, T., Lengyel, B., & Kertesz, J. (2019). Social capital predicts corruption risk in towns. *Royal Society Open Science*.

Wensink, W., & Vet, M. (2013). Identifying and Reducing Corruption in Public Procurement in the EU. *European Commission. Bruxelles*.

Whiteman, R. (2019). Fraud and corruption tracker. *The Chartered Institute of Public Finance and Accountancy – CIPFA.*

World_Bank_Group. (2017). A fair adjustment: efficiency and equity of public spending in

Zhu, J., Wang, B., Li, L., & et, a. (2020). Bidder Network Community Division and Collusion Suspicion Analysis in Chinese Construction Projects. *Advances in Civil Engineering*.

REFERENCES

Abidi, W., Daoudm, M., Ihnaini, B., & Khan, M. (2021). Real-Time Shill Bidding Fraud Detection Empowered with Fussed Machine Learning. *IEEE Access*.

Adjei-Bamfo P., M.-N. T. (2019). The role of e-government in sustainable public procurement in developing countries: a systematic literature review. *Resour Conserv Recycl*, p. 142:189. doi:203

Ahmed, M., Mahmood, A. N., & Islam, M. R. (2016). A survey of anomaly detection techniques in financial domain. Future Generation Computer Systems, 55, 278-288.

Albert, R., & Barabás, A.-L. (2002). Statistical mechanics of complex networks. *Rev Mod Phys, 30*(15). doi:47

Alencar, E. C. N. (2021). Assessing the potential for detecting fraud in the Brazilian public procurement using Latent Class Analysis. University of California, Irvine.

Alexopoulos, C., Charalabidis, Y., Androutsopoulou, A., Loutsaris, M. A., & Lachana, Z. (2019). Benefits and obstacles of blockchain applications in e-government.

Almendra, V. (2013). Finding the needle: A risk-based ranking of product listings at online auction sites for non-delivery fraud prediction. *Expert Systems With Applications*.

Almendra, V., & Enachescu, D. (2014). Using Self-Organizing Maps for fraud prediction at online auction sites. *15th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*.

Alzahrani, A., & Sadaoui, S. (2020). Clustering and Labeling Auction Fraud Data. *International Conference on Data Management, Analytics and Innovation (ICDMAI)*.

Amaral, M., Saussier, S., & Yvandre-Billon, A. (2013). Expected Number of Bidders and Winning Bids. *Journal of Transport Economics and Policy, 47:1*.

Amundsen, I. (2019). Extractive and power-preserving political corruption. In *Political Corruption in Africa* (pp. 1-28). Edward Elgar Publishing.

Andreucci, D. (2018). Populism, hegemony, and the politics of natural resource extraction in Evo Morales's Bolivia. *Antipode, 50*(4), pp. 825-845.

Angulo Garzaro, N. (2018). Eu competition law and public procurement: competition-driven limits imposed to public bodies when they source goods, works and services. *Works and Services*.

Anysz, H., Foremny, A., & Kulejewski, J. (2018). Comparison of ANN Classifier to the Neuro-Fuzzy System for Collusion Detection in the Tender Procedures of Road Construction Sector. *3rd World Multidisciplinary Civil Engineering, Architecture, Urban Planning Symposium (WMCAUS)*.

Aquino, A. C. B. D., Lino, A. F., & Azevedo, R. R. D. (2021). The embeddedness of digital infrastructures for data collection by the Courts of Accounts. Revista Contabilidade & Finanças, 33, 46-62.

Atkinson, C. (2019). Full and open competition in public procurement: values and ethics in contracting opportunity. *International Journal of Public Administration, 43*(13), pp. 1169-1182.

Bauhr, M., Czibik,, A., Licht, J., & Fazekas, M. (2019). Lights on the shadows of public procurement: Transparency as an antidote to corruption. *Governance, 33*(3), pp. 495-523.

Beittel, J., Meyer, P., Seelk, C., Taft-Morales, M., & Gracia, E. (2019). Combating corruption in Latin America: congressional considerations. *Congressional Research Service, Report number R45733*.

