Government Transparency Institute

Mihály Fazekas [1], *Bence Tóth* [2], *Johannes Wachs* [3]

Public procurement cartels: A large-sample testing of screens using machine learning[1]

Working Paper series: GTI-WP/2023:02

April 2023, Budapest, Hungary

[1] Central European University and Government Transparency Institute: corresponding author, mfazekas@govtransparency.eu

[2] University College London and Government Transparency Institute

[3] Corvinus University of Budapest and Centre for Economic and Regional Studies

[1] The work was generously supported by the Swedish Competition Authority (grant nr. KKV 72/2018-Ab-1) and the World Bank (Bulgaria - Country Economic Memorandum. 2023). The authors are immensely grateful to contributors and research assistants in the research project, in particular Isabelle Adam, Peter Horn, Yuliia Kazmina and Zsombor Teremy.

Abstract

Cartels in public procurement impose high costs on public budgets. Precisely measuring them has a prominent policy and academic importance. The literature so far used data which is not widely available, aimed to identify specific behaviours in isolation, and considered few cases to generalise from. By implication, it has not produced comprehensive and generalisable knowledge able to support policy. We address these gaps in the literature by simultaneously measuring multiple cartel behaviours, drawing on data for 78 cartels from 7 countries during 2004-2021. We apply state-of-the-art machine learning methods to combine diverse cartel screens in a predictive model. As expected, no single indicator or group of indicators can predict a wide set of cartel behaviours. Combining many indicators in a random forest algorithm achieves 77-91% prediction accuracy across countries. Most individual cartel screens contribute to prediction in line with theory. Policy implications are profound, offering to improve cartel investigations and policy making.

JEL codes: C21; C45; C52 ; D22; D40; K42; L41

1 Introduction

Public procurement, that is governments buying goods and services, account for about 12 percent of global GDP or 11 trillion USD per year (Bosio et al., 2020). However, anticompetitive behaviours such as collusion among bidding firms represent a major problem, as collusion is more likely to arise and operate for a longer period in public procurement than in traditional markets (World Bank, 2009, 2011). The extra costs of collusion are borne directly by the public budget, hence the public. Given the large volumes of spending, even a small percentage increase in prices translates into substantial budgetary implications and welfare losses. Hence, there is a great need for accurate detection techniques supporting effective investigations and prevention.

Massive amounts of data are generated by public procurement transactions which present competition authorities with great opportunities. It is likely that groups of firms engaging in collusion leave traces of their activity in the data. These signals offer investigators a novel source of valuable leads. However, combing through vast amounts of data is time consuming and difficult. It is made even harder by widespread data quality problems, including missing records and values, as well as incorrect data. Recognising the opportunities offered by large-scale public procurement data, an academic literature has grown proposing risk indicators which signal potential cartel behaviour, i.e. cartel screens. Competition authorities around the world have also launched exiting projects aiming to capitalize on these advances in at least 3 main ways: 1) mapping weaknesses in the competitive environment which calls for policy intervention; 2) identifying new investigative leads or ranking them; and 3) improving investigations' targeting and, in rare cases, offering relevant evidence for courts.

Even though, several data-driven initiatives by competition authorities have sprung up, they have produced considerable disappointment. Lack of initial success is due to a range of reasons such as large start-up costs for building a reliable data pipeline and training staff. Among the many such challenges, we find those particularly problematic which have their roots in the state of the academic literature. First, most of the times, the datasets used by the academic literature include variables not widely available and tend to be extremely high-quality compared to standard large-scale public procurement datasets. For example, academic studies include independently sourced cost estimates to quantify markups (Abrantes-Metz et al, 2006) or have low missing rates on key variables such as the cartel screens. Second, most academic research aims at identifying one specific cartel behaviour in isolation from other behaviours. For example, some studies look at cartels using bid rotation to rig a series of tenders without being noticed (Kawai et al, 2022). This focus produces screens with strong internal validity, but external validity remains problematic in the presence of unknown and diverse types of cartel behaviours as different cartel strategies give opposing signals. Third, most of the literature so far considered too few cartel cases to serve as a basis for generalizing to many cartel types, markets and countries. To the best of our knowledge, the most comprehensive study considered 1 cartel strategy, 1 sector, 2 countries and 5 proven cases which represents the approach closest to ours (Huber et al, 2022, Huber & Imhof, 2019). Finally, a number of studies apply collusion risk indicators (or screens) knowing the exact products of the collusive market with precise market boundaries. Sometimes even the goods procured are homogeneous and comparable whose prices can be modelled with high accuracy (e.g. Ohio school milk contracts in Porter & Zona (1999)). However, for large-scale cartel risk estimation, precisely estimating market boundaries necessary for applying a host of indicators such as market concentration is either imprecise or impractical due to resource needs.

Recognizing the major advances reached by the literature so far while also considering the above gaps, **we develop a general model for cartel detection** and validity test a wide range of cartel risk indicators. We hope to advance both the academic debate and supporting policy applications. This article sports such ambitions by addressing each of the 4 limitations listed above. We use readily available large-scale public procurement datasets accessible in public repositories and government websites without laborious manual data collection and correction. We combine a wide variety of cartel risk indicators used in the literature which indicate many different cartel behaviours. Moreover, our learning algorithms draw on a large number of proven cartel cases from multiple markets in multiple countries. This allows us to better understand external validity of individual indicators as well as their combinations, that is establish in which cases they are accurate and when they are not. Finally, we only make use of screens which are defined on the contract or company levels, hence avoiding the pitfalls of having to define market boundaries a priori.

We evaluate the validity and accuracy of many cartel screens to identify potential cartels in large-scale public procurement data from Bulgaria, France, Hungary, Latvia, Portugal, Spain and Sweden from the 2007-2020 period. We combined this administrative data with judicial records on proven cartels from official government sources. We test individual cartel screens in simple bivariate comparisons of proven cartel contracts versus likely non-cartel contracts (same market after the cartel has been busted). Moreover, we also compare regression and random forest algorithms to find the model with highest predictive accuracy. The best model is selected based on prediction accuracy on the test set (i.e. random sample of proven cartel and non-cartel cases not used for estimating the model).

We find that no single indicator or group of indicators performs well when tested against a variety of proven cases approximating a real-life problem of competition authorities. In line with our theory, different screens are better suited to detect specific types of cartels, while failing to indicate others. On the one hand, no one screen can accurately detect even the half of our cartels. On the other hand, each indicator is valuable to detect at least some cartels which underlines our assessment of high internal but low external validity for most screens addressed in the literature. Crucially for our claim about a general cartel prediction model, when a battery of screens is considered, their combined accuracy is very high across all proven cases. Our best random forest model achieves 77-91% prediction accuracy across all countries, markets, and proven cases. For most individual indicators, we identify an impact function consistent with theory in spite of the high degrees of flexibility and non-linear relationships in random forest models.

Our contributions are both academic and policy. We re-confirm the validity of a wide set of cartel screens put forward by the literature, while we also establish the limits of single-indicator, single-cartel type approaches. When cartels rig tenders in diverse ways, sometimes increasing the number of bidders, sometimes decreasing it, combining different indicators in a non-linear model produces the most accurate estimation. Another important finding of our research is that public procurement data quality in limits prediction accuracy even in high income countries with strong track record of e-government reforms in Europe. Moreover, several key variables for cartel screening are not collected systematically across a wide set of countries (e.g. information of bid prices of losing bidders). Nevertheless, we also contribute to policy discussions confirming that large-scale cartel risk estimation is possible with reasonable precision using readily available datasets and analytical tools. We demonstrate this point by predicting the risk of collusion to the more than 3.3 million observed public procurement contracts in the 7 countries in question. This should, hopefully, give further impetus to competition authorities considering the implementation of cartel risk measurement tools as part of their investigation support repertoire.

The rest of the article is organized as the following: first, we set out the conceptual framework discussing multiple cartel types. Second, we spell out our empirical strategy, including the data used, the cartel risk indicators tested, and the methodology applied for indicator testing and prediction. Third, we enumerate our main findings from the a) bivariate testing and b) predictive modelling, comparing different models. Fourth, we show the prediction of cartel risk scores for complete procurement markets and discuss policy implications. Finally, we discuss our findings and suggest future improvements to our approach.

2 Conceptual framework

Economists have long been interested in detecting signals of collusion in markets using data on firms, bids, and prices. An early example is the case of the Oklahoma asphalt market in the late 1950s and early 1960s, in which firms submitted identical bids (Funderbuck 1974). Cases of bid-rigging and market allocation in various US school milk markets in the 1980s and 1990s spawned a series of papers contrasting behavior and prices under collusion and competition (Hewitt, McClave & Sibley 1996, Lanzillotti 1996, Porter & Zona 1999, Pesendorfer 2000, Scott 2000). Localized markets for road construction and repair (Porter & Zona 1999, Bajari & Ye 2003, Ishii 2008) and global markets for highly specific chemicals such as lysine (Connor 2001, Evenett, Levenstein & Suslow 2001) have also been examined from this perspective. Harrington (2006) provides an excellent review of such studies.

Many of these works can present convincing evidence of anti-competitive behaviour because prices and data in these markets are highly comparable. For instance, it is possible to model the price of milk per litre using standardized raw milk prices and transportation costs. With such a model, unreasonable prices stick out. In many of these cases a whistleblower also provided information on how the collusive scheme worked.

Nevertheless, competition authorities are interested in casting a wider net and detecting cartels working in more heterogeneous markets. The scope and scale of markets studied in papers on collusion has grown in recent years (Kawai & Nakabayashi 2022, Chassang & Ortner 2015, Conley & Decarolis 2016), though many of these works exploit nuances of particular auction formats (i.e. average bid auctions, constrained bids) to highlight suspicious patterns suggestive of collusion. Reliable price and bid data are also essential to these approaches. Fine-grained data also enables the application of novel methods from network science (Morselli & Ouellet 2018, Wachs & Kertesz 2019) and machine learning (Vadász et al. 2016, Schwalbe 2018, Huber & Imhof 2019).

We argue that all of these approaches, though valuable, may have a tendency to overfit their methods to specific cases. A cornerstone of modern machine learning practice is the evaluation of predictive algorithms on unseen data. Perhaps owing to the rarity of clean data on proven cartel cases, few research papers apply cartel screening methods to multiple examples (for a recent exception see Huber et al, 2022). We see an opportunity to widen the scope of cartel screening to large, heterogeneous markets with varying data quality by ensembling, or combining, multiple cartel screens (extending earlier attempts by Toth et al. 2014). We borrow the term ensembling from the machine learning literature, which has long recognized that combining weaker predictive signals can produce a much stronger predictive model (Breiman 2001). Here, we examine whether such an approach can overcome the challenges of noisy data and heterogeneous markets to produce effective methods of detecting cartels and anti-competitive behaviour. This section introduces our conceptual framework cataloguing diverse collusion types which, in turn, can be used to guide measurement.

