Version prepared for the 2025 Bahamas research conference on financial crime. To cite the paper:
M. Siino, S. Iezzi, and M. Gara. "Corruption risk indicators in public procurement: A proposal using Italian open data." Quaderni dell'antiriciclaggio 23 (2024).
Link: https://uif.bancaditalia.it/pubblicazioni/quaderni/2024/quaderno-23-2024/index.html

CORRUPTION RISK INDICATORS IN PUBLIC PROCUREMENT: A PROPOSAL USING ITALIAN OPEN DATA

by Mario Gara, Stefano Iezzi and Marianna Siino*

Abstract

This paper leverages open data from Italy's Central Anti-Corruption Authority (ANAC) and relevant literature to propose a multi-layered system of risk indicators for detecting potential corruptive conducts in public procurement. The development and use of such indicators are widespread among national and international organizations due to the significant resources allocated to public procurement globally, making this sector particularly vulnerable to corruption. This vulnerability is particularly critical in Italy, where corruption is reportedly connected to criminal infiltration, and European institutions are currently disbursing an unprecedented amount of funds for infrastructures and structural reforms. The proposed individual indicators are auction-specific, each addressing a different aspect of a tender. Using these indicators, we also compute a composite risk measure at auction level and, by further aggregation, develop indicators at the level of contracting authorities and winning bidders. A significant contribution of this work is the use of confidential data from the Financial Intelligence Unit (UIF) to validate these indicators, providing evidence suggestive of their effectiveness. The potential operational applications of these indicators include monitoring public tenders, risk-ranking of awarding authorities and contractors, prioritizing investigative and anti-money laundering activities.

JEL Classification: D73, H57.
Keywords: public procurement, tender, corruption, risk indicators, contracting authority, firms.

* Bank of Italy, Financial Intelligence Unit.

1. Introduction[1]

According to the World Bank, public procurement represents, on average, 13% to 20% of a country's GDP, with global expenditure in procurement estimated at nearly 9.5 trillion US dollars.[2] The United Nations Office on Drugs and Crime warns that 10% to 25% of this amount may be lost to corruption (UNODC, 2013).

In Italy, public procurement amounts to about 12% of GDP[3], slightly less than the world average. However, Italy's public procurement sector is of particular concern due to the high level of corruption the country has allegedly suffered from. Despite criticisms regarding its reliability and empirical robustness (Andersson and Heywood, 2009), Transparency International's Corruption Perception Index systematically ranks Italy well above its peer countries in terms of perceived corruption.[4]

The strong ties between the management of public resources and illegal conducts have surfaced on many occasions when Italian central and local governments had to face catastrophic events: major episodes of embezzlement of public funds emerged after the earthquake in the Naples region in 1980 and more recently during the Covid-19 pandemic of 2020-2021.

It is also acknowledged, based on evidence from criminal investigations, that in Italy the relationship between organized crime and public officials has become an increasingly crucial feature of corruptive schemes (Cantone and Caringella, 2017). Corruption has evolved from occasional payments of bribes and kickbacks for individual enrichment to a central stage of a far-reaching criminal strategy, based on stable and long-lasting connections between public officials, entrepreneurs and organized crime.

Within this context, despite many reform attempts, the Italian public procurement system, also due to its fragmentation, has been deemed as particularly vulnerable to collusion, corruption and inefficiencies (De Carolis et al., 2011). For these reasons, an effective system for monitoring the allocation of public resources appears crucial not only for increasing the efficiency of public administration and government, but also for curtailing organized crime's grip on the economy and its threat to the state. Any instrument for detecting corruption is essential for increasing the quality and integrity of state institutions at political, social, juridical and economic levels (Gnaldi, 2018).

These arguments gain additional relevance now as Italy is receiving an unprecedented amount of financial resources from the European institutions within the so-called National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza – PNRR) for funding public projects and infrastructures. The success of this ambitious plan relies on an effective system of controls to ensure the lawful allocation of these resources and to prevent misappropriation and embezzlement.

This study proposes a system of risk indicators for detecting potential corruptive conducts based on data from Italy's Central Anti-Corruption Authority (ANAC), which is extensive in both time span covered and information provided. Indicators are auction-specific and defined in line with the relevant

---

[1] The views and the opinions expressed in this paper are those of the authors and do not necessarily represent those of the Institutions they are affiliated with. We thank Ilaria Cosenza for the normative support and the colleagues who participated at the internal UIF seminar for their very useful comments and suggestions. We would also like to thank the Italian Anti-Corruption Law Authority (ANAC) for the comments received and the profitable discussion carried out during a dedicated seminar.

[2] Cfr. The World Bank Global Public Procurement Database.
https://www.worldbank.org/en/topic/governance/brief/global-public-procurement-database

[3] Cfr. Organization for Economic Cooperation and Development (OECD) - General government procurement spending as a percentage of GDP, 2021
https://www.oecd.org/en/topics/public-procurement.html

[4] Corruption may be measured by both subjective (perception-based) and objective indicators, each with their own limitations (Gnaldi, 2018; ANAC, 2017). Subjective measures, such as the Corruption Perception Index (CPI), rely on surveys and are useful for international comparisons but face criticism for their accuracy and consistency. Objective measures include experiential data from real experiences, market-based proxies, and discrepancies in administrative data. Judicial measures track corruption through complaints, convictions, and penalties, but these are delayed and can indicate inefficiencies within the judicial system or under-reporting of corruption due to the nature of the crime.

literature in the field; they address specific features of tenders, such as the criteria applied for awarding the contracts, the number of bidders, and some characteristics of the latter. Based on these individual indicators, it is possible to rate auctions according to their risk level and, through additional aggregation, to build indicators at the level of contracting authorities and bidders. By exploiting confidential data available at the UIF, we also test the robustness and relevance of the indicators, obtaining encouraging results.

The paper is structured as follows. Section 2 illustrates the research motivation, set against the relevant contributions in the literature. Section 3 provides some background on the data used and the regulatory framework they are based on. Section 4 describes the data in more detail, including the issue of missing information. Section 5 illustrates the individual indicators, which are then used to build composite indicators at the award and contractor levels in Section 6, which also provides a validation test for their significance. Section 7 constructs a composite indicator at the level of awarding authority, while Section 8 reviews the potential operational applications of all indicators. Some brief concluding remarks are provided in Section 9.

2. Research motivation and literature review

Given the sheer amount of resources involved, public procurement, including the management of tenders and subsequent supply contracts, features a particularly high risk of corruption. This substantial involvement of resources provides a strong motivation for our research. Therefore, any system of early warning or red flags for detecting contracts and tenders that may involve underlying misconducts is particularly useful in preventing and fighting corruption, as well as in stopping considerable public resources from being unduly allocated or appropriated by criminals.

This study aims at defining a system of risk indicators of potential corruption for the monitoring of public procurement contracts, contractors and contract-awarding authorities.

The literature has consistently pointed to the definition and application of such indicators as valuable and efficient tools for fighting corruption (Ferwerda et al., 2017). The deployment of such systems of indicators is widely spread in international organizations for monitoring their projects in developing countries (Kenny and Musatova, 2010). World Bank (2013) illustrates an extremely comprehensive and ready–to–use system of micro red flags: indicators are diversified for each phase of the public procurement procedure and for different type of illicit behavior, such as collusion, fraudulent bids, irregular implementation of contracts, or illicit financial management. Similarly, Ferraris et al. (2016) compile a list of tender-specific risk indicators, distinguishing between pre-tender and post-tender ones, based on a survey identifying the most relevant vulnerabilities in the tender procedures adopted by most EU member states. However, they stress that the actual deployment of such indicators, which in some countries are fed to algorithms and automated decision-making procedures, is restricted to a tiny minority of the countries surveyed. The report from the European Anti-Fraud Office (OLAF, 2013) goes a step forward by defining 27 individual indicators based on the analysis of public procurement contracts in eight EU member countries across five different sectors and providing an estimate of the risk of corruption in each sector and of the direct costs associated to it. Moreover, OLAF (2017), after detailing a comprehensive set of red flags for detecting fraud in public procurement, tackles the issue from a pragmatic viewpoint by providing case examples, solutions, and best practices. Finally, Del Sarto et al. (2021) contend that corruption indicators, in addition to helping identify murky tenders and potentially colluded authorities and bidders, may also support the design of more effective procurement policies.

Hence, the exercise undertaken in this study is widely recognized as a good practice for the enhancement of the efficiency in a country's public procurement procedures, especially given the proven inefficiency of Italy's public procurement system and its allegedly vulnerability to corruption and organized crime, as stressed in the previous section. Moreover, the timing of this exercise is particularly fitting, given the substantial resources that are soon to be made available for infrastructure projects in Italy.

One factor identified in the literature as a potential facilitators of corruption is the discretionary powers that contracting authorities may exert when determining various aspects of the tender procedure, such as the starting bid, the content of the tender, and the criteria for awarding a contract. Decarolis et al. (2021) analyze the features of public contracts regarding which authorities make use of these powers; based on this analysis the authors establish a measure of corruption risk both at the single contract level and, through aggregation, at the awarding authority level. They also test the robustness of the indicators and find a slight correlation, both at the auction and auctioning authority levels, with evidence proving the involvement of either bidding companies or their administrators in corruption investigations. They conclude that the increase in corruption risk from the use of discretionary powers is more than balanced by the gains in terms in higher efficiency in the awarding procedures.

Along the same line of research, Decarolis and Giorgiantonio (2022) widen the set of indicators to include the exhaustiveness of the information submitted by companies participating in a tender and the presence of unusual obstacles to taking part in a tender. They find that satisfactory validation results can be obtained only by deploying a machine learning approach, while traditional regressions show that only a few indicators are correlated with opaque winning bidders identified by police investigation for corruption-related crimes. Using a battery of 28 indicators, Ferwerda et al. (2017) reach similar conclusions, stressing that only few indicators are effectively correlated with the occurrence of corruption based on a sample of approximately 200 EU procurement contracts, equally divided between regular and tainted ones.

Other studies provide further evidence that the use of discretionary powers is strongly correlated with other potential measures of corruption, such as inefficiency measured by delays in project completion or cost overruns (Coviello et al., 2018), and the share of contracts assigned to politically connected companies (Baltrunaite et al., 2021). Additionally, it is shown that, when discretionary procedures are applied, companies infiltrated by crime have a greater probability of winning a contract than other bidders (Mirenda et al., 2022), with this probability increasing following natural disasters such as earthquakes (Marcolongo, 2023).

Another prominent indicator deployed in these studies is the lack of competition in tender procedures, which is proxied by the number of bidders. Wachs et al. (2021) applies a graph analysis approach to the European procurement market, relying solely on the indicator signaling tenders with a single bidder, and provide insights into the structure of the relationship between contract-awarding authorities and bidding companies.