Blondel, V., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J Stat Mech: Theory Exp, 2008*(10). doi:10008

Bodin, O., & Norberg, J. (2005). Information network topologies for enhanced local adaptive management. *Environ Manag, 35*(2), pp. 175-193.

Borras, J., Cousins, B., Liu, J., McKay, B., & Milhorance De Castro, C. (2015). Emerging trends in global commodities markets: the role of Brazil and China in contemporary agrarian transformations. *International Institute of Social Studies, The Hague*.

Bosio, E., Djankov, S., Glaeser, E., & Shleifer, A. (2022). Public procurement in law and practice. *Am Econom Rev, 112*(4), pp. 1091-1117.

Bouman, J. (2021). Preventing Corruption of Civil Servants in France, Romania and the Netherlands: A Comparative Study of Law and Policy. *Wolters Kluwer*, pp. 1-307.

Carneiro, D., Veloso, P., Ventura, A., Palumbo, G., & Costa, J. (2020). Network Analysis for Fraud Detection in Portuguese Public Procurement. *Conference on Intelligent Data Engineering and Automated Learning.*

Carson, L., & Prado, M. M. (2014). *Brazilian anti-corruption legislation and its enforcement: Potential lessons for institutional design* (p. 18). IRIBA Working Paper: 09. Manchester. UK, 2014.

Chang, Z. (2018). Understanding the corruption networks revealed in the current Chinese anti-corruption campaign: a social network approach. *J Contemp China, 27*(113), pp. 735-747.

Charron, N., Dahlström, C., Fazekas, M., & Lapuente, V. (2017). Careers, connections, and corruption risks: Investigating the impact of bureaucratic meritocracy on public procurement processes. *J Polit, 79*(1), pp. 89-104.

Cheng, T., Liu, T., & Meng, L. (2017). The Analysis of Water Project Bid Rigging Behavior Based on Complex Network. *International Conference on Applied Mathematics, Modeling and Simulation (AMMS)*.

Chiu, C., Ku, Y., Lie, T., & Chen, Y. (2011). Internet auction fraud detection using social network analysis and classification tree approaches. *International Journal of Electronic Commerce*.

Chung, N., Miasojedow, B., Startek, M., & Gambin, A. (2019). Jaccard/Tanimoto similarity test and estimation methods for biological presence-absence data. *BMC Bioinform, 20*(15), pp. 1-11.

Clark, A. (2017). Measuring corruption: transparency international's "corruption perceptions index". *In Corruption, Accountability and Discretion, 29*, 3-22. Publishing, Emerald Limited.

Costa, G. A., Machado, D., & Martins, V. (2020). The Efficiency of Social Control in Municipal Bidding: A Study in Social Observatories. *Sociedade Contabilidade e Gestão 14(4):112.*

Curado, A., Damásio, B., Encarnação, S., Candia, C., & Pinheiro, F. (2020). Scaling behavior of public procurement activity. *arXiv preprint arXiv:2007.15276*.

Czibik, Á., Fazeka, s. M., Sanchez, A., & J., W. (2021). Networked Corruption Risks in European Defense Procurement. *Corruption Networks: Concepts and Applications*, 67-87.

Dadfarnia, M., Adibnia, F., Abadi, M., & Dorri, A. (2020). Incremental collusive fraud detection in large-scale online auction networks. *Journal of Supercomputing*.

Dávid-Barrett, E., & Fazekas, M. (2020). Grand corruption and government change: an analysis of partisan favoritism in public procurement. *European Journal on Criminal Policy and Research, 26*(4), 411-430.

Davis, K. E. (2019). Between Impunity and Imperialism: The Regulation of Transnational Bribery. *Oxford University Press*.

Davydenko, V. I., Morozov, N. V., & Burmistrov, M. I. (2017). Adaptation of cluster analysis methods in respect to vector space of social network analysis indicators for revealing suspicious government contracts. *IEEE 5th International Conference on Future Internet of Things and Cloud (FiCloud)*.

Dawar, K. and Oh, S.C. (2017), “The role of public procurement policy in driving industrial development, no. 8”, Inclusive and Sustainable Industrial Development Working Paper Series, Vienna.

De Catalunya, B. G. O. A. On Public Data Availability, Interoperability and Reusability.