2.1 Collusive strategies

Collusion in public procurement aims to coordinate companies' decisions regarding price, quantity, quality or geographical presence to eliminate competition. This strategy can be only sustained if a) companies can coordinate; it is b) internally (credible punishment system, effective detection of cheating), c) externally sustainable (ability to exclude new market entrants); and d) the scheme can go undetected (i.e. no fines).

We follow the categorization of procurement collusion schemes introduced in Tóth et al. (2015) and Fazekas and Tóth (2016) which map the most important choices faced by colluding firms. We identify three dimensions: a) elementary collusion techniques, b) forms of rent-sharing, and c) resulting market structure. Elementary collusion techniques describe companies' bidding behaviour that ensures that contracts are won by the agreed supplier. These are a) withheld bids, b) non-competitive bidding, and c) joint bidding. First, companies can withhold their bids, to put less competitive pressure on the other companies and eventually raise contract prices. Second, companies can mimic competition by either submitting deliberately losing bids at inflated bid prices or erroneous bids.[2] This is considered to be the most common form of public procurement collusion by expert practitioners (OECD, 2014). Third, companies can submit a joint bid, that can be a sign of a special collusion scheme that also establishes the method of rent allocation.

The second dimension of collusion schemes is their rent allocation mechanism. This choice is strongly influenced by whether companies are active or passive participants of public tenders (Pesendorfer, 2000). Rent allocation is straightforward for active members of a scheme. For example, a consortium can easily formalize rent allocation through their contracts. Companies can also agree to allocate geographical markets or to win contracts cyclically, which makes rent allocation straightforward. However, if companies are not participating in tenders directly, rent reallocation has to happen in alternative ways. For example, they might subcontract each other or give each other informal side-payments.

The third dimension is the market structure that follows from the various collusive strategies. First, coordination can lead to highly concentrated market structures. For example, if collusion involves splitting the markets by geographic or product markets, then companies will end up with a very high share of contracts at a regional or sub-sectoral level.[3] Coordination can also lead to high market shares when passive participants get paid in alternative ways (Levenstein & Suslow, 2006; Pesendorfer, 2000) – few companies winning all contracts whereas smaller ones 'get paid' through sub-contracts or side-payments. Second, prior research also suggests that colluding suppliers can effectively imitate competitive market structure (Athey, Bagwell, & Sanchirico, 2004; Mena-Labarthe, 2012; Pesendorfer, 2000; World Bank, 2011). The cartel uses time to evade competition, with individual companies deferring profits and waiting their turn. Companies winning cyclically will not face competitive pressure and their market share will not show any timely changes.

Each combination of a) elementary collusion techniques, b) rent allocation mechanisms, and the c) resulting market structures forms a distinct collusion strategy (Table 2.1).[4] As strategies vary by these measurable dimensions, we can combine (group) indicators by these theoretical

[2] Public tenders very often award companies based on a combination of price and quality. Therefore, losing bids might just offer significantly lower quality at the same price as the winner supplier agreed by the collusive scheme in practice.

[3] Note, that splitting geographical- or product- submarkets will not have detectable signs in the 'higher-level' market shares, that would be relevant in a competitive set-up. For example, a road construction market of a country with two big regions and several companies will look competitive if we look at market shares at the country level. If they start to collude and split the contracts so that half of the companies win all contracts from one region and the other half of the companies from the other, then we would observe an increase in market concentration in the regional sub-markets. However, when looking at the country-level picture, the market shares would be unchanged.

[4] Note, that not every combination is conceptually meaningful, while some dimensions are not possible to measure with indicators based on public procurement or company data (e.g. informal side-payments are hard to observe).

scenarios. Note, that whereas strategies even within the same cartel can change, many contract level dimensions are exclusionary. For example, we cannot observe single-bidding and extreme bid price ranges at the same time.

As an example, we discuss strategy B. The main features of this strategy is that companies submit losing bids (or they might withdraw them or submit false bids), while they share rents through subcontracts, which leads to a concentrated market structure.[5] First, there is no clear theoretical expectation on the number of submitted bids or probability of single bidding. Second, as many of the bids have to be losing bids, we expect either a) the number of withdrawn bids or faulty bids to increase, or b) an extreme distribution of bid prices. Bid prices might be both very closely aligned together or dispersed. Third, other traces of coordinated bidding that are harder to be found in an automated way - such as identical mistakes or having the same author of the bidding documents - are also expected to occur. Fourth, if subcontracting is indeed the dominant rent-reallocation mechanism, then public procurement data might have traces of it in terms of increased probability of subcontracted contracts.[6] Alternatively, these contractual or informal relationships are outside the procurement domain, hence traces of exchanges should come from alternative sources. Fifth, procurement spending should become concentrated, a few companies should have high market shares.

TABLE 2.1. MAIN CHARACTERISTICS OF COLLUSION TYPES AND THE AVAILABILITY OF INDICATORS

Resulting market structure	**Elementary collusion technique**	**Form of rent sharing**			
		Sub-contractor	Consortia/ joint ownership	Coordinated bidding	Informal side-payments
Concentrated market structure	Withheld bids	A			
	Losing bids	B			
	Joint bids		C		
Stable market structure	Withheld bids	D		F	
	Losing bids	E		G	
	Joint bids				

Notes: every dimension is measured, some dimensions are measured, conceptually non-existent type

Source: Fazekas and Toth (2016)

[5] Note that we would see concentrated market structure based on the share of public contracts won and not necessarily based on the turnover of the participating companies due to subcontracting.
[6] Most procurement systems collect information on whether a particular supplier won a contract with explicit mention of subcontracting parts of it.

3 Empirical strategy

3.1 Data

In this section we give a bird's eye view of the data we collected for this report. We give a high-level overview of the procurement data we used, outline our search for data on cartel cases, explain how we matched proven cartel cases with public contracting datasets, and present summary statistics by countries that show the number of matching contracts used for the bivariate and multivariate models. We provide a more detailed description on procurement data, cartel case collection and data linking in the Appendix.

Selection based on procurement data

As a first step, we screened several countries with sufficient procurement data quality for the analysis. We used two main criteria for shortlisting: availability and quality of public procurement data (especially the availability of bidder information) and the number of cartel cases that overlap with procurement data.

Based on this screening process and additional data quality checks - e.g. availability of long enough historical data, availability of key variables[7], we shortlisted seven countries for the current report, namely: Bulgaria, France, Hungary, Latvia, Portugal, Spain, and Sweden[8].

We use contract-level public contracting data to test the calculable collusion indicators. For all analysed countries - except Sweden - we use data collected by DIGIWHIST[9]. The DIGIWHIST project collects contract-level data on European public procurement contracts covering both above and below-EU-threshold contracts. It contains data on tender level information, such as key dates (call for tender publication, bidding deadlines, award date etc.), procedure type, product market, regions, estimated contract prices; information on buyers (name and address); and bids (such as company names, contract prices). We have implemented several data quality improvements - for example, we have improved variables that are key for the analysis, such as fixing missing company names, contract values and also connecting the same organizations through matching and filtered irrelevant data, such as direct contracts where coordination cannot (or hardly can) take place.

Cartel case collection

We collected information on the proven cartel cases[10] manually from country specific sources of court rulings by following three search strategies: a) searching competition authorities' online repositories of proven cases, b) screening competition authorities and courts annual reports, c) contacting competition authorities of selected counties and requesting and overview of proven cases (see Appendix for more details).

---

[7] For example, we analysed the availability of company names, bidder numbers, dates, that are all key for the analysis.
[8] Note that we identified a couple of key data errors in Lithuanian public procurement data that we plan to fix for an updated report.
[9] The data is published on http://opentender.eu/. For a technical explanation of the database building, see: https://github.com/digiwhist/wp2_documents/blob/master/d2_8.pdf. Note, we use data provided by Visma Opic for analyzing the Swedish proven cartel cases, that was used by the authors in Fazekas and Tóth (2016).
[10] Note, that in all countries except for Bulgaria we used only proven cartel cases. In Bulgaria, only five out of the ten analysed collusion cases (that are part of the learning) were prosecuted and we could not find matching contracts for three of the five proven cases. We include the five investigated cases in the analysis.

We collected cartel-level information manually into a data template we developed for storing all relevant information of the case documents[11]. The key information extracted in this process included: the names of the companies involved, the public authority that conducted the public procurement process(es) in question, the time period in which the cartel operated[12], information related to the relevant public tender(s) (e.g. tender IDs, product types), and the location.

Data linking

As explained above, we set out to collect the most important dimensions of each proven case into a structured dataset and match on as many dimensions as possible. However, matching cases based on all available dimensions proved hard and impractical. Most often the number of contracts that were awarded to one of the cartel companies, were awarded or advertised during the cartel period and managed by a public buyer that is explicitly mentioned in the court rulings were very small and often zero. Therefore, we had to apply a more lenient approach and only match by company names and the proven cartel time period that are explicitly mentioned in the cartel documents[13]. While identifying all rigged contracts unambiguously would be clearly important to find statistically meaningful patterns in the indicators, we also accept that the number of truly rigged contracts vs. the ones that could be proven at the court can (occasionally very significantly) differ. We discuss this issue in more detail in the Methods section.

Final dataset

As Table 3.1 shows, we have 78 cartel cases in total from the six analysed countries – out of which we could only test 44 with at least one indicator due to missing data. We have 10,206 contracts won by cartel members after the collusive period and 6,845 during the cartel period. Note that the number of contracts used for testing varies cartel by cartel. While three Bulgarian cartels had no matching contracts in the procurement data, there were other cartels with simply too few contracts that did not allow for testing at all. For the extrapolation (see Multivariate models section) we use all available contracts – around 3.3 million across the seven analysed countries.

[11] We used the Google Document translator to understand the case documents in countries where the source files were only available in the national language (e.g. Swedish, Latvian and Lithuanian).
[12] The case documents contained varying detail on the start and end date of the cartels. In some cases, precise dates or the months of the start and end of cartel activity were defined (e.g. in most of the Swedish, Lithuanian, and Latvian cases, and some of the French cases). In other cases only the years were given (e.g. in most of the Spanish and Portuguese cases, and some of the French, Latvian, Lithuanian, and Estonian cases), hence we simply marked the full year, i.e. 12 months, as an estimation of the cartels' length, which might overestimate the length of the cartel activity.
[13] Note, that in some cases (for example, for all Hungarian cases), the case data was often not clear enough for assigning monthly values, hence we marked whole years that were mentioned in the court documents.