On the other hand, the set of indicators defined by Troìa (2020), though limited in number, includes quite diverse dimensions of a bid that may be interpreted as signals of irregularity and thus of anomaly. The analysis relies on the same database used in this study (the Italian Open-Access Database of Public Contracts, or BDNCP from the Italian denomination), but the indicators are estimated only at the awarding authority level. Moreover, no evidence of their robustness is provided.

By using the same database, ANAC has created a comprehensive collection of procurement indicators to detect corruption at the territorial level (ANAC, 2021). These indicators are obtained by combining red flags at the provincial level to warn authorities of potential corruptive conducts, support prevention and integrity, and promote transparency in public administration.

Along similar lines, the aim of this study is to define a set of corruption risk indicators at the level of individual public procurement tenders, which are then appropriately combined into composite indicators at the award, contractor and contracting authority levels. Consistently with the literature, in constructing these indicators we account for the discretionary features of tenders, the extent to which awarding authorities comply with mandatory requirements, the degree of competitiveness in the procurement procedure, and some characteristics of the winners. By applying a bottom-up computational approach, we aggregate individual indicators to build a multi-layered system of indicators, resulting in composite indicators for awarding authorities, contractors and geographical areas.

The strength of our approach lies in that indicators are computed using an extremely wide

sample of tenders and their robustness is tested relying on confidential evidence on connections between bidders and criminal organizations.

Also, whilst previous studies solely concentrate on available data, we also take into account the pattern of data unreported by contracting authorities, regarding the awarding process and the winning bidders, which prove valuable in offering additional insights on the transparency, accountability, and integrity of such authorities. This approach provides further hints for interpreting and using the estimated indicators.

3. The regulatory framework and the Italian open-access database of Public Contracts

In order to better illustrate the data available in the Italian open-access portal of Public Contracts and the risk indicators described in the following sections, it is essential to provide a general overview of the European and Italian legislation regarding public contracts. This includes a wide array of laws, directives, and regulations governing the procurement process and the awarding of contracts by public authorities and other entities.

Public procurement refers to the process whereby public authorities and other entities purchase work, goods or services. The European Union regulation[5] establishes harmonized rules for tenders to ensure a level playing field for businesses across Europe. As detailed in Table 1, these rules apply to tenders exceeding certain monetary thresholds,[6] which are revised every two years and depend on the sector (ordinary and special), the type of contract (works and concessions, goods and services, design contests, social services, etc.), and the awarding entity (central governmental authorities, sub-central contracting authorities). EU rules require contracting authorities to issue EU-wide calls for tender above the EU thresholds, while tenders below the thresholds are governed by national rules. These latter rules must nonetheless adhere to the general principles of EU law, ensuring fairness, equity, transparency, and non-discrimination in the awarding of contracts.[7]

Table 1
European Union thresholds for public contracts between 2018 and 2023

Sector	Years	*Public works contracts and concessions*	*Public goods and services contracts and public design contests awarded by central governmental authorities*	*Public goods and service contracts and public design contests awarded by sub-central contracting authorities*	*Social services and other specific services contracts covered in Annex IX of the Public Contract Code*
Ordinary	**2018-19**	5,548,000	144,000	221,000	750,000
	2020-21	5,350,000	139,000	214,000	750,000
	2022-23	5,382,000	140,000	215,000	750,000
Special	**2018-19**	5,548,000	443,000		1,000,000
	2020-21	5,350,000	428,000		1,000,000
	2022-23	5,382,000	431,000		1,000,000

In Italy, contracts below EU thresholds are governed by the national Public Contract Code. Until June 2023, public procurement was regulated by Legislative Decree No. 50/2016 and its subsequent amendments. In July 2023, the new Public Procurement Code (Legislative Decree No. 36/2023) came

[5] Directive No. 24/2014.
[6] In the previous Italian Public Contract Code, the Legislative Decree No. 50/2016 and subsequent amendments, Article 35 describes EU thresholds and methods for calculating the value of contracts. Whereas in the new Public Contract Code, Legislative Decree No. 36 of 2023, EU thresholds are set by Article 14.
[7] Calls for EU tenders are published on the Tenders Electronic Daily (TED) at https://ted.europa.eu. Conversely, for contracts below EU thresholds, publicity takes place only at the national level.

into force. The "new Code" introduced significant changes, including the so-called "principle of result",[8] rules for direct contracting, new subcontracting procedures, price revision clauses, and new digitization standards.[9] Although our analysis, which encompasses data from January 2018 to June 2023, does not cover the period when the new procurement code took effect, the proposed methods can be expanded and applied to the latest data, ensuring their relevance from a policy perspective.

The rules set by the Public Contract Code[10] concern contracts awarded through competitive tendering and direct procedure, distinguishing between ordinary and special sectors.[11] Various competitive procedures are available, such as: open procedure, negotiated procedure, restricted procedure, competitive dialogue, and innovation partnership. The Code also lists the type of entities that may be considered as awarding bodies: authorities such as the State, any regional or local authorities, national or local public entities, and bodies governed by public law; contracting entities operating in a special sector; and specific private entities subject to the Code.

In addition to the regulation governing the actual functioning of procurement procedures and public contracts, the menace of corruption and the harm it causes to countries and societies has created worldwide momentum towards the introduction of a wide set of rules promoting the transparency of public procurements data. Such disclosure is rightly considered as an effective tool to fight misconduct in the management and allocation of public resources. To this end, various initiatives have been put in place, both at a transnational and a national level.

The World Bank has established the Global Public Procurement Database, which collects and discloses data on public procurement across 218 countries and independent territories. The initiative also aims to develop a web-based tool providing diagnostics for procurement-related data, for instance, allowing the comparison across countries in terms of their profile, procurement practices, laws and regulations, and performance.[12]

Also the European Union, through its funding programme for research and innovation, has funded the project Digiwhist, a repository collecting data on public procurement contracts covering 35 jurisdictions (EU member states plus some associate countries, including Switzerland).[13]

In this endeavour, the private sector is playing a relevant role as well, through non-governmental and non-profit organizations, such as Open Contracting Partnership, which works in over 50 countries with the aim of defining a common data model for enabling the disclosure of public procurement data and documents at all stages of the contracting process.[14]

As for Italy, in 2010 the BDNCP was established pursuant to the amendments made to the so-called Digital Administration Code introduced in that year.[15] The database contains the information that the responsible of the procedures of the awarding authorities is required to file to ANAC.[16] The database

[8] This principle establishes that tenders be carried out timely and with the aim of obtaining the best possible quality-to-cost ratio, in compliance with the principles of legality, transparency and competition (Article 1, Legislative Decree n. 36/2023).
[9] For details on the innovations introduced by the decree see the documentations available on the ANAC web-site, for example https://www.anticorruzione.it/-/appalti-sotto-soglia-e-subappalto-le-novit%C3%A0-del-codice-analisi-dei-princ%C3%ACpi-generali-della-legge#p1
[10] Article 36 of Legislative Decree No. 50/2016 outlines the procedures to use and the minimum number of economic operators to consult for tenders based on contract value, sector type (ordinary or special), and tender object (work, service, or goods contracts). During the analysis period, this provision was frequently modified to simplify the regulatory and administrative framework for public contracts.
[11] The special sector includes utilities, such as gas and heat, electricity, water, transport services, ports and airports, postal services, and extraction of oil and gas and exploration for, or extraction of, coal or other solid fuels. Moreover, the Code provides a specific regulation for tenders in the social services and others referred in the IX Annex.
[12] https://www.globalpublicprocurementdata.org/gppd/
[13] https://digiwhist.eu/about-digiwhist/
[14] Cfr. https://www.open-contracting.org/. In spite of the initiatives being mentioned and others around the world, Open Contracting Partnership estimates that data concerning only 3% of the overall public procurement budget are disseminated.
[15] Legislative Decree No. 82/2005, as amended by Legislative Decree No. 235/2010.
[16] https://dati.anticorruzione.it/opendata

is composed of several datasets each containing a wide array of information, those considered in our analysis are:

- **Tenders with ordinary Tender Identification Code (so-called CIG)**,[17] which contains essential data on all tenders with a value greater than 40,000 euros, or for which an ordinary CIG has been required.
- **Contracting authority**, containing information on the contracting authority, which is responsible for providing all data.
- **Award**, describing the information of the awarding of a lot; one tender may be split in different contracts for various reasons (such as dividing the original tender into different lots to encourage competition or having multiple phases or sub-projects), but all contracts correspond to the same identifier (CIG).
- **Successful bidder**, identifying the economic operators winning the tender; several bidders or groups of companies may correspond to one award.
- **Start and end of contract**, which are communicated by contracting authorities.
- **Sources of funding**, which contracting authorities are required to detail, including the origin of the economic resources used to pay the suppliers.

Given the extensiveness of the information provided, the structure of the database is extremely complex; hence the analysis of the data is sometimes cumbersome and requires compound matching exercises, which often results in the emergence of missing or incomplete information In addition to providing an overview of the data used for the computation of the indicators, the following section also examines the impact and potential determinants of the extensive data incompleteness featured in the database.

4. Data description and analysis of missing data

We analyse Italian public contracts with an ordinary CIG (i.e., the tender identifier) published between January 2018 and June 2023 in the BDNCP. These amount to approximately 2,1 million tenders (excluding those cancelled, revoked, deserted or irregular) for a total value of contracts equal to around 2,8 billion euros. Table 2 shows the distribution of tenders and the total amount in millions of euros based on the sector and type of contract. The ordinary sector accounts for 89.2% of the tenders, of which 73.0% pertain to the provision of goods and services and the 16.2 % works contracts.

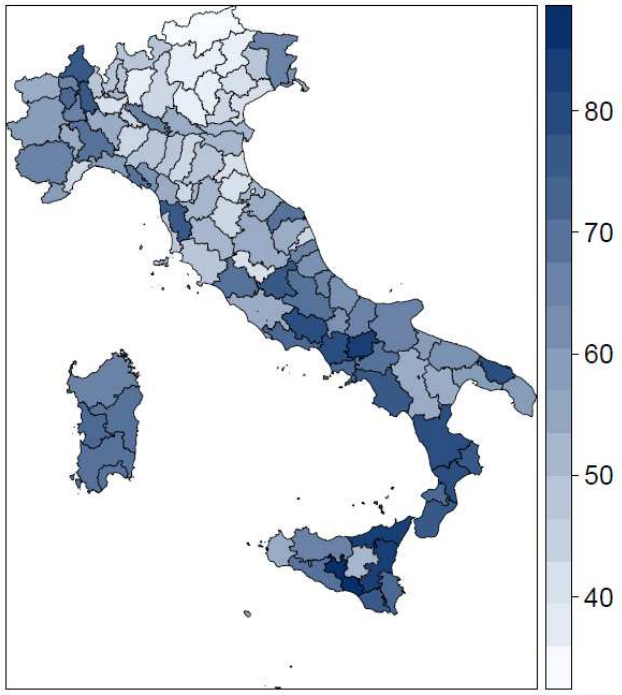
Table 2
Distribution of number and amount of tenders with an ordinary CIG by sector and type of contract (Jan 2018 - Jun 2023).