De Maria, M., & Howai, N. (2021). The role of open data in fighting land corruption: evidence, opportunities and challenges.

Dener, C., Nii-Aponsah, H., Ghunney, L. E., & Johns, K. D. (2021). GovTech maturity index: The state of public sector digital transformation. World Bank Publications.

Dewi, H. R., Mahmudi, M., & Maulana, R. (2021). An analysis on fraud tendency of village government officials. *Jurnal Akuntansi dan Auditing Indonesia*, pp. 33-44.

Dhurandhar, A., Graves, B., Ravi, R., & Maniachari, G. (2015). Big data system for analyzing risky procurement entities. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Diesing, L., Unit, P. P., Magina, O. P., & Head, P. P. U. (2019). Innovation in the Evaluation of Public Procurement Systems. In *Joint public procurement and innovation: lessons accross borders* (pp. 349-372). Bruylant.

Ding, X., & Yang, Z. (2020). Knowledge mapping of platform research: a visual analysis using VOSviewer and CiteSpace. *Electronic Commerce Research*.

Dong, S., Wang, H., Mostafizi, A., & Song , X. (2020). A network-of-networks percolation analysis of cascading failures in spatially co-located road-sewer infrastructure networks. *Physica A: Statistical Mechanics and its Applications, 538, 122971*.

Duguay, R., Rauter, T., & Samuels, D. (2022). The impact of open data on public procurement. Available at SSRN 3483868.

Faría, H. J., Montesinos-Yufa, H. M., Morales, D. R., Aviles B, C. G., & Brito-Bigott, O. (2013). Does corruption cause encumbered business regulations? An IV approach. *Applied Economics, 45*(1), pp. 65-83.

Fazekas, M., & Kocsis, G. (2020). Uncovering High-Level Corruption: Cross-National Objective Corruption Risk Indicators Using Public Procurement Data. *British Journal of Political Science, 50*(1), pp. 155-164.

Fazekas, M., & Tóth, I. J. (2016.2). From corruption to state capture: A new analytical framework with empirical applications from Hungary. *Political Research Quarterly, 69 (2), 320–334.*

Fazekas, M., & Wachs, J. (2020). Corruption and the network structure of public contracting markets across government change. *Politics and Governance, 8*(2), pp. 153-166.

Fazekas, M., Cingolani, I., & Tóth, B. (2018). Innovations in objectively measuring corruption in public procurement. *Governance indicators. Approaches, progress, promise*, pp. 154-185.

Fazekas, M., Sberna, S., & Vannucci, A. (2021). The extra-legal governance of corruption: Tracing the organization of corruption in public procurement. *Governance*.

Fazekas, M., Skuhrovec, J., & Wachs, J. (2017). Corruption, government turnover, and public contracting market structure–insights using network analysis and objective corruption proxies. *Budapest: Government Transparency Institute*.

Fazekas, M., Toth, I. J., & King, L. P. (2016.1). An Objective Corruption Risk Index Using Public Procurement Data. *European Journal on Criminal Policy and Research, 22*(3), pp. 369-397.

Fierăscu, S. (2017). The networked phenomenon of state capture: network dynamics, unintended consequences, and business-political relations in Hungary. *Central European University, Budapest*.

Fortini, C., & Motta, F. (2016). Corrupção nas licitações e contratações públicas: sinais de alerta segundo a transparência Internacional. *Revista de Direito Administrativo e Constitucional*.

Gallego, J., Rivero, G., & Martinez, J. (2021). Preventing rather than punishing: An early warning model of malfeasance in public procurement. *International Journal of Forecasting, 37*(1), pp. 360-377.

Ganguly, S., & Sadaoui, S. (2018). Online Detection of Shill Bidding Fraud Based on Machine Learning Techniques. *31st International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems (IEA/AIE)*.

Gasparikova, S. (2012). The Czech Republic's experience with competition policy and law. In *Evolution of competition laws and their enforcement* (pp. 26-42). Routledge.

Gomes, C. (2016). Dispensa Indevida de Licitação: O Dano ao Erário na Caracterização do Ato de Improbidade Administrativa. *Universidade Federal de Campina Grande – UFCG*.

