SSE Riga Student Research Papers
2023 : 9 (261)

TRAGEDY OF THE COMMON LATVIAN: LATVIAN MUNICIPALITY CORRUPTION RISK RATING

Authors: Jokūbas Vaičiulis
Kristers Šavalgins

ISSN 1691-4643
ISBN 978-9934-623-11-0

May 2023
Riga

Tragedy of the Common Latvian:
Latvian Municipality Corruption Risk Rating

Jokūbas Vaičiulis

and

Kristers Šavalgins

Supervisor: Dominik Gerber

JEL codes: D72, D73

May 2023
Riga

COPYRIGHT DECLARATION AND LICENCE

Name(s) of the author(s) in full: Jokūbas Vaičiulis, Kristers Šavalgins

Title of the Thesis: Public procurement corruption drivers in Latvian municipalities

I (We) hereby certify that the above-named thesis is entirely the work of the persons named below, and that all materials, sources and data used in the thesis have been duly referenced. This thesis – in its entirety or in any part thereof – has never been submitted to any other degree commission or published.

In accordance with Section 1 of the Copyright Law of Latvia, the person(s) named below is (are) the author(s) of this thesis.

Pursuant to Article 40 of the Copyright Law the author(s) hereby agree(s) and give an explicit licence to SSE Riga to deposit one digital copy of this thesis in the digital catalogue and data base at SSE Riga Library for an unlimited time and without royalty. The licence permits SSE Riga to grant access to the duly deposited thesis to all users of the catalogue and data base without royalty and limitations to downloading, copying and printing of the digital thesis in whole or in part provided I am (we are) indicated as the author(s) of the thesis according to Clause 4 Section 1 Article 14 of Copyright Law. I (We) assert my (our) right to be identified as the author(s) of this thesis whenever it is reproduced in full or in part.

Signed

/signed/ /signed/

Kristers Šavalgins Jokūbas Vaičiulis

Date

………………………………………

Table of Contents

Abstract **5**

1. Introduction **6**

2. Literature Review **7**

2.1. Short Overview of Corruption 7

2.1.1. Definition of corruption 7

2.1.2 Theoretical framework 8

2.1.3. Corruption Drivers 8

2.2. Public Procurement 9

2.2.1. European Union regulations 10

2.2.2. Latvian regulations 11

2.3. Risk-Based Measurement 11

2.3.1. Measuring corruption 11

2.3.2. Red flag method 14

2.3.3. Risks of Red Flags 14

3. Methodology **17**

3.1. Data Collection and Manipulation 17

3.2. Variables 19

3.2.1 Red flags (explanatory variables) 19

3.2.2. Dependent variables of the first regression analysis 22

3.2.3. Control (explanatory) variables of the first regression analysis 23

3.2.4. Independent variables for the second regression analysis 24

3.2.5 Descriptive statistics 26

3.3. Research Methods 28

3.3.1 Red flag regression 28

3.3.2 Creation of the municipal score 29

3.3.3 Validation of corruption risk scores 30

3.3.4 Municipal regression 31

4. Findings **31**

4.1 Regression Results 31

4.2 Calculation of Contract-Level Scores and Analysis 34

4.3 Contract Score Validation 35

4.3.1. Municipal corruption cases in media 35

4.3.2. Financial performance validation 35

4.4 Municipal Score 36
4.5 Municipal Regression 38
4.6 Robustness 38
5. Discussion 39
5.1. Red Flag Regression 39
5.2. Municipal scores 40
5.3. Municipal regression 41
5.4. Limitations 41
6. Conclusion 42
7. References 44
8. Appendices 50
Appendix A. Tables and figures 50
Appendix B. VIF score results 65

Abstract

This thesis aims to check whether variance in corruption risk exists among Latvian municipalities. We use the Red Flag method and a database of all public procurement tenders performed by municipalities in Latvia between 2013 and 2021 to create a public procurement risk score. Additionally, we provide time-based and geographic descriptive statistical analysis of what drives municipal corruption risk. The results indicate that the risk of municipal corruption is geographically dispersed. The municipality with the most corruption risk is Varakļāni and Riga, while Līgatne - has the least. We find faint links between corruption risk scores and distance from the capital and political party ties in the government. Since there has historically been a gap in the research surrounding municipal corruption, the results and the ranking ascertain the variance of corruption in Latvia and help shape further research in the field of fighting municipal corruption.

1. Introduction

Public procurement refers to purchasing goods or services by the government or a government-owned entity (OECD, 2022). Corruption in this context would mean corrupt actors taking advantage of the inefficiencies in the system to influence contracts and get a desirable result for themselves. Usually, when measuring corruption, the result is homogeneously attributed to the whole country (Transparency International, n.d.), which leaves out an important dimension - municipal variations.

Public procurement makes up around 12% of GDP in OECD countries (OECD, 2019). It might sound small, but it still comprises billions of Euros spent annually. Even if a small portion of this is lost due to corruption, millions of taxpayers' Euros are not reaching their intended goal. Wensik and de Vet (2013) estimated that roughly 13% of public procurement costs are lost due to corruption. If we apply these EU and OECD average numbers to Latvia, that will mean around 500 million euros lost to corruption each year.

Corruption has long been present in the again independent Latvia. Corruption was rampant in Latvia just after the fall of the USSR, and while the situation undoubtedly has improved, the rule of law has never truly gained trust. The Latvian parliament and ministries regularly rank amongst the least trusted institutions in Latvia (KNAB, 2022). Interestingly, municipal institutions often rank high in trust surveys, while at least from a theoretical point of view, the more dispersed scrutiny might be a breeding ground for corruption.

Most countries employ a public procurement system that has elements of both centralised and decentralised systems. A municipal - decentralised system has advantages and disadvantages - voters might be more in tune with what their politicians are doing, which could decrease corruption risks. However, it also decreases oversight potential, increasing corruption risks simultaneously. Since 2010, Latvian municipalities and municipally owned enterprises have procured around 37 thousand contracts worth more than 5 billion euros, roughly 11% of the money spent in total (IUB, n.d.). Despite the volume of deals, Latvian municipality public procurement deals are often found to be involved in corruption, like the various scandals in the Riga municipality over the years (Puriņa, 2018). The understanding of local-level corruption risks remains weak and is mainly derived from the national level, which limits policymakers' possibilities to introduce new and innovative monitoring or mitigating measures.

Research question nr. 1: Does the public procurement corruption risk vary across Latvian municipalities?

Research question nr. 2: What drives corruption risk in Latvian municipalities?

In this paper, we will use the academically widely used Red Flag method to analyse corruption risk on a municipal level. The chosen corruption indicators or red flags, among control variables were regressed onto single bid using the Ordinary Least Squares regression method. Using the regression coefficients created in the first part of the analysis, we attributed a risk score to each contract and aggregated it to the municipal level. Afterward, we ran a regression using municipal variables to understand the factors influencing our created corruption index.

2. Literature Review

2.1. Short Overview of Corruption

2.1.1. Definition of corruption

Corruption is an important and relevant problem in general and especially in the case of Latvia and its municipalities. While the concept of corruption is often portrayed in the media without much thought, academically, the definition of corruption and the factors affecting it are vague and the subject of intense debate. Many definitions like Transparency International's are pretty broad: "... abuse of entrusted power for private gain." (Transparency International, n.d.). For our research, we need to clearly define the forms of abuse that come with it, define the variables, and draw their impacts. Forms of abuse related to corruption can be split into two: "bribery" and "extortion" (Morris, 2011). Bribery can be defined as corrupt actors making illegal payments to change the outgoing legislation (Morris, 2011).

On the other hand, extortion refers to the abuse of power by political actors to coerce an illegal payment from an affected party (Morris, 2011). Additionally, corruption may occur on several levels of authority - from a policeman/woman taking a bribe to the head of the state - corruption that is happening on the highest levels of the bureaucratic apparatus and the lowest.

For the scope of this study, we will look at the possible corruption occurring on the municipal level, so the actors of the act of corruption will have to be defined as the people responsible for approving and issuing public procurement contests and the companies that take on the project. We define corruption in this as corrupt actors taking advantage of the inefficiencies in the system to influence contracts and gain personal benefit.

2.1.2 Theoretical framework

Public procurement pits two interests against each other, which can be described as an extended principal-agent relationship (Dahlstrom, 2009). The principal - the municipality or other governing body performing the tender, grants power to a decision-maker unit - usually a panel of bureaucrats or elected representatives, which evaluates bids of the agents - entities supplying goods or services. Per our definition, in this scenario, the decision maker would mutually agree with the agent to exchange goods or services in favour of abusing power granted by the principal. While the principal’s aim most of the time is to receive a set of goods or services for the lowest possible price, the agent’s aim is to offer the goods or services for the highest possible price to extract the corruption rents. Petrovska (2008) uses a simplified principal-agent model and introduces public procurement contract monitoring costs to find break-even points where an increased penalty would increase corruption savings. While Petrovska proposes to increase penalties through ex-ante assessment of public tenders, our approach focuses on increasing the productivity of existing monitoring agencies by pinpointing the riskiest contracts to check first.

The mutual agreement between the decision maker and the agent can be executed in multiple ways. Fazekas, King, and Toth (2016) propose three ways of influencing tenders: limiting bidders in the submission phase, excluding bidders in the evaluation phase, or modifying conditions ex-post.

2.1.3. Corruption Drivers

Besides defining corruption, it is also essential to understand what influences it to prescribe the correct solutions for the identified problems. For this study, we will refer to corruption drivers as the leading force impacting the presence of corruption. Corruption drivers are the factors that knowingly or unknowingly influence the political actors and tender participants to act corrupt. There are national and municipal level corruption drivers.

The national drivers of corruption affect the entire country, as to say, on a macro level, as opposed to impacting only the given municipality, for example. In previous literature, there are various factors explaining the presence of corruption. Mishra A. (2006) finds a pattern that corrupt activity is linked to how much corrupt activity is currently happening in society, meaning that corruption levels are affected by historical corruption. In addition to that, national corruption levels are also found to be affected by three factors: frequency of changes in the constitution, possession of wealth deriving from public goods, and the concentration of authority in the hands of a single individual (Goutte, Péran & Porcher, 2022). Additionally, unlimited and exclusive access to information and access to endowments from abroad can be thought of as a corruption driver (Soreide, 2014). Unlimited access to information can induce the risk of corruption, as some actors in the procurement would be acting with knowledge that is not available to the public, giving them an unfair advantage in securing favorable deals. The corruption-driving nature of the funds from abroad mainly refers to the natural resource-rich nations, depending on the revenues generated from selling the said resources. The link between the availability of funds from abroad and corruption risk in the form of oil can also be seen in the work of Shaxson, N. (2007).

Unlike the national drivers, municipal corruption drivers stem mainly from the differences between the studied municipalities. Charron et al. (2016) find that the risk of corruption is lowered if the bureaucratic apparatus enables bureaucrats' careers to be less reliant on their political connections. This suggests that lower corruption on the municipal level is achieved by enforcing the professional competencies to rule over the connections that a given bureaucracy has. The research on municipal corruption drivers is relatively scarce, indicating we could contribute to closing a gap in research.

2.2. Public Procurement

Public procurement is the purchase by governments and state-owned enterprises of goods, services, and works. (OECD, 2022). Latvian law, specifically, has different monetary thresholds on what it considers public procurement and what it defines as spending.

Following the proposed definition, it is important to look at the actors of the two-sided exchange - the principal or government enterprises and the companies or agents providing said services. From the government's perspective, public procurement often falls into 3 venues - procurement by the central government, municipal authorities, and by state-controlled or owned enterprises that fall under the jurisdiction of the government. From the other side of the coin, the fulfillers of the contract can be all entities, disregarding the country's origins or sector. For the purposes of this research, the receivers of the procurement will be the municipalities, i.e. we will only be looking at the tenders procured by the municipalities, disregarding all other governmental contests.

2.2.1. European Union regulations

Before delving into the legal intricacies of the country of interest, Latvia, it is important to understand the roots of the Latvian regulation - European public procurement directives. With 14% of EU GDP being the cause of public procurement activities (European Commission, n.d.), even a 1% improvement in efficiency, would result in savings of more than EUR 20 billion.

In 2017 the European Union issued a communication on how to improve the public procurement process in the member states. These recommendations included boosting the digitalization of public procurement and improving the availability of access to public procurement markets.

To establish the influence of European Union public procurement laws, it has to be stated that the European Union does not impose a joint public procurement basis in all member states of the union, but rather imposes that the context of the EU public procurement directives must be carried through to the national directives (Sigma, 2016). Additionally, the European legislative framework does not set out to directly influence all public procurement on the national level, but rather the contests of a significant size, which could attract parties from other member states. Following this, the basis of EU legislature is to ensure that the notion of a market without barriers is followed. Thus, the main role of the directives is to ensure equal competition of procurement actors of all nationalities, to evade the notion of choosing a national supplier of goods and services.

2.2.2. Latvian regulations

In Latvia, public procurement is regulated by the public procurement law (Saeima, 2022) based on European Union and Parliament directives. The law outlines the classification of various contracts, the procedure, and other exceptional cases.

Several state organizations are involved in public procurement's facilitation and regulatory process. The drafting of the law is the responsibility of the Ministry of Finance. The State Regional Development Agency is responsible for managing the e-procurement service. However, more applicable to our research is the Procurement Monitoring Bureau, which under the supervision of the Ministry of Finance, is responsible for compliance and guidelines for public procurement. The Bureau also acts as the central data aggregator and first-instance review in case of disputes. In cases of misconduct, police and prosecution are involved, but the Corruption Prevention and Combating Bureau (KNAB) or the Competition Council may get involved if appropriate. Furthermore, the State Audit Office audits public procurement, especially EU-co-funded projects (EU, 2014).

In general, the public procurement system in Latvia is plagued by similar problems as other sectors – low pay and subpar execution. State institutions often struggle to attract and maintain high-quality staff, which can lead to the low-quality implementation of even the best legislature. Execution deficiencies in many cases happen due to various irregularities, as the law is often vague and highlights the administrative focus on detection rather than prevention of wrongdoings (EU, 2014).

2.3. Risk-Based Measurement

2.3.1. Measuring corruption

Corruption measurement often is not straightforward. The major problem is that corruption is an accumulation of activities like small-level bribery, influencing public procurement contracts, extortion and abuse of official power, and more. The general view is that no one approach can capture all types of corruption single-handedly. However, when narrowing it down to public procurement corruption, it is possible to address the positives and drawbacks of each approach. As such, multiple measurements have been employed to measure corruption, the most popular of which are: 1) perception indices, 2) investigation of specific cases, and 3) the big data approach.

Most national corruption levels are explained using the first perception type, like the Corruption Perception Index (CPI) developed by Transparency International (TI). The CPI combines 13 different data sources, which ask experts or citizens of a country questions aimed at discovering the frequency, amount, and direction of corruption (Transparency International (TI), 2020). The major problem with perception indices is that they are quite good at capturing low-level bribery, the involved parties, and frequency but lack in detecting grand corruption, which involves significantly fewer actors, like in public procurement. Moreover, survey or interview-based research is time-consuming and requires much funding, illustrated by the fact that TI enjoys a near-monopoly over corruption perception index publishing. Moreover, although not unique to the perceptions index solely, surveys require good sampling; otherwise, the results might be swayed by personal interests or experiences, i.e., perceptions do not always reflect reality due to reverse causality or time lags (Andersson & Heywood, 2009). We find the perception approach unsuitable for our research, as building a representative sample in each Latvian municipality would be unfeasible during the allocated time frame.