Sector	Type of contract	Tenders		Amount in millions of euros			
		Number	%	Sum	%	Mean	S.d.
Ordinary	Goods	918,450	43.3	1,893,895	66.3	2.0	8.0
	Works	343,352	16.2	172,977	6.1	0.5	2.4
	Services	631,552	29.7	496,797	17.4	0.8	4.0
	Total ordinary sector	*1,893,354*	*89.2*	*2,563,669*	*89.8*	*1.4*	*6.1*
Special	Goods	850,88	4.0	82,749	2.9	1.0	4.4
	Works	472,05	2.2	98,091	3.4	2.1	6.1
	Services	967,99	4.6	110,376	3.9	1.1	4.7
	Total special sector	*229,092*	*10.8*	*291,216*	*10.2*	*1.3*	*4.9*
Total		**2,122,446**	**100**	**2,854,885**	**100**	**1.4**	**6.0**

[17] Generally, contracts with an amount less than 40,000 euros and those not subject to the reporting obligations by the contracting authorities receive a so-called SmartCIG. Consequently, information on the latter contract category is narrower and basically relates to data contained in the call for tender. However, there are exceptions to these rule: for instance, an ordinary CIG is required for the tenders financed with resources from the PNRR and the National Cohesion Plan (Piano Nazionale per gli Investimenti Complementari - PNC) regardless of their value.

Despite the considerable amount of data provided by the BDNCP, our preliminary analyses reveal that information concerning single tenders is often incomplete. For example, up to approximately 58% of the tenders lack information about the award procedure. Likewise, roughly the same share of tenders have an unspecified winning bidder.[18] Moreover, in the vast majority of cases, it is the same tenders that lack both sets of data.[19]

Figure 1
Missing data: Percentage of tenders for which the award procedure was not communicated



Given the sheer size of missing or incomplete data, analysing this feature thoroughly is crucial for developing risk indicators and ensuring the interpretability of our research findings. To this end, Figure 1 shows the percentage of tenders for which the award procedure was not communicated, according to the province of the contracting authority. Figure 1 indicates a higher presence of missing information for tenders in Southern Italy.

To examine this further, we assess the hypothesis of Missing Completely at Random (MCAR) using a model-based approach. The MCAR assumption refers to a scenario where the probability of a data point being missing is completely unrelated to any observed or unobserved data. Specifically, we use a logistic regression model to detect patterns or trends of missing data, aiming to identify any attributes of the tender and the contracting authority that may contribute to this phenomenon.

[18] Other glaring examples of the problem are: for 77% of tenders there is no indication of the starting date and for 90% of the end date; finally, for only 32% of contracts information on the sources of funding is provided. For these last three missing linkages, we define proper risk indicators as described in Section 5.

[19] The Legislative Decree 14 March 2013, n. 33 establishes general principles of transparency and mandates the publication of various administrative documents online. Furthermore, the D. Lgs. n. 50/2016, particularly in Article 29, established the principles of transparency, mandating that all acts related to public procurement be publicly disclosed on the profiles of contracting authorities and the ANAC digital platform. The new Public Contracts Code (D. Lgs. n. 36/2023) builds upon these principles in Articles 29 to 31, emphasizing the importance of transparency and digitalization. Also in that code there are references of sanctions related to lack of communications.

The response variable of our logistic model is binary, indicating whether each tender lacks data on the awarding process (1) or otherwise (0).[20] The predictor variables, detailed in Table 3, include features of the tender, such as the sector and the object of the contract, along with some characteristics of each contracting authority, including its type, geographical area, and size proxied by the overall value of ordinary tenders it has awarded in the period of analysis. The model also includes a linear trend capturing the variability of the missing pattern across the years.

Table 3
Missing data: Predictor variables in the logistic model

Variable name	Description	Benchmark in the regression
Year	Year of publication of the tender from 2018 to 2023	-
Amount	Logarithm of the tender amount	-
Sector	Type sector of the awarding procedure: "Ordinary" or "Special"	Ordinary
Object	Object of the public contract: "Works", "Goods" or "Services"	Works
Type C. A.	Classification of contracting authority in 19 macro-types	Other
Geography C. A.	Geographical area of residence of the contracting authority: "North", "Center" and "South and Islands"	North
Dimension C. A.	Total value of tenders with ordinary CIG by contracting authority from January 2018 to June 2023 divided into 4 classes according to the quartiles (very small, small, medium, large)	Very small

The estimated coefficients indicate the direction and relevance of the relationship between the predictor variables and the incompleteness of data. Figure 2 plots the value of the estimated coefficients and their 95% confidence intervals. Confidence intervals help distinguish those predictor variables that have statistically significant effects on missing data from those that are not influential. When confidence intervals do not include zero, coefficients are statistically significant, indicating a robust relationship between the corresponding predictor variable and missing data. Comparing confidence intervals across different predictor variables lets us identify those factors with the most relevant influence on missing data.

As shown in Figure 2, all the estimated coefficients are statistically significant, indicating a robust relationship between the corresponding predictor variable and the presence of missing data. Specifically, the positive relationship between missing data and the year of publication shows that data for more recent years are more incomplete. Additionally, the logistic regression indicates a lower probability of missing data in the special sector (compared to the ordinary sector, which is used as the benchmark of the regression). As for contracting authorities, those located in the Northern regions exhibit a lower probability of incomplete information compared to all the others, which feature a positive and statistically significant probability of missing data. Furthermore, authorities belonging to the highest operational activity bracket (i.e., dimension of the contracting authority "Large") are more prone to provide complete information.

[20] As it is described in ANAC, 2021, failure to comply with the award communications might reflect poor administrative conduct and might be linked to an increased risk of corruption.

Figure 2
Estimated coefficients and their 95% confidence intervals associated to the predictor variables in the logistic regression (type of contracting authority not included)

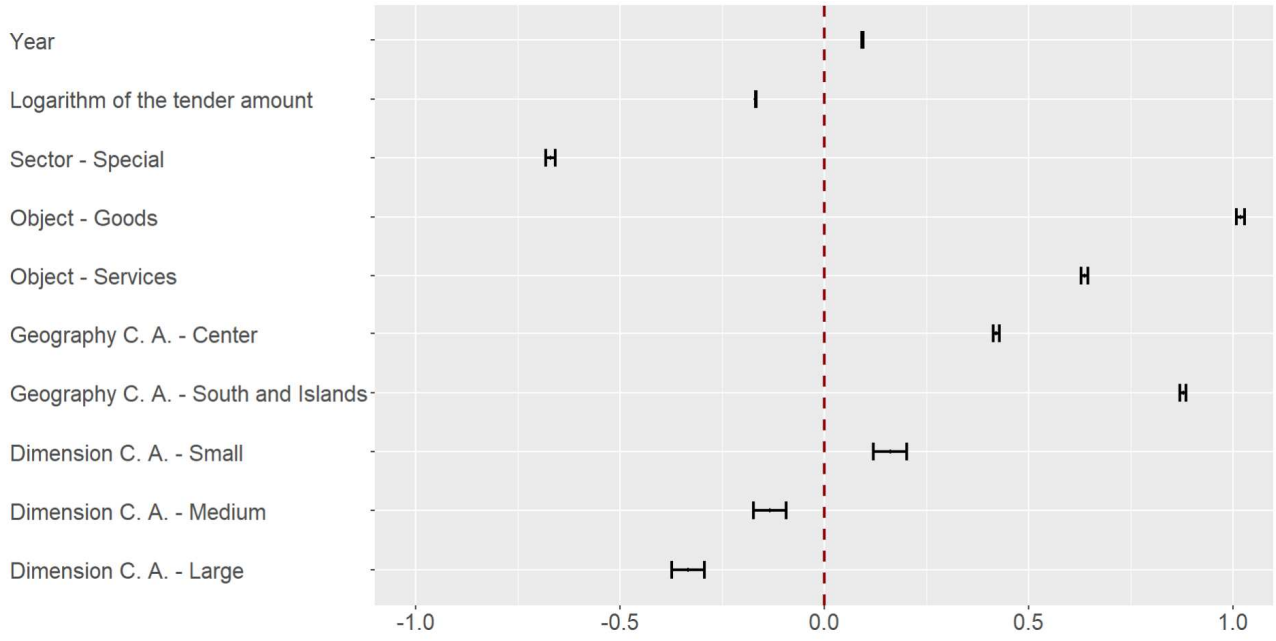


Table 4 provides the estimated coefficients of the different types of contracting authorities, in ascending order by value, together with their confidence intervals.

Table 4
Correlation coefficients of contracting authorities with missing data

Type of contracting authority	Logistic regression coefficients	Lower bound 95% CI	Upper bound 95% CI
Individual company	1.25	0.46	2.05
Education and training	0.48	0.45	0.51
Local government	0.22	0.20	0.25
Regions, autonomous provinces and their consortia and associations	0.19	0.17	0.21
Social security assistance and professional associations	0.00	-0.04	0.05
Environmental protection/industrial development of the territory	0.00	-0.04	0.04
Central purchasing body	-0.01	-0.05	0.02
Company in consolidated income statement	-0.09	-0.11	-0.07
Municipalities, mountain communities, metropolies, union of municipalities, consortia and associations	-0.14	-0.15	-0.13
Scientific, research and educational development	-0.17	-0.19	-0.15
Healthcare	-0.26	-0.27	-0.25
Provinces and their consortia and associations	-0.27	-0.30	-0.25
Services of general interest	-0.28	-0.30	-0.27
Welfare and cultural services	-0.29	-0.35	-0.23
Residential construction/real estate sector	-0.46	-0.52	-0.40
Non-economic public bodies	-0.48	-0.54	-0.43
Central bodies	-0.48	-0.51	-0.46
Market regulation	-1.07	-1.15	-0.99

The estimated coefficients indicate that the behaviour of contracting authorities varies depending on their type, with the most significant influence on missing data observed in "Individual companies", followed by authorities operating in "Education and training".

These results show how the pattern of missing data in the dataset is strongly connected to the type of procurement and characteristics of the awarding authority. This evidence highlights how the level of transparency and efficiency in procurement procedures can significantly vary across contracting authorities. Overall, these findings seem to suggest a correlation between the missing data and any observable or unobservable factors that are themselves correlated with the risk of corruption, and therefore a potential distortion in the estimation of the corruption risk indicators. In other terms, the analysis of missing data presented in this Section indicates that the proposed risk indicators could be affected by a selection bias, the extent of which may vary based depending on the geographical location, size and type of authority. Specifically, the risk indicators might be underestimated for small contracting authorities located in the Center and South of Italy. This emphasizes the importance of comparing risk indicators within homogenous peer groups of contracting authorities, for instance, within the same geographical macro area.