Goryunova, N., Baklanov, A., & Ianovski, E. (2022). A Noisy-Labels Approach to Detecting Uncompetitive Auctions. *7th International Conference on Machine Learning, Optimization, and Data Science (LOD) / 1st Symposium on Artificial Intelligence and Neuroscience (ACAIN), 13163*, pp. 185-200.

Gotelli, N. (2000). Null model analysis of species co-occurrence patterns. *Ecology, 81*(9), pp. 2606-2621.

Governo_Federal_do_Brasil. (2022). Decreto aprimora regras do pregão eletrônico. From https://www.gov.br/compras/pt-br/acesso-a-informacao/noticias/decreto-aprimoraregras-do-pregao-eletronico. Accessed on 08/04/2022

Grassi, R., Calderoni, F., Bianchi, M., & Torriero, A. (2019). Betweenness to assess leaders in criminal networks: New evidence using the dual projection approach. *Social Networks, 56 , 23–32*.

Hafner, M., Taylor, J., Disley, E., Thebes, S., Barberi, M., Stepanek, M., & Levi, M. (2016). The cost of non-Europe in the area of organised crime and corruption: annex II-corruption. *RAND Corporation, Santa Monica, CA*.

Hall, D. (2012). Corruption and public services. *Public Services International Research Unit (PSIRU)*, *1*(1), 1-50.

Hanák, T., & Serrat, C. (2018). Analysis of construction auctions data in Slovak public procurement. *Adv Civ Eng*, pp. 1-13.

Hawken, A., & Munck, G. L. (2009). Measuring corruption: A critical assessment and a proposal. *Perspectives on Corruption and Human Development, Macmillan*.

Heggstad, K., & Heggstad, M. (2011). The basics of integrity in procurement. *Bergen, Norway: Anti-Corruption Resource Centre Chr. Michelsen Institute (CMI)*.

Herrera, M. (2019). Using social network analysis in open contracting data to detect corruption and collusion risks. *Eindhoven University of Technology - Department of Mathematics and Computer Science*.

Horvat, T., Havaš, L., & Logozar, R. (2015). The Analysis of Keyword Occurrences Within Specific Parts of Multiple Articles — the Concept and the First Implementation. *SSN 1846-6168 UDK 001.8*.

Hosseini, M., Martek, I., & Banihashemi, S. (2019). Distinguishing Characteristics of Corruption Risks in Iranian Construction Projects: A Weighted Correlation Network Analysis. *Science and Engineering Ethics.*

Huber, M., & Imhof, D. (2019). Machine learning with screens for detecting bid-rigging cartels. *International Journal of Industrial Organization*.

IMPIC. (2018). Contratação Pública em Portugal. *Relatório anual 2018. Lisboa: IMPIC*.

Kartika, D., Kebijakan, I., Barang, P., & Pemerintah, J. (2020). The impact of e-procurement implementation on public procurement's corruption cases – evidences from Indonesia and India. *J Kajian Wilayah, 11*, pp. 193-212.

Kenny, C., & Musatova, M. (2010). Red Flags of Corruption in World Bank Projects: Na Analysis of Infrastructure Contracts. *Policy Research Working, Artc. 5243, World Bank, Washington DC*.

Kertész, e., & Wachs, J. (2021). Complexity science approach to economic crime. *Nat Rev Phys, 3*(2), pp. 70-71.

Khomnotai, L., & Lin, J. (2015). Detecting fraudsters in online auction using variations of neighbor diversity. *International Journal of Engineering and Technology Innovation.*

Kolan, N., Jailani, N., Abu Bakar, M., & Latih, R. (2018). Trust model based on Islamic business ethics and social network analysis. *International Journal on Advanced Science, Engineering and Information Technology, 8*(6), pp. 2323-2331.

Kuncoro, M. (2013). Research methods for business and economics. *Indonesia Erlangga Publisher*(4).

Lauchs, M., Keast, R., & Yousefpour, N. (2011). Corrupt police networks: uncovering hidden relationship patterns, functions and roles. *Policing & Society 21 (1), pp.110-12.*

Lei, M., Yin, Z., Li, S., & Li, H. (2017). Detecting the collusive bidding behavior in below average bid auction. *13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD).*

Lemke, M., Mezalis, E., Piasta, D., & Smith, S. (2020). Central public procurement institutions in the Western Balkans: With selected EU country examples.