TABLE 3.1. FINAL DATASET SCOPE BY COUNTRY

Country	Number of cartels	During vs. after Number of contracts		Extrapolation	
		during	after	Number of contracts	Years
BG	7	460	1,279	219,808	2007-2021
ES	15	4699	3,161	360,510	2005-2019
FR	10	178	1,619	2,248,488	2004-2018
HU	18	670	736	65,099	2005-2012
LV	20	524	2,784	148,143	2006-2020
PT	2	56	106	62,607	2009-2018
SE	6	258	497	245,609	2009-2018
TOTAL	**78**	**6,845**	**10,182**	**3,350,264**	

3.2 Indicators

This section briefly enumerates all indicators widely used in the literature and discusses in depth those which we use in the subsequent analysis. Table 3.2 lists all potential indicators with their brief definitions. All three types of indicator groups capture a different outcome of anti-competitive behaviour - either by measuring how prices are consistently at odds with competitive pricing, bidding patterns showing companies strategically losing (or unrealistically winning) contracts on specific markets, and market structures being unlikely stable or concentrated over time as a consequence. In the final analysis, we include the ones that are most consistently available between countries and allow for cartel risk predictions without market definitions, that are the relative price, number of bidders, single bidding, missing bidders per buyer and per market, subcontracting and consortia indicators.[14]

As mentioned above, several cartel indicators cannot be calculated across countries due to their demanding data requirements, that is information on losing bids, accurate organization identifiers and accurate market definition. Furthermore, while some indicators can indicate anti-competitive patterns of proven cases, they cannot be straightforwardly used for predictions. Among the indicators based on price distribution[15], the difference between lowest and second lowest price, relative price range, relative standard deviation of bid prices cannot be calculated in most countries due to the lack of information on losing bid prices. Calculating winning probability and identifying suppliers in cut-point position is not feasible between countries due to lack of data on losing bids.[16] While bid rigging is expected to affect market

---

14 See the Appendix's Indicator section for indicator availability across countries. Note that for the multivariate models we use company size, that is the number of contracts won by company in a year, and sector, that is the 2-digit CPV code, as control variables.

15 Extreme or unusual offer price distributions were found to signal collusion by previous research, see Abrantes-Metz et al. (2006), Oxera (2013), Padhi and Mohapatra (2011). Another price-based indicator of anti-competitive pricing is the Benford's law, which posits that the first, second etc. digits of naturally occurring sets of data (such as prices emerging from a competitive process) follows a specific distribution (Berger and Hill, 2015, Fewster, 2009). The artificial process of fake bid submission is expected to produce a distribution that would not follow Benford's law. However, as markets cannot be reliably defined across countries (see below), we do not use it neither for testing nor for prediction.

16 Winning probability and cut-point position indicators cannot be calculated across countries due to the lack of losing bidder information in most countries. Companies having extremely high (close to 100%) winning probabilities over a long period on an theoretically competitive procurement market ought to be

shares – by making market structure artificially stable (Athey and Bagwell, 2001; Athey et al., 2004; Mena Labarthe, 2012; Harrington, 2006) or concentrated (Levenstein & Suslow, 2006; Pesendorfer, 2000), as markets are hard to define, we do not use market structure based indicators for predictions as markets cannot be reliably defined over time.[17]

---

a sign of a competitive anomaly, that can be an outcome of an collusive agreement (Fazekas and Tóth, 2016; Harrington and Joseph, 2005).
The bidding patterns subgroup of collusion risk indicators also includes an indicator based on the position of a company in a market. One way to represent a market is as an economic network, specifically as a network of firms that are connected when they bid on the same tenders. Examining the whole economic network that cartels are part of rather than focusing on the individual behaviour of cartel members provides a complex view on the interaction of firms and allows to assess the role and importance of a market player.
Cut-points of a network are key nodes that play an important role in connecting a graph. Their removal would result in the system becoming divided into disconnected elements. Such nodes are critical elements of a network that act as channels, brokers, agents between otherwise unrelated subsystems (Waseserman and Faust, 1994; McGloin, 2005).
In the context of the public procurement market, when examining co-bidding interactions of firms, companies in cut-point positions would be the ones that extensively bid against numerous competitors, while those competitors do not necessarily interact among themselves. Therefore, a company in a cut-point position would be the only connecting element of the graph and the removal of such a company from a market would lead to a situation when firms (or groups of firms) do not compete for the same contracts. The theoretical expectation for the cut-point analysis is that strawman companies supporting a bid-rigging collusion arrangement will often submit intentionally losing bids (Tóth et al., 2014). Once a firm has been allocated a contract in a collusive agreement, the firm has incentive to obscure the arrangement by creating fake competitors. In the co-bidding network, this would manifest as a cut-point centered on the winning cartel firm: the artificial firms submitting losing bids would only be connected to the rest of the co-bidding network through the winner.

[17] Markets can be defined based on product markets codes (CPV codes), geographic location and supplier identifiers. However, CPV codes do not overlap with de facto markets, and regional codes to not necessarily capture the geographical scope of markets – i.e. only the product and regional codes of proven bid rigging companies' contracts can be informative for defining markets on which concentration can be calculated. Furthermore, the lack of good enough identifiers and missing contract prices make is hard to reliably calculate market shares of individual suppliers.

TABLE 3.2. TESTED INDICATORS

Category	Nr	Indicator	Level of observation	Description	Included in analysis
Prices	1	Difference between lowest and second lowest price	Contract	Relative difference between the lowest and second lowest bid price (1%, 5%, 10% etc differences)	No
	2	Relative price range	Contract	Relative price range based on the lowest and highest bid price is less than 10% or more than 90% of the distribution	No
	3	Relative standard deviation	Contract	Relative standard deviation of bid prices is less than 10% or more than 90% of the distribution	No
	4	Relative price	Contract	Ratio of the final price and the estimated price	Yes
	5	Benford's law	Market-period	Whether first digits of contract prices of a given market in a given period follow Benford's law	No
Bidding patterns	6	Single bidding	Contract	Contract receiving a single bid during the tendering process.	Yes
	7	Number of bids	Contract	Number of bids received per contract	Yes
	8	Missing bidders by buyer	Company-period	The number of unique buyers companies submitting a bid at.	Yes
	9	Missing bidders by market	Company-period	The number of unique markets companies submitting a bid at.[18]	Yes
	10	Subcontracting	Contract	Whether a contract has a subcontractor.	Yes
	11	Consortia	Contract	Whether the winning bid was a consortia.	Yes
	12	Cut-point position	Market-period	Whether there are companies in a cut-point position in a given market and time period.	No
	13	Winning probability	Market-period	The average winning probability of companies of a given market and time period	No
Market structure	14	Concentrated market structure	Market-period	HHI change from during to after the cartel period	No
	15	Stable market structure	Market-period	Average absolute market share changes during vs. after the cartel period	No

Relative price

Relative price is defined as the final contract price divided by its initial estimate.[19] Healthy competition ought to lead to lower prices (i.e. bigger discounts) compared to the initial estimate - hence relative price can proxy competition. As collusion is about generating rents - either

[18] We define product markets by using the first 4 digits of the common procurement vocabulary (CPV) codes assigned to each tender.

[19] Note that the rules differ between countries on initial estimate calculation. For example, it can be an average estimation of market prices, but also an upper-bound estimation so that public buyers choose a more competitive procedure type. The higher the estimated tender value is, the more competitive (or at least more regulated) the to be applied tendering procedure is in most regulatory regimes.

through higher prices or lower quality - an increase in relative contract values can be a by-product of bid-rigging schemes.

However, relative price can be affected by a range of factors. Public buyers might lack the capacity to assess market prices accurately (i.e. there is noise in the estimated price), they might be incentivised to underestimate their tender prices for administrative reasons (e.g. bad incentives for budgetary planning). Furthermore, other anti-competitive practices, such as corruption, also can bias the estimated prices: buyers might deliberately overestimate prices to avoid clear signs of overpricing. All of these factors affect both what we find in the collusive tenders, and of course in the control tenders. For example, if half of the market is captured by a cartel, but the comparison group is corrupt, then we do not expect to find significant relative price differences between these contract groups.

With the limitations kept in mind the literature on collusion also uses prices, and relative price in particular as well, to analyse bid rigging schemes. Odd price increases that cannot be explained by costs as well as long term price stability at unusually high levels indicate market performance problems OECD (2014) and Oxera (2013). Prior research has found that tenders with large discounts (relative price below 90%) have a significant relationship with the number, capacity and experience of bidding suppliers, whereas these dimensions are unrelated to prices if discounts are small (relative price is above 90%) (Morozov and Podkolzina, 2013). Others have used relative winning price (in combination with low bid price variance) to distinguish between collusive vs. competitive tenders for modelling favour exchanges among bid-rigging suppliers (Ishii, 2009).

Single bidding and number of bidders

Withholding bids is one of the most straightforward ways to rig a tender, which results in a low number of bids and higher probability of single-bidder contracts by definition. While empirical research focusing on single-bidding as a collusion indicator is slim (Barrus 2011; Tóth et al, 2016), competition policy guidelines cite it as one possible elementary technique (OECD 2014; SCA 2015). Submitting fake bids is time-consuming, costly, and poses its own risks for the cartel members (for instance if the same language is used in multiple bids or if such an effort requires additional communication and coordination).

The reliability of bidder number based collusion indicators – especially single bidding - are affected by three possible confounding factors. First, it is a bluntly obvious signal of anti-competitive risks, that collusive companies might want to hide - especially if they are participating in markets with historically many bidders. Second, single-bidding is also a potential side-effect of corruption in public contracting, as favouring well-connected suppliers can exclude outsider companies entirely from the bidding process (Fazekas et al., 2018). While the first issue suggests that many cartels could operate on a basis that cannot be captured by this indicator, the second warns us about the limitations of this indicator used for indicating collusion specifically - as it might capture other anti-competitive behaviours instead. Third, a market with many single-bidder contracts attracts not only the attention of the competition and anti-corruption authorities, but also of potential competitors who would naturally see such a market and its high markups as a target for expansion, making the sustainability of the collusive agreement less viable.

Nevertheless, some of the bid-rigging schemes can be picked up even by this simple indicator. One relative strength of this approach to collusion is that it is easy to organize. Indeed, Barrus' (2011) study of the Kentucky highway construction market links single-bid contracts to tacit collusion. Such behaviour may be highly visible, but it is difficult to legally prove that it is the result of illegal coordination.

Missing bidders

The missing bidder indicators, similarly to the ones based on the number of submitted bids discussed above, capture how colluding companies withhold their bids from specific tenders of the market. Companies can withhold their bids by different market dimensions that can make coordination easier, hence lower the costs of cartel maintenance: from specific sub-markets (e.g. based on CPV codes), from specific (group of) buyers, or geographical location. Such techniques are also commonly quoted as a possible cartel strategy (OECD 2014; SCA 2015).

In our analysis we calculate two versions of the missing bidder indicators: a) number of unique buyers companies win a contract at, b) number of unique markets companies win a contract on. For example, in the bivariate tests, we expect that companies win contracts from more buyers and more markets (i.e. measured as the number of unique product codes assigned to their contracts) after the cartel period has ended.