Less often used are the other - individual case review & convictions. This approach uses investigative methods to uncover suspicions about an individual selection of procurement cases or analyse ongoing trials and accusations. The main problems of this approach include 1) dependence on the limited size of a sample leading to swayed or non-representative results (Kenny, 2017); 2) difficulties in comparing data across countries or, in our case, across municipalities due to data availability or legal framework differences; & 3) consistency over time. Moreover, the complexity of case review investigation makes it hard to produce comparatively regular results like the perceptions index, often released yearly (Bello y Villarino, 2021). Similarly to the perception approach, we find case reviews unfeasible as our goal is to compare Latvian municipalities.

An emerging type of measuring corruption is the so-called Big Data approach. The issues of inconsistency, regularity, and comparability are fixed using public data on procurement. A Big Data approach is more reliable, as it mostly takes out the subjectiveness of the interviewees or the reviewers. It also meets our goal of making the measurement of corruption more consistent over time and across countries. Similar datasets can be quickly adjusted to fit the research method used (Fazekas, Cingolani, & Toth, 2018).

As for data-based approaches, there are several, but we highlight the two most widely used - the two-stage residual approach involving data envelopment analysis and the red flag method. Most of these methods use some regression and variables to calculate the index. Both have several advantages and disadvantages.

Generally, the Big Data approaches still suffer from subjectivity in selecting variables. Regression models can be overloaded or undervalued based on the variables selected. Data models also need data that is still not always available freely, depending on the country. The two-stage residual approach is more advanced, as it also looks to split the risk impact into corruption or active waste and other impacts of passive waste like low-quality procurement inefficiencies (Lisciandra, Milani, & Millemaci, 2022). Compared to the two-stage residual approach, the red flag method is more straightforward, thus more replicable, and easier to describe to key stakeholders. On the other hand, data approaches still do not measure corruption, only the risk involved (Kenny & Musatova, 2010).

Another set of literature (Heywood (2016), Søreide (2006), & Cobham (2013)) criticises the idea of ranking itself, as it creates perverse incentives. Indices like TI's CPI generally check perception and might indicate risk, which does not paint a complete picture. As poorer countries generally have poorer governance, they can be found at the bottom of corruption indicator lists, but this does not directly imply wrongdoing. Instead, the conception of heavy corruption further corroborates a vicious cycle in which foreign direct investment might be pulled due to low corruption scores, removing funds from strengthening governance. Furthermore, wrongdoing can be hidden instead of addressed properly to inflate scores and attract foreign direct investment. This puts the investment in danger and does not contribute to the welfare of the country's inhabitants, as problems are not fixed but hidden. While these are appropriate concerns, we believe such problems are much more prominent while comparing countries, as more considerable differences come down to national-level institutions. When comparing municipalities, they all are under the same national-level institutions. Thus, differences in our results should suggest some inefficiency or risks directly associated with implementation problems.

2.3.2. Red flag method

The Red Flag method of analysis or quantitative indicator method collects theory-based characteristics that might impact the public tender in the form of corruption risks, which are then regressed onto independent variables following regression analysis. The regression analysis, as will be seen from the research studied by other authors, can come in the form of many regression forms, as the red flag method is not limiting in parameters. However, the overall presence of red flags does not grant the presence of corruption - it is only an indicator that the studied subject might be linked to a higher risk of corruption. The coming section will discuss the flags or the risks associated with corruption. This section will overlook the usage of the red flag method work by other authors researching corruption.

Historically, the red flag method has been primarily used to study the national corruption level. Decarolis, F. and Giorgiantonio, C. used the red-flag method to assess the corruption levels in the roadwork industry in Italy. Their findings concluded that there is a link between a higher risk of corruption and the use of multi-parameter criteria for awarding the public tender. The red flag method is also employed in other parts of research, excluding corruption. Popovic et al. (2018) have used quantitative indicators or red flags that can be used to determine social sustainability. The work of Vargas and Schlutz (2016) also uses the red flag method, as they create an index of the effort of financial disclosure of a given country to study corruption levels. Red flags have also been used to study corruption in World Bank projects. A paper from 2010 found that most of the World Bank projects analysed contracts displayed a level of red flags, meaning that there was a risk of corrupted contracts. (Kenny, Musatova, 2010). Lastly, quantitative indicators were used to study corruption by Fazekas & Kocsis. In their 2017 paper, the authors create a composite score of procurement red flags from the contract level, which can be used to study national corruption. (Fazekas & Kocsis, 2017).

2.3.3. Risks of Red Flags

The list of red flags used in academia for public procurement corruption research is long and regularly updated. Looking at the red flags or indicators of corruption, we will base our division of variables on Fazekas, Cingolani, and Toth (2016), who provide a comprehensive review of corruption proxies. The authors of the review state that the proxies, or red flags in this case, that previous researchers employed are contract lengthening, contract modification, and length of eligibility criteria, among a plethora of other indicators.

The red flag indicators can be grouped into four major groups: 1) tendering risks; 2) political connections; 3) supplier risk; 4) contracting body risk (Fazekas, Cingolani, & Toth, 2016). In the upcoming sections, we will discuss these indicators in the context of municipal procurement.

Tendering risk

Tendering risk pertains to indicators of the public procurement contracts themselves. Such indicators are often the most used, as they are easy to derive from the public procurement contract databases. Tendering risks can further be divided into bidder number, procedure, changes in delivery, and price.

The number of bidders is one of the more classical red flags, linking to the basic economic theory of competitiveness – less competitive markets allow firms to push prices up and exert pressure on authorities and other firms through rising entry costs. The number of bidders is often also used as the independent variable. Coviello and Gagliarducci (2010) proved a link between one additional term as a mayor and deteriorating public spending, decreasing the number of bidders, and winning rebates.

The procedure relates to how the contract is announced. If the procedure type is closed, decreased transparency levels increase corruption risk. Similarly, the complexity and length of the tendering description could indicate swaying the contract towards a particular firm. Also, an expedited and shorter submission process can lead to firms being unable to launch their bids at all. For example, public procurement corruption flourished due to emergency tendering during the Covid-19 crisis (Derby, & Wright, 2020).

Changes in delivery are easy to interpret as red flags. Changes to the requirements, price, or delivery term could mean that there is corruption involved, as such modifications often benefit the firm. Modifications can also happen due to unforeseen events like increasing material costs or economic downturns.

The final group of tendering risks is related to the price. Olken (2007) used independent engineer calculations of the price of Indonesian road projects to measure the differences between independent and actual project costs. Although highly efficient, such methods are also highly costly, especially in developed countries where thousands of public procurement projects are executed each year.

Political connections

The connection between the key agents – firms, municipalities or national institutions – can be present in several ways, but most often through direct ties and revolving doors. Direct ties imply that the firm's decision-makers, be owners or high-position managers, are tied by kinship, friendship, or other membership to the key decision-makers in the procuring institution, be that in a voted-in position or, for example, on the procurement board. Similarly, the revolving door describes the process where a former politician or high-ranking government worker takes a role in a firm to facilitate a relationship with the institutions they were working at. Luechinger and Mose (2014) look at the defence sector in the US and find a link between former politicians' corporate appointments and increased public procurement.

Supplier risk

There is an abundance of supplier risk indicators due to the extensive literature on firm attributes and, in many cases, the availability of a wide array of firm-level information. We divide the supplier risk attributes similarly to Fazekas, Cingolani, & Toth (2016) into attributes of the registry, the financials, the owners, and governance.

Basic company register information is available in many countries. Furthermore, such variables are often easily comparable and interpretable, making company register variables perfect for cross-country comparisons. Company registers variables include the number of companies on the same address, and the company's incorporation age. The number of companies on the same address often indicates that the company might be a shell company. Similarly, the number of years since the company incorporation also indicates shell companies, as firms are often incorporated for the specific reason of participating in a tender without having the necessary resources to fulfil the contract. Caneppele, Calderoni, and Martocchia (2009) find that both variables impact their criminological model significantly in Italian regions.

Company financial evaluation is a heavily researched topic on its own. Regarding public procurement, company financial risk variables can be divided into growth and concentration of public procurement. The growth variables deal with sudden spikes in income, specifically after an ownership or political party in charge changes. The overall idea is that severe outliers are risky. Growth changes are a red flag as companies often grow naturally. The concentration of public procurement is also a red flag, as it might show favouritism. Fazekas and David-Barrett (2016) use sales growth as one of their variables to explain firm favouritism differences between Hungary and the UK.

Several red flags regarding ownership and governance have been used, like complex ownership structures, incorporation in tax havens, and general transparency of the firm of their ownership. Ownership and governance red flags would indicate that there is some reason that the firm is hiding its actual owners, which might be related to political connections. Less to say, such red flags are often hard to aggregate on a large population focused on one country due to data availability.

Contracting body risk

Contracting body risk indicators are used for national research. Most often, contracting body risk variables are other aggregates of corruption, governance, or transparency that could be linked to the country's corruption risk. As such indicators are often not developed for municipalities, it leaves a gap in research to investigate municipalities on other fronts to help make decisions regarding the corruption risk at a more regional level.

3. Methodology

3.1. Data Collection and Manipulation

The central database used for the study on public procurement was retrieved and combined with Opentender data (Opentender Latvia, 2022). Opentender is a data source that gathers publicly available tender information from more than 20 sources for a country using government data (Opentender Latvia, 2022). The timespan of our extracted database is from 2013-2021 to capture two complete election cycles. We are also limited by the Latvian municipality reform carried out in 2021, which changed the borders of most municipalities.

The created database has information about all public procurement tenders that took place in Latvia between 2013 and 2021, with variables like the monetary size of the contract, year of the tender, length of the contract, and the tender type and whether the tender was EU-funded, among other factors. The dataset has 66056 unique data entries that provide information about the tender buyer and seller, in our case - the municipality and the entity which won the contract.

As many tenders had several smaller lots, we combined the data for a single tender by calculating the average or the sum of the values for the lots, resulting in a dataset with only information about a single tender. We do this to minimise noise from contracts, some of which had more than 20 different lots.

Additionally, we also add several more databases. First, a significant problem with the Opentender database is that it does not include the Latvian entity registration codes, thus making it challenging to combine our base data with other databases. We mitigate this by introducing the Entity Register database (LR Uzņēmumu reģistrs, n.d.). We are unable to identify the entity for 3546 contracts.

We extract information about all elected public officials in municipalities and the central government from the website of the Central Election Commission of Latvia (Central Election Commission of Latvia, n.d.). Then we add the beneficial owners of the tender-winning entities from the Entity Register database (LR Uzņēmumu reģistrs, n.d.). Lastly, we extract all relatives of the elected municipal public officials from the State Revenue Service income declarations website (Valsts amatpersonu deklaracijas, 2022). We then can combine this database to check whether the beneficial owners of the winning entity are related to the elected official in the municipality. The relatives most often listed in the declarations were spouses, children, and siblings. This means the potential risk is underestimated, as connections through close friends, parents, or more distant relatives could also be used to influence tender results. From this dataset, we create a dummy variable, nepotism, described in more detail in the next section. We also use the elected official database to extract information about the political party representation in Latvian municipalities, like whether a party has more than 50% of the seats.

We also add information about donations to political parties from the Corruption Prevention and Combating Bureau (KNAB, 2023). We use a similar method; we reconcile whether the tender-winning entity's beneficial owner has donated to a political party and whether that party has representation in the respective municipality.

We extract NACE industry codes from a Lursoft catalogue (Lursoft, n.d.)). We reconcile them with the registration codes and assign the first and second NACE levels. We are unable to identify NACE codes for 12 981 contracts.

We also obtain information about the financial data for registered entities from the Entity Register (LR Uzņēmumu reģistrs, n.d.). We combine the data by using the tender and financial data years and the entity registration number. We are unable to retrieve financial information for 21 495 tenders.

The final dataset used for this study is a database describing differences between the studied municipalities. To fit the scope of this study, the datasets were amended to match the time span of the central database, taking the municipal characteristics for the period of 2013-2021. We extract the municipal area, public sector wages, and population from the Central Statistical Bureau's databases (Official Statistics of Latvia, n.d). We extract the municipal centres' coordinates using Google Maps (Google, 2023). We also add municipality budgets from the Financial Ministry's website (Finanšu ministrija, 2020). Additionally, we retrieve a territorial development index from the State Regional Development Agency of Latvia (VRAA, 2020).

3.2. Variables

3.2.1 Red flags (explanatory variables)

In our research, we use red flags - variables that indicate corruption risk and are used in our regression. The creation and usage of red flags will be outlined in this section.

The OpenTender database uses three sets of red flags - integrity, administrative, and transparency. Integrity indicators include single bid for the lot, calls for tender publication, advertisement period sufficiency, procedure type, decision period length, whether the bidder is from a tax haven, and whether the bidder is a new company.

Administrative indicators include whether the contract had centralised procurement, whether the process was electronic, whether it is covered by the World Trade Organization's Agreement on Government Procurement, whether the agreement was based on a framework, whether the tender was also published in English, and whether there were any discrepancies between the data presented at the notice and awarding stages.

Transparency indicators include variables that indicate whether any information is missing from the tender, like bidder name, value, address of implementation, award criteria, duration, funding, selection method, and subcontractor information.

From the indicators created by the authors of the Opentender databases, based on data availability and economic theory outlined in the literature review, we continue with a selection of red flags: nepotism, tender publication, description length, procedure type, missing subcontractors, area match, and donations.

Nepotism

The variable Nepotism was created by us, as described in the section above, with data collection on the related persons of municipal electives. The variable was created as a dummy variable, taking a value of 1 if the person relating to the tender matched with the created related person database and taking a value of 0 otherwise. The risk-inducing nature of this red flag on the output of the tender transaction is relatively straightforward, as persons relating to the municipal electives won the tender. The choice for this indicator was also based on the paper by Stéphane Straub, in which the indicator defining firms' connections to political parties was used (Straub, 2014).

Tender Publication

The indicator was created by the OpenTender database for tender data for Latvia. The indicator is a binary variable, taking the value of 0 if the occurrence of the tender was not announced and taking the value of 1 if the announcement was published to the public. The reasoning for the risk-inducing nature of this variable follows the logic that if the tender was not publicised, possibly fewer parties made their bids, creating a suspected leniency towards informed bidders. The choice for this variable stems from the paper by Fazekas, Cingolani, and Toth (2016), in which objective risk indicators or corruption proxies, were overlooked. Among them, the announcement of the tender was outlined as a valuable corruption proxy.

Description length

The indicator was created by the OpenTender database for tender data for Latvia. This particular red flag shows how long the description of the tender was in characters. The usage of this indicator in our regression analysis was based on the findings of the paper by Fazekas, Cingolani, and Toth (2016), in which they found that the length of the description for the eligibility criteria captures tendering risks.