Li, Z., Shen, H., & Sapra, K. (2013). Leveraging social networks to combat collusion in reputation systems for peer-to-peer networks. *IEEE Transactions on Computers*.

Lin, J., & Khomnotai, L. (2016). Improving Fraudster Detection in Online Auctions by Using Neighbor-Driven Attributes. *Entropy, Vol. 18, Ed:1, N:e18010011*.

Lin, S., Jheng, Y.-Y., & Yu, C. (2012). Combining ranking concept and social network analysis to detect collusive groups in online auctions. *Expert Systems With Applications.*

Lisciandra, M., Milani, R., & Millemaci, E. (2021). A corruption risk indicator for public procurement. *European Journal of Political Economy, 73, 102141*.

Locatelli, G., Mariani, G., Sainati, T., & Greco, M. (2017). Corruption in public projects and megaprojects: There is an elephant in the room! *Int. J. Proj. Manag, 35*, pp. 252-268.

Luna-Pla, I., & Carlock, N. (2020). Corruption and complexity: a scientific framework for the analysis of corruption networks. *Applied Network Science*.

Lyra, M. S., Damásio, B., Pinheiro, F. L., & Bacao, F. (2022). Fraud, corruption, and collusion in public procurement activities, a systematic literature review on data-driven methods. *Applied Network Science, 7*(1), p. 83.

Lyra, M., Bacao, F., & Pinheiro, F. (2021.1). Public Procurement Fraud Detection – A Review using Network Analysis. *Complex Network 2021.*

Lyra, M., Curado, A., Damásio, B., Bacao, F., & Pinheiro, F. (2021.2). Characterization of the Firm-Firm Public Procurement Co-bidding Network from the State of Ceara (Brazil) Municipalities. *Applied Network Science, 6 (1)*.

Mahdillou, H., & Akbary, J. (2014). E-procurement adoption, its benefits and costs.

Mainali, K., Bewick, S., Thielen, P., Mehoke, T., Breitwieser, F., Paudel, S., . . . Karig, D. (2017). Statistical analysis of co-occurrence patterns in microbial presence-absence datasets. *PLoS ONE, 12*(11). doi:0187132

Majadi, N., Trevathan, J., & Bergmann, N. (2017). Analysis on bidding behaviours for detecting shill bidders in online auctions. *Proceedings - 2016 16th IEEE International Conference on Computer and Information Technology, CIT 2016, 2016 6th International*.

Mamavi, O., Meier, O., & Zerbib, R. (2017). How do strategic networks influence awarding contract? Evidence from French public procurement. *International Journal of Public Sector Management*.

Marsden, P. (2005). Network Analysis. *In Encyclopedia of Social Measurement.*

Martins, D. G. (2008). O estado da arte da capacidade institucional: uma revisão sistemática da literatura em língua portuguesa. *EBAPE.BR, 19*(1), pp. 165-189.

McAlister, D. T., & Ferrell, O. C. (2016). Corporate governance and ethical leadership. In Business ethics: New challenges for business schools and corporate leaders (pp. 68-93). Routledge.

Merigó, J. (2016). Academic research in innovation: A country analysis. *Scientometrics 108, 559–593*.

Merigó, J., Mas-Tur, A., Roig-Tierno, N., & Ribeiro-Soriano, D. (2015). A bibliometric overview of the Journal of Business Research between 1973 and 2014. *J. Bus. Res., 68, 2645–2653.*

Miasato, C. S. (2017). Licitação: Da Perfeição Idealizada à Parcial Ineficácia do Sistema. Faculdade de Ensino Superior e Formação Integral – FAEF

Modrušan, N., Rabuzin, K., & Mršic, L. (2021). Review of Public Procurement Fraud Detection Techniques Powered by Emerging Technologies. International Journal of Advanced Computer Science and Applications, 12(2).

Mlondo, N. (2013). Effectiveness of knowing your customer policy in combating money laundering in commercial banks in Tanzania: a case of bank of Africa. *(T) Limited (Doctoral dissertation, Mzumbe University)*.