5. Individual risk indicators

Using the data available in the BDNCP and taking stock of the indications provided in the literature, we define a set of 12 indicators as proxies for assessing corruption risk in tender procedures. These indicators are related to different features of the award procedure and the winning bidder. Furthermore, these indicators can be properly aggregated at different levels, depending on the dimension of analysis: contracting authorities, geographical areas or contractors. All our indicators are binary, that is, they are equal to 1 if the underlying condition is observed, and 0 otherwise.

At this stage, however, a preliminary caveat is imperative. Each individual indicator aims to flag anomalies in different features of the awarding process, which do not necessarily imply abuse or irregular conducts. At times, the adoption of procedures which are not fully compliant with mandatory requirements may be associated to emergency situations or other exceptional conditions (for instance, the tender concerns the supply of an extremely specialised service, which can be satisfied by a restricted number of qualified firms, which cannot ensure an adequate level of competition). In other cases, the flagged indicators may be due to inaccuracies in the underlying data. However, the concurrent observation of different anomalies in a single tender or the recurrent flagging of a specific risk indicator for the same contracting authority or contractor across different tenders may conversely signal a consolidated propensity to adopting unconventional practices, which may be associated to irregularities or illegal conducts. Risk indicators should make staff and managers more vigilant and prompt them to take the necessary action; they signal that the situation needs to be checked and monitored with due diligence (European Commission, 2021).

Our set of indicators are divided into groups, each considering a different feature of the tendering process, as illustrated below.

a. Characteristics of the tender

a.1 Discretionary procedure (NoOpenAward)

The Code establishes specific situations when contracting authorities have the discretion to choose the contractor and may opt for alternative procedures to the open tender, such as direct award, negotiated procedure, or competitive dialogue. As according to the literature (Coviello, et al., 2018; Decarolis et al., 2020, 2022; Fazekas et al., 2016), in these procedures there is a higher risk of abuse and irregularities, a specific indicator has been built so as to account for the adoption of non-open procedures.

a.2 Direct award (DirectAward)

Direct award is a method of awarding a public contract in which the contracting authority directly selects a contractor without conducting a public tender. This procedure might be more opaque and increases the risk of irregularities, also compared to other discretionary procedures; the indicator flags the tenders in which the contractor is chosen using this method and it is a subcase of the indicator a.1.

a.3 Anomalous non-open procedure (AnomalousNoOpen)

The indicator flags tenders in which a non-open procedure was adopted in a way which appears inconsistent[21] with the general criteria established by the regulations for the EU open procedure. It has been computed by comparing the procedure adopted for selecting the contractor and the tender characteristics - type, sector, amount, and publication date – with the general references described in Table 1.

a.4 Anomalous direct procedure (AnomalousDirect)

For tenders below EU thresholds, the conditions under which the direct award criterion can be employed are defined by law.[22] The indicator flags tenders in which the direct award was adopted in a way which appears inconsistent with the general rule by comparing the procedure adopted for selecting the contractor and the tender characteristics - type, sector, amount, and publication date - with the maximum thresholds established for direct award by the law.

a.5 Most economically advantageous tender (MEAT)

In selecting the winning bid, contracting authorities can apply either of two criteria: the lowest price and the most economically advantageous offer. The latter leaves wider room for discretion, as the contracting authority is required to assess not only the price but also non-quantitative factors (such as quality, innovation, sustainability, and the overall value of the offer).[23] Therefore, the application of this criterion might hamper transparency and at the same time may not necessarily safeguard the efficiency of the selection (ANAC, 2017; Decarolis et all., 2022; OLAF, 2017). The indicator highlights whether the procedure is based on this criterion.

b. Characteristics of the awarding process

b.1 Single offer (SingleBid)

One widely recognised indicator of a high-risk condition for corruption in the context of public procurement is the presence of a single offer, which signals the lack of competitive bidding[24] (Abdou et al., 2021; Fazekas et al., 2016). The indicator flags tenders with a single bid.

b.2 Negotiated procedure with a number of invited companies less than five (NPInvitedLow)

[21] Various cases have been identified concerning revocation from general law framework, such as contract modifications (e.g., continuation, extension, or renewal as described in Article 106 of the Legislative Decree No. 50/2016), cases of extreme urgency and civil protection (Article 163), design and related assignments (Article 157), and situations outlined in Articles 63 and 125 pertaining to the use of negotiated procedures. Additionally, in some cases discrepancies have been noted between the contractor selection criteria and the description of tender subjects, likely due to inaccuracies in the data entry.

[22] Article 36 of the Legislative Decree No. 50/2016 and subsequent amendments describes when to use the direct award depending the combination of the tender amount, the type of tender and object of the contract.

[23] In practice, the most economically advantageous tender (MEAT) criterion is finding an increasingly prominent role. The European Directive No. 24/2014 on public procurement establishes that the MEAT should be the primary criterion for awarding public contracts. From an economic literature perspective, the use of this criterion is held to be more suitable for complex procurements, whereas the lowest price criterion is considered preferable for standardized procurement components.

[24] Though the presence of multiple participants supports competitiveness, Carbone et al., 2024, show that bid-rigging schemes may emerge also in the context of competitive auctions with several participants. On the basis of judicial evidence, they examine bidder's behaviours and their patterns of interactions underlying collusive dynamics. By defining appropriate company-specific indicators, they reveal that cartels frequently exploit subcontracts and price similarity, plus their tactics vary depending on the number of companies involved.

In the case of a negotiated procedure, the contracting authority may award the contract only after inviting a statutorily-set minimum number of companies to place their bid.[25] These companies are identified through market surveys or official lists; authorities need also allow eligible bidders to rotate in their participation to tenders. The indicator flags those tenders where the number of bidders which are invited to take part falls below the mandatory threshold. In particular, we set a threshold equal to 5, which represents the minimum among the various thresholds specified by the law for the negotiated procedures during the period of analysis.[26]

b.3 Subcontract request (SubRequest)

Subcontracting occurs when the contractor, who was awarded a tender, delegates part of or all the contract execution to another company (or subcontractor).[27] The use of subcontracting introduces an element of opacity in the relationship between the contracting authority and the winning bidder, limiting the former's ability to oversee the actual execution of the contract. Additionally, it can be a means to conceal corruption practices, such as bribery.[28] The proposed indicator highlights whether the main contractor requested the possibility of subcontracting (part of or all) the contract execution during the bidding phase. The indicator does not identify the effective use of subcontractors but solely the request for subcontracting, as data on the actual introduction of subcontractors are missing or incomplete.

c. Characteristics of the winning bidder

c.1 Recurrent winner (ReWinner)

A bidder which happen to be awarded contracts by the same contracting authority with a particularly high frequency may indicate a low degree of competition in the bidding process or conflict of interest (European Commission, 2021). The indicator flags if the winning bidder was awarded multiple tenders by the same contracting authority during the period of analysis.

d. Missing communications

The obligation to report to ANAC the information related to the several stages of the tender process is essential to ensure transparency and integrity. The failure to comply with this obligation is an indication of misconduct by contracting authorities that could be associated to inefficiencies, but could also signal a greater risk of corruption. For each award, we identify three risk indicators that measure the incompleteness of the data regarding the beginning of the contract (***MissingStart***), its conclusion (***MissingEnd***), and the funding sources used to finance the tender (***MissingFunds***).

***

The set of indicators as proposed above seems to capture nicely the various dimensions that are typically held to be relevant for the detection of potential corruptive conducts, as laid out in Section 2 above: some address the issue of contracting authorities' discretionary powers (e.g., a.1, a.2, a.5) and their actual compliance with regulatory requirements (e.g., a.3, a.4, b.2); others concern the degree (or lack) of

[25] Under normal circumstances, negotiated procedures require a minimum number of invitations. Article 36 of Legislative Decree No. 50/2016 and subsequent amendments establish the minimum number of economic operators to be consulted for the negotiated procedure, specifying the tender amount and the object of the award procedure (work, service, or goods contracts). Instead, in the new contract Code, Legislative Decree No. 36/2023, Article 50 regulates the negotiated procedure. Fewer than the legally mandated number of invitations does not automatically indicate abuse, rather it may instead result from a contract's contract's urgency or a lack of qualified economic operators.

[26] With an exception between April 2019 and June 2019, during which period Decree Law No. 32/2019 "Sblocca Cantieri" indicated the consultation of at least three economic operators for the negotiated procedure for works.

[27] For complex awards involving multiple diverse projects, subcontracting is allowed to enhance efficiency. However, Italian regional administrations have varying legislation on subcontracting. Decarolis and Giorgiantonio (2015) note that while several Northern regions and autonomous provinces have laws that extend subcontracting beyond national regulations, some Southern regions, such as Sicily, enforce stricter standards than the national level.

[28] Sub-contractors may act as channels of bribes through, for instance, false invoicing.

competition characterising the awarding procedures (e.g., b.1, c.1); and finally also the issue of transparency is dealt with (e.g., the three indicators under d.).

6. Empirical Analysis: award and contractor levels

6.1. Award level

The individual risk indicators presented in the preceding Section are first computed for each award reported in the BDNCP, including 2,316,231 awards.[29] Table 5 provides the basic descriptive statistics (mean and standard deviation) of the indicators. Five out of the twelve indicators have missing data primarily due to incomplete information regarding the tender procedure and the winner.

The majority of awards (85%) are associated with non-open procedures, with 53% specifically related to direct awards. Approximately 18% of the awards appear to deviate from the prescribed legal criteria for open procedures or direct awards.

In 55% of the awards where the data on the number of offers is provided, only one offer is observed. Regarding negotiated procedures, 14% of cases involve fewer companies than the fixed threshold of 5. In approximately 22% of tenders a subcontract has been requested and approximately 70% of tenders were awarded to a bidder that won at least another tender with the same contracting authority in the period of analysis.

Data on the start of the contract is available for 25% of the awards, that on the end only for 11%. Information about funding sources is provided for 35% of the awards.

Table 5
Descriptive statistics of the individual risk indicators at the award level.