Consortia

Instead of withholding bids – captured by the previous indicators – companies can also decide to submit joint bids, that is another elementary collusion technique. By joint bidding, companies lower competition and facilitate communication therefore it can be used as a price-fixing tool (Albano et al., 2009). Joint bidding also acts as an enforcement mechanism, as rent sharing is agreed in a formal contract. Due to the nature of procurement data, calculating the consortium indicator requires additional data processing[20], and once we had a good indication of consortia status, we connected the already known cartel member names based on a simple string matching to decide whether cartel members were also part of the joint bidding.

Subcontracting

Rent division between cartel members is a challenge (Asker, 2011), as transferring money between cartel members is risky - receiving money from a competitor is a signal of potential collusion. Another simple way to reallocate rents is through subcontracts. The prevalence of subcontracting in public contracts is contentious. While it can increase competition and efficiency through cooperation and knowledge exchange (Albano, Spagnolo, and Zanza, 2009; Estache and Limi, 2008), it can signal a collusive arrangement and serve as a tool for rent-reallocation (Fazekas and Tóth, 2016; Tóth, et al., 2014, Alexander, 1997). Therefore, while subcontracting in itself – as neither of the indicators - is not a strong enough indicator of collusion, it can indicate a form of rent sharing if other red flags are also present. We analyse subcontracting at the contract-level and calculate the share of contracts using subcontracting in all collusive vs. competitive contracts.[21]

3.3 Methods

This section explains our methodology. It outlines how we test collusion risk indicators by exploiting the differences between proven cartel and likely non-cartel contracts. First, we offer a simple, bivariate test of each collusion risk indicator which is aimed to underline the need for combining individual indicators for better measurement accuracy. Second, we test all individual indicators in competing multi-variate models such as logistic regression, random forests and gradient boosting machines. We select the best model based on test-set accuracy.

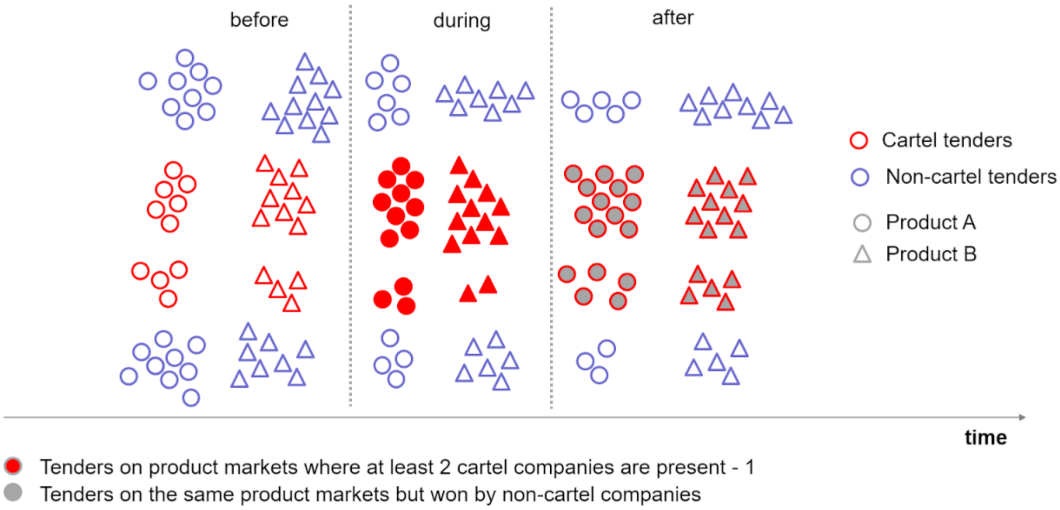
We start by conducting bivariate tests of elementary collusion risk indicators comparing cartels with non-cartels, that is contracts from the period the cartel was most likely in operation,

[20] For example, there is no clear indicator in the source data on whether a given bid is submitted by a group of companies, hence we need to find them based on an algorithm – detailed in the Appendix.
[21] Note, that the indicator could be calculated in the following countries: Latvia, Spain, and France.

compared to contracts from the period it was most likely not in operation. We compare elementary collusion risk indicators defined on the contract level using two samples t-tests or proportion tests.[22] Contracts are flagged as collusive when they have been awarded to companies explicitly reported as collusive in publicly available court records, irrespective of the particular product group, that is companies are allowed to collude on a range of product markets (filled red symbols in Figure 3.1). Conversely, contracts are flagged as non-collusive when they are awarded to the same companies, but after the cartel was busted (grey filled symbols).

The contract grouping covers contract-level indicators, we had to use a different logic for the missing bidder indicators, where the indicator is defined at a company-time period level. As we cannot apply a statistical test we accept a test as confirming if the unique number of buyers or markets companies are winning contracts at is lower on average based on the collusive contracts vs the control group.

FIGURE 3.1. DURING-AFTER COMPARISON[23]



As cartels are diverse and public procurement datasets are generally noisy (e.g. widespread missing, erroneous data), it is expected that elementary collusion risk indicators are imprecise outside of very well delineated and homogeneous collusive markets. The validity and reliability of a cartel risk detection framework can be increased if elementary indicators are combined into either a sequence of tests (e.g. Tóth et al, 2014) or screens or into a composite risk score (e.g. Huber and Imhof, 2019). In the absence of sufficiently precise theory guiding the methods for combining elementary collusion risk indicators, including prescribing indicator weights, we turn to data-driven approaches. The main advantage of such approaches is that they are able to directly learn from the patterns in the data in order to identify the accurate combination of individual indicators. We compare 3 different methods to identify the model with the highest prediction accuracy:

- Binary logistic regression;
- Random forests; and

[22] We only calculate t-tests for cartels that won at least 6-6 contracts both during and after the cartel period and the indicator value varies.
[23] Each symbol represents an awarded contract.

- Gradient boosting machines.

Such supervised machine learning-based approaches, make use of a wide set of known cartel and non-cartel cases by labelling contracts won by cartel and non-cartel members as the outcome variable. The algorithm then learns how best to predict that label in the training set, making use of a set of individual collusion risk indicators and control variables, serving as predictors or features (Huber et al, 2020). For model comparison, we calculate the accuracy of model prediction on an unseen, test dataset, that is dataset which was not used to fit the model. **Accuracy** is defined as the percent of correctly classified contracts over all contracts in the test set. We use 3 different definitions of test datasets to reflect different types and degrees of difficulties for prediction tasks.

- Random sample of contracts (30%): This is considered the easiest prediction task as contracts belonging to cartel cases used to train the model are used to test prediction accuracy. It is considered a policy relevant scenario, i.e. identifying further cartel contracts for already known cartel cases, however, it is not necessarily the most typical one. This test-train split is the most comparable set-up to prior studies.
- Random sample of cartels (20%): We consider this set-up as the most relevant albeit more challenging than the previous one. This is because in a typical competition policy use case, our algorithm would be used to identify previously completely unknown cartels based on patterns in known cases. Considering the mode demanding nature of this test-train split, we increased the proportion of the training sample.
- Leave 1 country out: Given our main goal of developing a general model accurately predicting cartel behaviour, we are particularly interested whether cartels from one country can predict cartels in another. We consider this as the most demanding test to our models, given cross-country differences in data, cartel and non-cartel behaviours.

While we compare 3 generic types of models, as outlined above, we tune each of these first so that the best variant of the model can be compared to the other methods.[24] Binary logistic regression models are considered as the baseline as they are the least variable model, that it they are the least sensitive to different test-train splits. However, we expect binary logit models to have the lowest accuracy too, given the more restrictive constraints the model imposes on parameters.

In the context of fraud, corruption, and collusion, various studies have used the random forest and gradient boosting machine algorithms. We also use these methods because of their ability to model a diverse array of different collusive strategies and the markers they leave. Random forests and gradient boosting machines are supervised machine learning methods which predicts the output by constructing multiple decision trees with given features (Breiman, 2001). They are particularly well suited for datasets with many explanatory variables and where the same outcome may be the result of multiple different combinations of predictor values (James et al, 2015). In spite of their flexibility and suitability of the complex prediction problem we aim to develop, these models lead to results which are harder to interpret within our theoretical framework. In other words, to achieve high prediction accuracy, we have to sacrifice some degree of interpretability. Nevertheless, we will explore the relationships identified by the best model below in order to compare our theoretical expectations and best empirical results.

[24] We implemented data preparations and modeling in R (version 4.1.2). For logistic models, we used stats library and glm function; for Random Forests we used the randomForest library and randomForest function; and for Gradient Boosting Machines we used the gbm library and gbm function.

Predicting cartel risks

Once the highest accuracy model is selected, it can be used to make predictions to the full universe of contracts across all the countries analysed in the study, covering nearly 4 million contracts. This extrapolation is predicated on the assumption that cartel behaviours in the whole economy are comparable to the uncovered, proven cases and that the underlying data points are also comparable (e.g. variable distributions, missing rates, variable availability). Given the wide range of proven cartels we analyse and the use of standard public procurement datasets harnessed from government publication portals, we consider these preconditions met. However, the different elementary collusion risk indicators can be defined on different levels of observations such as contracts, companies or markets. Indicators defined on the market level require a precise definition of markets matching to the boundaries set by the cartel itself which is very hard to reliable calculate.[25] Hence, only those indicators can be used to model building and extrapolation which can be calculated without such additional parameters, that is indicators defined on the contract or company levels.

[25] It is possible to define market IDs, for example based on procurement classifications such as CPV (Common Procurement Vocabulary) and NUTS (Nomenclature of Territorial Units for Statistics), however, it is likely to be imprecise given the diversity and specificity of cartel agreements (Fazekas and Tóth, 2016).

4 Results

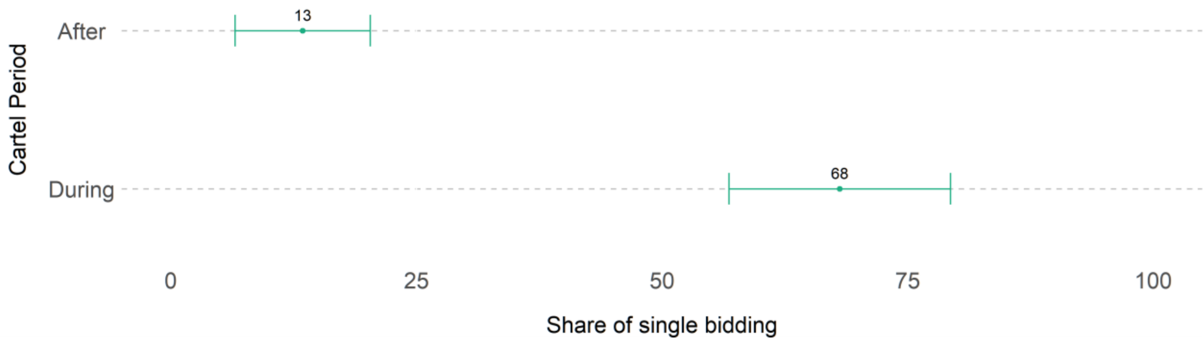
In this section, we present our results, first in a bivariate set-up testing each indicator separately, second looking at predictive models combining all our indicators. Finally, once the best model is selected, we explore the relationships between individual collusion risk indicators and predicted collusion risks to verify fit with theoretical predictions.

4.1 Bivariate tests

Out of the total 78 cartels we collected from court documents, we could test 44 for anti-competitive behaviour with at least one of our screens in this binary set-up. The remaining cartels could not be tested even by a single indicator due to lack of data, either not finding (enough) awarded public contracts, or missing key variables required for indicator testing (e.g. missing prices or number of bids). Second, out of the 44 cartels for which we could run tests, we found 32 that were caught by at least one indicator (72%), 19 were caught by at least three (43%), and 10 were caught by at least three (22%). Overall, these suggest that by casting a broad net, we can observe signals of anti-competitive behaviour at scale albeit with varying precision by country, cartel and indicator.

Figure 4.1 shows the logic of a binary test based on cartel number 9 from Spain. Single bidding was 68% during the collusive period based on contracts won by cartel companies and it dropped to 13% after the cartel ended. Following this logic, we carry out and report binary indicator test results for each indicator and each cartel – i.e. the same cartel is typically tested by multiple risk indicators. For each test we report how many total cartels could be screened, and how often a significant signal in line with theory was observed.

FIGURE 4.1. SINGLE BIDDING BIVARIATE TEST EXAMPLE (CARTEL 9 – SPAIN)[26] FOR DURING VS. AFTER COMPARISONS. ERROR BARS REPRESENT 95% BOOTSTRAPPED CONFIDENCE INTERVALS



[26] Each symbol represents an awarded contract.

The number of cartels identified by at least 1 individual cartel risk indicator is relatively high, 32 out of 44. However, only a few cartels are identified by multiple indicators, 10 out of 44 cartels are indicated by 3 or more indicators. Countries differ due to the number of matching collusive contracts, b) the share of available indicator values, and c) applied cartel strategies that the tested indicators can capture. For example, many French and Hungarian cartel members had only a couple of matching contracts, hence only 3 out of the 10 and 7 out of the 19 could be tested respectively. Furthermore, while 16 out of the 17 cartels in Spain could be tested by at least one indicator, none of the indicators captured five of them.

TABLE 4.1. CARTEL IDENTIFICATION SCOPE: CARTELS BY THE NUMBER OF SIGNALING INDICATORS

Country	Cartels tested by at least one test*	Number of cartels caught by		
		at least 1 indicator	at least 2 indicators	at least 3 indicators
BG	7	7	5	3
ES	16	11	8	4
FR	3	3	0	0
HU	7	5	2	1
LV	6	3	3	1
PT	2	2	1	1
SE	3	1	0	0
Total	44	32	19	10

*Note: * We only calculate bivariate tests for cartels that won at least 6-6 contracts both during and after the cartel period and the indicator value varies. Hence, the number of cartels considered in the bivariate tests is lower than the number of cartels in the multivariate analysis below.*

Turning to the performance of individual indicators, Table 4.2 shows that most indicators have relatively low accuracy individually - they indicate 10-30% of the cartels. For example, we find only 5 out or the 41 cartels a statistically lower share of single bidder contracts won by cartel members after the cartel period ended at a 10% significance level. The only two indicators indicating the majority of cartel (65-75%) are based on missing bids – i.e. companies winning contracts from more buyers and more product codes after the cartel period.[27]

In some cases, weak indicator accuracy may indicate data quality and reliability issues. For example, in many cases, relative price does not drop more significantly after the cartel period which suggests that a) estimated prices have too much noise, b) buyers might be involved in the cartel activity too by overestimating tender prices, or c) some of the cartel activity is mimicking competition by offering close to competitive prices but the delivered quality is lower.

[27] As explained above, for the missing bids indicators, we can only assess the sign of the change and no significance tests are calculated. Therefore, a test is deemed confirming simply when the number of unique markets or unique buyers a cartel member was submitting a bid for is higher after the cartel period.

TABLE 4.2. INDICATOR ACCURACY: BINARY (BEFORE-AFTER) TEST RESULTS BY INDICATOR

Group	Indicator name	Number of tests conducted	Pct. of sign. tests at			
			5%	10%	5%	10%
Prices	Relative price	26	4	6	15%	23%
Bidding patterns	Single bidding	41	4	5	10%	12%
	Number of bidders	33	6	6	18%	18%
	Missing bidders (buyer)	24	18	18	75%	75%
	Missing bidders (market)	17	11	11	65%	65%
	Subcontracting	18	4	5	22%	28%
	Consortium	29	8	9	28%	31%

The table highlights in particular the issue of data quality heterogeneity across countries and over time. Note for instance that only for 26 cartels could we calculate relative price and conduct statistical testing. We emphasize that this is a sign that investment in data quality should be a top priority of competition authorities, not that these indicators require an unrealistic level of detail.

4.2 Multivariate models

This section describes 3 different predictive models - logistic regression, random forest and gradient boosting machines – and it compares their prediction accuracy on 3 different test datasets – random contracts sample, random cartel sample, and leave 1 country out sample. Before making such comparisons, we tweaked each model to achieve highest prediction accuracy on the most widely used test set definition (30% random sample of contract). The binary logit model was not tuned in any meaningful way as it has a high degree of inflexibility. We simply included all individual collusion risk indicator and the control variables described above. As for random forest, we optimised the number of trees (optimal parameter is 500 trees) and the number of variables used in each tree (optimal parameter is 4 variables). Regarding boosting, we optimised the number of trees (optimal number of trees is 3000) and interaction depth (optimal depth is 4), keeping shrinkage parameter at the default 0.1.

The 3 models perform differently on the 3 different test sets reflecting the different challenges of each set-up. Considering the most policy relevant scenarios in which we either already know or suspect some cartel contracts and also when we predict to new cartel cases from known cases, the random forest model performs best. It achieves 91% accuracy on a random sample of contracts and 77% accuracy on a random sample of cartels (Table 4.3). As expected, accuracy drops for all 3 models when we extrapolate across countries, however the drop is lowest for logistic regression which is the least sensitive model to variations in test-train split.

TABLE 4.3. MODEL ACCURACY COMPARISONS: PERCENT CORRECTLY CLASSIFIED

Model/test set	Random sample of contracts (30%)	Random sample of cartels (20%)	Leave 1 country out*
Logit	73.1%	75.1%	**67.4%**
Random forest	**91.5%**	**77.1%**	54.7%
Boosting	90.5%	75.3%	59.2%

** Setting 1 country from the list (PT, SE, HU, BG, FR, or LV) aside as test set. Please note that ES is not considered as a test set as it is too large, about half of the total sample,*

The best random forest model while achieving 91.5% accuracy, its performance is also relatively balanced. Its precision is 91.8% and recall is 87.7%, that is when it predicts cartel it is nearly always correct, while it misclassifies some cartel cases as non-cartel (Table 4.4).

TABLE 4.4. CONFUSION MATRIX, BEST RANDOM FOREST MODEL, ALL COUNTRIES COMBINED, TEST SET: 30% OF CONTRACTS

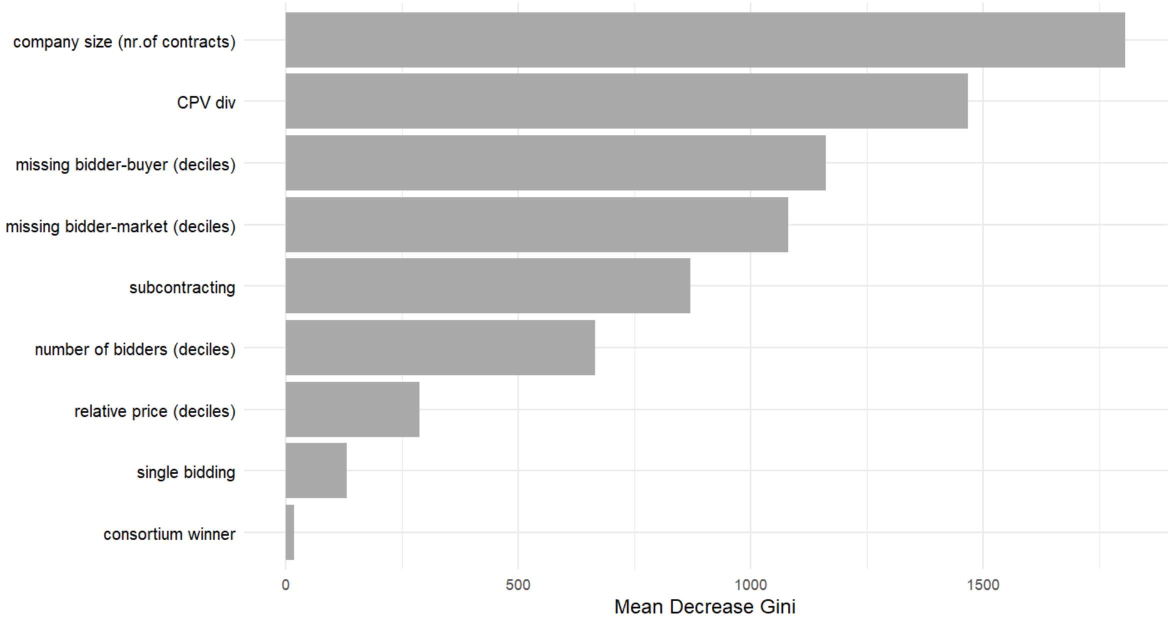
		Reference	
		No	Yes
Prediction	No	2796	265
	Yes	167	1881

4.3 Description of the final best model

While the overall prediction performance of the model is of central importance, from a scientific and policy perspective, it is also crucially important to open up the black box of machine learning and investigate the estimated response functions and their fit with theory. We consider any predictive model useful for policy uses and supporting further scientific work if it is not only accurate, but the relationships between predicted cartel risk and individual collusion risk indicators it estimates correspond to theoretical expectations. Hence, we review the most important predictors and their impact on predicted cartel risks.

We can conclude that knowing the size of the company and the sector of the contract (2-digit CPV codes corresponding to broad sectors such as construction or healthcare) are the most impactful contributors to prediction accuracy (Figure 4.2). Among the individual collusion risk indicators, the most important are missing bidders (on the level of buyers and markets), then subcontracting, numb of bidders and relative price, with single bidding and consortium being the least important for prediction accuracy.

FIGURE 4.2. VARIABLE IMPORTANCE CHART, BEST RANDOM FOREST MODEL, ALL ELEMENTARY COLLUSION RISK INDICATORS PLUS COMPANY SIZE AND CONTRACT SECTOR, ALL COUNTRIES, FULL DATASET

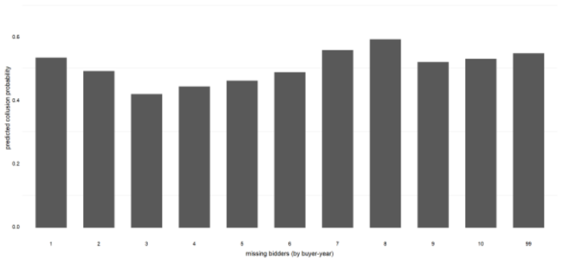


We look at the predicted impact of each indicator in the model to verify their fit with established theories of collusive behaviour. Specifically, we look at partial dependence plots of the 4 most important individual collusion risk indicators (Figure 4.3). Overall, we find a complex and varied picture, as expected, given the flexibility of the random forest model and the diversity of cartels investigated.

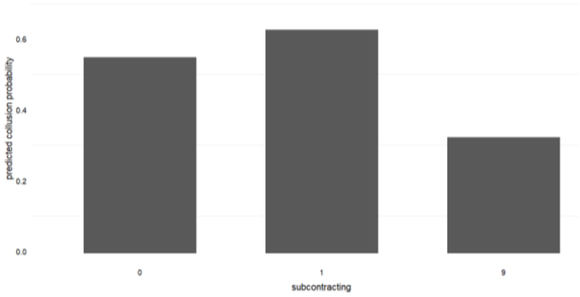
More specifically, the most important individual collusion risk indicator, missing bidders by buyer, displays a notable U-shaped relationship with predicted cartel probability (Figure 4.3, panel A). This means that the model predicts high cartel risk when there are notably few bidders per buyer in a year, but also when there are many bidders per year and buyer. This is in line with our understanding that there are 2 distinct cartel strategies at play, either withholding bids (i.e. low number of bidders per buyer) or submitting losing bids (i..e high number of bidders per buyer). Identifying such a U-shaped relationship shows the strength of a flexible machine learning approach capable of tracking complex and varied behaviours. The second most important cartel risk indicator is subcontracting (yes versus no, plus a missing category (9)) (Figure 4.3, panel B). As predicted by our theory, we see a higher predicted cartel probability when part of the contract was subcontracting which is indicative of paying off losing bidder members of the consortium with subcontracts. The third most important elementary cartel risk indicator is the number of bidders (deciles plus a missing category (99)) (Figure 4.3, panel C). Here, we see a largely positive pattern, that is a higher than usual number of bidders are predicted to have higher cartel probability, especially the 9th and 10th deciles. This indicates the theoretically predicted cover bidding behaviour being the prominent behaviour captured by our model. The fourth most important individual cartel risk indicator is relative price which is expected to work similarly across all cartel types (deciles plus a missing category (99) (Figure 4.3, panel D). Again, we can see a pattern in line with our theory. The 10th decile of relative price between 1 and 1.3 (contract value being at or above the auction reference price) is predicted to have the highest cartel probability. Interestingly, we see a drop in predicted cartel probability for the 7th decile (in essence when the contract value is the same as the auction reference price) but again a higher predicted probability for the 5th and 6th deciles (relative prices between 0.909 and 0.999). This pattern may make sense as a large portion of contracts come in just on the auction reference price so the 7th decile probably includes a range of competition violations other than cartel behaviour. While a small discount, a contract value close to the auction reference price is again risky, often consistent with cartel behaviour.

FIGURE 4.3. PARTIAL DEPENDENCE PLOTS FOR SELECTED VARIABLES, BEST RANDOM FOREST MODEL, ALL COUNTRIES, FULL DATASET

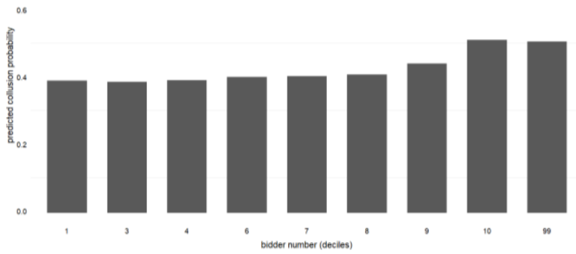
Panel A. Missing bidder (by buyer) deciles and missing (99)



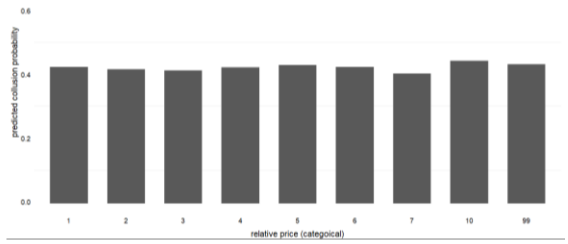
Panel B. Subcontracting: yes, no, or missing (9)



Panel C. Bidder number deciles and missing (99)



Panel D. Relative price deciles and missing (99)



5 Discussion

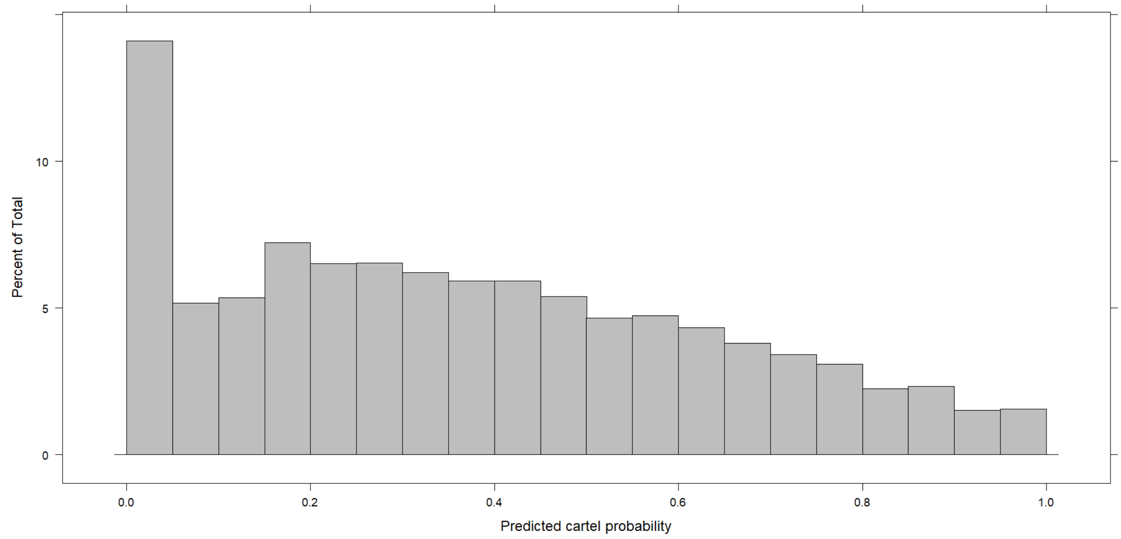
Having reviewed the main results of the bivariate and multivariate modelling and having identified the best predictive model, we can turn to extrapolating to contracts with unknown cartel status and discuss the policy implications of our results.

5.1 Extrapolation

Taking the best random forest model identified above, we predict the risk of collusive behaviour in the full database consisting of more than 3.3 million observed public procurement contracts in the 7 countries during 2007-2020 (Figure 5.1). Overall, the model predicts a 37% average cartel risk score. About 15% of all contracts observed in the 7 countries receives virtually zero cartel risk score, while about 2/3rd of contracts receive less than 50% cartel risk prediction. Nevertheless, the upper 1/3rd of contracts is predicted to have more than 50% cartel probability, implying a strong indication of cartel behaviour. This is roughly in line with findings in the literature using different data and indicators (e.g. Kawai & Nakabayashi, 2022). However, there is a strong cross-country variation in both the distribution of predicted cartel risks (Figure 5.1, panel B) and the average risk per country (Figure 5.2). With the exception of Spain, all countries are predicted to have only a minority of their public procurement markets likely collusive.

FIGURE 5.1. HISTOGRAMS OF THE PREDICTED CARTEL PROBABILITY USING THE BEST RANDOM FOREST MODEL, ALL COUNTRIES, FULL DATASET (INCLUDING CONTRACTS WITH UNKNOWN CARTEL STATUS)

Panel A: Combined histogram



Panel B: Histograms by country

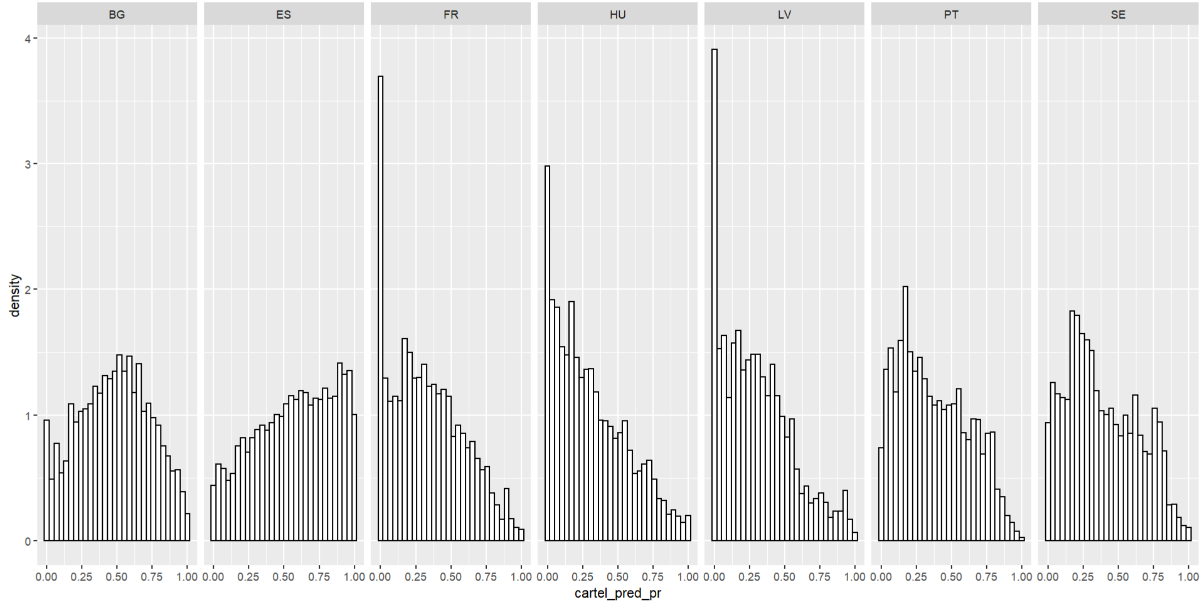
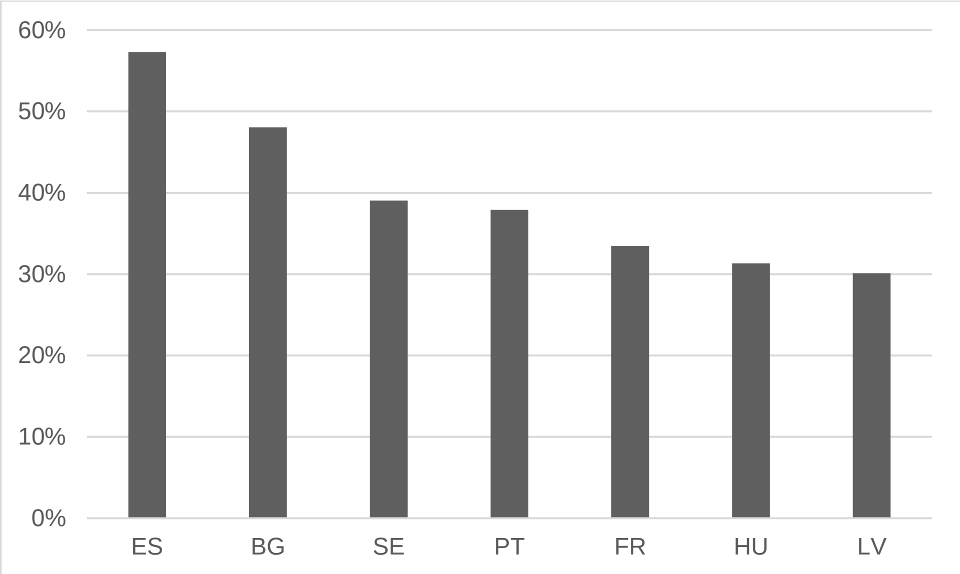


FIGURE 5.2. AVERAGE PREDICTED CATREL PROBABILITY BY COUNTRY, USING THE BEST RANDOM FOREST MODEL, ALL COUNTRIES, FULL DATASET (INCLUDING CONTRACTS WITH UNKNOWN CARTEL STATUS)



5.2 Policy implications

Based on the exceptionally high prediction accuracy of our predictive models, surpassing any other known model, to the best of our knowledge, we see profound policy implications. There are 2 main policy uses of our models and predictions: i) supporting investigation; and ii) informing preventive and pro-competition policy interventions.

In order to increase investigation targeting of public procurement cartels, we recommend the adoption of models similar to the ones developed in this article. We can conclude that both the theoretical and empirical advances in the literature have reached a level which warrants real life use. Applying such predictive models to regularly updated data, markets can be scanned regularly to find highly likely cartels. In addition, ongoing investigations could also be supported by predictive modelling where models could recommend further contracts of the investigated firms or further accomplice firms.

As we find that about 1/3rd of the public procurement markets of 7 European countries has high cartel risk, there is a strong argument for pro-competition, preventive policies. These policies could be tailored to address market entry and other barriers to competition especially in markets where collusion risks are found to be high.

Nevertheless, the limitations of our modelling exercise have also revealed the need to improvements in public procurement data and cartel risk modelling. There is a strong case for improving public procurement data quality and scope. In particular, the rate of missing data should be lowered in publicly available public procurement data repositories and publication websites. Moreover, a range of key missing fields have prevented our models from more comprehensively measuring collusion risks. Among these, crucial variables such as losing bidder and bid price information were missing. When it comes to data scope, several countries under investigation only publicly disclose higher value contracts which means that real-time monitoring of lower value contracts is very costly, if not impossible. Hence, recording all or nearly all public procurement transactions in centrally maintained e-procurement systems is of great value, not only for cartel risk detection but also for other uses of public procurement data (e.g. spending efficiency measurement).

There is a similarly strong case for building risk prediction and monitoring systems which continuously improve. That is, starting from the currently available best predictive models, competition authorities and other law enforcement bodies should feed the latest investigative results back into the learning models allowing them to adapt and improve. It is expected that cartels learn from past, especially recent, enforcement actions hence predictive models should reflect changes in cartel behaviours in order to stay relevant and accurate.

6 Conclusions

We set out to develop a high accuracy predictive model capable of tracking diverse cartel behaviours across many countries and over years. We drew on publicly available data for 84 cartels from 7 countries during 2004-2021 and applied state-of-the-art machine learning methods, such as random forests, to combine diverse cartel screens into a theoretically sound model. Our best model achieves 77-91% prediction accuracy across countries on unseen, teste dataset of proven cartels. This model is highly parsimonious, making use of only 5 elementary cartel risk indicators (with some variations on formulation to some) and 2 control variables (company size and contract sector). Unusually to many black-box machine learning approaches, our most impactful predictors are estimated to have an impact on predicted cartel probabilities in line with theory.

We also used the most accurate model to predict cartel probability for over 3.3 million contracts in our 7 European countries. This extrapolation suggests that about 1/3rd of public procurement contracts was awarded to likely cartelling firms. As our models were built on readily available, large-scale public procurement datasets, they can be readily applied to support competition enforcement, supporting both investigations and pro-competition policy interventions.

Our approach nevertheless suffers from a range of limitations. The administrative data we could gather from publicly available sources while large-scale, it often has quality and scope issues. Many variables suffer from high missing rates and key variables suggested by the literature were not available at all (e.g. losing bid prices). Moreover, any learning model can only learn from known cases which may represent a biased sample of the true range of cartel behaviours. For example, if more sophisticated cartels are harder to detect they may not show up in our learning dataset at all, so we under-estimate their presence in the data.

Future research could address the shortcomings of our approach. In particular, adding more countries, indicators and cartel cases should improve on many limitations listed above. As more known cartel cases and countries are added to the model, we get closer to a genuinely generic cartel risk detection model which should be of great scientific and policy value.

References

Abrantes-Metz, Rosa M., Luke M. Froeb, John Geweke, & Christopher T. Taylor. (2006). A Variance Screen for Collusion. International Journal of Industrial Organization 24(3): 467–86.4

Albano, G. L., G. Spagnolo, and M. Zanza. 2009. "REGULATING JOINT BIDDING IN PUBLIC PROCUREMENT." Journal of Competition Law and Economics 5(2): 335–60. http://jcle.oxfordjournals.org/content/5/2/335.short (May 13, 2016).

Alexander, B. J. (1997). Mechanisms for rent transfers: Subcontracting among military aircraft manufacturers. Public Choice, 91(3), 251-269.

Asker, J. (2010). A study of the internal organization of a bidding cartel. American Economic Review, 100(3), 724-62.

Athey, S., Bagwell, K., & Sanchirico, C. (2004). Collusion and price rigidity. The Review of Economic Studies, 71(2), 317-349.

Bajari, P., & Ye, L. (2003). Deciding between competition and collusion. Review of Economics and statistics, 85(4), 971-989.

Banal-Estañol, A., Newham, M., & Seldeslachts, J. (2021). Common Ownership in the U.S. Pharmaceutical Industry: A Network Analysis. The Antitrust Bulletin, 66(1), 68-99.

Barrus, D. R. (2011). Firm Bidding Behavior in Highway Procurement Auctions: An Analysis of Single-Bid Contracts, Tacit Collusion, and the Financial Impact on Kentucky.

Berger, A., & Hill, T. P. (2015). An introduction to Benford's law. Princeton University Press.

Bosio, Erica, Simeon Djankov, Edward Glaeser, and Andrei Shleifer. 2022. "Public Procurement in Law and Practice." American Economic Review, 112 (4): 1091-1117.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Buccirossi, P., & Spagnolo, G. (2008). Corporate Governance and Collusive Behavior. In Issues in Competition Law and Policy (pp. 1219-1240). American Bar Association.

Chassang, S., & Ortner, J. (2019). Collusion in auctions with constrained bids: Theory and evidence from public procurement. Journal of Political Economy, 127(5), 2269-2300.

Conley, T. G., & Decarolis, F. (2016). Detecting bidders groups in collusive auctions. American Economic Journal: Microeconomics, 8(2), 1-38.

Connor, J. M. (2001). "Our customers are our enemies": the lysine cartel of 1992–1995. Review of Industrial Organization, 18(1), 5-21.

de Haas, S., & Paha, J. (2020). Non-Controlling Minority Shareholdings and Collusion. Review of Industrial Organization, 1-24.

European Commission. (1997). Commission Notice on the Definition of Relevant Market for the Purposes of Community Competition Law. Brussels.

European Commission. (2014). Commission Staff Working Document Accompanying the White Paper "Towards More Effective EU Merger Control". Brussels.

Evenett, S. J., Levenstein, M. C., & Suslow, V. Y. (2001). International cartel enforcement: lessons from the 1990s. The World Bank.

Fazekas, M., & Tóth, B. (2016). Assessing the potential for detecting collusion in Swedish public procurement. Stockholm: Swedish Competition Authority.

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

Fewster, R. M. (2009). A simple explanation of Benford's Law. The American Statistician, 63(1), 26-32.

Funderburk, D. R. (1974). Price Fixing in the Liquid-Asphalt Industry: Economic Analysis versus the Hot Document. Antitrust L. & Econ. Rev., 7, 61.

Harrington Jr, J. E., & Chang, M. H. (2009). Modeling the birth and death of cartels with an application to evaluating competition policy. Journal of the European Economic Association, 7(6), 1400-1435.

Harrington, Joseph. (2008). Detecting Cartels. In Handbook of Antitrust Economics, MIT Press, 796.

Hewitt, C., McClave, J., & Sibley, D. (1996). Incumbency and bidding behavior in the dallas-ft. worth school milk market. Unpublished manuscript.

Huber, M., & Imhof, D. (2019). Machine learning with screens for detecting bid-rigging cartels. International Journal of Industrial Organization, 65, 277-301.

Huber, M., Imhof, D. & Ishii, R. (2022) Transnational machine learning with screens for flagging bid-rigging cartels. Journal of the Royal Statistical Society: Series A (Statistics in Society), 1074– 1114.

Ishii, R. (2008). Collusion in repeated procurement auction: a study of a paving market in Japan.

Ishii, R. (2009). “Favor Exchange in Collusion: Empirical Study of Repeated Procurement Auctions in Japan.” International Journal of Industrial Organization 27(2): 137–44.

Kawai, Kei, Jun Nakabayashi, Juan M. Ortner, and Sylvain Chassang, 2022. "Using Bid Rotation and Incumbency to Detect Collusion: A Regression Discontinuity Approach," NBER Working Papers 29625, National Bureau of Economic Research.

Kawai, Kei and Jun Nakabayashi. 2022. “Detecting Large-Scale Collusion in Procurement Auctions.” Journal of Political Economy 130 (5): 1364–1411.

Lanzillotti, R. F. (1996). The great school milk conspiracies of the 1980s. Review of Industrial Organization, 11(4), 413-458.

Levenstein, M. C., & Suslow, V. Y. (2006). What determines cartel success?. Journal of Economic Literature, 44(1), 43-95.

Levenstein, M. C., & Suslow, V. Y. (2011). Breaking up is hard to do: Determinants of cartel duration. The Journal of Law and Economics, 54(2), 455-492.

McGloin, J. M. (2005). Policy and intervention considerations of a network analysis of street gangs. Criminology & Public Policy, 4(3), 607-635.

Mena-Labarthe, C. (2012). Mexican experience in screens for bid-rigging. Antitrust Chronicle, 3.

Morozov, I. & Podkolzina E. (2013) Collusion detection in procurement auctions. No.
Morselli, C., & Ouellet, M. (2018). Network similarity and collusion. Social Networks, 55, 21-30.

OECD (2014) Detecting Bid Rigging in Public Procurement. Downloaded: 2021.01.15.

OECD (2017). Common Ownership by Institutional Investors and its Impact on Competition. Paris.

Oxera (2013). Hide and Seek: The Effective Use of Cartel Screens. Downloaded: 2021.01.19. http://www.oxera.com/getmedia/210bc5bc-0cc9-40ea-8bc9-6c8b2406b485/Cartelscreens.pdf.aspx?ext=.pdf

Padhi, Sidhartha S., & Pratap KJ Mohapatra (2011). Detection of Collusion in Government Procurement Auctions. Journal of Purchasing and Supply Management 17(4): 207–21.

Pesendorfer, M. (2000). A study of collusion in first-price auctions. The Review of Economic Studies, 67(3), 381-411.

Porter, R. H. & Zona, J. D. (1999) Ohio school milk markets: An analysis of bidding. The Rand Journal of Economics 263–288.

SCA (2015). Honesty Pays: How to Detect and Prevent Bid-Rigging Cartels.

Schwalbe, U. (2018). Algorithms, machine learning, and collusion. Journal of Competition Law & Economics, 14(4), 568-607.

Scott, F. A. (2000). Great school milk conspiracies revisited. Review of Industrial Organization, 17(3), 325-341.

Tóth, B., Fazekas, M., Czibik, Á., & Tóth, I. J. (2014). Toolkit for detecting collusive bidding in public procurement: With examples from Hungary. Budapest: Government Transparency Institute.

Vadász, P., Benczúr, A., Füzesi, G., & Munk, S. (2016). Identifying Illegal Cartel Activities from Open Sources. In Open Source Intelligence Investigation (pp. 251-273). Springer, Cham.

Wachs, J., & Kertész, J. (2019). A network approach to cartel detection in public auction markets. Scientific reports, 9(1), 1-10.

Wasserman, S., & Faust, K. (1994). Social Network Analysis: Methods and Applications. Cambridge University Press.

Wasserman, S., & Faust, K. (1994). Social network analysis: Methods and applications.

World Bank. (2009). Fraud and Corruption. Awareness Handbook. Washington DC: World Bank.

World Bank. (2011). Curbing Fraud, Corruption, and Collusion in the Roads Sector. Washington DC: World Bank.

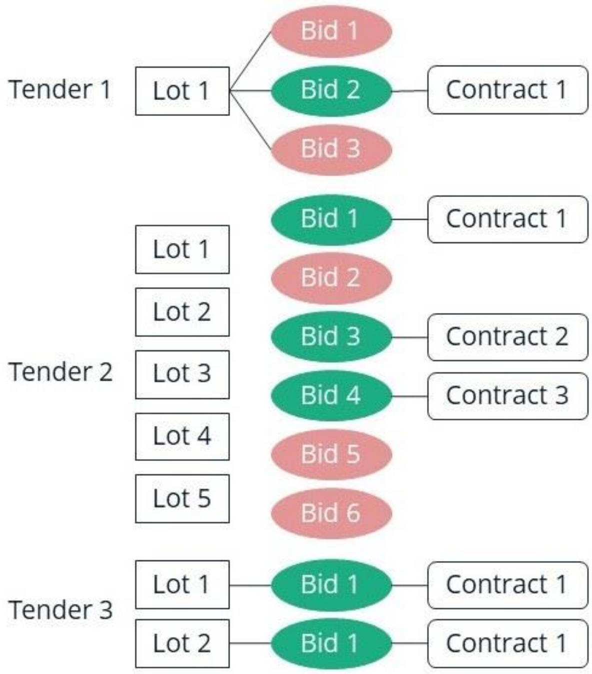
7 Appendix

Data

Public procurement data

Going into the technical details of procurement datasets is beyond the scope of this study, however, we want to summarise the dataset structure we use in this report. First, in an optimal scenario we unambiguously identify separate lots per each tender and all submitted bids to individual lots. However, due to the nature of most procurement data, we can only separate bids (winning and losing) for tenders with a single contract (Figure 7.1**Error! Reference source not found.** tender 1). Second, tenders can have multiple awarded companies - that are presumably the result of having multiple lots per tender. However, the competing bids per lot cannot be separated - i.e. we are aware of the losing bids 2, 5, 6 of tender 2 but we do not know which winning bid (1, 3 or 4) they competed against. This grouping would be necessary to calculate meaningful indicators (such as bid price range or relative range of bid prices), therefore, we have to exclude these ambiguous tenders from the analysis[28]. Third, the French and Spanish datasets that we analyse do not have information on losing bidders, hence each observation in the dataset corresponds to an awarded lot (tender 3).

FIGURE 7.1. DATA STRUCTURE



Cartel case collection

First, we looked up the national competition authorities' websites and searched for their repositories with documentation of proven cartel cases. If we found such a repository, we searched for proven cases of public procurement collusion from the past 10-15 years. We used a range of search terms, such as "public procurement", "public contract", "public tender",

[28] This ambiguity affects all datasets where bidding information is available.

"tendering procedure" etc. in the national languages. We processed the shortlisted case documents manually to find the ones that are indeed related to public procurement.

Second, we searched alternative sources, such as the authorities' annual reports or the Court of Justice's website where the competition authorities did not offer a case repository. We used a similar strategy of combining search terms such as "public procurement", "cartel", "collusion" in the national languages, to identify all the relevant cases and then processed them manually.

At last, we contacted the competition authorities of the selected countries and requested an overview of proven bid rigging cases in order to verify or extend our case collection. As a result of the three strategies, we have collected over 156 cartel cases. The country-level sources and results of the case collection process are detailed in the next section.

Following this process, we drew on sources provided by the competition authorities of the selected countries and hence identified the cartel cases related to public procurement. The below table gives an overview of the relevant competition authorities bodies and their case repositories we used as our final source.

TABLE 7.1. OVERVIEW OF SOURCES FOR CARTEL CASES[29]

Country	Competition authority	Case repository resource	Number of cartel cases
Bulgaria	Комисия за защита на конкуренцията	https://cpc.bg/en/homepage	10*
Portugal	Autoridade da Concorrencia	https://extranet.concorrencia.pt/PesquisAdC/Results.aspx?EntryClass=1	2
Sweden	Konkurrensverket	https://www.konkurrensverket.se/en/Competition/decisions/horizontal-anticompetitive-cooperation/	8
Hungary	Gazdasági Versenyhivatal	https://www.gvh.hu/dontesek/birosagi_dontesek/kereses-a-birosagi-dontesekben	19
France	Autorité de la Concurrence	https://www.autoritedelaconcurrence.fr/fr/liste-des-decisions-et-avis	11
Spain	Comisión Nacional de los Mercados y de la Competencia	https://www.cnmc.es/en/acuerdos-y-decisiones	17
Latvia	Konkurences padome	https://www.kp.gov.lv/decisions	23

We initially planned to categorize cartels by their strategies: whether they withheld bids, submitted fake bids, used subcontractors, divided markets by geography or product, etc. Unfortunately, this level of detail regarding the inner-workings of the cartel was rarely if ever reported in publicly available court documents.

Data matching

Linking entities that are explicitly mentioned in the cartel documents and suppliers from the public procurement records was a major challenge of the data preparation stage. As company names and addresses are not standardized in court documents and public procurement data, we had to clean company names and addresses and match them with the court cases with country specific codes. Company names and addresses – as published in the procurement records – could have either additional information that is irrelevant for the matching task or could include important information which was not represented consistently. For instance, one of the cleaning steps was to distinguish between company names, their legal forms, and other unnecessary information in this context. Some redundant information such as hyperlinks,

[29] Note that we could process the Hungarian cases ourselves, and we did not get back an answer from the French competition authority.

procurement-related terms, punctuation, and accents were removed. A great variety of legal form representation in the public procurement data has been simplified and standardized (e.g. for a legal form “LTD” we would account for variations such as “Limited”, “PVT Limited”, “PVT LTD”, “Private Limited”, etc.). The same had to be done for cartel case data.

We then applied machine learning methods using the Dedupe software library (Forest and Derek, 2019) to identify most likely matches in company names and addresses in the procurement data. The Dedupe algorithm is based on string metrics that represent the level of similarity between strings and performs a comparison field by field that allows treating differences in fields with individual weights (for instance, in this matching task we would want company names to be as similar as possible while allowing for a greater variation in company addresses). Once company names and addresses were standardized and matched in tender records, we manually identified and matched those to cartel members by searching for relevant company names. The algorithm implements active learning: by asking the analyst to manually verify a handful of potential matches that are difficult for the algorithm to distinguish, it learns the optimal subset of features to use in the deduplication, balancing precision and recall.

Indicators

Table 7.2 summarizes the testable indicators introduced above by our analysed countries.

TABLE 7.2. AVAILABILITY OF ELEMENTARY COLLUSION INDICATORS BY COUNTRY

Countries			Bulgaria	France	Hungary	Latvia	Portugal	Spain	Sweden
Price based	1	Relative price	Yes	No	Yes	Yes	Yes	Yes	No
Bidding patterns	2	Single bidding	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	3	Number of bidders	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	4	Consortia	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	5	Subcontracting	No	Yes	No	Yes	No	Yes	No
	6	Missing bidders (market)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	7	Missing bidders (buyer)	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Defining the consortia indicator

To find consortium bids, bidder names were cleaned by standard text cleaning steps (for example, removing special characters, setting them lowering case etc.). After cleaning, bids were classified as consortia by the following set of rules:

- A company name contains more than 1 country specific legal forms
- The company name contains the language specific words for consortium
- Company name that starts digits and matches the pattern: digit between 0-999 followed by a dot or a left bracket or colon or a semicolon or a space, which is followed by any three characters (including spaces) followed by at least three consecutive numbers
- Company name has more than 95 characters
- Company name has more than 1 dash, slash or percentage sign
- Company name matches the pattern: a semicolon or an ‘and’ (local language) followed by a space but not followed by any of the legal forms more than once
- In the original tender the contract was marked as consortia