Morselli, C. (2008). Inside Criminal Networks. Studies of Organized Crime. *Springer New York*.

Morselli, C., & Ouellet, M. (2018). Network similarity and collusion. *Social Networks, 55*, pp. 21-30.

Motta, F. (2003). *Teoria das Organizações: evolução e crítica* (2 ed.). São Paulo: Pioneira Thomson Learning.

Mufutau, G., & Mojisola, O. (2016). Detection and prevention of contract and procurement, fraud Catalyst to organization profitability. *Journal of Business and Management*, pp. 9-14.

Mundra, A., & Rakesh, N. (2014). Online hybrid model for online fraud prevention and detection. *Advances in Intelligent Systems and Computing*.

Murray, J. (2014). Procurement fraud vulnerability: a case study. *EDPACS EDP Audit Control Secur Newsletter, 49*(5), pp. 7-17.

Mvelase, T. (2015). The impact of non-compliance with Eskom procurement policies. *Nelson Mandela Metropolitan University - Doctoral dissertation*.

National Bureau of Statistics (NBS) (2019), "National Accounts Statistics, Popular Version 2018", Dar es salaam, Tanzania

Newman, M., Cantwell, G., & Young, J. (2020). Improved mutual information measure for clustering, classification, and community detection. *Phys Rev E, 101*(4). doi:042304

Nicolás-Carlock, J., & Luna-Pla, I. (2021). Conspiracy of Corporate Networks in Corruption Scandals. *Frontiers in Physics*.

Nikolova, E. V. (2019). Using Data Analysis and Instrumental Variables to Study the Drivers of Corruption. *SAGE Research Methods Cases*.

Nowrousian, B. (2019). Combatting public procurement criminality or simple rules for complex cases. *J Financ Crime, 26*, pp. 203-210.

O Povo (2017). Operação Fraternidade é deflagrada pela Polícia Federal no Ceará. Jornal O Povo. 13 de setembro de 2017. Consultado em 20 de novembro de 2022. http://blog.opovo.com.br/politica/operacao-fraternidade-deflagrada-pf/

Oberoi, R. (2014). Mapping the matrix of corruption: Tracking the empirical evidences and tailoring responses. Journal of Asian and African Studies, 49(2), 187-214.

OECD. (2015). Preventing corruption in public procurement. From OECD.org: https://www.oecd.org/gov/public-procurement/integrity/

OECD. (2017). *Government at a glance—2017 edition: public procurement.* From OECD.Stat: https://stats.oecd.org/Index.aspx?QueryId=78413. Accessed: 14 March 2021

OECD. (2021). Government at a Glance 2021. *OECD Publishing, Paris*.

Olsson, P., Folke, C., & Berkes, F. (2004). Adaptive comanagement for building resilience in social-ecological systems. Environ Manag. *Environ Manag, 34*(1), pp. 75-90. From https://doi.org/10.1007/s00267-003-0101-7

Owusu, E., Chan, A., & Wang, T. (2021). Tackling corruption in urban infrastructure procurement: Dynamic evaluation of the critical constructs and the anti-corruption measures. *Cities. Hong Kong Polytechnic University, 119. N. 103379*.

Padhi, S., & Mohapatra, P. (2011). Detection of collusion in government procurement auctions. *J. Purch. Supply Manag. 17, 207–221.*

Paganetto, L., & Scandizzo, P. (2015). Governance, moral and economic values in achieving dynamism in an anaemic Europe. *Springer, Cham*.

Page, M., McKenzie, J., & Bossuyt, P. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International Journal of Surgery, 88, N. 105906*.

Pavlopoulos, G. A., Kontou, P. I., Pavlopoulou, A., Bouyioukos, C., Markou, E., & Bagos, P. G. (2018). Bipartite graphs in systems biology and medicine: a survey of methods and applications. GigaScience, 7(4), giy014.

Perevezentceva, E., Osipova, K., & Skrynnikova, K. (2021). Corruption in the Implementation of Public Procurement from Small and Medium Businesses. *Lecture Notes in Networks and Systems*.

Piccolo, S., Lehmann, S., & Maier, A. (2018). Design process robustness: a bipartite network analysis reveals the central importance of people. *Des Sci, 4:e1*.