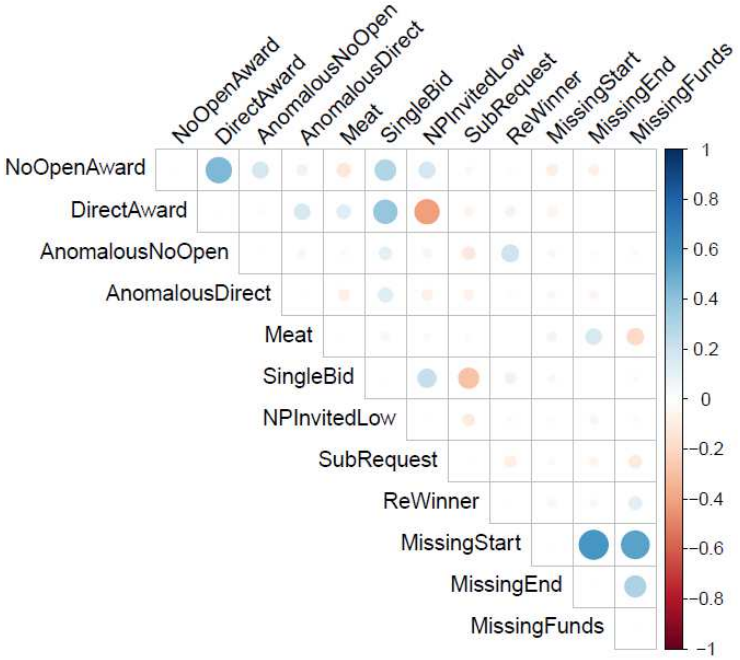
Indicator	Mean (%)	S.d. (%)	Missing (%)	N
NoOpenAward	84.9	35.8	-	2,316,231
DirectAward	52.7	49.9	-	2,316,231
AnomalousNoOpen	14.7	35.4	-	2,316,231
AnomalousDirect	3.0	17.0	-	2,316,231
MEAT	83.0	37.6	71.0	671,930
SingleBid	55.0	49.8	69.2	712,406
NPInvitedLow	14.0	34.7	54.1	1,063,857
SubRequest	22.2	41.5	52.9	1,090,237
ReWinner	68.6	46.4	53.1	1,087,004
MissingStart	75.2	43.2	-	2,316,231
MissingEnd	89.0	31.3	-	2,316,231
MissingFunds	65.0	47.7	-	2,316,231

Figure 3 presents a matrix illustrating pairwise association between the indicators computed by the so-called Phi coefficient. The Phi coefficient is a measure of association for two binary variables. It ranges from -1 to 1. Positive values indicate positive association, in particular a value equal to +1 indicates perfect agreement. Negative values indicate negative association and a value equal to -1 indicates perfect

[29] As a robustness check, the analysis has been also undertaken using a subset of tenders selected by applying the criteria employed by ANAC in its annual report. This subset includes tenders valued at over 40,000 euros and excludes certain types of awards, such as those pertaining to insurance and financial services, direct awards adhering to framework agreements or conventions, and awards granted to in-house companies. The results obtained are consistent with the main findings presented in the paper.

disagreement. Finally, a value equals to 0 indicates no relationship between the two variables. We emphasize that computationally the Pearson correlation coefficient reduces to the Phi coefficient. In the plot, a circle represents the association for each pair of variables: in particular, its size is proportional to the absolute value of the association coefficient. Additionally, its colour is a function of both the magnitude and the polarity of the relationship (colours on the red scale represent the negative association instead colours on the blue scale represent the positive one).

Figure 3
Pairwise association matrix of risk indicators (Phi-coefficient).



It is worth noting that some indicators are weakly correlated by construction: for instance, *NoOpenAward* and *DirectAward* exhibit a positive association of 0.45, given that the latter is a subset of the former; likewise, *DirectAward* and *SingleBid* show a positive association (0.38), reflecting reduced competitiveness in direct awards. Moreover, the three indicators concerning missing data are positively associated to each other, suggesting that normally contracting authorities fail to provide data on multiple aspects of the tender. Finally, *DirectAward* and *NPInvitedLow* are negatively associated (-0.41) because they refer to alternative procedures and thus cannot be flagged simultaneously.

Individual risk indicators can be aggregated into a composite indicator providing a synthetic measure of the degree of potential corruption in procurement process, thus facilitating evaluation and decision-making. Among the several methods for constructing a composite indicator available in the literature (OECD, 2008), we compute ours as a weighted average of the individual risk indicators. The weights assigned to each indicator take into account the statistical association between them so as to encompass to what extent different indicators capture distinct facets of the same phenomenon.[30] The composite indicator can be computed only for awards having the complete set of the individual indicators.

Therefore, the formula for the composite risk indicator is as follows:[31]

[30] Before computing the overall average, we compute the mean of the two indicators connected with the direct pro (*DirectAward* and *AnomalousDirect*), as well as for the three indicators addressing data deficiencies (*MissingStart, MissingEnd MissingFunds*).

[31] The denominator is equal to the maximum number of indicators that can be activated at the same time, as *ForDirectAward* and *NPInvitedlow* are alternative to each other.

[1]

$$CIAward =$$

$$(NoOpenAward + ForDirectAward\ + AnomalousNoOpen + Meat + SingleBid + NPInvitedLow + SubRequest + ReWinner + Missing\ )/8$$

where:

$$ForDirectAward = (DirectAward + AnomalousDirect)/2$$

$$Missing = (MissingStart + MissingEnd + MissingFunds)/3$$

Figure 4 shows the density plot of the composite indicator, computed on 621,974 awards; it assumes 44 distinct values, the ones with the highest frequency located at the centre of the distribution. Table 6 presents the summary statistics by sector and contract type. As shown by the mean and quantile values, the distribution of the composite indicator in the "Special sector" exhibits higher values compared to the "Ordinary sector". In particular, the mean is the highest for services and works contracts in the "Special sector".

Figure 4

Density plot of the composite indicator at the award level

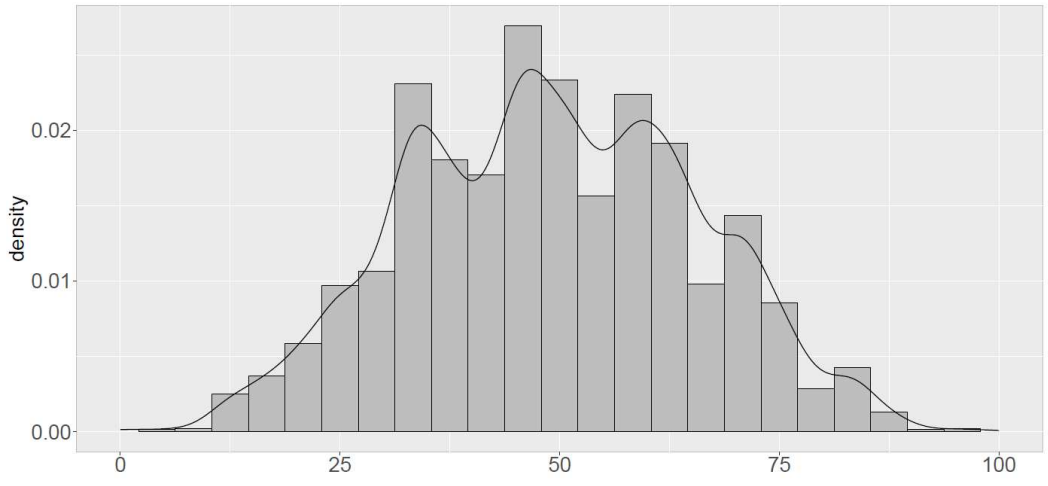


Table 6

Summary statistics of the composite risk indicator *CIAward*: overall and by sector and contract type. (percentage values)

Sector	Type of contract	*Summary statistics*							
		Min	**1st Qu.**	**Median**	**3rd Qu.**	**Max.**	**Mean**	**S.d.**	**%missing**
Ordinary	Goods	0	37.5	52.1	66.7	100	52.5	16.9	82.4
	Works	0	33.3	45.8	56.3	95.8	44.7	14.8	57.5
	Services	0	33.3	47.9	60.4	100	47.4	16.2	72.9
Total ordinary sector		***0***	***35.4***	***47.9***	***60.4***	***100***	***48.2***	***16.3***	***74.5***
Special	Goods	4.2	37.5	50	62.5	100	52	15.7	65.3
	Works	0	45.8	54.2	66.7	100	55.3	15.7	67.3
	Services	4.2	45.8	58.3	68.8	100	55.4	16.1	57.4
Total special sector		***0***	***45.8***	***54.2***	***64.6***	***100***	***54.4***	***16.1***	**62.6**
Total		***0***	***37.5***	***50.0***	***62.5***	***100***	***49.2***	***16.4***	***73.1***

In addition to helping compare macro groups of awards (as in Table 6), the composite indicator is particularly useful in ranking awards within each group according to their risk. That could be used as an operational tool for identifying most opaque tenders with a view of inquiring whether the procedure

adopted complies with the regulatory requirements. For instance, according to the divisions of the Common Procurement Vocabulary (CPV) classification system[32], for awards in the ordinary sector and related with goods, the higher mean value of the indicator is for petroleum products, and other sources of energy and medical equipment, pharmaceuticals and personal care products. Instead, for awards in the special sector and related with goods, the higher mean value of the indicator is for machineries and apparatus, and software and information systems.

6.2. Contractor level

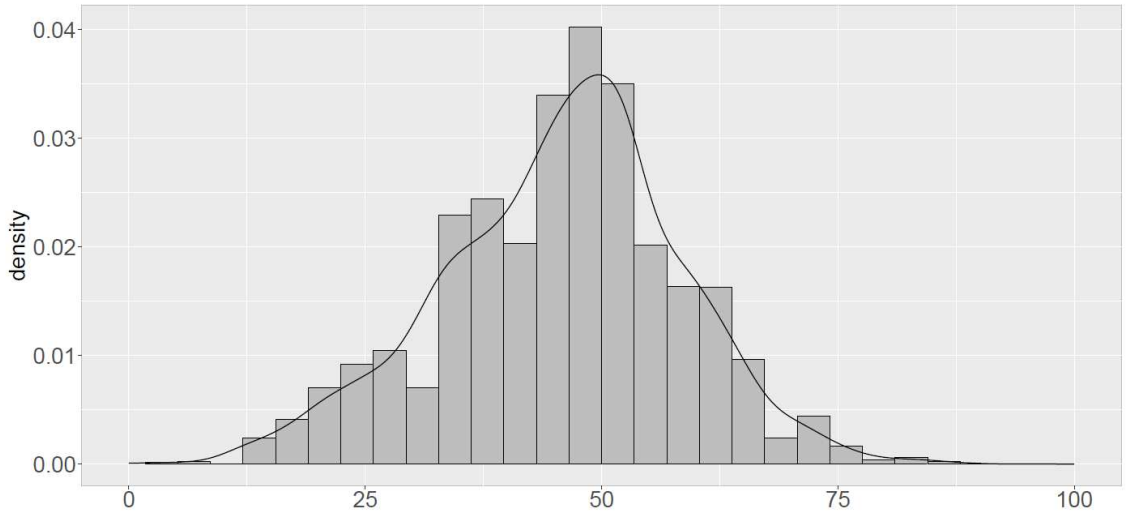
From the composite indicator, as illustrated in the previous Section, one can obtain an aggregate risk indicator at contractor level by simply averaging the composite indicator for the tenders each bidder has been awarded.

In the dataset under analysis, there are a total of 136,065 unique contractor identification codes. Based on the availability of data, the composite indicator can be computed for approximately 83% of the contractors. Table 7 provides the key summary statistics of the contractor indicator, while Figure 5 illustrates its distribution. Based on this indicator, contractors can be risk-ranked. As for awards, this ranking can be done for the whole group of contractors or distinguishing them in sub-groups depending, for instance, on geographical area and economic activity.