Procedure type

The indicator was created by the OpenTender database for tender data for Latvia. The indicator looks at whether the procedure of winning the bid on the tender is non-open, meaning not available for the public to participate, has a higher probability of becoming a single bid, or is open to competition. This binary red flag takes the value of 0 if the procedure type was deemed not open to other bidders and 1 if the bid was open to others. The openness of the tender procedure as a corruption (risk) indicator has been shown by many researchers. Chong, Klien, and Saussier find that public contracts, which are negotiated without any publication or competition in mind, are associated with lower quality of governance.

Missing subcontractor

The authors of the OpenTender database for Latvia created this indicator. It denotes whether information regarding a subcontractor is present in the description of a tender or not. The logic of including this variable as a red flag to our regression analysis and the municipal score calculation stems from the fact that excluding information about a subcontractor creates a basis for the risk of corruption for a tender. The binary variable can be described as having a value of 0 if a contract has no information about a subcontract and a value of 1 if such information exists in its description.

Area Match

We created the indicator by matching the bidder's location that won the contract against the municipality. This binary variable checks whether the winning bidder is from the respective municipality where the contract occurred. The reasoning for including this variable was to check whether geographical favouritism impacts the corruption score. The binary variable can be described as having a value of 1 if the municipality where the tender took place matches the corresponding municipality of the bidder, and a value of 0, if it does not.

Donation

The variable Donations was created by us, using the combined data from the Opentender database along with the KNAB database (KNAB, 2023). The creation of this variable started with extracting information about all of the political donations made in the studied timespan.

After this, we matched the person that has donated to a political party with tender information in our main database, whilst simultaneously checking if the political party that has received the donation has a seat in the corresponding municipality the tender took place. If all these conditions were satisfied, the variable took a value of 1, and if any of the conditions were not satisfied, it took a value of 0. The choice of using this variable stems from the fact that political donations have long been the target of anti-corruption research, with many linking donations with possible corruption risks.

3.2.2. Dependent variables of the first regression analysis

The basis of the usage of the dependent indicators follows Fazekas & Kocsis (2017), who created a high-level corruption revealing proxy. We pick variables that are associated with competition and include both objective reasons and possible corruption risk impacts.

Single bid

The indicator was created by the OpenTender database for tender data for Latvia. The indicator looks at the number of participants in the tender - if a tender only has one participant, it deems the tender a single bid. The variable single bid is 0 for tenders with only one bidder and 1 otherwise.

Number of bidders

Number of bidders also is derived from the OpenTender database. The variable describes the number of bidders who submitted a bid on the tender.

Market share

Market share is calculated based on the second NACE code hierarchical step. Market share denotes the decimal amount of contracts won in the specific municipality by the winning entity, based on the bid price.

3.2.3. Control (explanatory) variables of the first regression analysis

Price

This continuous variable has values of the contract price in EUR. The basis for using this variable stems from a comprehensive review of corruption risk factors, in which the contract's size was shown to be a good indicator of corruption (Fazekas, Cingolani & Tóth, 2016).

Tender year

This variable shows the year of the tender for the studied period of 2013-2021. Using this variable in the regression was since corruption risks differ by year, as major infrastructural contracts can obscure the risks linked to the tenders.

EU funding

The indicator was created by the OpenTender database for Tender data for Latvia. The binary indicator can be described as having a value of 0 if the tender did not receive EU funding and 1 if it had. The link between an EU-funded tender and possible corruption has literature backing, as a 2019 study by Fazekas & King found that EU funding raises the risk of corruption by 34% compared to other forms of funding. (Fazekas & King, 2019)

NACE

We created the indicator to denote the industry in which the tender is taking place. The variable was split into 19 dummy variables, representing the tender companies' various industries. The choice to include this variable in the regression comes from the differences in the industries participating in public tenders, as some industries have larger contracts than others. Additionally, corruption is attributed more commonly in some industries compared to others, as Transparency International indicates that extractive industries are more prone to corruption risks when compared to other industries (Transparency International, n.d.)

Distinct winner

The distinct winners variable was created by counting the number of entities of the second hierarchical step of NACE which have won any tenders during the period of our database. We introduce this variable to account for the objective measure of having a small number of companies working in a specific industry.

3.2.4. Independent variables for the second regression analysis

Based on economic theory and potential hypothesis, considering the availability of information, we select area, distance from the capital, population, public sector wage, budget per capita, development index, absolute power, governmental ties, and period as our municipal variables.

Municipal score

The calculation specifics of the municipal score variable can be found in section 3.2.3. The variable can be described as a risk score given to a single municipality based on the scores its tenders received from the calculation in the first regression. It was calculated as the average of all the tender-level calculated scores for the specific municipality. This variable will be used as the dependent variable in the second regression, as, in line with the research question, we want to see what influences the municipality score.

Area

This independent variable is the size of the municipality expressed in km2. The choice for this variable stems from the fact that larger municipalities by area tend to have larger monetary public tenders, as they might have larger infrastructural projects, among other factors.

Distance from the capital

This independent variable is the calculated difference between a given municipality and the capital of Riga, calculated in a straight line expressed in kilometres. The choice for this variable comes from the study by Kopczewska (2013), which shows that the benefits of social-economic policies differ for central and outer Polish municipal bodies.

Population average

This variable denotes the average population in each municipality in the studied period. An average of 2 values was taken: the population in 2013 and the population in 2021. The choice for this variable follows the logic that larger municipalities attract larger contracts, which in turn are commonly captured by single bidding companies, such as in the roadwork industry.

Average public sector municipal wage

This variable can be defined as the average wage of the municipal workers in the respective municipality. The wage in the public sector has long been discussed in the literature on corruption. As found by Cornell & Sundell, higher compensation in the national sector, when compared to all other industries, results in lower corruption risk. (Cornell & Sundell, 2019).

Budget per capita

This variable can be defined as the municipal budget for 2020 divided by the number of people in the municipality. The basis of using this variable in the second regression model is similar to that of the average municipal wage, as higher municipal budgets might indicate increased corruption risks or the bidder being a singular entity.

Development index

The Ministry of Environmental Protection and Regional Development creates the territorial development index. It considers economically active entities for 1000 people, level of unemployment, people below the poverty line, crime rate, natural migration, long-term migration, population above the working age, and the personal income tax level. Our index was calculated as an average of the observed development indexes for a given municipality between 2013 and 2020. We use this variable as a proxy for governance and development. If the development of a municipality is high, it could also have the needed processes and controls in place to prevent or minimise corrupt deals.

Absolute power

We constructed this municipal dummy variable by checking whether a single party in the municipality has 50% of the seats or more. The reasoning being this variable is a theoretical assumption that parties that have most of the power do not have to take into account the views of other parties, by such increasing the risk of corruption-like behaviour.

Government

We created this dummy variable by checking whether the party that has won the majority of seats in a municipality has representation in the national government or "Saeima." In the case that the following is true, the dummy variable takes a value of 1, 0 otherwise. Including this variable follows the theoretical assumption that the parties with both municipal and national presence might suffer from bias, for example, when monitoring or making decisions about questions directly relating to the corresponding municipalities.

Period

We introduced a dummy variable for the observed period. The dummy variable took a value of 0 if the observed period was from 2013-2017 and one if the observed period was from 2017-2021. The inclusion of this variable assumes that there is a difference between the observed variables in the two periods, either for political reasons, as the municipal election was in 2017, or for general reasons, such as the development index following a naturally increasing trend because of country-wide effects.

3.2.5 Descriptive statistics

The descriptive statistics for the central database used in the regression analysis can be seen in Appendix A, table 1. The dataset contains 66056 observations of Latvian public tenders. Single Bid has a mean of approx. 0.72, indicating that most tenders had multiple bidders (or multiple lots), though there still are some single-bidding tenders, visible from the minimal value of 0. The same observation can be made for all of our red flags used in calculating the municipal scores, as they also had varying results around the mean. Lastly, the mean of the variable EU funded, which is approx. 0.11 shows that the number of tenders that only around 1 in 10 tenders that have happened in Latvia in the studied period have been funded by the European Union.

When checking years, seen in appendix A, table 2, we can see that the number of contracts has fallen over the years, closing out recent years at around the 5-7 thousand mark. The average values have fluctuated from around 7 thousand to 23 thousand. The percentage of multiple-bidder contracts also fluctuates around 70%, reaching its highest point in 2015 and the lowest in 2020. Similarly, for average bidders, which fluctuates around 2, but had 3.45 bidders on average in 2013 and only 2.48 in 2019. The average market share is consistently low, around 0.02 to 0.03.

Looking at summary statistics by the bidder industry (Appendix A, table 3), we first note that for 12 981 contracts we could not identify the industry, due to the fact we could not reconcile it with the company register or the NACE code databases. The most contracts are in trade and transportation, roughly 13 and 17.5 thousand each. Only seven contracts have been registered under public administration & defence, as most likely, the majority of these are procured by the central government instead. The average value of contracts by industry varies greatly, with the highest, more than 900 thousand euros in the water supply. The lowest value of 20 thousand euros is again for the public administration & defence industry. From average bidders, we can observe that human health and recreation contracts get the least number of bidders per contract. At the same time, the competition is highest for other service contracts, where almost four bidders apply on average. The average market shares are highest in the mining, electricity & gas, water supply, and construction industries. The human health and information & communications industries have the lowest market shares. This is roughly in line with our understanding of market share, as we expect lower market shares in industries where resources can be moved across municipalities easier or where the companies have more national scope of operations. Distinct winners are approximately in line with the number of contracts, where the highest number of distinct winners are in the construction and trade industries but the lowest in public administration & defence.

When analysing the summary statistics of the municipal variables (Appendix A, table 4), a few trends can be spotted. The average municipal budget per capita varies quite a lot in different respective municipalities, as seen by the standard deviation of approximately 288 EUR per capita. Similarly, the average municipal wage is not constant for all the municipalities, indicated by the standard deviation of 140 EUR. This can be explained by the fact that the size of the studied municipalities differs, which respectively entails a different number of municipal workers in them, also seen by the minimum being around 500 EUR.

3.3. Research Methods

3.3.1 Red flag regression

As the purpose of this study dictates that we calculate and evaluate the risk measures for Latvian municipalities, the study's goal is to create a so-called municipal corruption indicator or "score." Although the term for the indicator is arbitrary, we draw from the long-standing studies of credit scoring seen in use by financial institutions. In a 2010 study, Dong, Lai, and Yen outlined how the most used logistic regression with random coefficients can be used to improve the results of the industry-standing logistic regression (Dong, Lai, Yen, 2010). In this study will not use a complicated model such as a Scorecard model or employ machine learning approaches. However, we are going to calculate the municipal score using linear regression.

Public procurement corruption can be summarised as extracting corruption rents. To extract corruption rents, the contract-winning company has to be pre-selected to maximise the profit, and the price has to be either higher or the quality has to be lower than in a competitive market. As estimating competitive market prices and quality for every contract is impractical, and comparisons can only be made between homogenous deliveries, we must rely on other contracts, financial, or nepotism variables associated with corruption risks through literature.

We can identify potential risks using our datasets of public procurement contracts, company financials, and nepotism. We use regression to link inputs associated with corruption, i.e., red flags, to outputs associated with corrupt rent extraction like single contract bidder. We also look to control for other variables like years, which illustrate inflation effects, the contract price, and other variables.

The chosen regression model is a linear Ordinary Least Square model (OLS). The regression's goal is to determine whether a red flag should be included in our index, as well as its assigned weight. Furthermore, regressions also account for non-linearity, as we identify values where our corruption indices significantly increase. By doing this, we define market norms ranges and outliers that can be linked with corrupt outcomes (Kenny & Musatova, 2010).

The benefits of using this model for the analysis are that it is simple in parameters and fits the type of databases collected for the study, cross-sectional databases. We will employ several techniques to account for all the assumptions of OLS. As we have the whole population, we will not run into the problem of biased sampling. By averaging out red flags in the open tender database, we will minimise the impact of multicollinearity.

The form of the red flag OLS regression can be seen below:

$$Single\ bid = \beta_0 + \beta_1 * Nepotism + \beta_2 * Donation + \beta_3 * Publication + \beta_4 * Procedure + \beta_5 * missing\_subcontractor + \beta_6 * descr\_length + \beta_7 * area\_match + \beta_8 * log(Sum.of.bid\_price\_EUR) + \beta_9 * NACE + \beta_{10} * tender\_isEUFunded + \beta_{11} * tender\_year + \beta_{12} * distinct\_win + u_{i,t}$$

3.3.2 Creation of the municipal score

The equation of how the risk score was calculated can be found below:

$$Risk_{i,t} = \beta_1 * descr\_length + \beta_2 * Publication + \beta_3 * Procedure + \beta_4 * area\_match + + \beta_5 * missing\_subcontractor$$

To calculate the risk score, we multiply the regression coefficients with the respective variables and sum the results. In other words, we calculate the contribution of our red flag variables. We do not use the explanatory variables for this step, as we view the impact as objective instead of a corruption risk. By summing, we also maintain the contribution of the variables provided by the regression in terms of the beta coefficients.

We weigh our initial base scores on each municipality's total sum of the tender values. Thus, we do not run into a problem when comparing large and small municipalities and prioritize the interest of the local municipality residents instead of a more national level. The bid price is a proxy for the impact of the municipal risk. We assume that residents would be more concerned when large contracts have more corruption risk, as the tenders are directly financed through their taxes.

With the calculated weights being the bid price, we then aggregate the corruption risk score to a municipal level by calculating the average of the base score and by summing up our weighted scores. The resulting municipal score thus consists of the coefficients extracted from the regression outlined in 3.3.1 and the estimate from the same regression. For the weighted score, it is also multiplied by the municipal-level weight of the bid price.

3.3.3 Validation of corruption risk scores

To check tender-level scores, we implement two validation methods - a selection of corruption cases portrayed in the media and a company financial validation. We select 8 cases from the Latvian national news portal LSM.lv for the media case validation by searching for municipal corruption (Appendix A, table 5). Not many municipal corruption cases have been reported, but we select cases in varying geography and years to obtain objective results.

We extract the scores of the picked cases and perform a simple analysis, comparing the cases to the average contract and the top 25% based on the corruption risk score.

There are several problems with validation from cases portrayed by the media. First, the cases portrayed by the media may constitute only a tiny part of the real corruption cases happening. Second, we might have an unintended bias when selecting the cases. Third, cases portrayed in the media are often based on assumptions and correlations and may not reflect actual corruption risk, which can only be proven in court. Fourth, our model could be better at finding certain corruption as it is based on a handful of picked red flags.

We perform a similar analysis to Fazekas, King, & Toth (2016). We calculate the net profit margin by dividing net profit by total revenues for the year after the contract has been recorded in our dataset. We use net profit margin, as corrupt companies extract corrupt rents, which should make these corrupt companies more profitable than their peers, providing an unfair advantage. We then split our dataset into thirds based on the net profit margin, calculate the means of the corruption risk scores and compare the scores in the three samples by performing a t-test to make sure they are distinct.

Nevertheless, this approach is also not fool-proof, as corrupt companies could be, for example, less well managed, thus also providing a disadvantage at the same time. An example of this can be found in the study of Vu, Tran, Nguyen & Lim (2018), who found that although cases of bribery might not be linked to the worse financial performance of a firm, it can be impacted negatively if it has illicit expenses to secure government contracts or licences. Even so, we believe validation is an essential step in testing our results and might provide insights into how to improve the model going forward.

3.3.4 Municipal regression

To understand the municipal-level impacts of the corruption risk scores, we regress the variables outlined in section 3.2.1. onto the municipal risk score and the weighted municipal risk score. The choice of regressing the chosen variables on both the weighted and non-weighted risk scores stems from the theoretical assumption that the impact differs for municipalities when accounting for the sizes of the contracts. The chosen form of the regression model was OLS because of its simplicity and other merits.

$$(Weighted)\ Municipal\ score_{i,t} = \beta_0 + \beta_1 * area_{i,t} + \beta_2 * distance\ from\ capital_{i,t} + \\ +\beta_3 * average\ municipal\ wage_{i,t} + \beta_4 * budget\ per\ capita_{i,t} + \beta_5 * development\ index_{i,t} \\ +\beta_6 * average\ number\ of\ parties_{i,t} + \beta_6 * absolute\ power_{i,t} + \beta_6 * government_{i,t} \\ +\beta_6 * period_{i,t} + u_{i,t}$$

4. Findings

In this section, we first analyse the results of the first step red flag regressions. Then, we proceed with contract-level data validation using cases portrayed in the media and company financial data. After, we aggregate contract-level data with the municipal-level data and proceed with further analysis. We conclude this section with various robustness tests to validate the findings that we gathered.

4.1 Regression Results

We first analyse the results of all our collected variables using the three independent variables of single bidder, number of bidders, and market share. These results will decide which variables to include in our corruption risk score calculation. While there is no agreed-upon threshold for variable acceptance, we generally look for statistical significance <0.001 and coefficient signs in line with our expectations. The regression results can be seen in Appendix A, table 6.

While the results align with our previous expectations, there are some significant outliers. First, our nepotism dummy variable is statistically significant only in the market share regression. Nevertheless, it also changes the sign from the expected negative to a positive effect.

We decided to exclude it from our final corruption risk calculation for these reasons. Although hard to prove, our theory is that we lack observations for the nepotism variable to be significant. We cannot collect more data at this stage due to the incomplete political relationship databases by not including identification codes.

Secondly, we decided to exclude the created donation variable. We expected the coefficient to be negative to reflect the bidders' ties with the municipality's winning party. However, the coefficient is only negative in the number of bidder regression and is not significant. We expect that this variable could be refined further by introducing a time dimension of how recently the bidder has donated to the party and linking the winning party in the municipality and others. Overall, we believe these links could be studied separately as they would be too time-consuming to address in the scope of this thesis.

We will use area match as a proxy variable to still include some nepotism links in our score. Area match is statistically significant across all of our independent variables and consistent with our expectation of it having a negative effect. While there might also be objective reasons for a bidder from a local area to win, we believe it still poses a risk, as local communities are often close-knit, especially between local politicians and businesspeople.

There are also some inconsistencies with the publication and description length variables. The publication variable changes its sign in the number of bidders and market share regressions, while the description length variable is not statistically significant in the market share regression. We still decided to keep them, as we have a solid theoretical and economic understanding outlined in the literature review part and a whole population.

We also decide to proceed with the single bid regression for further analysis, as we note several conceptual issues with the number of bidders and market share variables. Although they might serve a purpose as an extra robustness check in the initial stage, the number of bidders might be problematic due to how we created it. To avoid unwanted noise in our results, we averaged the number of bidders between contract lots and assigned this value to the overall contract. This left us with some high outliers that might skew the results. Similarly, when creating the market share variable, we calculate the percentage of contracts which the entity has won in a certain period, thus already absorbing some of the risks, as corrupt entities are likely to have inflated measures.

Based on the analysis above, we remove the nepotism and donation variables and proceed with the single bidder as the only independent variable (Appendix A, table 7). As these coefficients will also be used to calculate our municipal risk scores, we can interpret the beta coefficients of our risk variables. Most of our variables are binary. As such, an increase of 1 in our variable increases the probability of a single bidder contract by beta coefficient *100 (percent).

The red flags are in line with our expectations. If the tender is not published, the probability of a single bidder increases by 26%. Similarly, if the tender procedure is deemed closed, the probability of a single bidder also increases by 25%. If a subcontractor is not reported, the probability of a single bidder increases by 12%. Furthermore, if the area matches between the bidder and the municipality, the probability of a single bidder also increases by 5%. The only non-binary red flag is description length. As we apply a natural logarithm due to skewness towards short descriptions, the interpretation changes to straight percent. As such, a 1% increase in the description length corresponds to a 0.01% increase in the possibility of a single bidder.

Aside from industry and year, we also have three additional explanatory variables. We use the log of bid price as a proxy for complexity. Hence as the bid price increases, the probability of a single bidder also increases. A 0.03% increase in the value corresponds to a 1% increase in the probability of a single bidder. The more complex the contract, the fewer companies can undertake it, supporting our reasoning that the bid price is an objective measure.

We also use a binary variable of a project being EU funded. We use this dummy variable as a proxy for supervision. If the project is EU funded, it must undergo a more rigorous certification process and be checked by more overseers during production. As such, we suppose that the likelihood of corruption risk would be decreased, which is supported by our regression. If the tender is EU funded, it is 7% less likely to be a single bid.

We use distinct winners as a proxy for competition. We construct this using 2nd step NACE codes to ensure that the results are not correlated with the NACE dummies used in the regression. If there are more competitors on the market, it is less likely that the tender will have a single bidder. If there is one more distinct winner, the probability of the tender being a single bidder falls by 0.02%.

4.2 Calculation of Contract-Level Scores and Analysis

We calculate the contract-level corruption risk scores (Appendix A, table 8). We do this by applying the same logic as we explained the effects in the previous section and multiplying the beta coefficient effect with the true variable values present for each of the contracts in our database. We then sum the five red flag variable scores to obtain the contract-level corruption risk score.

We can also analyse the corruption risk by year and by industry based on our calculated scores. Summary statistics for years reveal a positive trend in decreasing scores over time, with 2020 being the lowest and 2016 - the highest (Appendix A, table 9). We also notice that years have some explanatory power in our regression. If a contract took place in 2015, the probability of a single bidder would be 14% higher, even though 2015 was the year with the lowest on-record contracts with single bids.

Looking at the industries (Appendix A, table 10), we can identify the riskiest and least risky industries based on our corruption risk scores. Human health, accommodation & food, and recreation are the most risky industries. While least risky industries are electricity & gas, and manufacturing, closely followed by finance & insurance, trade, and water supply. It is hard to make conclusions about these results, as many reasons could explain their ranking. One possible explanation might be the regulation of an industry, which could explain why utilities and finance would be among the lowest ranked. However, this conclusion also immediately contradicts itself, as human health is among the riskiest. There is an argument to be made about competition affecting our scores - both industries with the highest number of unique winners in construction and trade are amongst the lowest on our corruption risk scores. We also see that the dummy variables have some explanatory power, possibly accounting for competition, complexity, or other unexplained factors. The most significant coefficients are for the recreation, human health, and information industries, where the industry increases single bidder chances by 25-30%, possibly explainable by the specialization of these industries and relatively small competition between the specialised vendors.

4.3 Contract Score Validation

Although we see statistical significance between our variables and the independent variables, to ensure our results are indicative of actual risk, we can perform simple validation analysis using municipal corruption cases portrayed in the media and company financials.

4.3.1. Municipal corruption cases in media

We extract the cases based on the compiled list. Sometimes, the media portrayed a link between municipal politicians and a company. In these instances, we extract all contracts signed between the parties from our dataset (Appendix A, table 11).

Only one case concerning the purchase of legal services by the city of Riga falls in the top 25% of our calculated scores. All these contracts are single bids, which might partly explain why the corruption risk is rated so highly. At the same time, this is also positive news, as our model can pick up on corruption in more minor value cases.

Half of the cases have a below-average score, indicating that there might be some bias involved either in our model or the portrayal of the cases by the media. There seems to be a faint link between the percentage of single bid contracts and the score, which is consistent with our model as the single bid is the independent variable used. Apart from this, we do not see any connection between the scores, contract number, total value, or average value.

4.3.2. Financial performance validation

We can draw several conclusions after performing financial validation (Appendix A, table 12). First, we can say that at the 95% confidence interval, the scores in our three samples are genuinely different. Second, the corruption risk score is slightly higher for more profitable companies, but even more curiously, it is not far behind the average score of the least profitable companies. This might support one of our initial theories, that companies engaging in corrupt rent extraction might improve profitability and decrease it either due to worse governance and operations or to paying corrupt officials for the corrupt rent extraction. Third, we also see that companies in the middle sample, with the lowest corruption risk score, are the largest by net turnover and employee count on average. This might indicate that larger companies either have better governance practices or a higher level of supervision.

4.4 Municipal Score

We calculate the municipal corruption risk score by aggregating contract-level scores as an average per municipality. For the weighted score, we weigh by the bid price of the contract, where the weights in one municipality sum up to one, thus accounting for the size of the municipality and the possible impact of the costlier tenders on the local residents. We then rank the scores where the higher the rank, the higher the corruption risk score. We also calculate the differences between the base score and the weighted score. Results of all the municipal scores can be seen in table 13, Appendix A.

We first extract the top 10 and bottom 10 municipalities by the average score (Appendix A, table 14 & table 15). The least corrupt municipality is Līgatnes, consistently having the lowest corruption risk in both scores and with a significant margin from the closest competitor of Varakļānu municipality. Tukuma municipality is the 10th least corrupt municipality. Tukuma municipality is of significant size, with total bid prices of over 100 million euros and around 900 contracts over nine years. However, we can also observe that Tukums' score does improve after weighting.

The municipality with the most corruption risk is Kandava municipality, followed by Baldone, Mazsalaca, Olaine, and Dundagas municipalities. Just behind, in sixth place, is the city of Riga, which falls in the ranking significantly after we apply weights. Rounding out the top 10 is Cēsu municipality, which is also one of the bigger municipalities in Latvia with 1650 contracts of 91 million euro value in total over nine years.

There are also several municipalities where the weighting seriously changed the ranking (Appendix A, table 16 & 17). Most of these are small municipalities with few contracts, but, for example, the city of Ventspils & Mārupes municipality falls in our corruption risk rating significantly when weighted. This could signal a lot of minor corruption instead of grand corruption.

Looking at the grouped municipal scores, a few tendencies can be spotted. From the results in table 18 (Appendix A) the city of Riga has the highest municipal corruption risk, followed by Vidzeme and Kurzeme. The region of Zemgale has the lowest municipal corruption risk.

The trend of Riga having the highest municipal corruption risk follows when analysing the risks of the biggest cities (Appendix A, table 19). An observation could be made that the size of the municipality might impact the municipal score of Riga, but this is disproven as the combination of the various municipalities not included, having a similar amount of contract size, has a lower municipal corruption risk than Riga. This shows that the size of the contract is not the underlying reason for bigger municipal units, namely Riga, having a higher corruption score. This is further disproven, with Vidzeme having the lowest corruption risk score but not the lowest in contract size.

By comparing cities, we can exhibit that Riga and Ventspils, with the highest scores, also have the highest percentage of single-bidder contracts. This might show some bias of our model toward valuing single-bidding contracts more harshly. Furthermore, we can also see that other cities apart from Riga and Venstpils have less corruption risk and roughly similar scores. These scores are also similar to the average municipality scores, even though the average contract size and the single bidder percentages differ.

To further analyse the scores, we extract the tender level scores for municipalities of Riga and Ventspils, which had the highest difference in score and the weighted score of the big cities, as well as Valmiera, which vice versa had the lowest difference between the two scores, and Kandava, which has the highest corruption risk according to our model. We divide the samples in increments of 10% by the contract value. We also perform the same summary for the whole dataset, including averages of our score components at each increment. The results of this analysis can be seen in tables 20-24 (Appendix A)

We see several things. Scores tend to fall for larger contracts. Most components are consistent over the denary, except the procedure and missing subcontractor components. Procedure rapidly loses its contribution when reaching the top 10% of the contracts by size, which can be explained as larger contracts often come under much more scrutiny from authorities and the media. Missing subcontractor components, on the other hand, significantly increase when reaching the top 20-10% range. This could be explained by larger contracts having more subcontractors. Thus, more of them might also not be mentioned, whether done by mistake or knowingly.

By comparing Valmiera to Ventspils and Riga, we conclude that the largest 10% of contracts decide the change the weighted score will have against the base score. This difference is much more prominent in Ventspils and Riga, by around 7 and 5 points, respectively, compared to the 3 points in Valmiera. The opposite can be said about Kandava, whose base score increases for the largest 10% of contracts. Although there are significantly fewer contracts in Kandava municipality, we suppose that the largest contracts have not been announced or have a closed procedure, thus ranking higher in our corruption score. Another thing we notice is that the distribution of the contracts by value is very much skewed, with minor differences in contract values between the bottom 90% of contracts.

4.5 Municipal Regression

We regressed the chosen variables for the municipal regression outlined in section 3.2.4 onto both the weighted and unweighted municipal risk scores. (Appendix A, Table 25). From the output of the regressions, municipalities had a municipal risk score that was lower by 4.37 compared to 2013-2017, with under 1% statistical significance, keeping everything else constant. Additionally, if a municipal government has a seat in the national government, the risk of corruption is raised by 0.7, under 10% significance, keeping everything else constant. Furthermore, an increase in the average municipal wage by one euro does not yield conclusive results, as the sign of the impact changes whether the municipal score is weighted or not. Lastly, increasing the distance of the given municipality by 1 kilometre from the capital decreases the corruption risk by 0.01, under a 10% significance level, keeping everything else constant. Interestingly, this particular result differs from the research conducted by Kopczewska (2013) and our hypothesis.

4.6 Robustness

To test the robustness of the regression model, we used the rlm() function for the MASS package (Venables & Ripley, 2002) and calculated the robust standard errors for each regressions. In essence, the test checks the robust standard error of each regressions to check which model fits the selected database better. The RSE calculated for the Robust regression is lower than the RSE calculated for our regression models. For example, for the Red Flag model, the calculated RSE was 0.42 compared to 0.05, leading us to conclude that our selected models are robust and better fit the studied database. Additionally, further robustness checks regarding the data were not possible due to the studied dataset including the entire population, namely all the public tenders in Latvia from 2013 to 2021, so checking robustness by subsetting would not yield constructive results.

To test the heteroskedasticity of the regression results, we employed the Breush-Pagan test. The results of the Breush-Pagan tests revealed that the null hypothesis, of all error term variances being 0, had to be rejected. Thus, we see that heteroskedasticity is present in the primary regression. However, since we observe the entire population of public tenders in Latvia, the issue cannot be fixed, as removing outliers would mean removing large municipal tenders, or other important tenders, which would obscure the municipal results even more.

Further on, we performed the Variance inflation factor test to test the correlation between the studied variables. In essence, the VIF test checks the multicollinearity of the chosen regression variables or whether the variables in the regression are correlated with each other. The results of the VIF tests for the Red Flag regression, as well as the municipal regressions, can be seen in table O. From the results, it is visible that none of the variables crosses a value of 10, which would indicate that multicollinearity is a severe issue in our regression analyses.

Subsequently, when tested with the VIF test, the municipal regression variables also did not cross the value of 10. Although sometimes a rule of thumb value of 5 is used to test VIF, based on the literature review and the economical and logical understanding of variables, we believe the variables are different enough to remain in our regression analysis.

5. Discussion

5.1. Red Flag Regression

Looking at the results gathered from the red flag regression, namely the chosen regression model of using single bid as a dependent variable, as seen in Table 7, most results make economic sense. Firstly, we can see that all of the red flags used in calculating municipal scores (Tender publication, Missing subcontractor, Area match, Description length & Procedure type) impact single bid negatively, meaning that they raise the risk of a contract having only one bid. When checking the robustness of our main regression, it was concluded that the chosen model does fit the data.

Nevertheless, we see the potential for creating and testing other red flags. Further effort could be put into the nepotism and donation variables, completing the datasets and incorporating more rigid controls where possible. Our nepotism database had a small number of observations, whilst donations were clouded by noise, as not all political donations can be attributable to corruption risks.

We also find the variance between the scores by industries. This can pave new ways of checking, for example, cartel likelihood or industry competition. Similar research on corruption risk can also be adapted not for municipalities but for industries, which can lead to introducing a different set of variables that check variables that may lead to higher-risk industries.

The calculated tender-level scores could also be used for case analysis or further investigation. Case analysis can be introduced as validation or further analysis. Either way, we see the potential for using the scores in academia and corruption prevention and detection.

5.2. Municipal scores

Analysing the findings of the municipal scores, we can turn to tables 7-24 (Appendix A) and the created heatmap (Appendix A, figure 1 & 2). Summarising the results, it can be seen that the municipalities with the most corruption risk do not follow a geographic trend. This partly disproves the correlation of the distance from the capital as an indicator of corruption, which we inferred from Kopczewska (2013), whose study did show that distance from the capital had a negative impact.

We also identify the potential for further analysis by doing interviews. Interviews of government and municipal officials and experts could also provide more insights into our research's input and output sides. Officials and experts could guide new red flag variable introduction and explain the potential ranking and what impacts the score itself. Our created score could be the starting point, as we have identified the municipalities with the most and least corruption risk, which could be investigated first.

While weighing could be considered arbitrary, it might also point to a relative presence of small-scale and grand corruption in a municipality. If the results significantly change between the base and weighted scores, as they did in several municipalities, we could hypothesise the possibility of either small-scale or grand corruption being more present than the other. Although it is unclear how this could be researched further, the tenders could be sorted by elected officials and municipal worker involvement or by size and complexity.

5.3. Municipal regression

To summarise the results of the municipal regression, we can see that most of the results are not statistically significant, except for financial indicators such as public sector wage and distance from the capital and governmental ties. However, public sector wage has an inconsistent sign between the two regressions. Governmental ties increase corruption risk, as the municipal parties might have more favour in the central government and its institutions, thus decreasing supervision. Distance from capital seems to decrease corruption risk, which might be impacted by high scores in and around Riga. We can also infer from the results that the corruption risk decreases over time. As for the other results of the regression, we cannot conclude or make inferences about the results, as the results are not statistically significant. We acknowledge that the results of our regression might have been impacted by the fact that many more factors can impact corruption on the municipal level than the ones we have chosen in our regression model.

Although in our paper we try to introduce variables that could explain our scores on a municipal level, we conclude that there is a lack of governance metrics which could be used. We know that attempts to create governance proxies like Delna (2021) take significant time and effort. Nevertheless, we think municipal governance indicators like Delna's or, for example, the amount and experience of procurement workers in the municipality could meaningfully explain corruption risk variance in Latvian municipalities.

We also checked whether the political parties in charge of the municipality impacted our scores. As the results were not significant, we decided not to include them. Most likely, our sample for such analysis is too small. We identify the possibility of sorting the political parties by political ideologies, but we also acknowledge that this might introduce researcher bias. Alternatively, the political party impact could also be amended to include the number of seats each party holds and the impact it has on the municipality decision-making process through the coalition or high-ranking municipal workers. However, this would require an extensive data collection effort.

5.4. Limitations

The first limitation in our work comes from the availability of information. Although the OpenTender database has much information about tenders in Latvia from 2013, many variables are absent for most tenders. Furthermore, the OpenTender database proved challenging to combine with other datasets due to it not having locally recognised entity registration numbers.

The usage of the secondary regression, which explains which variables impact the municipal score, could be further looked at and developed. Within the time constraints of this research, we could not fully find the correct variables of interest to explain the municipal scores. The complexity of corruption itself could also explain this, and it is caused not only by the variables chosen by us but by a plethora of other factors, which could prove to have a higher explanatory power.

Lastly, it must be noted that it is extremely difficult to assess corruption itself, as it may take many forms. This paper introduces several decisions regarding variables that could impact the results. As such, it is possible that our model can explain only a part of the corruption present.

6. Conclusion

This paper contributes to the existing literature surrounding corruption risk measurement and is novel in measuring municipal corruption specifically. Even though national-level corruption has been researched extensively, municipal-level corruption has been eluded by the broader research community. We used the Red Flag method to calculate the municipal scores and analysed potential corruption risk drivers at the municipal level.

Answering our first research question, "**Does the public procurement corruption risk vary across Latvian municipalities?"** we can state that the results do differ for municipalities across Latvia, and the risk score can be seen as being entirely dispersed, not following a geographical trend. Based on our red flag regression, we have created a ranking of corruption risk in Latvian municipalities. Based on the unweighted scores, Riga has the highest corruption risk, while the municipality of Līgatne - has the lowest. When weighing the scores, Kandava municipality has the highest corruption risk, while Līgatne remains at the bottom of the ranking.

We could not fully answer the second research question, "**What drives corruption risk in Latvian municipalities?**". We identify that the distance from the capital slightly decreases corruption risk while election-winning political party representation in the central government - increases the corruption risk. Despite this, other variables were statistically insignificant or inconsistent between regressions on unweighted and weighted scores. Our corruption scores also differ by industry. The highest risk is present in human health and accommodation & food industries, while the lowest is in electricity & gas.

We identify multiple avenues for further research. Researchers looking into municipal corruption could employ detailed analysis, analysis by industry, and develop further risk variables. We believe that introducing more municipal-level variables has the most potential, as confirming and understanding corruption risk drivers would potentially have important implications not only for Latvian municipalities but, if replicated correctly, for municipalities of other comparable countries.

7. References

Andersson, S., & Heywood, P. M. (2009). The Politics of Perception: Use and Abuse of Transparency International's Approach to Measuring Corruption. *Political Studies*, 57(4), 746–767. https://doi.org/10.1111/j.1467-9248.2008.00758.x

Bello y Villarino, J.-M. (2021). Measuring Corruption: A Critical Analysis of the Existing Datasets and Their Suitability for Diachronic Transnational Research. *Social Indicators Research*, 157(2), 709–747. https://doi.org/10.1007/s11205-021-02657-z

Caneppele, S., Calderoni, F., & Martocchia, S. (2009). Not only banks: Criminological models on the infiltration of public contracts by Italian organized crime. *Journal of Money Laundering Control, 12(2), 151–172. https://doi.org/10.1108/13685200910951910*

Central Election Commision of Latvia, (n.d.). Local Elections. *Retrieved January 26th, 2023 from https://www.cvk.lv/en/local-elections*

Charron, N., Dahlström, C., Fazekas, M., & Lapuente, V. (2015). Careers, Connections and Corruption Risks in Europe *(SSRN Scholarly Paper No. 2711956).* https://doi.org/10.2139/ssrn.2711956

Cobham A. (2013, July, 23). Corrupting Perceptions: Why Transparency International's Flagship Corruption Index Falls Short. *Center For Global Development*. Retrieved from https://www.cgdev.org/blog/corrupting-perceptions-why-transparency-international%E2%80%99s-flagship-corruption-index-falls-short

Coviello, D., & Gagliarducci, S. (2010). Building Political Collusion: Evidence from Procurement Auctions. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1631074

Darby S.,Wright T. (2020, July 1). Opinion: How to tackle corruption in health procurement. *Devex*. Retrieved from https://www.transparency.org.uk/covid-19-has-created-conditions-which-corruption-health-procurement-can-flourish-heres-how-open

Decarolis, F., & Giorgiantonio, C. (2022). Corruption red flags in public procurement: New evidence from Italian calls for tenders. *EPJ Data Science, 11*(1), Article 1. https://doi.org/10.1140/epjds/s13688-022-00325-x

Delna (2021, November 29). PAŠVALDĪBU ATKLĀTĪBAS INDEKSS. Retrieved from *https://delna.lv/lv/2021/11/29/pasvaldibu-atklatibas-indekss/*

Dong, G., Lai, K. K., & Yen, J. (2010). Credit scorecard based on logistic regression with random coefficients. *Procedia Computer Science*, *1*(1), 2463–2468. https://doi.org/10.1016/j.procs.2010.04.278

European Commision (n.d.). Public procurement. Retrieved 26 March 2023, from *https://single-market-economy.ec.europa.eu/single-market/public-procurement_en*

Fazekas, M., Cingolani, L., & Tóth, B. (2016). A Comprehensive Review of Objective Corruption Proxies in Public Procurement: Risky Actors, Transactions, and Vehicles of Rent Extraction. *SSRN Electronic Journal*. *https://doi.org/10.2139/ssrn.2891017*

Fazekas, M., & King, L. P. (2019). Perils of development funding? The tale of EU Funds and grand corruption in Central and Eastern Europe. *Regulation & Governance*, *13*(3), 405–430. https://doi.org/10.1111/rego.12184

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), 155–164. https://doi.org/10.1017/S0007123417000461

Fazekas, M., Tóth, I. J., King, L. P. (2016). An Objective Corruption Risk Index Using Public Procurement Data. *European Journal on Criminal Policy and Research*, *22*(3), 369–397. https://doi.org/10.1007/s10610-016-9308-z

Finanšu ministrija (2020). Pašvaldību finanšu rādītāju analīze [Data file]. Retrieved March 2, 2023 from https://www.fm.gov.lv/lv/pasvaldibu-finansu-raditaju-analize

Goutte, S., Péran, T., & Porcher, T. (2022). Corruption, economy and governance in Central Africa: An analysis of public and regional drivers of corruption. *Finance Research Letters*, *44*, 102086. https://doi.org/10.1016/j.frl.2021.102086

Heywood P. (2016, February 3). The Corruption Perceptions Index (CPI): The Good, the Bad and the Ugly. *The British Academy*. Retrieved from https://www.thebritishacademy.ac.uk/blog/corruption-perceptions-index-cpi-good-bad-and-ugly

Iepirkumu uzraudzības birojs (2020). Atvērtie dati. Retrieved 26 March 2023, from https://www.iub.gov.lv/lv/atvertie-dati

Iepirkumu uzraudzības birojs, (n.d.). Publikāciju datu vizualizācija. Retrieved 26 March 2023, from https://info.iub.gov.lv/lv/visual

Kenny, C. (2017, January, 23). How Much Aid is Really Lost to Corruption ? *Center For Global Development.* Retrieved from https://www.cgdev.org/blog/how-much-aid-really-lost-corruption

Kenny, C., Musatova, M. (2010, March). 'Red Flags of Corruption' in World Bank Projects: An Analysis of Infrastructure Contracts. Policy Research working paper. No. WPS 5243. Retrieved from https://documents1.worldbank.org/curated/en/790591468321562564/pdf/WPS5243.pdf

Kopczewska, K. (2013). The spatial range of local governments: Does geographical distance affect governance and public service? *The Annals of Regional Science*, *51*(3), 793–810. https://doi.org/10.1007/s00168-013-0567-z

Kupčs, E. (2016, August 23). Aizdomas par negodīgu «Rīgas satiksmes» iepirkumu mikroautobusu pārvadājumiem. Retrieved 26 March 2023, from *https://www.lsm.lv/raksts/zinas/latvija/aizdomas-par-negodigu-rigas-satiksmes-iepirkumu-mikroautobusu-parvadajumiem.a197549/*

Kupčs, E. (2017, September 22). Jelgavas autobusu iepirkumā uzvarēt var tikai pilsētas mēra Rāviņa dēla firma. Retrieved from *https://www.lsm.lv/raksts/zinas/latvija/jelgavas-autobusu-iepirkuma-uzvaret-var-tikai-pilsetas-mera-ravina-dela-firma.a251058/*

Kupčs, E. (2020, April 4). Jelgavā aizdomas par ielu uzturēšanas iepirkumu konkrēta uzņēmēja interesēs. Retrieved from https://www.lsm.lv/raksts/zinas/latvija/jelgava-aizdomas-par-ielu-uzturesanas-iepirkumu-konkreta-uznemeja-intereses.a354654/

Kupčs, E. (2019, October 30). Jūrmalā pie miljoniem eiro vērtiem līgumiem arvien biežāk tiek domes galvenās pilsētplānotājas brāļa firma / Raksts. Retrieved from *https://www.lsm.lv/raksts/zinas/latvija/jurmala-pie-miljoniem-eiro-vertiem-ligumiem-arvien-biezak-tiek-domes-galvenas-pilsetplanotajas-brala-firma.a336776/?utm_source=lsm&utm_medium=theme&utm_campaign=theme*

KNAB. (2022).Dāvinājumi, ziedojumi, biedru naudas un iestāšanās naudas. Retrieved from: *https://info.knab.gov.lv/lv/db/ziedojumi/?recordsPerPage=all*

Krenberga, O. (2014, September 7). Rīgas domei aizdomīgi iepirkumi. Retrieved from https://www.lsm.lv/raksts/zinas/latvija/rigas-domei-aizdomigi-iepirkumi.a97502/

Latvijas Republikas Uzņēmumu reģistrs. (n.d.). Uzņēmumu reģistrs [Data file]. Retrieved February 16, 2023 from https://data.gov.lv/dati/lv/dataset/uz

Latvijas Republikas Uzņēmumu reģistrs. (n.d.). Patiesie labuma guvēji [Data file]. Retrieved February 16, 2023 from https://data.gov.lv/dati/lv/dataset/patiesie-labuma-guveji

Latvijas Republikas Uzņēmumu reģistrs. (n.d.). Gada pārskatu finanšu dati [Data file]. Retrieved February 20, 2023 from https://data.gov.lv/dati/lv/dataset/gada-parskatu-finansu-dati

LSM.lv. (2020, March 30). Konkurences uzraugi: Rēzeknē dome un namsaimnieks iepirkumos deformē konkurenci / Raksts. Retrieved from *https://www.lsm.lv/raksts/zinas/ekonomika/konkurences-uzraugirezekne-dome-un-namsaimnieks-iepirkumosdeforme-konkurenci.a353867/?utm_source=lsm&utm_medium=theme&utm_campaign=theme*

LETA. (2017, November 26). Raidījums: Juristu un uzņēmēju grupa radījusi shēmu, kas ļauj pašiem gatavot pašvaldību iepirkumu konkursus un tajos uzvarēt. Retrieved from *https://www.lsm.lv/raksts/zinas/latvija/raidijums-juristu-un-uznemeju-grupa-radijusi-shemu-kas-lauj-pasiem-gatavot-pasvaldibu-iepirkumu-konkursus-un-tajos-uzvaret.a258797/*

Lisciandra, M., Milani, R., & Millemaci, E. (2022). A corruption risk indicator for public procurement. *European Journal of Political Economy*, *73*, 102141. https://doi.org/10.1016/j.ejpoleco.2021.102141

Luechinger, S., & Moser, C. (2014). The value of the revolving door: Political appointees and the stock market. *Journal of Public Economics*, *119*, 93–107. https://doi.org/10.1016/j.jpubeco.2014.08.001

Lursoft (n.d.). Statistical Classification of Economic Activities NACE Rev. 2. Retrieved January 20, 2023, from https://nace.lursoft.lv/activities/tree?l=en

Mishra, A. (2006). Persistence of corruption: Some theoretical perspectives. *World Development*, *34*(2), 349–358. https://doi.org/10.1016/j.worlddev.2005.03.010

Morris, S. D. (2011). Forms of Corruption. *CESifo DICE Report*, *9*(2), 10–14. Retrieved from https://www.econstor.eu/bitstream/10419/167031/1/ifo-dice-report-v09-y2011-i2-p10-14.pdf

OECD. (2021). Government at a Glance 2021. *OECD Publishing*. https://doi.org/10.1787/1c258f55-en

OECD. (n.d.). Public procurement. Retrieved 26 March 2023, from *https://www.oecd.org/governance/public-procurement/*

OECD. (2011). Public Procurement in the EU: Legislative Framework, Basic Principles and Institutions. *SIGMA Public Procurement Briefs*, Vol. 1. *https://doi.org/10.1787/5js4wzvq4m36-en*

Official Statistics of Latvia (n.d.). Statistics Portal. Retrieved 26 March 2023, from *https://stat.gov.lv/en*

Olken, B. A. (2007). Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy*, *115*(2), 200–249. https://doi.org/10.1086/517935

Opentender Latvia (n.d.). About Opentender. Retrieved March 5, 2023, from https://opentender.eu/lv/about/about-opentender

Pārresoru koordinācijas centrs (n.d.). Attieksme pret korupciju Latvijā | Pētījumu un publikāciju datu bāze. Retrieved 26 March, 2023, from *http://petijumi.mk.gov.lv/node/3921*

Popovic, T., Kraslawski, A., Barbosa-Póvoa, A., & Carvalho, A. (2017). Quantitative indicators for social sustainability assessment of society and product responsibility aspects in supply chains. *Journal of International Studies*, *10*(4), 9–36. https://doi.org/10.14254/2071-8330.2017/10-4/1

Puriņa, E. (2018, December 17). 12 skaļākie korupcijas skandāli Ušakova vadītajā Rīgas pašvaldībā. Retrieved 26 March 2023, from *https://www.lsm.lv/raksts/zinas/zinu-analize/12-skalakie-korupcijas-skandali-usakova-vaditaja-rigas-pasvaldiba.a303226/*

Shaxson, N. (2007). Oil, Corruption and the Resource Curse. *International Affairs (Royal Institute of International Affairs 1944-)*, *83*(6), 1123–1140.

Søreide, T. (2006). Is it wrong to rank? A critical assessment of corruption indices. *CMI Working Paper WP 2006*: *1*, 13 p. Retrieved from https://www.cmi.no/publications/2120-is-it-wrong-to-rank

Straub, S. (2014). Political Firms, Public Procurement, and the Democratization Process. *IDEI Working Papers,* Article 817. https://ideas.repec.org//p/ide/wpaper/27861.html

Transparency International. (n.d.). The ABCs of the CPI: How the Corruption Perceptions Index is Calculated. Retrieved March 4, 2023 from *https://www.transparency.org/en/news/how-cpi-scores-are-calculated*

Transparency International. (2022, January 25). 2021 Corruption Perceptions Index—Explore the results. Retrieved from: https://www.transparency.org/en/cpi/2021

Transparency International. (n.d.). Extractive industries—Our priorities. Retrieved February 16, 2023, from *https://www.transparency.org/en/our-priorities/extractive-industries*

Transparency International (n.d.). Global Corruption Barometer. Retrieved March 2, 2023, from https://www.transparency.org/en/gcb

Transparency International. (n.d.). What is corruption? Retrieved 26 March 2023, from https://www.transparency.org/en/what-is-corruption

Valsts ieņēmumu dienests. (n.d.). Valsts amatpersonu deklarācijas. Retrieved 26 March 2023, from https://www6.vid.gov.lv/VAD

Van Vu, H., Tran, T. Q., Van Nguyen, T., & Lim, S. (2018). Corruption, Types of Corruption and Firm Financial Performance: New Evidence from a Transitional Economy. *Journal of Business Ethics*, *148*(4), 847–858.

Valsts reģionālās attīstības aģentūra. (2020). Teritorijas attīstības indekss [Data file]. Retrieved 1 March 2023, from https://www.vraa.gov.lv/lv/teritorijas-attistibas-indekss

Vidzemes televizija. (2014, September 25). Sūdzības par iepirkumu Valmierā kavē peldbaseina būvniecību. Retrieved 26 March 2023, from https://www.lsm.lv/raksts/zinas/latvija/sudzibas-par-iepirkumu-valmiera-kave-peldbaseina-buvniecibu.a99910/

8. Appendices

Appendix A. Tables and figures

Table 1

Summary statistics of regression variables.

	n	mean	sd	median	min	max	skew	kurtosis
single_bid	66056	0.718	0.450	1.000	0	1	-0.968	-1.063
publication	66056	0.149	0.356	0.000	0	1	1.975	1.901
Nepotism_na	66056	0.003	0.051	0.000	0	1	19.635	383.556
donation	66056	0.031	0.173	0.000	0	1	5.431	27.491
procedure	66056	0.667	0.471	1.000	0	1	-0.708	-1.499
missing_subcontractor	66056	0.355	0.478	0.000	0	1	0.608	-1.631
descr_length	66056	75.816	53.541	67.000	1	975	4.704	52.520
Area_match	66056	0.245	0.430	0.000	0	1	1.187	-0.592
Sum.of.bid_price_EUR	66056	136196.735	1447021.986	19039.520	100	154539758	58.052	4697.213
NACE_A	66056	0.029	0.167	0.000	0	1	5.639	29.795
NACE_B	66056	0.003	0.056	0.000	0	1	17.736	312.571
NACE_C	66056	0.075	0.263	0.000	0	1	3.233	8.449
NACE_D	66056	0.010	0.099	0.000	0	1	9.939	96.788
NACE_E	66056	0.008	0.090	0.000	0	1	10.934	117.557
NACE_F	66056	0.198	0.398	0.000	0	1	1.519	0.308
NACE_G	66056	0.265	0.441	0.000	0	1	1.063	-0.870
NACE_H	66056	0.015	0.122	0.000	0	1	7.950	61.202
NACE_I	66056	0.011	0.104	0.000	0	1	9.380	85.995
NACE_J	66056	0.019	0.135	0.000	0	1	7.143	49.029
NACE_K	66056	0.014	0.116	0.000	0	1	8.352	67.761
NACE_L	66056	0.008	0.091	0.000	0	1	10.801	114.672
NACE_M	66056	0.080	0.272	0.000	0	1	3.084	7.511
NACE_O	66056	0.000	0.010	0.000	0	1	97.124	9431.286
NACE_P	66056	0.005	0.073	0.000	0	1	13.628	183.731
NACE_Q	66056	0.007	0.083	0.000	0	1	11.937	140.500
NACE_R	66056	0.013	0.115	0.000	0	1	8.495	70.160
NACE_S	66056	0.014	0.118	0.000	0	1	8.235	65.812
tender_year_2014	66056	0.125	0.331	0.000	0	1	2.265	3.130
tender_year_2015	66056	0.113	0.316	0.000	0	1	2.452	4.015
tender_year_2016	66056	0.139	0.346	0.000	0	1	2.087	2.354
tender_year_2017	66056	0.089	0.285	0.000	0	1	2.890	6.350
tender_year_2018	66056	0.095	0.293	0.000	0	1	2.772	5.685
tender_year_2019	66056	0.085	0.278	0.000	0	1	2.983	6.901
tender_year_2020	66056	0.097	0.295	0.000	0	1	2.731	5.459
tender_year_2021	66056	0.104	0.305	0.000	0	1	2.596	4.739
tender_isEUFunded_yes	66056	0.111	0.314	0.000	0	1	2.483	4.167
Distinct_win	66056	420.973	390.760	417.000	0	1086	0.522	-1.081
Average.of.lot_bidsCount	66056	2.984	2.857	2.000	1	74	5.879	71.600
market_share	66056	0.032	0.087	0.003	0	1	4.798	29.316

Note. Created by the authors.

Table 2

Summary statistics of tender database by year.

tender_year	contracts	avg_value	single_bid	avg_bidders	avg_msh
2013	10218	75574	75.661	3.45	0.02
2014	8270	112390	70.326	2.98	0.02
2015	7432	130400	61.800	3.18	0.06
2016	9184	102373	69.523	3.02	0.02
2017	5870	142773	73.049	2.66	0.04
2018	6243	169469	69.165	2.48	0.03
2019	5594	235095	74.437	2.78	0.04
2020	6381	162224	79.345	3.24	0.04
2021	6864	165975	73.747	2.70	0.02

Note. Created by the authors.

Table 3

Summary statistics of tender database by industry.

NACE	contracts	avg_value	single_bid	avg_bidders	avg_msh	distinct_win
0	12981	147554	67.029	2.99	4.81	1914
A	1900	26369	75.579	2.79	1.46	517
B	208	67970	71.635	2.57	5.26	31
C	4941	43748	81.137	2.87	1.40	530
D	649	366051	75.809	3.03	7.29	37
E	539	927998	75.696	2.53	6.32	66
F	13049	269925	77.025	3.40	6.75	1187
G	17525	60354	74.128	2.90	1.28	1352
H	998	215418	65.230	2.40	2.95	186
I	726	193577	56.061	2.78	2.56	171
J	1223	90092	41.047	1.91	0.39	262
K	908	247576	80.286	2.91	1.79	33
L	552	143212	70.833	2.62	3.68	99
M	5317	71789	73.086	3.11	0.90	757
N	1916	113629	70.042	2.84	1.63	324
O	7	20510	57.143	1.86	0.49	5
P	350	32211	75.143	3.04	2.63	68
Q	454	64360	24.009	1.49	0.34	104
R	879	44151	33.106	1.68	0.74	225
S	933	37177	65.059	3.88	1.52	255
U	1	4500	100.000	2.00	0.00	1

Note. Created by the authors.

Table 4

Summary statistics of municipal variables.

	n	mean	sd	median	min	max	range	skew	kurtosis
score	236	-24.56	3.12	-24.80	-34.70	-13.30	21.40	0.15	0.36
score_w	236	-20.08	3.64	-19.55	-34.59	-10.89	23.70	-0.81	0.94
area	236	526.77	475.29	363.10	17.48	2,524.10	2,506.62	1.86	3.65
distance	236	109.92	63.48	103.65	0.00	253.21	253.21	0.21	-0.99
wage	236	739.17	140.39	719.75	498.80	1,203.75	704.95	0.68	0.06
budget_pc	236	1,428.31	288.22	1,387.66	896.36	2,454.88	1,558.52	0.78	0.44
period	236	0.50	0.50	0.50	0.00	1.00	1.00	0.00	-2.01
absolute	236	0.54	0.50	1.00	0.00	1.00	1.00	-0.15	-1.99
gov	236	0.53	0.50	1.00	0.00	1.00	1.00	-0.12	-1.99

Note. Created by the authors.

Table 5

Summary of media validation cases.

Case	Municipality	Year	Industry	Reference
1	Rīga	2014	Legal services	Krenberga, 2014
2	Valmiera	2014	Construction	Vidzemes televīzija, 2014
3	Jelgava	2017	Public transport	Kupčs, 2017
4	Rīga	2016	Public transport	Kupčs, 2016
5	Talsi municipality	2017	Construction	Nekā Personīga, 2017
6	Jelgava	2020	Road maintenance	Kupčs, 2020
7	Rēzekne	2020	Building maintenance	LSM.lv, 2020
8	Jūrmala	2019	Construction	Kupčs, 2019

Note. Created by the authors.

Table 6

Regression output of the initial regression with three dependant variables.

	Dependent variable:		
	Single Bid	Bidders	Market Share
	(1)	(2)	(3)
Constant	1.32 (0.02)***	6.01 (0.12)***	−0.05 (0.003)***
Nepotism_na	−0.04 (0.04)	−0.34 (0.16)	0.03 (0.01)***
donation	0.03 (0.01)***	−0.07 (0.06)	0.01 (0.003)***
publication	−0.26 (0.01)***	0.01 (0.05)	0.01 (0.001)***
procedure	−0.25 (0.01)***	−2.62 (0.13)***	−0.005 (0.002)***
missing_subcontractor	−0.12 (0.01)***	−2.42 (0.13)***	0.002 (0.002)
log(descr_length)	−0.01 (0.003)***	−0.07 (0.01)***	−0.0001 (0.0005)
Area_match	−0.05 (0.004)***	−0.18 (0.03)***	−0.01 (0.001)***
log(Sum.of.bid_price_EUR)	−0.03 (0.001)***	0.04 (0.02)***	0.01 (0.0003)***
NACE_A	−0.04 (0.01)***	−0.76 (0.06)***	−0.02 (0.001)***
NACE_B	0.01 (0.03)	−0.56 (0.13)***	−0.001 (0.01)
NACE_C	0.04 (0.01)***	−0.36 (0.04)***	−0.03 (0.001)***
NACE_D	0.03 (0.02)*	−0.12 (0.08)	0.01 (0.004)***
NACE_E	0.05 (0.02)***	−0.53 (0.07)***	0.01 (0.01)**
NACE_F	−0.02 (0.01)***	−0.55 (0.06)***	0.02 (0.002)***
NACE_G	−0.15 (0.01)***	−1.44 (0.06)***	−0.02 (0.002)***
NACE_H	−0.08 (0.02)***	−0.94 (0.07)***	−0.02 (0.003)***
NACE_I	−0.01 (0.02)	−0.11 (0.10)	−0.02 (0.003)***
NACE_J	−0.25 (0.01)***	−1.16 (0.06)***	−0.04 (0.001)***
NACE_K	0.09 (0.01)***	−0.09 (0.06)	−0.03 (0.002)***
NACE_L	−0.01 (0.02)	−0.65 (0.09)***	−0.01 (0.003)***
NACE_M	−0.04 (0.01)***	−0.53 (0.05)***	−0.03 (0.001)***
NACE_N	0.003 (0.01)	−0.28 (0.06)***	−0.03 (0.002)***
NACE_O	−0.09 (0.20)	−1.21 (0.39)	−0.05 (0.01)
NACE_P	0.07 (0.02)***	−0.02 (0.12)	−0.02 (0.002)***
NACE_Q	−0.29 (0.02)***	−0.98 (0.13)***	−0.04 (0.001)***
NACE_R	−0.30 (0.02)***	−1.55 (0.06)***	−0.04 (0.002)***
NACE_S	−0.05 (0.02)***	0.67 (0.19)***	−0.03 (0.002)***
tender_isEUFunded_yes	0.07 (0.01)***	0.75 (0.05)***	0.02 (0.001)***
Distinct_win	0.0002 (0.0000)***	0.001 (0.0001)***	−0.0000 (0.0000)***
year_effects			
Observations	66,056	66,056	66,056
R^2	0.13	0.08	0.14
Adjusted R^2	0.13	0.08	0.14

Note: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Note. Created by the authors.

Table 7

Municipal scores by year.

	Dependent variable:
	single_bid
Constant	1.32 (0.02)***
publication	−0.26 (0.01)***
procedure	−0.25 (0.01)***
missing_subcontractor	−0.12 (0.01)***
log(descr_length)	−0.01 (0.003)***
Area_match	−0.05 (0.004)***
log(Sum.of.bid_price_EUR)	−0.03 (0.001)***
NACE_A	−0.04 (0.01)***
NACE_B	0.01 (0.03)
NACE_C	0.04 (0.01)***
NACE_D	0.03 (0.02)*
NACE_E	0.06 (0.02)***
NACE_F	−0.02 (0.01)**
NACE_G	−0.15 (0.01)***
NACE_H	−0.08 (0.02)***
NACE_I	−0.01 (0.02)
NACE_J	−0.25 (0.01)***
NACE_K	0.09 (0.01)***
NACE_L	−0.01 (0.02)
NACE_M	−0.04 (0.01)***
NACE_N	0.005 (0.01)
NACE_O	−0.09 (0.20)
NACE_P	0.07 (0.02)***
NACE_Q	−0.29 (0.02)***
NACE_R	−0.30 (0.02)***
NACE_S	−0.04 (0.02)***
tender_year_2014	−0.04 (0.01)***
tender_year_2015	−0.14 (0.01)***
tender_year_2016	−0.03 (0.01)***
tender_year_2017	−0.03 (0.01)***
tender_year_2018	−0.08 (0.01)***
tender_year_2019	−0.05 (0.01)***
tender_year_2020	−0.01 (0.01)**
tender_year_2021	−0.03 (0.01)***
tender_isEUFunded_yes	0.07 (0.01)***
Distinct_win	0.0002 (0.0000)***
Observations	66,056
R^2	0.13
Adjusted R^2	0.13
Note:	*p<0.1; **p<0.05; ***p<0.01

Note. Created by the authors.

Table 8

Summary statistics of calculated score components.

	n	mean	sd	median	min	max	skew	kurtosis
sc_pub	66056	-3.84	9.20	0.00	-25.85	0.00	-1.98	1.90
sc_pro	66056	-16.50	11.66	-24.74	-24.74	0.00	0.71	-1.50
sc_sub	66056	-4.41	5.95	0.00	-12.44	0.00	-0.61	-1.63
sc_length	66056	-1.04	0.73	-0.91	-13.31	-0.01	-4.70	52.52
sc_area	66056	-1.31	2.29	0.00	-5.33	0.00	-1.19	-0.59

Note. Created by the authors.

Table 9

Municipal scores by year.

year	score	single bid %	coefficient	significance
2013	-29.06	75.66	-	-
2014	-28.62	70.33	-0.04	***
2015	-27.44	61.80	-0.14	***
2016	-30.08	69.52	-0.03	***
2017	-27.49	73.05	-0.03	***
2018	-25.90	69.17	-0.08	***
2019	-23.51	74.44	-0.05	***
2020	-22.47	79.34	-0.01	**
2021	-25.91	73.75	-0.03	***

Note. Created by the authors.

Table 10

Municipal scores by industry.

NACE	Industry	score	single bid %	Distinct winners	coefficient	significance
0	None	-30.74	67.03	1914	-	-
A	Agriculture	-25.16	75.58	517	-0.04	***
B	Mining	-25.65	71.63	31	0.01	
C	Manufacturing	-23.76	81.14	530	0.04	***
D	Electricity & Gas	-21.29	75.81	37	0.03	*
E	Water supply	-24.64	75.70	66	0.06	***
F	Construction	-25.60	77.03	1187	-0.02	**
G	Trade	-24.36	74.13	1352	-0.15	***
H	Transportation	-26.77	65.23	186	-0.08	***
I	Accomodation & Food	-41.21	56.06	171	-0.01	
J	Information & Comms	-32.67	41.05	262	-0.25	***
K	Finance & Insurance	-24.47	80.29	33	0.09	***
L	Real Estate	-26.69	70.83	99	-0.01	
M	Professional & Science	-28.44	73.09	757	-0.04	***
N	Admin & Support	-28.72	70.04	324	0.005	
O	Public Admin & Defence	-30.59	57.14	5	-0.09	
P	Education	-30.06	75.14	68	0.07	***
Q	Human Health	-44.12	24.01	104	-0.29	***
R	Recreation	-36.55	33.11	225	-0.30	***
S	Other Services	-31.55	65.06	255	-0.04	***

Note. Created by the authors.

Table 11

Corruption risk scores of media cases.

validation	score	single_bid	contracts	value	avg_value
Riga law	-55.58	0.00	8	278,989	34,874
Valmiera pool	-25.39	55.56	9	9,424,937	1,047,215
Jelgavas busses	-19.45	100.00	2	5,336,008	2,668,004
Riga busses	-44.30	0.00	1	96,000,000	96,000,000
Talsi construction	-20.65	66.67	9	2,844,487	316,054
Jelgava maintenance	-30.85	46.77	124	51,285,677	413,594
Rēzekne maintenance	-25.67	84.48	58	952,443	16,421
Jūrmala construction	-30.89	65.71	35	15,378,064	439,373
Average	-27.09	71.79	66,056	8,996,611,512	136,197
top 25%	-45.49	52.91	16,514	1,300,708,750	78,764

Note. Created by the authors.

Table 12

Financial validation of media cases risks.

tert	score	net_profit_margin	net_profit_avg	net_turnover_avg	employee_avg	pvalue
1	-25.77	-0.48	-35,551.27	1,945,653	30.27	-
2	-24.77	0.03	170,280.26	5,230,520	55.14	0
3	-26.05	0.25	414,351.63	2,142,749	25.79	0.04832

Note. Created by the authors.

Table 13

Corruption risk scores of all municipalities.

Municipality	Score	W_Score	Municipality	Score	W_Score	Municipality	Score	W_Score
Aglonas novads	-23.37	-17.96	Jaunpils novads	-25.25	-17.55	Raunas novads	-24.50	-24.34
Aizkraukles novads	-25.84	-18.99	Jelgava	-26.49	-18.32	Rēzekne	-26.78	-19.23
Aizputes novads	-26.34	-22.33	Jelgavas novads	-24.24	-16.16	Rēzeknes novads	-24.42	-21.04
Aknīstes novads	-23.70	-16.00	Jēkabpils	-25.68	-18.61	Riebiņu novads	-22.63	-18.46
Alojas novads	-25.06	-22.30	Jēkabpils novads	-26.48	-22.79	Rīga	-32.94	-21.97
Alsungas novads	-22.63	-19.66	Jūrmala	-27.16	-19.48	Rojas novads	-26.73	-26.48
Alūksnes novads	-29.28	-18.11	Kandavas novads	-28.97	-32.66	Ropažu novads	-26.89	-18.54
Amatas novads	-22.82	-19.16	Kārsavas novads	-25.39	-20.98	Rucavas novads	-26.09	-21.19
Apes novads	-25.91	-21.54	Kocēnu novads	-27.49	-23.51	Rugāju novads	-21.66	-16.91
Auces novads	-22.15	-20.96	Kokneses novads	-24.60	-24.93	Rundāles novads	-24.74	-18.60
Ādažu novads	-26.98	-20.67	Krāslavas novads	-27.42	-18.53	Rūjienas novads	-25.76	-21.66
Babītes novads	-25.41	-20.35	Krimuldas novads	-24.82	-19.23	Salacgrīvas novads	-25.47	-19.12
Baldones novads	-27.67	-27.62	Krustpils novads	-24.41	-22.74	Salas novads	-22.89	-16.13
Baltinavas novads	-25.80	-24.46	Kuldīgas novads	-25.56	-19.65	Salaspils novads	-25.64	-22.72
Balvu novads	-22.34	-18.28	Ķeguma novads	-26.08	-23.65	Saldus novads	-27.68	-17.76
Bauskas novads	-25.14	-21.05	Ķekavas novads	-26.62	-18.20	Saulkrastu novads	-27.25	-24.08
Beverīnas novads	-21.93	-19.62	Lielvārdes novads	-24.44	-19.15	Sējas novads	-24.64	-17.40
Brocēnu novads	-22.59	-20.31	Liepāja	-26.03	-19.08	Siguldas novads	-26.18	-19.47
Burtnieku novads	-27.08	-21.07	Limbažu novads	-26.15	-19.88	Skrīveru novads	-22.16	-17.92
Carnikavas novads	-25.63	-18.75	Līgatnes novads	-14.62	-13.57	Skrundas novads	-25.33	-22.25
Cesvaines novads	-22.60	-16.42	Līvānu novads	-22.84	-16.35	Smiltenes novads	-23.98	-20.58
Cēsu novads	-27.40	-21.53	Lubānas novads	-22.75	-17.86	Stopiņu novads	-24.10	-17.23
Ciblas novads	-23.59	-16.91	Ludzas novads	-22.56	-15.87	Strenču novads	-27.45	-18.98
Dagdas novads	-20.96	-18.42	Madonas novads	-26.50	-19.79	Talsu novads	-26.33	-20.20
Daugavpils	-26.34	-19.52	Mazsalacas novads	-27.41	-24.94	Tērvetes novads	-21.90	-20.26
Daugavpils novads	-24.88	-21.08	Mālpils novads	-25.59	-20.96	Tukuma novads	-23.58	-16.07
Dobeles novads	-26.18	-19.37	Mārupes novads	-26.17	-16.93	Vaiņodes novads	-19.79	-17.57
Dundagas novads	-27.33	-26.55	Mērsraga novads	-25.69	-18.11	Valkas novads	-26.23	-18.55
Durbes novads	-21.97	-19.44	Naukšēnu novads	-23.92	-16.82	Valmiera	-25.51	-19.77
Engures novads	-22.27	-18.71	Neretas novads	-21.83	-20.01	Varakļānu novads	-20.71	-16.34
Ērgļu novads	-23.18	-20.06	Nīcas novads	-23.17	-18.80	Vārkavas novads	-23.67	-24.16
Garkalnes novads	-29.12	-19.10	Ogres novads	-25.59	-19.66	Vecpiebalgas novads	-24.08	-20.27
Grobiņas novads	-26.88	-20.04	Olaines novads	-27.83	-23.99	Vecumnieku novads	-24.01	-14.49
Gulbenes novads	-25.57	-18.64	Ozolnieku novads	-26.33	-21.74	Ventspils	-26.82	-13.66
Iecavas novads	-23.10	-17.81	Pārgaujas novads	-23.24	-21.42	Viesītes novads	-25.13	-20.37
Ikšķiles novads	-29.13	-16.48	Pāvilostas novads	-22.66	-19.36	Viļakas novads	-23.58	-18.69
Ilūkstes novads	-26.67	-19.14	Pļaviņu novads	-22.97	-18.57	Viļānu novads	-24.30	-22.16
Inčukalna novads	-24.15	-17.36	Preiļu novads	-24.08	-17.12	Zilupes novads	-25.94	-21.57
Jaunjelgavas novads	-28.73	-20.63	Priekules novads	-23.28	-18.47	NA	NA	NA
Jaunpiebalgas novads	-27.36	-20.70	Priekuļu novads	-24.36	-19.52	NA	NA	NA

Note. Created by the authors.

Table 14

Top 10 municipalities with least corruption risk.

Municipality	score	w_score	rank_score	rank_score_weighted	rank_avg	rank_diff
Līgatnes novads	-14.62	-13.57	118	118	118.0	0
Varakļānu novads	-20.71	-16.34	116	110	113.0	6
Ludzas novads	-22.56	-15.87	105	115	110.0	-10
Rugāju novads	-21.66	-16.91	114	104	109.0	10
Vaiņodes novads	-19.79	-17.57	117	97	107.0	20
Cesvaines novads	-22.60	-16.42	103	108	105.5	-5
Salas novads	-22.89	-16.13	96	112	104.0	-16
Līvānu novads	-22.84	-16.35	97	109	103.0	-12
Dagdas novads	-20.96	-18.42	115	86	100.5	29
Skrīveru novads	-22.16	-17.92	108	93	100.5	15
Tukuma novads	-23.58	-16.07	88	113	100.5	-25

Note. Created by the authors.

Table 15

Top 10 municipalities with most corruption risk.

Municipality	score	w_score	rank_score	rank_score_weighted	rank_avg	rank_diff
Kandavas novads	-28.97	-32.66	5	1	3.0	4
Baldones novads	-27.67	-27.62	9	2	5.5	7
Mazsalacas novads	-27.41	-24.94	13	5	9.0	8
Olaines novads	-27.83	-23.99	7	11	9.0	-4
Dundagas novads	-27.33	-26.55	16	3	9.5	13
Rīga	-32.94	-21.97	1	21	11.0	-20
Kocēnu novads	-27.49	-23.51	10	13	11.5	-3
Saulkrastu novads	-27.25	-24.08	17	10	13.5	7
Rojas novads	-26.73	-26.48	25	4	14.5	21
Cēsu novads	-27.40	-21.53	14	26	20.0	-12

Note. Created by the authors.

Table 16

Highest increases of corruption risk due to weighing.

Municipality	score	w_score	rank_score	rank_score_weighted	rank_avg	rank_diff
Vārkavas novads	-23.67	-24.16	85	9	47.0	76
Auces novads	-22.15	-20.96	109	34	71.5	75
Tērvetes novads	-21.90	-20.26	112	44	78.0	68
Neretas novads	-21.83	-20.01	113	48	80.5	65
Pārgaujas novads	-23.24	-21.42	91	27	59.0	64
Kokneses novads	-24.60	-24.93	69	6	37.5	63
Brocēnu novads	-22.59	-20.31	104	42	73.0	62
Raunas novads	-24.50	-24.34	70	8	39.0	62
Krustpils novads	-24.41	-22.74	73	15	44.0	58
Beverīnas novads	-21.93	-19.62	111	55	83.0	56

Note. Created by the authors.

Table 17

Largest decreases of corruption risk due to weighing.

Municipality	score	w_score	rank_score	rank_score_weighted	rank_avg	rank_diff
Ikšķiles novads	-29.13	-16.48	3	107	55.0	-104
Ventspils	-26.82	-13.66	23	117	70.0	-94
Alūksnes novads	-29.28	-18.11	2	91	46.5	-89
Saldus novads	-27.68	-17.76	8	96	52.0	-88
Krāslavas novads	-27.42	-18.53	12	83	47.5	-71
Garkalnes novads	-29.12	-19.10	4	69	36.5	-65
Mārupes novads	-26.17	-16.93	38	103	70.5	-65
Ķekavas novads	-26.62	-18.20	27	89	58.0	-62
Ropažu novads	-26.89	-18.54	21	82	51.5	-61
Strenču novads	-27.45	-18.98	11	72	41.5	-61

Note. Created by the authors.

Table 18

Municipal corruption risk by historic Latvian regions.

Historic	score	w_score	score_diff	single_bid	contracts	value	avg_value
Kurzeme	-26.17	-20.68	-5.49	69.66	10,049	1,310,870,378	130,448
Latgale	-25.06	-19.39	-5.67	75.81	12,134	1,011,327,647	83,347
Rīga	-32.94	-21.97	-10.97	62.09	12,406	3,463,770,439	279,201
Sēlija	-25.53	-19.22	-6.31	78.37	2,436	201,038,799	82,528
Vidzeme	-26.14	-20.05	-6.09	74.48	22,492	2,250,180,704	100,044
Zemgale	-25.08	-18.28	-6.80	74.28	6,539	759,423,546	116,138

Note. Created by the authors.

Table 19

Municipal corruption risk by biggest Latvian cities.

Municipality	score	w_score	score_diff	single_bid	contracts	value	avg_value
Daugavpils	-26.34	-19.52	-6.82	72.89	3,349	434,942,152	129,872
Jelgava	-26.49	-18.32	-8.17	70.77	2,412	372,236,648	154,327
Jēkabpils	-25.68	-18.61	-7.07	76.18	1,335	151,578,241	113,542
Jūrmala	-27.16	-19.48	-7.68	68.58	2,190	396,714,781	181,148
Liepāja	-26.03	-19.08	-6.95	67.98	2,011	408,784,166	203,274
Rēzekne	-26.78	-19.23	-7.55	69.82	1,819	219,106,424	120,454
Rīga	-32.94	-21.97	-10.97	62.09	12,406	3,463,770,439	279,201
Valmiera	-25.51	-19.77	-5.74	73.00	1,815	242,966,364	133,866
Ventspils	-26.82	-13.66	-13.16	60.74	2,802	522,832,432	186,593
Municipalities	-25.41	-19.90	-5.51	76.25	35,917	2,783,679,864	77,503

Note. Created by the authors.

Table 20

Risk score analysis by division in 10% increments on tender value, Riga.

tert	score	w_score	score_diff	weight	single_bid	contracts	value	avg_value
1	-34.13	-0.01	-34.12	0.00	82.11	1,241	1,457,622	1,175
2	-49.41	-0.06	-49.35	0.00	37.07	1,241	4,369,471	3,521
3	-35.42	-0.08	-35.34	0.00	58.42	1,241	8,097,131	6,525
4	-34.28	-0.14	-34.14	0.00	59.39	1,241	13,730,978	11,064
5	-34.02	-0.21	-33.81	0.01	62.77	1,241	21,461,976	17,294
6	-31.76	-0.29	-31.47	0.01	62.45	1,241	31,878,889	25,688
7	-31.41	-0.42	-30.99	0.01	58.31	1,240	46,439,300	37,451
8	-28.87	-0.55	-28.32	0.02	66.21	1,240	66,450,904	53,589
9	-27.33	-1.23	-26.10	0.04	61.94	1,240	155,509,925	125,411
10	-22.71	-18.98	-3.73	0.90	72.26	1,240	3,114,374,244	2,511,592

Note. Created by the authors.

Table 21

Risk score analysis by division in 10% increments on tender value, Ventspils.

tert	score	w_score	score_diff	weight	single_bid	contracts	value	avg_value
1	-30.48	-0.05	-30.43	0.00	63.35	281	929,408	3,308
2	-27.79	-0.10	-27.69	0.00	58.72	281	1,851,634	6,589
3	-27.58	-0.14	-27.44	0.01	57.86	280	2,724,392	9,730
4	-27.77	-0.20	-27.57	0.01	56.79	280	3,732,557	13,331
5	-29.87	-0.29	-29.58	0.01	61.79	280	5,034,798	17,981
6	-29.49	-0.40	-29.09	0.01	56.07	280	7,179,596	25,641
7	-27.87	-0.56	-27.31	0.02	55.71	280	10,507,085	37,525
8	-25.39	-0.85	-24.54	0.03	64.64	280	17,583,979	62,800
9	-24.42	-1.74	-22.68	0.07	62.86	280	37,360,193	133,429
10	-17.55	-9.32	-8.23	0.83	69.64	280	435,928,792	1,556,889

Note. Created by the authors.

Table 22

Risk score analysis by division in 10% increments on tender value, Valmiera.

tert	score	w_score	score_diff	weight	single_bid	contracts	value	avg_value
1	-29.69	-0.02	-29.67	0.00	75.82	182	155,105	852
2	-26.20	-0.08	-26.12	0.00	81.87	182	713,994	3,923
3	-25.17	-0.15	-25.02	0.01	78.57	182	1,489,156	8,182
4	-27.38	-0.26	-27.12	0.01	74.18	182	2,268,022	12,462
5	-27.27	-0.36	-26.91	0.01	73.63	182	3,202,178	17,594
6	-25.68	-0.46	-25.22	0.02	70.72	181	4,372,150	24,156
7	-28.03	-0.71	-27.32	0.03	67.40	181	6,105,346	33,731
8	-24.21	-0.89	-23.32	0.04	67.40	181	9,070,367	50,113
9	-22.23	-1.62	-20.61	0.07	66.85	181	17,497,440	96,671
10	-19.24	-15.22	-4.02	0.82	73.48	181	198,092,607	1,094,434

Note. Created by the authors.

Table 23

Risk score analysis by division in 10% increments on tender value, Kandava.

tert	score	w_score	score_diff	weight	single_bid	contracts	value	avg_value
1	-27.23	-0.13	-27.10	0.00	81.82	22	69,065	3,139
2	-28.30	-0.29	-28.01	0.01	95.45	22	153,284	6,967
3	-34.31	-0.55	-33.76	0.02	86.36	22	241,581	10,981
4	-27.06	-0.56	-26.50	0.02	72.73	22	313,568	14,253
5	-26.68	-0.73	-25.95	0.03	80.95	21	416,703	19,843
6	-25.90	-0.98	-24.92	0.04	57.14	21	579,134	27,578
7	-29.09	-1.61	-27.48	0.05	66.67	21	834,222	39,725
8	-28.01	-1.76	-26.25	0.06	57.14	21	946,074	45,051
9	-30.27	-3.50	-26.77	0.11	71.43	21	1,692,121	80,577
10	-32.85	-22.55	-10.30	0.65	80.95	21	9,945,264	473,584

Note. Created by the authors.

Table 24

Risk score analysis by division in 10% increments on tender value, Latvia.

perc	score	sc_pub	sc_pro	sc_sub	sc_len	sc_area	single_bid	contracts	value	avg_value
1	-29.33	-5.67	-17.58	-3.96	-0.99	-1.12	81.65	6,606	9,524,589	1,442
2	-31.60	-6.11	-19.64	-3.67	-0.97	-1.22	70.30	6,606	27,585,032	4,176
3	-28.50	-4.16	-19.66	-2.66	-0.96	-1.07	71.75	6,606	47,370,958	7,171
4	-28.11	-3.75	-19.66	-2.57	-0.97	-1.15	71.62	6,606	73,131,300	11,070
5	-27.89	-3.49	-19.74	-2.52	-1.00	-1.16	70.86	6,606	106,065,013	16,056
6	-27.61	-3.25	-19.49	-2.61	-1.01	-1.24	68.88	6,606	148,917,382	22,543
7	-27.43	-3.19	-18.91	-2.94	-1.04	-1.35	67.12	6,605	213,062,725	32,258
8	-26.70	-3.15	-16.78	-4.21	-1.04	-1.52	65.40	6,605	303,147,939	45,897
9	-23.69	-2.77	-10.74	-7.56	-1.12	-1.50	71.98	6,605	627,383,786	94,986
10	-20.07	-2.88	-2.77	-11.43	-1.26	-1.73	78.32	6,605	7,440,422,788	1,126,483

Note. Created by the authors.

Table 25

Results of the municipal regression.

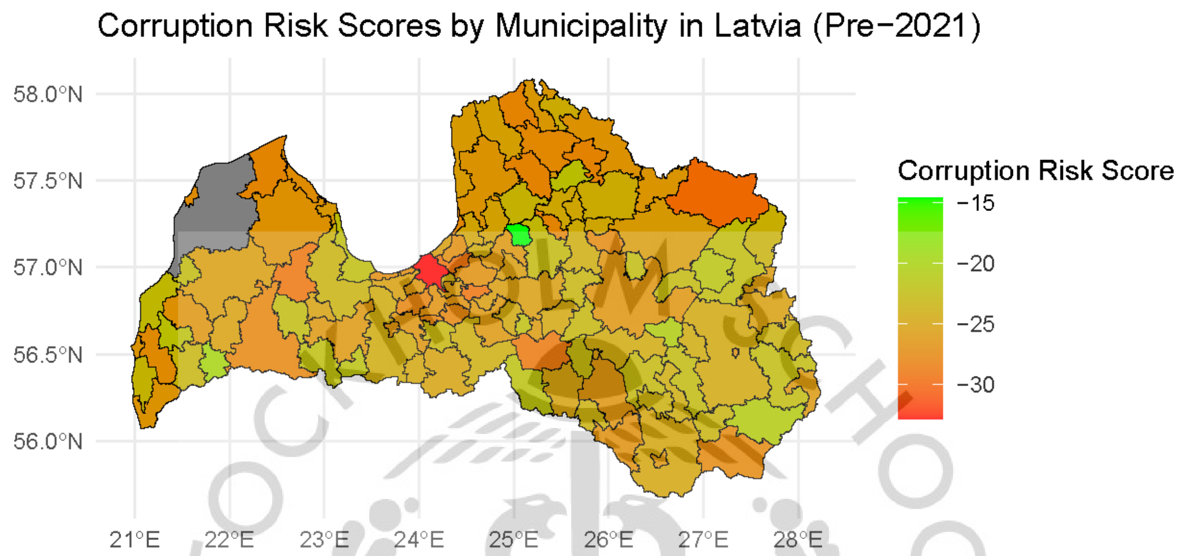
	Dependent variable:	
	score	score_w
	(1)	(2)
Constant	-22.46 (1.76)***	-28.36 (2.37)***
area	-0.001 (0.0003)	0.0003 (0.0004)
wage	-0.01 (0.002)*	0.01 (0.003)*
budget_pc	0.0003 (0.001)	0.001 (0.001)
distance	0.003 (0.004)	0.01 (0.005)*
period	4.37 (0.55)***	1.02 (0.75)
absolute	-0.01 (0.34)	0.15 (0.45)
gov	-0.70 (0.34)*	-0.82 (0.46)
Observations	236	236
R^2	0.37	0.18
Adjusted R^2	0.35	0.15

Note: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Note. Created by the authors.

Figure 1

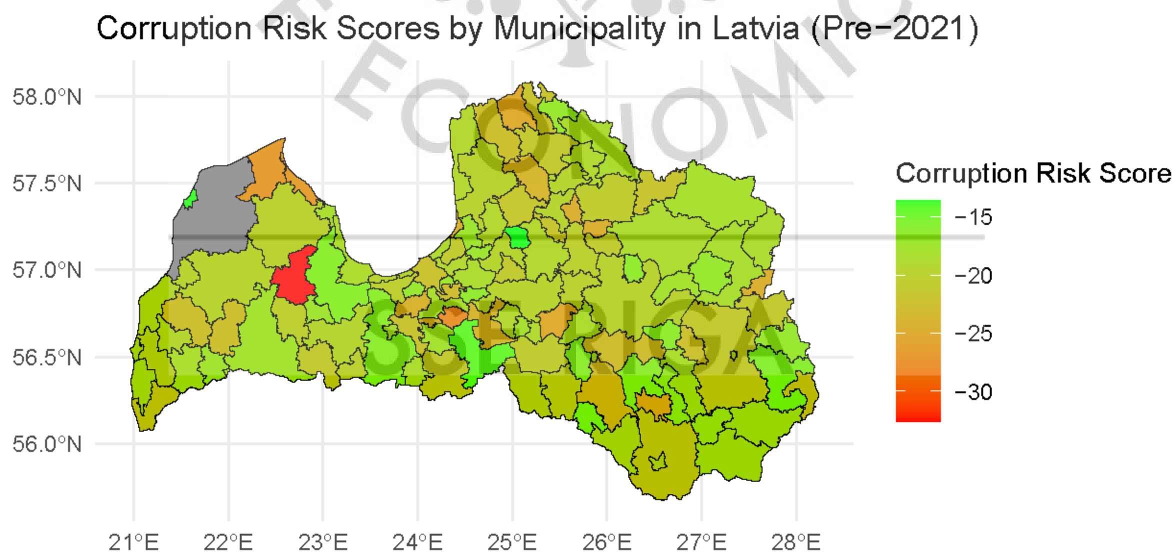
Heatmap of Latvian municipalities by risk score.



Note. Ventspils municipality was merged into Ventspils by OpenTender database. *Created by the authors.*

Figure 2

Heatmap of Latvian municipalities by weighted risk score.



Note. Ventspils municipality was merged into Ventspils by OpenTender database. *Created by the authors.*

Appendix B. VIF score results

Table 26

VIF scores of main regression variables.

	x
publication	1.28
procedure	5.85
missing_subcontractor	5.62
log(descr_length)	1.08
Area_match	1.08
log(Sum.of.bid_price_EUR)	1.43
NACE_A	1.28
NACE_B	1.02
NACE_C	1.35
NACE_D	1.06
NACE_E	1.04
NACE_F	4.16
NACE_G	8.82
NACE_H	1.09
NACE_I	1.07
NACE_J	1.09
NACE_K	1.07
NACE_L	1.04
NACE_M	1.81
NACE_N	1.15
NACE_O	1.00
NACE_P	1.02
NACE_Q	1.04
NACE_R	1.07
NACE_S	1.07
tender_year_2014	1.60
tender_year_2015	1.61
tender_year_2016	1.66
tender_year_2017	1.47
tender_year_2018	1.51
tender_year_2019	1.47
tender_year_2020	1.54
tender_year_2021	1.60
tender_isEUFunded_yes	1.10
Distinct_win	6.59

Note. Created by the authors.

Table 27

VIF scores of main regression variables.

area	1.21
wage	3.74
budget_pc	1.46
distance	1.83
period	2.53
absolute	1.07
gov	1.14

Note. Created by the authors.

Memo on AI tool usage

We have not used any AI based tools for content creation in the thesis.