Popa, M. (2019). Uncovering the structure of public procurement transactions. *Business and Politics*.

Porter, R. H., & Zona, J. D. (1993). Detection of bid rigging in procurement auctions. *Journal of political economy, 101*(3), pp. 518-538.

Prado, M. M., & Cornelius, E. (2020). Institutional Multiplicity and the Fight Against Corruption: A Research Agenda for the Brazilian Accountability Network. *Revista Direito GV*, p. 16.

Rabuzin, K., & Modrusan, N. (2019). Prediction of public procurement corruption indices using machine learning methods. *11th International Conference on Knowledge Management and Information Systems (KMIS)*.

Ramalho, H., Almeida, A., & Fraga, A. (2020). Detecção de Casos Suspeitos de Conluio em Licitações Públicas: uma Aplicação do Algoritmo a Priori de Aprendizado de Máquina para o Estado da Paraíba. *Teoria e Prática em Administração, 10*(2), pp. 5-22.

Ratcliff, C., Martinello, B., & Kaiser, K. (2022). *Factsheets on the European Union: Public Procurement Contracts*. From https://www.europarl.europa.eu/factsheets/en/sheet/34/public-procurement-contracts

Rauen, A., & Paiva, B. (n.d.). *Impacts of Public Procurement on Business R&D Efforts: The Brasilian Case*. From Instituto de Pesquisa Econômica Aplicada – IPEA. Report number: 0249.

Reeves-Latour, M., & Morselli, C. (2017). Bid-rigging networks and state-corporate crime in the construction industry. *Social Networks, 51*, pp. 158-170.

Robson, C. (2002). Real world research: a resource for social scientists and practitioner-researchers. *Wiley-Blackwell*.

Roca, T., Orme, W., & Brown, J. (2010). Fear and Loathing of the corruption Perception Index: Does Transparency International Penalize Press freedom? *Available at SSRN 1694211*.

Rustiarini, N., Sutrisno, T., Nurkholis, N., & Andayani, W. (2019). Why people commit public procurement fraud? the fraud diamond view. *Journal of Public Procurement 19(4), 345–362*.

Sandström, A., & Carlsson, L. (2008). The performance of policy networks: the relation between network structure and network performance. *Policy Stud J, 36*(4), pp. 497-524.

Sedita, S., & Apa, R. (2015). The impact of inter-organizational relationships on contractors' success in winning public procurement projects: The case of the construction industry in the Veneto region. *International Journal of Project Management*.

Shaffril, H. A., Samsuddin, S. F., & Samah, A. A. (2020). The ABC of systematic literature review: The basic methodological guidance for beginners. *Quality & Quantity, 1–28*.

Sharkawy, M. H. D. (2022). Potential applications of collaborative intelligence technologies in manufacturing: study of applicability of collaborative intelligence technologies in manufacturing small-and-medium enterprises, collaborative intelligence frameworks, application benefits and adoption barriers.

Sharma, S., Sengupta, A., & Panja, S. (2019). Mapping Corruption Risks in Public Procurement: Uncovering Improvement Opportunities and Strengthening Controls. *Public Performance & Management Review*.

Silva Filho, J. B. (2017). A eficiência do controle social nas licitações e contratos administrativos. *Master's thesis - Universidade Nove de Julho, São Paulo*.

Sloane, M., Chowdhury, R., Havens, J. C., Lazovich, T., & Rincon Alba, L. (2021). AI and Procurement-A Primer.

Smith, C. D. (2022). A Study of Public Procurement Fraud: Framing the Study. In White-Collar Crime and the Public Sector: An Interdisciplinary Approach to Public Procurement Fraud (pp. 47-64). Cham: Springer International Publishing.

Soni, V. D. (2019). Role of artificial intelligence in combating cyber threats in banking. International Engineering Journal For Research & Development, 4(1), 7-7.

Søreide, T. (2002). Corruption in public procurement. Causes, consequences and cures. *Chr. Michelsen Institute*.

Spagnolo, G. (2012). Reputation, competition, and entry in procurement. *Int J Ind Organ, 30*(3), pp. 291-296.