Table 7
Summary statistics of the mean composite risk indicator at the contractor level and contractor province
(percentage values, except absolute numbers in N)

Mean by	*Summary statistics*								
	Min	**1st Qu.**	**Median**	**3rd Qu.**	**Max.**	**Mean**	**S.d.**	**missing**	**N**
Contractor	0	37.5	47.0	53.3	100	45.7	12.6	16.7	136,065
Contractor province	38.3	45.3	48.2	50.0	55.8	47.6	3.0	-	107

Figure 5
Density plot of the mean composite risk indicator at the contractor level.



Taking into account the province where each contractor is based, one can obtain a geographical outlook of risk by computing the mean by province of the composite indicator of the previous Section.

[32] The CPV system is used in public procurement processes within the European Union to standardize the references for various categories of products, services, and works. It has a hierarchical structure based on nine digits that indicate divisions, groups, classes, and categories, allowing for precise identification of the contract's subject matter and avoiding ambiguities. The first two digits of the code represent the "divisions", values between 0 to 44 and 48 are associated with goods, codes 50 to 98 pertain to services, and code 45 represents works.

The bottom line in Table 7 above provides the key summary statistics for this indicator, while Figure 6 maps its distribution across Italian provinces, showing that there is a higher risk associated with procedures involving contractors based in the North, though the highest value is observed in a Central province, just south of Rome.

However, the evidence provided in Section 4 on missing data (see Figure 1 in particular) shows that contracting authorities located in Southern Italy are more prone to provide incomplete information to ANAC than those in the North; this suggests that the corruption risk of Southern provinces may be underestimated due to the paucity of data. Hence, the indications emerging from the map in Figure 6 have to be interpreted with extreme caution. In this respect, it is advisable to compare provinces within the same geographical area. To this end, Table 8 indicates the riskiest provinces separately for each macro area.

Figure 6
Contractor's risk indicator at the provincial level (percentage values)

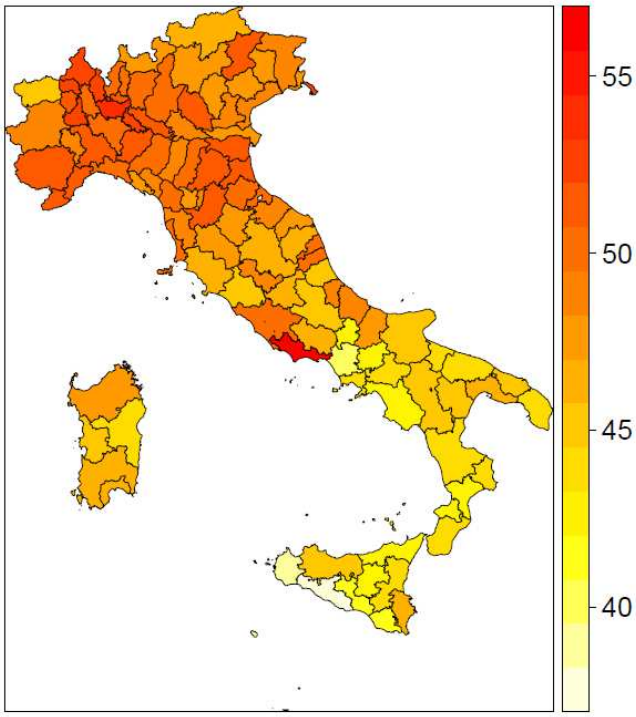


Table 8 – 5 riskiest contractor's provinces in each macro area

Rank	North	Center	South and Islands
1	Milano	Latina	Chieti
2	Verbano-Cusio-Ossola	Firenze	Pescara
3	Varese	Fermo	Campobasso
4	Lodi	Livorno	Sassari
5	Trieste	Roma	Taranto

6.3. Validation of the award risk indicators

For tenders with available information about the winning bidders, a validation exercise is conducted by matching the risk indicators at award level computed in Section 6.1 with a sample of Italian firms potentially connected to organized crime (OC), based on a mapping carried out by the Italian FIU

(UIF, 2021, pages 47-48) [33]. This analysis aims to support our results, thereby enhancing the reliability of the indicators, and sheds light on the relation between contractors' OC connections and corruption risk.

During the period of analysis, out of a total of 136,065 distinct contractors, 2,447 emerged as potentially OC-linked (1.8% of the total number of winning bidders); the latter won 70,676 awards (7.5% of the total number of awards with a known winning bidder). We computed the average of each individual risk indicator for tenders won by potentially OC-linked bidders and compared it to the average for tenders won by allegedly "healthy" contractors (see the left-hand columns of Table 9).

The difference between the two means is statistically tested using a two-pronged approach: 1) a logistic regression model where each risk indicator is the response variable and the regressor is a dummy variable indicating whether the winning firm is OC-linked; 2) the same logistic regression model estimated with the inclusion of additional control variables, in order to account for differences in contracting authority characteristics and tender features.[34] The coefficients of the dummy variable for OC-linked firms are displayed in the right-hand columns of Table 9 along with their significance levels, both with and without control variables.

Results show the following features of tenders won by potentially OC-linked contractors:

- a higher prevalence of discretionary procedures, particularly non open ones;
- more cases of awards failing to meet the legal criteria for non-open procedures;
- increased likelihood of using direct awards, even when the specified conditions for their application are not met;
- as for the adoption of the most economically advantageous offer (MEAT) criterion, results are ambiguous: on average its application appears to be slightly less frequent, whilst, after controlling for authority and tender characteristics, a higher likelihood of adopting such criterion emerges;
- higher incidence of single bids in tendering processes;
- in negotiated procedures, the number of invited bidders less frequently falls below the threshold of 5;

[33] The firms included in the sample have been selected as those whose directors and other corporate officers include i) persons of interest on the basis of information exchanges with the DNA, ii) persons investigated for mafia crimes who appear in business archives, or iii) persons named in information requests from the judicial authorities regarding organized crime.

[34] The covariates in the logistic regression are:

- characteristics of the tender: total value; year (from 2018 to 2023); sector (Ordinary or Special); object (goods, services or works);
- features of the contracting authority: type (19 macro-types, for details see Section 4 above); macro geographical area (North, Center, South and Islands);
- macro geographical area of the contractor (North, Center, South and Islands).

Table 9
Statistical comparison of individual risk indicators between awards won by OC-linked and "lawful" contractors

Response variable	Average value of risk indicator for awards won by *(percentage values)*		Logistic regression coefficients of OC-link status of the contractors				N
	"lawful" contractors	OC-linked contractors	without control variables		with control variables		
NoOpenAward	84.1	85.0	0.07	***	0.14	***	945,384
DirectAward	47.6	59.8	0.49	***	0.61	***	945,384
AnomalousNoOpen	19.8	25.3	0.34	***	0.56	***	945,354
AnomalousDirect	1.2	1.3	0.26	***	0.15	***	945,371
MEAT	82.1	81.0	-0.07	***	0.08	***	578,731
SingleBid	53.0	56.8	0.15	***	0.18	***	607,190
NPInvitedLow	14.6	11.7	-0.25	***	-0.27	***	922,413
SubRequest	23.0	33.1	0.50	***	0.49	***	945,384
ReWinner	70.7	76.0	0.27	***	0.23	***	945,384
MissingStart	46.8	47.2	0.01		-0.16	***	945,384
MissingEnd	75.7	76.8	0.07	***	-0.02	***	945,384
MissingFunds	24.7	32.3	0.37	***	0.08	***	945,384

Note: (*) p-value less than 0.05, (**) p-value less than 0.01, and (***) p-value less than 0.001.

- more frequent requests for subcontracting during the bidding phase;
- winning contractors with a higher probability of having previously won other tenders with the same contracting authority;
- lower rate of missing information on the beginning and the end of the contract; data regarding the funding sources typically less available.

Most of the results are consistent with what one could expect and signal that tenders awarded to potentially OC-linked contractors are more opaque, less competitive and more frequently entail the use of discretionary powers by the contracting authority, even when the conditions would not allow that.

We conduct a similar analysis for the composite indicator, comparing its average value for tenders won by potentially OC-linked bidders with that of tenders won by allegedly "healthy" contractors (Table 10). To assess whether the difference between the two means is statistically significant, we estimate a fractional logit model[35] and consider the significance and direction of the coefficient associated with the dummy variable indicating connections with OC. As for individual indicators, the fractional logit model is estimated with and without control variables[36]. Consistently with what emerged from the same analysis on individual indicators, results show that on average the composite indicator is significantly higher for tenders won by OC-linked bidders.

Table 10
Statistical comparison of the composite risk indicator between awards won by OC-linked and "lawful" contractors

Average value of composite risk indicator for awards won by *(percentage values)*		Fractional logit regression coefficient of OC-link status of the contractors				N
"Lawful" contractors	OC-linked contractors	without control variables		with control variables		
49.3	50.5	0.04	***	0.05	***	537,909

Note: (*) p-value less than 0.05, (**) p-value less than 0.01, and (***) p-value less than 0.001.

[35] Fractional regression is commonly used for modelling bounded continuous dependent variables. In our case, the response variable is the composite indicator that ranges between 0 and 1, making this approach more appropriate than the classic regression model.

[36] The control variables added in the fractional logit regression are the same as those described in footnote 29.

7. Risk indicators at contracting authority level

Having indicators at the contracting authority level is crucial for conducting an effective risk assessment of these entities. It provides insights into their behaviour and performance, as it allows for the identification of potential irregularities and inefficiencies and enables benchmarking and comparison between different authorities.

The indicators are obtained by aggregating each individual risk indicator defined in Section 5 at the contracting authority level. More specifically, for each contracting authority, we compute the frequency of awards undertaken by the authority that meet each indicator's conditions. Table 11 provides the details of how the indicators are built, whilst Table 12 displays their summary statistics.

Table 11
Risk indicators at the contracting authority level

Indicator	Numerator	Denominator
$NoOpenAward_{CA}$	N. non-open awards	N. awards
$DirectAward_{CA}$	N. direct awards	N. awards
$AnomalousNoOpen_{CA}$	N. anomalous non-open awards	N. open awards
$AnomalousDirect_{CA}$	N. anomalous direct awards	N. awards
$MEAT_{CA}$	N. awards with MEAT criterion	N. awards
$SingleBid_{CA}$	N. awards with a single bid	N. awards
$NPInvitedLow_{CA}$	N. negotiated procedures with less than 5 invited companies	N. negotiated procedures
$SubRequest_{CA}$	N. awards with a subcontracting request	N. awards
$ReWinner_{CA}$	N. awards with at least one recurrent winner	N. awards
$MissingStart_{CA}$	N. awards without communication of initial contract phase	N. awards
$MissingEnd_{CA}$	N. awards without communication of final contract phase	N. awards
$MissingFunds_{CA}$	N. awards without the communication of funding sources	N. awards

Table 12
Summary statistics of the risk indicators at the contracting authority level
(percentage values)

Indicator	*Summary statistics*							
	Min	1st Qu.	Median	3rd Qu.	Max.	Mean	S.d.	% missing
$NoOpenAward_{CA}$	0	90.0	100	100	100	90.7	18.9	0
$DirectAward_{CA}$	0	38.3	66.7	94.7	100	61.4	33.6	0
$AnomalousNoOpen_{CA}$	0	0	0	1.6	100	3.3	10.8	0
$AnomalousDirect_{CA}$	0	0	0	3.6	100	5.5	15.1	11.3
$MEAT_{CA}$	0	75.0	91.6	100	100	82.4	25.0	39.5
$SingleBid_{CA}$	0	40.0	62.7	90.0	100	60.6	32.0	34.3
$NPInvitedLow_{CA}$	0	0	33.3	69.2	100	40.8	36.5	53.6
$SubRequest_{CA}$	0	0	15.4	40.0	100	24.2	27.7	31.9
$ReWinner_{CA}$	0	0	20.0	44.4	100	25.5	27.5	31.9
$MissingStart_{CA}$	0	72.2	95.2	100	100	82.7	24.5	0
$MissingEnd_{CA}$	0	87.5	100	100	100	90.6	17.3	0
$MissingFunds_{CA}$	0	51.4	80.0	100	100	72.3	29.7	0

Aggregation of the individual indicators is carried out using the following formula:

$$CIAuthority = (NoOpenAward_{CA} + ForDirectAward_{CA} + AnomalousNoOpen_{CA} + Meat_{CA} + SingleBid_{CA} + NPInvitedLow_{CA} + SubRequest_{CA} + ReWinner_{CA} + Missing_{CA})/9 \quad [2]$$

where:

$$ForDirectAward_{CA} = (DirectAward_{CA} + AnomalousDirect_{CA})/2$$

$$Missing_{CA} = (MissingStart_{CA} + MissingEnd_{CA} + MissingFunds_{CA})/3$$

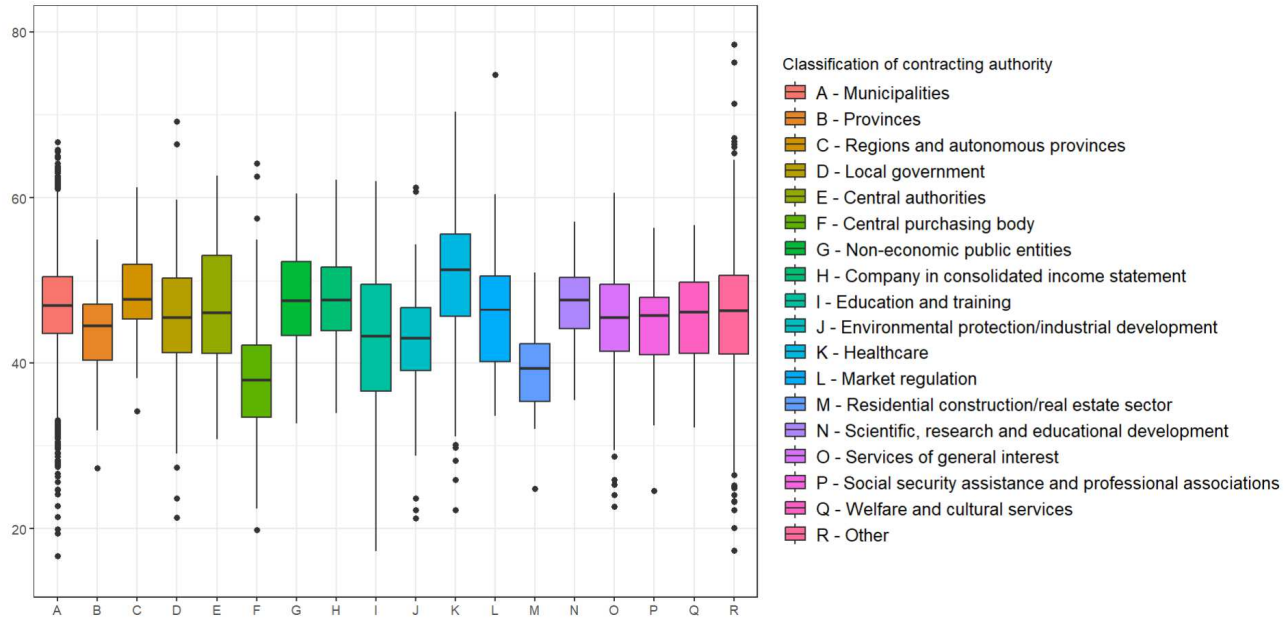
In the period under analysis the datasets include 24,814 distinct contracting authorities. However, the composite indicator can be computed only for 10,699 of them, due to the presence of missing data.

The indicator can thus be used to risk-rank the contracting authorities and identify those with the highest level of corruption risk. However, since the pattern of missing communications is strongly dependent on the type of authority, its size and location, we have classified contracting authorities according to their type and ranked them within their peer group.

Figure 7 displays boxplots of the composite indicator by type of contracting authority, showing that authorities in the "Health sector" exhibit, on average, the highest value for the composite indicator, whereas central purchasing bodies have the lowest value. However, the diagram also identifies the authorities within each group featuring the highest risk levels (indicated by the black spots) compared to their peers. For example, municipalities tend to have the largest number of authorities with extremely high values of the composite indicator. Furthermore, the group of central purchasing bodies shows a distribution of the indicator's value that is significantly lower than the other groups, likely because these authorities implement better methods and have a more efficient organization.

Figure 7

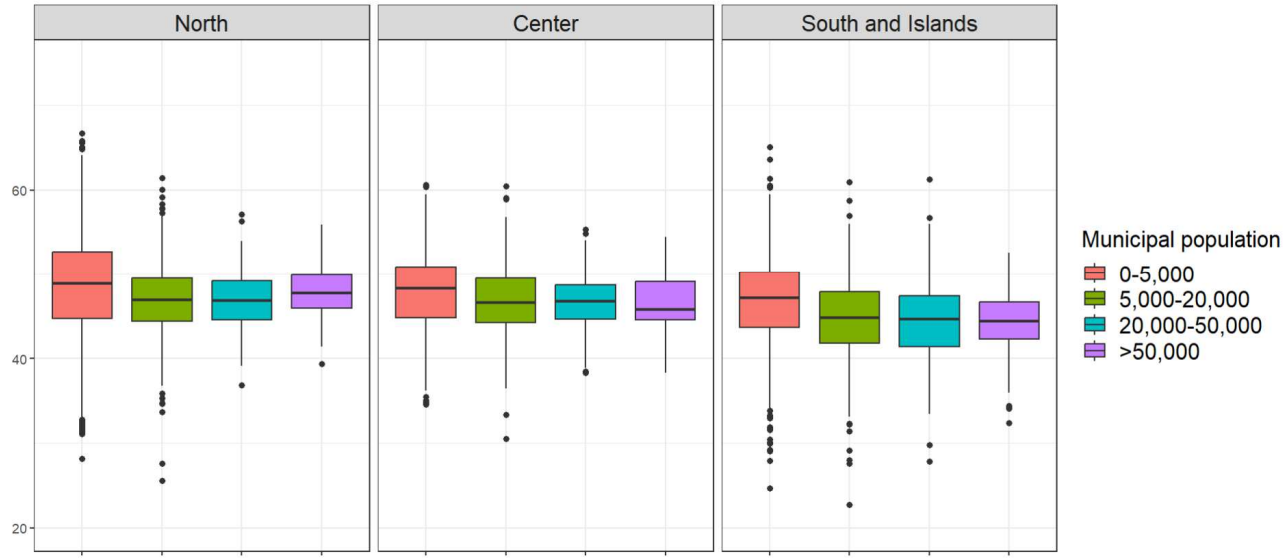
Boxplots of the composite risk indicator by type of the contracting authority



Clearly, considering the type of authority alone is not sufficient to eliminate potential biases arising from missing data. For this reason, it is important to jointly consider the type of authority, its size, and its location, and identify peer groups based on these three dimensions simultaneously. To this end, for example, Figure 8 displays boxplots of the composite risk indicator for municipalities only, categorized by geographic area and size, proxied by the number of residents. The diagram allows to risk-rank the municipalities and to appropriately assess the relative risk of each authority within its own peer

group. Regardless of the macro area, smaller municipalities appear to feature a higher risk, while the other dimensional classes do not seem to differ significantly.

Figure 8
Boxplots of the composite risk indicator for municipalities by geographic areas and population (percentage values)



In order to provide a geographical distribution of contracting authorities' risk, we compute an aggregate risk indicator at the provincial level by applying a methodology similar to the one explained at the beginning of this Section. For each province, where each contracting authority is headquartered, we compute the frequency of awards that meet the indicators' conditions and then aggregate using [2]. The indicator thus obtained is plotted in Figure 9, which shows that the Northern provinces (most notably in the North-West) exhibit a higher risk. However, as previously mentioned, missing data are significantly influenced by the geographic location of the authorities, particularly at macro-area level. Therefore, provinces can be appropriately compared only within the same area: Table 13 shows the riskiest provinces for each area.

Figure 9
Composite risk indicator at the province of the contracting authority (percentage values)

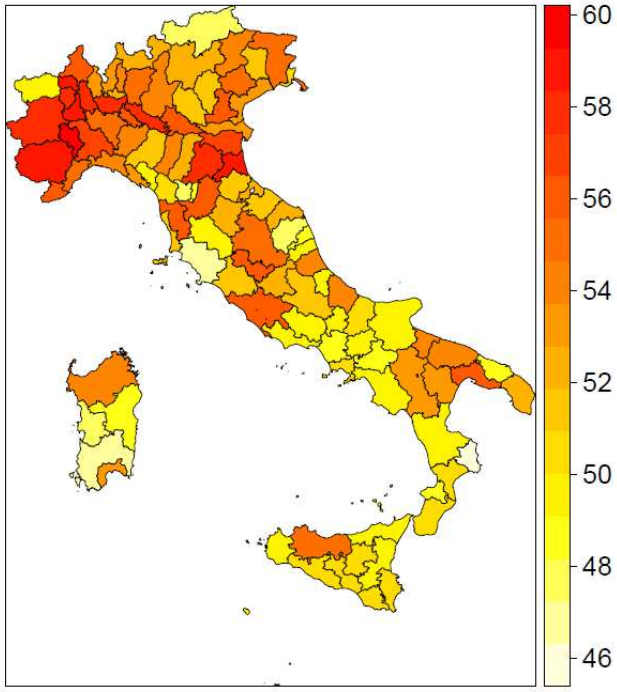


Table 13
Top-five contracting authorities' province according to the risk indicator by macro area

Rank	North	Center	South and Islands
1	Asti	Terni	Taranto
2	Vercelli	Pisa	Palermo
3	Ravenna	Roma	Bari
4	Cuneo	Firenze	Teramo
5	Bologna	Perugia	Sassari

8. Potential operational applications

The operational use of the risk indicators in public procurement proposed here spans across various governmental and institutional domains.

Primarily, they may be employed in monitoring public procurement, aiding in combating and preventing corruption by spelling out high-risk tenders. The assignment of a risk score to each contract enables the prioritization and targeting of tenders that necessitate closer scrutiny and oversight. By employing risk indicators, government agencies can proactively identify potential vulnerabilities within the procurement process and implement appropriate measures to mitigate the risk of malfeasance.

Additionally, risk indicators can serve as a tool for decision-making on resource allocation and implementing control mechanisms. By providing an extensive understanding of varying levels of contractual risks, the indicators could guide regulators in establishing where to deploy their limited resources for conducting audits of contracting authorities and winning contractors. This may help authorities allocate resources more effectively by setting tailored control measures to prevent misconduct.

Moreover, contracting authority risk indicators can have numerous applications for evaluating processes, procedures, and practices within each authority in order to identify potential weaknesses or

areas at risk of corruption, thereby enhancing internal control, risk management and audit in the public sector. By tracking each authority's indicator over time, it becomes possible to assess performance and measure improvements or deteriorations in procurement practices.

The indicators might also facilitate the prioritization and optimization of investigative efforts by law enforcement agencies, enhancing their ability to allocate resources effectively and focus on the most critical areas. This tool may prove particularly valuable within the framework of the PNRR, where substantial resources allocated to public tenders pose a tangible risk of attracting organized crime interests.

The indicator holds several potential applications for anti-money laundering purposes as well. The Italian FIU could leverage this indicator (or some refinements) in order to enrich the informational set that support its institutional functions. This may help to enable the UIF to focus intelligence resources on transactions potentially associated with corrupt practices or on more opaque contractors, thereby supporting efforts to combat money laundering stemming from illicit activities in the public procurement sector.

Furthermore, the indicator could be employed for strategic anti-money laundering analysis and to support the country's National Risk Assessment. This facilitates the identification of typologies of awarding authorities, types of goods and services, and geographical areas with a higher risk of corruption, allowing for more targeted allocation of investigative powers, as well as the development of specific anti-money laundering policies and strategies to mitigate these risks.

Finally, the results of the indicators analysis can be disseminated publicly to promote transparency and accountability in the public procurement sector, benefiting both the general public and citizenship.

9. Conclusions and future analysis

This study proposes a set of indicators for measuring corruption risk in public procurement. Given the substantial resources deployed in this sector, amounting to up to a fifth of a country's GPD, and its high vulnerability to misconduct and embezzlement, efforts to define red flags, early-warning indicators, and risk measures to prevent them are widely pursued both at national and international levels.

Extensive literature suggests relevant features that should be considered in constructing such indicators. One important feature is the discretion that regulations allow contracting authorities to use when awarding certain contracts, as long as these contracts meet specific criteria regarding value, type, or activity sector. Also the degree of competition characterizing the contract awarding procedure comes to the forefront.

An additional relevant facet is transparency. Several worldwide initiatives have promoted the establishment of publicly accessible procurement data repositories which enable governmental audit bodies, interested parties and the citizenship as a whole to exert some control on how procurement auctions are carried out. For this very reason, ANAC, the Italian central anti-corruption authority, has established since 2010 the BDNCP, an open database containing an extensive set of data on procurement contracts.

By exploiting this set of data with reference to the period from January 2018 to June 2023, including over 2,1 million contracts worth nearly 2,9 billion euros, and drawing on the relevant literature, this study builds a multi-layered framework of indicators for the measurement of corruption in public tenders.

We define a set of 12 individual indicators, each addressing specific aspects of awards that may be linked to irregularities and misconduct. A particular attention is devoted to missing data that contracting authorities fail to provide to ANAC, since specific patterns across geographical, dimensional and sectoral features emerge in authorities' compliance with the mandatory requirement to report. The

individual indicators are used to compute a composite award-specific risk indicator, measuring the overall potential vulnerability of each single contract to corruptive practices. By exploiting a dataset of firms potentially infiltrated by (or colluded with) OC, through an appropriate statistical procedure we test the robustness of both the individual and the award composite indicators, obtaining some evidence on their reliability and significance. The results show that tenders awarded to OC-linked contractors are more opaque, less competitive and more frequently entail the use of discretionary powers by contracting authorities, even under conditions where such practices are not warranted.

Alternative aggregation methods produce composite indicators for assessing risk associated with contractors, contracting authorities, and geographical locations. An additional outcome of the analysis is a proposed classification of contracting authorities, facilitating risk comparisons among similar entities.

The results of this work can be further developed along different lines of research. Correlations between the proposed composite indicators and others available in the literature could be examined to verify their coherence and further test their reliability. A more in-depth analysis by sector, type of tenders, economic sector, and type of contracting authority could be developed to garner more insights into the risk profile of each dimension. Enhanced methodologies could strengthen the indicators' precision and robustness.[37]

Despite potential improvements that can be made to the methodology applied in this study, the end result of the analysis — the set of indicators applicable at various relevant levels — provides a ready-to-use framework that might be deployed for various ends, from the auditing of contracting authorities, their profiling and the identification of potential inefficiencies in their management of public contracts, to possibly the actual prevention and repression of corruption, for the benefit of investigative bodies.

[37] One could assess the feasibility of using advanced statistical methods, such as the Heckman model, to adjust for non-random sample selection, thereby providing more accurate and unbiased estimates of corruption.

References

Abdou. A.. Á. Czibik. B. Tóth. and M. Fazekas. (2021). COVID-19 emergency public procurement in Romania: Corruption risks and market behavior. Government Transparency Institute. GTI-WP/2021:03.

ANAC (2017). "Analisi istruttoria per l'individuazione di indicatori di rischio corruzione e di prevenzione e contrasto nelle amministrazioni pubbliche coinvolte nella politica di coesione" - Rapporti e studi – Anticorruzione e trasparenza.

ANAC (2021). "Misurazione del rischio di corruzione a livello territoriale e promozione della trasparenza". - Programma Operativo Nazionale Governance e Capacità istituzionale 2014-2020 - https://www.anticorruzione.it/il-progetto

ANAC (2023). "Banca Dati Nazionale Contratti Pubblici: Manuale di gestione operativa degli OPEN DATA". Versione 1.3 – Autorità Nazionale Anticorruzione – Ufficio programmazione e sviluppo delle banche dati.

Andersson. S. and P. M. Heywood (2009). "The politics of perception: use and abuse of transparency international's approach to measuring corruption". Political studies 57(4). 746–767.

Baltrunaite. A. C. Giorgiantonio. S. Mocetti. and T. Orlando (2021). "Discretion and supplier selection in public procurement". The Journal of Law. Economics. and Organization 37(1). 134–166.

Carbone C. Calderoni. F. and Jofre. M. (2024). Bid rigging in public procurement: cartel strategies and bidding patterns. Crime, Law and Social Change, 1-33.

Cantone R. and Caringella F. (2017). *La corruzione spuzza*. Mondadori.

Coviello. D. A. Guglielmo. and G. Spagnolo (2018). "The effect of discretion on procurement performance". Management Science 64(2). 715–738.

Decarolis. F. and C. Giorgiantonio (2015). "Local public procurement regulations: the case of Italy". International Review of Law and Economics 43. 209–226.

Decarolis. F. C. Giorgiantonio and V. Giovanniello (2011). "The awarding of public works in Italy: an analysis of the mechanisms for the selection of contractors." Mercato Concorrenza Regole. 13(2): 235-272.

Decarolis. F. R. Fisman. P. Pinotti. and S. Vannutelli (2021). "Rules. discretion. and corruption in procurement: evidence from Italian government contracting". Technical report. National Bureau of Economic Research.

Decarolis. F. and C. Giorgiantonio (2022). "Corruption red flags in public procurement: new evidence from Italian calls for tenders". EPJ Data Science 11(1). 16.

Del Sarto. S. P. Coppola. and M. Troìa (2021). "Clusters of contracting authorities over time: an analysis of their behaviour based on procurement red flags". Book of short paper SIS 2021 – Pisa. 21 -25 giugno 2021 p. 205.

European Commission (2021). "Commission notice: Guidance on the avoidance and management of conflicts of interest under the Financial Regulation" (2021/C 121/01). Official Journal of the European Union. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021XC0409(01)

Fazekas. M. I. J. Tóth and L.P. King (2016). "An Objective Corruption Risk Index Using Public Procurement Data". European Journal on Criminal Policy and Research 22. 369–397.

Ferraris. V. C. Mazza and L. Scomparin. eds. (2016). "Warning on Crime. Preventing and combatting crime in public procurement". (www.warningoncrime.eu) Torino: Università degli Studi di Torino.

Ferwerda. J. Deleanu. I. and Unger. B. (2017). "Corruption in Public Procurement: Finding the Right Indicators". European Journal on Criminal Policy and Research 23. 245–267.

Gnaldi. M. (2018). "Indicatori di corruzione e nuovi indicatori di prevenzione della corruzione" (in M. Gnaldi and B. Ponti. "Misurare la corruzione oggi: obiettivi. metodi. esperienze". FrancoAngeli)

Kenny. C. and M. Musatova (2010). "Red flags of corruption in World Bank projects: an analysis of infrastructure contracts". The World Bank, Policy Research Working Paper.

Marcolongo. G. (2023). "Organized crime. earthquakes and local public procurement: evidence from Italy". Technical report. Working paper.

Mirenda. L. S. Mocetti. and L. Rizzica (2022). "The economic effects of mafia: firm level evidence". American Economic Review 112(8). 2748–73.

OECD (2008). "Handbook on constructing composite indicators: methodology and user guide". Joint Research Centre-European Commission.

OLAF (2013). "Procurement: costs we pay for corruption. Identifying and Reducing Corruption in Public Procurement in the EU". European Anti-Fraud Office - PwC EU Services.

OLAF (2017). "Fraud in Public Procurement. A collection of Red Flags and Best Practices". European Anti-Fraud Office - Fraud Prevention. Reporting and Analysis

Troìa. M. (2020). "Data analysis e costruzione di indicatori di rischio di corruzione per la banca dati nazionale dei contratti pubblici". Autorità Nazionale Anticorruzione ANAC Working paper (5).

UIF (2021). "Rapporto Annuale 2020 – Unità di Informazione Finanziaria per l'Italia – Anno 2020". UIF N. 13 – 2021. https://uif.bancaditalia.it/pubblicazioni/rapporto-annuale/2021/index.html

UNODC (2013). Good practices in ensuring compliance with article 9 of the United Nations Convention against Corruption. Vienna.

Wachs. J. Fazekas. M. and Kertész. J. (2021). Corruption risk in contracting markets: a network science perspective. International Journal of Data Science and Analytics 12. 45–60.

World Bank. (2013). "Fraud and corruption awareness handbook: a handbook for civil servants involved in public procurement".