Speck, B., & Ferreira, V. (2012). Sistemas de Integridade nos Estados Brasileiros. *Instituto Ethos de Empresas e Responsabilidade Social, 12*, pp. 30-82.

Spink, J. W. (2019). International Public and Private Response. *Food Fraud Prevention: Introduction, Implementation, and Management*, pp. 443-478.

Toth, B., Fazekas, M., Czibik, A., & Toth, I. (2014). Toolkit for detecting collusive bidding in public procurement with examples from Hungary. *Corruption Research Center Budapest (CRC-WP/2014:02)*.

Transparency International (2022). Corruption Perceptions Index. Transparency.org. Retrieved 2022-12-1

Tribunal de Contas do Estado do Ceará (2022). https://www.tce.ce.gov.br/cidadao/consulta-de-processos. Accessed on 07/07/2022.

Van_Eck, N., & Waltman, L. (2014). Visualizing bibliometric networks. *In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), Measuring scholarly impact: Methods and practice (pp. 285-320). Springer.*

Veech, J. (2013). A probabilistic model for analysing species co-occurrence. *Glob Ecol Biogeogr, 22*(2), pp. 252-260.

Velasco, R., Carpanese, I., Interian, R., & et, a. (2021). A decision support system for fraud detection in public procurement. *International Transactions in Operational Research*.

Volosin, N. (2015). Datos abiertos, corrupción y compras públicas. *RIGA - ILDA*.

Wachs, J., & Kertesz, J. (2019). A network approach to cartel detection in public auction markets. *Scientific Reports, 9*. doi:10818

Wachs, J., & Kertész, J. (2019). A network approach to cartel detection in public auction markets. *Sci Rep, 9*. doi:10818

Wachs, J., Fazekas, M., & Kertész, J. (2021). Corruption Risk in Contracting Markets: A Network Science Perspective. *International Journal of Data Science and Analytics*.

Wachs, J., Yasseri, T., Lengyel, B., & Kertesz, J. (2019). Social capital predicts corruption risk in towns. *Royal Society Open Science*.

Weigel, U., & Ruecker, M. (2017). e-procurement. *In: the strategic procurement practice guide*, 179-209.

Wensink, W., & Vet, M. (2013). Identifying and Reducing Corruption in Public Procurement in the EU. *European Commission. Bruxelles*.

Whiteman, R. (2019). Fraud and corruption tracker. *The Chartered Institute of Public Finance and Accountancy – CIPFA.*

World_Bank_Group. (2017). A fair adjustment: efficiency and equity of public spending in Brazil. *Brazil public expenditure review.[S. l.]: World Bank Group, 1*.

Wu Chebili, B., La Cascia, H., Collineau, F., Salomon, A., Calvet, B., & Moreau, Y. (2022). Electronic government procurement implementation types.

Xiao, L., Ye, K., Zhou, J., & Ye, X. (2021). A social network-based examination on bid riggers' relationships in the construction industry: A case study of China. *Buildings*.

Xiao, Q., Chang, H., Geng, G., & Liu, Y. (2018). An ensemble machine-learning model to predict historical PM2. 5 concentrations in China from satellite data. *Environ Sci Technol, 52*(22), pp. 13260-13269.

Yu, C., & Lin, S. (2013). Fuzzy rule optimization for online auction frauds detection based on genetic algorithm. *Electronic Commerce Research*.

Zeferino, L. (2020). A corrupção na construção de edifícios públicos no brasil: análise de instrumentos inibidores e facilitadores na etapa de projeto arquitetônico. *Universidade de De Brasília – UnB Master's thesis*.

Železnik, D., Vošner, H., & Kokol, P. (2017). A bibliometric analysis of the Journal of Advanced Nursing, 1976–2015. *J. Adv. Nurs, 73, 2407–2419.*

Zhu, J., Wang, B., Li, L., & et, a. (2020). Bidder Network Community Division and Collusion Suspicion Analysis in Chinese Construction Projects. *Advances in Civil Engineering*.

Zhu, Y., Zhang, W., & Yu, C. (2011). Detection of feedback reputation fraud in Taobao using social network theory. *Proceedings - 2011 International Joint Conference on Service Sciences, IJCSS*

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa