Government Transparency Institute

Isabelle Adam [1], Mihály Fazekas [2], Yuliia Kazmina [3], Zsombor Teremy [4], Bence Tóth [5], Isabela Rosario Villamil [6], Johannes Wachs [7]

Public procurement cartels: A systematic testing of old and new screens*

Working Paper series: GTI-WP/2022:01

6 March 2022, Budapest, Hungary

[1] Government Transparency Institute

[2] Central European University and Government Transparency Institute: corresponding author misi.fazekas@gmail.com

[3] University of Amsterdam, Government Transparency Institute

[4] Government Transparency Institute

[5] University College London and Government Transparency Institute

[6] Central European University

[7] Vienna Uni. of Economics and Business

* The work was generously supported by the Swedish Competition Authority (grant nr. KKV 72/2018-Ab-1).

Table of Contents

List of figures

List of tables

Executive summary

Though cartels are thought to be common, they tend to be hard to find. Successful prosecutions are even more rare, and usually begin with an exceptional event: when a cartel member makes a sloppy mistake or decides to blow the whistle. Researchers have long studied these cases to learn how cartels function, and how collusive behavior sends signals in data. These signals are used to build cartel screens, methods to scan data on prices and activity for evidence of collusion. Even though these screens are theoretically sound and tend to work very well on the cases they are designed to highlight, less is known about their external validity. As competition authorities are collecting more and more data, there is a growing need to evaluate the general applicability of cartel screens.

In this report we collect data on 156 proven cartel cases in public procurement from around the world. We link the firms involved in these cartels to procurement market activity in 77 cases from six countries, covering a rich variety of sectors. We test the efficacy of a suite of cartel screens across these contexts. We could apply at least one screen to 47 cases. We come to several conclusions:

- Data quality and availability are severe hinderance for testing and applying cartel screens. This finding is especially disconcerting given that public procurement data ought to be transparent and accessible to citizens.
- Most cartels set off one or more screens in our toolkit, some set off as many as five screens. Specifically, 33 of the 47 cartels set off one screen, 13 set off at least 3 screens, and 5 set off 5 or more screens. This suggests that some cartels will be highly visible under a multi-screen microscope.
- No single screen works in even a majority of cases, reflecting the rich variety of cartel types operating in the economy and proving that only a multi-screen approach could lead to effective indicator-based investigations.

Reflecting on our findings, we suggest that with improved data and additional examples, machine learning methods such as ensembling, in this context the pooling of cartel screens into a composite index, could be applied to rate groups of firms for collusive potential.

A rigorous cartel screening program, applying multiple tested screens to clean data has the potential to be a gamechanger for competition authorities working in procurement. By monitoring signals of collusion, cartels will be forced to collude in increasingly sophisticated ways, increasing costs and adding uncertainty to the long list of challenges of coordinating a cartel.

Chapter I

Introduction

The total public procurement market in the EU – i.e. the purchases of goods, services and public works by governments (excluding public utilities) – amounts to about €2 trillion, or about 13 percent of total GDP (European Commission, 2016) (data for 2014). The share of public procurement in GDP is likely to increase further in the next years, due to increased state intervention in the economy and greater investment in health care.

However, anticompetitive behavior is a major problem in these public procurement markets. The extra costs that collusion imposes are borne directly by the state, hence the public. Given the large volumes of spending, even a small percentage increase in prices translates into substantial budgetary implications and welfare losses.

Massive amounts of data generated from procurement outcomes present competition authorities with opportunities and challenges. On the one hand, it is likely that groups of firms engaging in anti-competitive (collusive) behavior leave traces of their activity in the data. These signals offer investigators a novel source of valuable leads. On the other hand, combing through this data is time consuming and difficult, made even more challenging by systematic data quality problems including missing records and values, as well as incorrect data.

Based on a steadily growing literature on cartel screens, that is risk indicators which signal potential cartel behavior, there is quite some excitement about the potential values of such quantitative indices. The uses of cartel screens in competition policy and enforcement are three-fold: 1) They help identify general weaknesses in the competitive environment which calls for policy intervention. 2) Cartel screens help identify new leads for investigation or rank numerous investigative leads. 3) Cartel screens, in relatively few cases, can support ongoing investigation with additional evidence, potentially accepted as additional evidence by courts.

As a result of this excitement, a number of data-driven initiatives by competition authorities have sprung up, however they have also 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. Crucially, the economics literature proposing screens is very strong on the internal validity of the metrics but typically lack thorough checks on external validity and the discussion of the precise conditions under which the proposed indicators can be deployed at larger scales, for instance economy wide. Previous work generally leverages peculiarities of exotic auction types (for example average bid auctions), extremely high quality data (for example including independently sourced cost estimates to quantify markups), or studies cases in which firms deliver homogeneous, comparable goods, the price of which can be modelled with high accuracy (i.e. school milk contracts).

While these efforts can very effectively identify mechanisms and consequences of collusion in procurement, the methods they develop rarely generalize to other settings. They do not give competition authorities the tools they need to evaluate large amounts of activity for collusion. Moreover, cartels are diverse in their strategies, sizes, and durations hence competition authorities also need to be able to select among a battery of cartel screens and look across them systematically.

This report sports the ambition to fill this gap as much as possible within the constraints of the data at hand. We do so by using ‘real-life data’, that is datasets readily available in public repositories and government websites at large scale without laborious manual data collection. We compare and combine a wide variety of cartel screens suggested by the literature, but belonging to different types such as those based on prices or networks of bidders. We also compare across a large number of cartels in multiple countries in order to genuinely start to understand external validity of individual indicators as well as their combinations.

By implication, this report evaluates the validity and applicability of various generic or high-level cartel screens to identify potential cartels in large-scale public procurement data from France, Hungary, Latvia, Portugal, Spain and Sweden. We gathered judicial data on proven cartels from official government sources and combined them with micro-level public procurement data on calls for tenders, contract awards and whenever possible individual bids too. We then apply a variety of cartel screens from the literature, testing if they can detect the known cartels. In general, different screens are better able to detect certain kinds of cartels, following different kinds of anticompetitive behavior. No one screen works significantly better than the others in even a majority of cases, though small sets of screens together can cover many cases. Our findings hence represent the first step towards generalizing to many cartel screens across many countries.

An important finding of our investigation is that data quality in this domain is severely limiting analytical precision. In a majority of the countries we considered for inclusion in this report, the lack of reasonable quality procurement data, the lack of information on proven cartel cases in procurement or both blocked our efforts. Although competition authorities generally have access to proprietary databases with higher quality information, in most cases the required procurement data is not collected systematically (e.g. the information is scattered across hundreds of thousands of scanned pdfs), implying that even competition authorities face considerable data constraints. We discuss specific recommendations for improving data scope as well as quality, such as publishing information on all bidders and their bid prices, identifying firms and buyers via unique identifiers, and the inclusion of procurement announcement identifiers in cartel case decisions.

This report gives stakeholders in this domain an overview of what is possible when parsimonious cartel screens are applied to big data on public procurement markets. It finds both reasons to be hopeful about these approaches, for instance by showing that known cartels are often detected by simple screens, and areas where more work is needed.

The subsequent discussion is organized as the following: first, we set out the conceptual framework and the main methodology for the study. Second, public procurement data is discussed in dept for all 6 case study countries. Third, we set out the findings from the main analysis and discuss what they mean, going indicator by indicator while looking across all our countries. Fourth, we show the first results of our novel analysis of company ownership-based cartel indicators. This part of the report sits as a separate section because the logic of the analysis is different and so far we have only covered 1 country. Fifth, we discuss the implications and limitations of our findings while also looking ahead to future work.

1 Conceptual framework

Economists have long been interested in detecting signals of collusion in markets using data on firms, bids, and prices for decades. 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 market 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 1997, Pesendorfer 2000, Scott 2000). Localized markets for road construction and repair (Porter & Zona 1997, 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. The review of Harrington (2006) provides an excellent overview of these examples.

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

Of course, competition authorities are interested in casting a wider net and catching cartels working in more heterogeneous markets. Indeed the scope and scale of markets studied in papers on collusion has grown in recent years (Kawai & Nakabayashi 2014, 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 is 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, Imhof & Ishii 2020).

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 in a suite of indicators (extending earlier attempts by Toth et al. 2014). We borrow the term ensembling from the machine learning literature, which has long recognized that combining many weaker predictive signals can produce a much stronger predictive model (Breiman 2001). In this report we will 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. In this chapter, we introduce a) the collusion type based framework and b) the indicators tested, and c) connect indicators to the theoretical cartel strategies.

1.1 Cartel strategies

Collusion in public procurement markets 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 is based on 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. The first consideration is 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 on the tenders directly, rent reallocation has to happen in alternative ways. For example, they might subcontract each other or give informal side-payments.

The third dimension is the market structure that evolves 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.

All the possible combinations of a) elementary collusion techniques, b) rent allocation mechanisms, and the c) resulting market structures form a distinct collusion strategy (Table 1.1).[4] As strategies vary by these measurable dimensions, we suggest combining (grouping) indicators by these theoretical scenarios (see A-G columns in the accompanying indicator list). 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.

[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).

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, companies should be in a cut-point position (see discussion above). Fifth, 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. Sixth, procurement spending should become concentrated, a few companies (i.e. the ones in cut-point position) should have increasing market shares.

TABLE 1.1. 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)

1.2 Indicator descriptions

This section briefly introduces all tested indicators. Table 1.2 lists all potential indicators with their brief definitions. Note, that indicators 1-3 are only calculable for one Swedish cartel as losing bid prices are not published in most countries, and the quality of bid prices in the Swedish data is also limited (i.e. the share of missing data is high), hence we only report the results in the Appendix for these three indicators. We explain the logic behind each indicator below (sections 1.2.1 – 1.2.3).

[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.

TABLE 1.2. TESTED INDICATORS

Category	Nr	Indicator	Level of observation	Description
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)
	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
	3	Relative standard deviation	Contract	Relative standard deviation of bid prices is less than 10% or more than 90% of the distribution
	4	Relative price	Contract	Ratio of the final price and the estimated price
	5	Benford's law	Market-period	Whether first digits of contract prices of a given market in a given period follow Benford's law
Bidding patterns	6	Single bidding	Contract	Contract receiving a single bid during the tendering process.
	7	Missing bidders	Company-period	The number of unique buyers/cities/regions/product codes companies submitting a bid at.
	8	Subcontracting	Contract	Whether a contract has a subcontractor.
	9	Consortia	Contract	Whether the winning bid was a consortia.
	10	Cut-point position	Market-period	Whether there are companies in a cut-point position in a given market and time period.
	11	Winning probability	Market-period	The average winning probability of companies of a given market and time period
Market structure	12	Concentrated market structure	Market-period	HHI change from during to after the cartel period
	13	Stable market structure	Market-period	Average absolute market share changes during vs. after the cartel period

1.2.1 Price-based indicators

In well-established competitive markets where companies regularly bid for similar contracts, bid prices are expected to randomly fluctuate around the market price with relatively few outliers. However, if companies coordinate their bidding behaviour, they are likely to leave traces in their bid prices, hence the bid price distribution can be used to identify cartels. Variance, range and skew can each signal a behaviour that is at odds with genuinely competitive behaviour (Fazekas - Toth, 2016).[7]

Relative price

Relative price is defined as the final contract price divided by its initial estimate.[8] Healthy competition is expected to lead to lower prices (i.e. bigger discounts) compared to the initial estimate, hence we expect that lower relative prices are good proxies for competition. As collusion is about generating rents - either through higher prices or lower quality - an increase in relative contract values is an expected 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 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).

Benford's law

[7] Extreme or unusual offer price distributions are found to signal collusion by academic literature Abrantes-Metz et al. (2006), Oxera (2013), Padhi and Mohapatra (2011). A number of characteristics of bid price distributions can be used to identify cartels, each of which follow a similar theoretical reasoning while being formulated in different ways: a) relative difference between the first and second lowest bid prices, b) relative standard deviation of bid prices, c) relative bid price range, d) difference between lowest and second lowest bid.
The difference between the lowest (winning) and second lowest (best losing) bid prices can capture artificial bidding patterns. Based on Abrantes-Metz et al. (2006), Oxera (2013), Padhi and Mohapatra (2011), both extremely small and large differences between the lowest and second lowest offer prices can signal collusive behavior. Another approach is the identification of suspiciously rounded bid price values. For example, the winning bid is strictly 10% less than the second lowest bid. Although it can be observed in competitive bidding by chance, a consistent difference suggests that cartel members agree on the exact bid difference in advance. We tested the second approach in the report. Thus, we transformed the indicator into a binary variable which has value 1 if the relative price difference is close to 5%, 10%, 15%, 20%, 25% or 30%, and 0 otherwise.

[8] 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.

The other indicator is Benford's law, which is a statistical rule commonly used in forensic accounting, election monitoring, and in the study of economic crimes including collusion and corruption (Berger and Hill, 2015). Benford's law posits that the first, second etc. digits of most naturally occurring sets of numerical data follows a specific pattern[9] (Fewster, 2009). In public procurement markets, for example the first digits of bid prices observed on a specific market ought to follow Benford's law as they are a result of a natural competitive process. However, if bids are generated in an artificial process (for example, cartel members deciding losing bid values in a bid-rigging scheme), they would not follow this expected distribution. Hence, fake bids can manifest in the distribution of these digits in a way that violates Benford's law.

1.2.2 Bidding patterns

Single bidding
Withholding bids is one of the most straightforward ways to rig a tender, which results in a 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 single-bidding as a collusion indicator is 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 - instead, it might capture other anti-competitive behaviours. 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' 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 bidders indicator, similarly to single-bidding, captures how colluding companies withhold their bids from specific parts of the market. Withholding bids from certain tenders lowers the costs of a cartel maintenance. This indicator captures the rotated bidding of firms therefore a lower number of bids are expected during the cartel period for the participating companies. This technique is also commonly cited as a possible cartel strategy (OECD 2014; SCA 2015).

Companies can withhold their bids by several dimensions that can ease the coordination: from specific sub-markets (e.g. based on CPV codes), from specific (group of) buyers, or geographical location.
Therefore, we test this indicator using several definitions. First, we calculate the number of unique markets companies submit a bid (or win a contract) in the collusive vs. competitive group of contracts.[10]

---

[9] The proportion of 1, 2, 3 etc. numbers as first digit should be proportional to the logarithmic difference between them.
[10] We analyse three different versions (2-, 3-, and 4-digit CPVs).

Second, we calculate the number of unique NUTS codes (2-digit) and buyer cities – i.e. this would capture whether colluding companies withhold their bids from certain geographical areas. Third, we calculate and compare the number of unique buyers between collusive and non-collusive group of contracts.
As we test the difference either during vs. after the cartel period, or during the cartel period, we expect that the number of unique CPVs, locations or buyers increase after the cartel period ends, and similarly, we expect that companies not involved in a collusive scheme during the cartel period are bidding on more unique CPVs, location or buyers.

Winning probability
Companies' winning probability is another indicator that can be used in triangulating a strategy based on withholding bids. If public procurement markets are competitive, companies having extremely high (close to 100%) winning probabilities over a long period ought to be a sign of a competitive anomaly, that can be an outcome of an collusive agreement (Fazekas and Tóth, 2016; Harrington and Joseph, 2005). We define winning probability as the share of contracts companies won out of all bids they have submitted on a given set of contracts (either collusive or competitive). Companies with a winning probability significantly higher based on the collusive contracts vs. with a lower probability in the competitive comparison group would be regarded as a confirming indicator – i.e. high winning probability is an outcome only due to coordinated bidding.[11]

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 the companies lower the 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[12], 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
The division of rents between cartel members is a challenge (Asker, 2011), transferring rent 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.[13]

Cut-point companies
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

[11] As data on all winning and losing bids are required for this indicator, we can only calculate it in Portugal, Sweden, and Hungary, where data on losing bids is available (with limitations).
[12] 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.
[13] Note, that the indicator could be calculated in the following countries: Latvia, Spain, and France.

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.
A network-based indicator investigated in this report is whether companies of interest - cartel members - are positioned as cut-points in a network. 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 (Wasserman 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.

1.2.3 Market structure

The association between collusion and market structure is ambiguous. On the one hand, explicit market division with relatively high market shares can emerge (Levenstein & Suslow, 2006; Pesendorfer, 2000). On the other hand, a cartel can imitate a competitive market structure if members agree on a given winning order leading to an artificially stable market (Athey and Bagwell, 2001; Athey et al., 2004; Mena Labarthe, 2012; Harrington, 2006). As cartel members have short term incentives to defect, cartels often adopt sophisticated agreements and decision-making processes to determine who should win which contract (Asker, 2010). These agreements tend to create an artificial stability in market shares, relative to market outcomes. Both concentration and stability are less likely to persist over long periods in a functioning market environment. Our definition of market structure-based indicators reflect these two ideas: a) a decrease in market concentration after the cartels' (assumed) collapse, and b) more stable market shares during vs after the collusive period. We only carry out during vs. after tests and assume that the cartel members define the whole relevant market.[14] Furthermore, these indicators are point estimates without confidence intervals since we have one market in each period per cartel.

Concentration
Concentration in a public procurement market refers to a situation in which few companies win many contracts while competing bidders are either entirely absent or only mimic participation by submitting fake bids. Having a concentrated market in the first place makes it easier for cartels to form. Potential cartel participants even have the incentive to buy up non-collaborating companies (Levenstein & Suslow, 2006), which manifests as higher concentration. Alternatively, market concentration increases if tenders are serviced by the most efficient companies (though not competitively) to reap the largest profits possible (Pesendorfer, 2000). By implication, an indicator built on concentration should be defined with reference to a competitive baseline, and a sudden increase or decrease in concentration can be the sign of a collusive practice. A clear-cut situation when concentration signals collusion is when a particular market turns from competitive to a concentrated one in a short period of time without

[14] See more in section.

any apparent alternative explanation such as changing regulations, technology, firm exit, or steep decline in total demand.

As the intersection of proven cartel cases and available public procurement data does not allow us to track down the increase in concentration for most markets, we analyse the change in concentration – an expected decrease – after the cartel period has ended. We analyse concentration at the level of proven cartel members (i.e. we assume them forming a complete market that is rigged), using the Hirschman-Herfindahl Index[15] (HHI) that can be formulated as the following:

$$HHI = \sum_{i=1}^{N} s_i^2$$

where s_i is the market share of company *i* on the market, where the whole market is formed by only cartel members, and N is the number of cartel participants. Our expectation is that concentration is higher during the cartel period. Significant HHI change is considered if the drop in HHI is more than 250, following the European Commission's guideline.[16] We only calculated these HHI values for cartels that have at least 2 companies winning at least 3 contracts altogether both during and after the collusive period.

Stability

An overly stable market structure, indicated by low variance of market shares of participating firms, can also suggest collusion risk. Athey and Bagwell (2001) and Athey et al. (2004) show that following a market share rule for allocating rent can be also an optimal way of allocating rents in collusion. Regarding empirical studies, Pesendorfer (2000) shows that if bid-rigging is used as a rent reallocating mechanism instead of side-payments, relatively stable market shares can be observed. Mena Labarthe (2012) also shows that the market shares of the colluding parties were practically the same in the collusive period. Furthermore, Harrington (2006) suggests two relevant collusion indicators based on market structure: highly stable market shares over time and highly stable market shares of a subset of firms.

Our stability test captures the idea that the market share changes during the collusive period are noticeably smaller compared to the period after. Th indicator is defined as one minus the average absolute change in market shares:

$$Stability = \left(1 - \frac{\sum_{i=1}^{N} |S_{i2} - S_{i1}|}{N}\right) - \left(1 - \frac{\sum_{i=1}^{N} |S_{i4} - S_{i3}|}{N}\right)$$

where s_{it} stands for the market share of the company *i* in the *t* period, N is the number of participants on the cartel market. The lower the indicator value is, the less stable the market is. For the calculations we split the analysed time period to four 2-year long intervals – two before and two after the cartel period. We compare the stability change within the during vs after the cartel period. To illustrate how this indicator definition would work, assume a market of 6 companies with the market leader having 50% share and the five others having 10% each. A 3.3% decrease in the market leader's share and a corresponding increase 2 other companies' market shares[17] would be exactly a 250 HHI change (the threshold we used for concentration), that would correspond to a 1.1% decrease in the stability (e.g. either the change from 1 to 2 or to 3 to 4 time periods). We count a test as confirming when the stability change is bigger after the cartel period than during collusion. We only calculated the stability indicator for cartels that have at least 2 companies winning at least 3 contracts altogether in each 4 time periods assessed.

[15] We calculate market shares for each company based on contract values on intersecting product markets that are transformed to Hirschman-Herfindahl Indices (HHI).
[16] See in the Official Journal of the European Union under the 20. paragraph.
[17] In the proportion to 2 to 3.

1.3 Indicators by cartel strategies

Elementary indicators individually, however, cannot capture the complexities of various cartel strategies. Instead, we analyse the co-occurrence of individual indicators by cartel types, that can unfold whether the administrative footprints of different cartel strategies point towards the same direction (Table 1.3). For example, cartel type B is based on participants submitting losing bids, uses subcontracts to share profits and eventually leads to a concentrated market structure. We expect that the bid price distribution (1) and the contract value (2) of rigged contracts will differ significantly from competitive tenders[18]. We also anticipate distinctive bidding patterns such as having a higher share of subcontracting (8) or companies being in a cut-point position (9). Finally, we also expect that the market will be more concentrated (10). Therefore, we do not expect all elementary indicators to work for all cartels, on the contrary: we anticipate a combination of them to signal collusion, with certain groups of screens working more effectively for certain organizational forms of cartels. The two theoretical exceptions from this rule are Benford's law and relative price, that should flag all cartels with enough contracts (for example, Benford's law does not apply if only a few dozen contracts are available).

The co-occurrence of these indicators suggests that cartel strategies do not vary over time. However, strategies can vary even if renegotiating the terms of collusion is costly, hence cartels having a pair of indicators that are theoretically exclusionary – for example, finding single bidding and companies in cut-point position at the same time as red flags – is possible. If companies change strategy too often, indicators might not work – as no significant difference could be identified in them.

TABLE 1.3. COLLUSION TYPES AND INDICATORS

Indicator group	Nr	Indicator name	Collusion types						
			A	B	C	D	E	F	G
Prices	1	Relative price							
	2	Benford's law							
Bidding patterns	3	Single bidding							
	4	Missing bidders							
	5	Subcontracting							
	6	Consortia							
	7	Cut-point position							
	8	Winning probability							
Market structure	9	Concentrated market structure							
	10	Stable market structure							

[18] It can go either way: it might show extremely low or high variation depending on the agreement details.

2 Methodology

This section explains our testing approach by discussing all main decisions concerning the definition of rigged vs. control contracts. First, we discuss the dimensions we used for identifying the relevant rigged contracts and our approach for defining the group of an appropriate control for each test. Second, we define the exact definitions of rigged vs. control contracts for all tests we used in this report. The overarching goal was to create a control group of contracts that was both similar enough to the rigged group and large enough to enable comparisons in the three implemented tests.

2.1 Finding rigged contracts

For all tests (during-after, cross-sectional), our focus was to precisely select a) rigged contracts (treatment) and b) an appropriate set of competitive ones (control group). First, we needed to find all tenders that can be marked as rigged based on the corresponding court decision. In an optimal scenario, each case should have a list of unique contract identifiers based on which rigged contracts could be unambiguously labelled. However, court documents rarely contain such technical details. For the vast majority of cases in our initial search and beyond we could not find such lists.

Therefore, we needed to identify the rigged tenders in a probabilistic way based on the tender characteristics mentioned in the legal texts. We have considered the following dimensions: a) company names, b) proven cartel time period, c) product market, d) cartel location, e) affected public buyers. In theory, the intersection of these dimensions should identify all rigged contracts. However, this matching process revealed that using all these dimensions would either leave us with a) an extremely small number of affected tenders or b) some dimensions would not be consistently available across tenders. In the former case we would likely be underestimating the tenders affected by collusion.

In the final tests reported in the main text, we marked contracts based on the **company names** mentioned in the court documents for all cartels (Panel A of Figure 2.1) – see more details on name matching in the Data section. The **active cartel periods** were available consistently in most countries at least per month[19] (Panel B). To identify contracts during the proven cartel period we used the publication date of the call for tender announcements or if it was missing, we have imputed an estimated start date for the procedure.[20]

We calculated each test considering only contracts from those product markets (based on CPV codes) where at least 2 cartel companies submitted a bid.[21] However, we do not report the results based on this restricted sample. Regarding the cartels' location and affected public buyers (d,e) we either have

[19] See more details on cartel periods in the section 3.1 Proven cases.

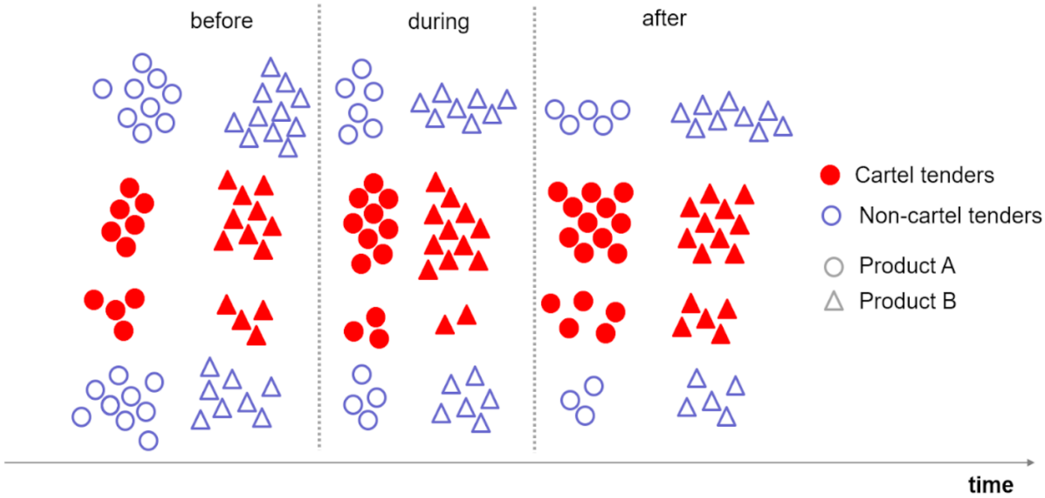
[20] For estimating the start date of a tendering procedure, we have calculated the median difference between call for tender and contract award date publications per each 4-digit CPV code product markets (https://simap.ted.europa.eu/cpv) that are included in the procurement data.

[21] While some court documents mentioned product markets, they were not referring to the product codes included in the procurement dataset (CPV nomenclature) but were only textual descriptions. Therefore, we have decided to create a product market filter that is only based on procurement records. We assumed that while colluding companies might be active on multiple product markets, their collusive behaviour would focus on those markets where at least two of them had won a public contract regardless of the time period. For example, if a cartel has two companies and one was winning contracts on market A and B, while the other on markets A, B and C, then only the intersecting A and B markets are marked as relevant from the cartels perspective (Panel C). Note, that marking rigged tenders by markets this way assumes that the collusive strategy itself was unrelated to the product markets themselves. Furthermore, depending on how detailed product market codes are, the number of intersecting contracts will differ. Nevertheless, we marked these intersecting product market wins based on 4-digit product codes in each country, assuming that they represent a reasonable outline of the cartel market.

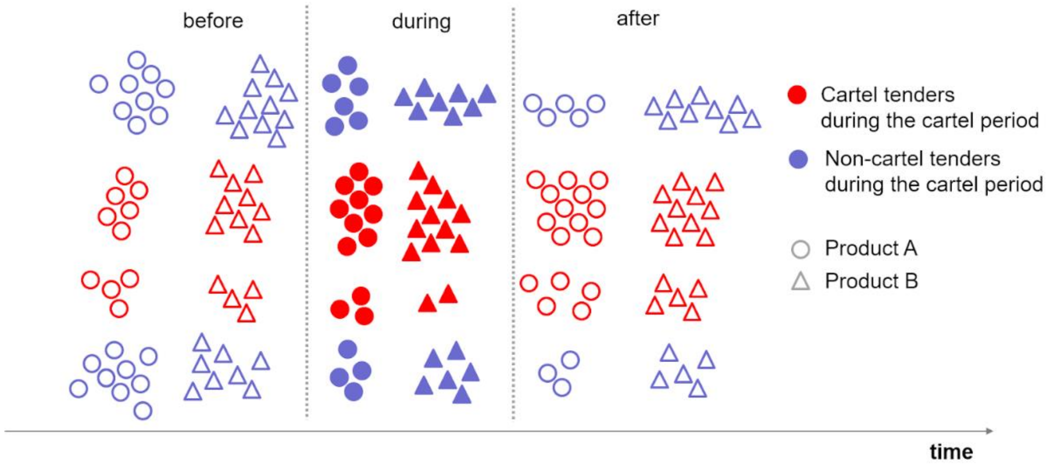
not been able to find consistent mentions in the court documents, or the intersection of company names with buyers or locations were either outside the proven cartel time-period or returned no contracts at all[22].

FIGURE 2.1. TRIANGULATING RIGGED TENDERS

A: Company names



B: Time period



Consequently, we used the first two dimensions - a) company names, b) cartel time period - to define the set of rigged contracts for the tests reported in the main text (see the discussion below on the exact combination of these for the final tests). Due to the challenges, we have faced during identifying even these ground truth cartel cases, our working assumption is that the proven cases should be interpreted as a lower-bound estimation of the real scope of each proven collusive agreement (see Levenstein and Suslow, 2011 and Harrington and Chang, 2009 for discussions on estimating cartel activity). Already by relying only on the two dimensions and disregarding product market, location or buyer names, we loosen up the set of tenders strictly defined by the court documents. However, even this lenient take on the cartelized tenders often led to very few treated contracts due to the extremely short cartel periods documented.

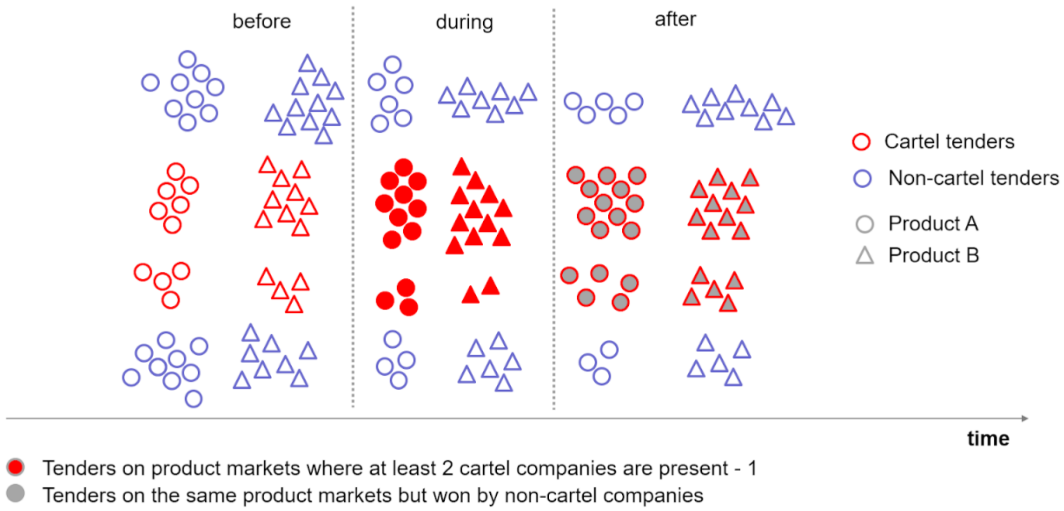
[22] Note that public procurement documents contain information on the buyers and the implementation locations, however, these are often not detailed enough - for example, implementation location is simply the whole country instead of a specific city or region.

2.2 Testing elementary indicators

Identifying rigged contracts is only the first step for indicator testing, we also needed to find a group of control tenders that form the comparison group for statistical testing. In the main text we report two tests on most elementary procurement indicators: a) during-after comparison, b) cross-sectional comparison. While these can be calculated for contract-level indicators (provided that we have enough contracts in the respective comparison groups), we apply a different logic for those that are defined at a more aggregated level, such as winning probability or market concentration or the co-bidding network based indicators (see below after introducing the tests used for contract-level indicators).

Our first test is a simple before-after comparison of the indicator values based on the subset of contracts won by cartel companies. We compare the outcome indicators of contracts won by the collusive companies during the proven cartel period (filled red symbols) to those won by them after the cartel period (filled grey symbols) - Figure 2.2. This comparison is straightforward for indicators that can be defined at the contract level with either a t-test or a proportion test. For example, we can compare the average share of single-bidder contracts, the share of contracts with extreme bid price distributions with a proportion test, while the average number of bids with a t-test of contracts during (red filled) and after (grey filled) the cartel period. 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.

FIGURE 2.2. DURING-AFTER COMPARISON[23]



Our second test is a cross-sectional comparison of contracts won by the cartel members (red filled circles) vs. non-cartel members (grey filled circles) during the cartel period (Figure 2.3). In order to find an appropriate comparison group, we restrict the data in two steps. First, we only keep those companies that submitted bids in greater than 50% on CPV codes where at least two cartel members won a tender. This excludes companies with a bidding profile that significantly differ from the cartel companies' – for example, if a company submitted a competitive bid but usually it does not operate on the product markets or an administrative error explains the product code that does not correspond to the actual market. Second, we apply coarsened exact matching (CEM) to identify tenders with similar market characteristics. In step one, the variables on which the matching is made are coarsened (transformed to discrete categories). In step two, all exact matches are made using the matching variables. CEM can be summarized in the following 3 steps (Iacus et al., 2009; Iacus et al., 2012):

[23] Each symbol represents an awarded contract.

1. It categorises all treatment and control contracts of the same contract value category (3 different contract sizes and a missing contract price category), 4-digit CPV code and tender year into groups.
2. It filters out contracts from any group that do not include at least one treated and one control unit[24].
3. Calculate the average effect on group level and take the average of averages[25].

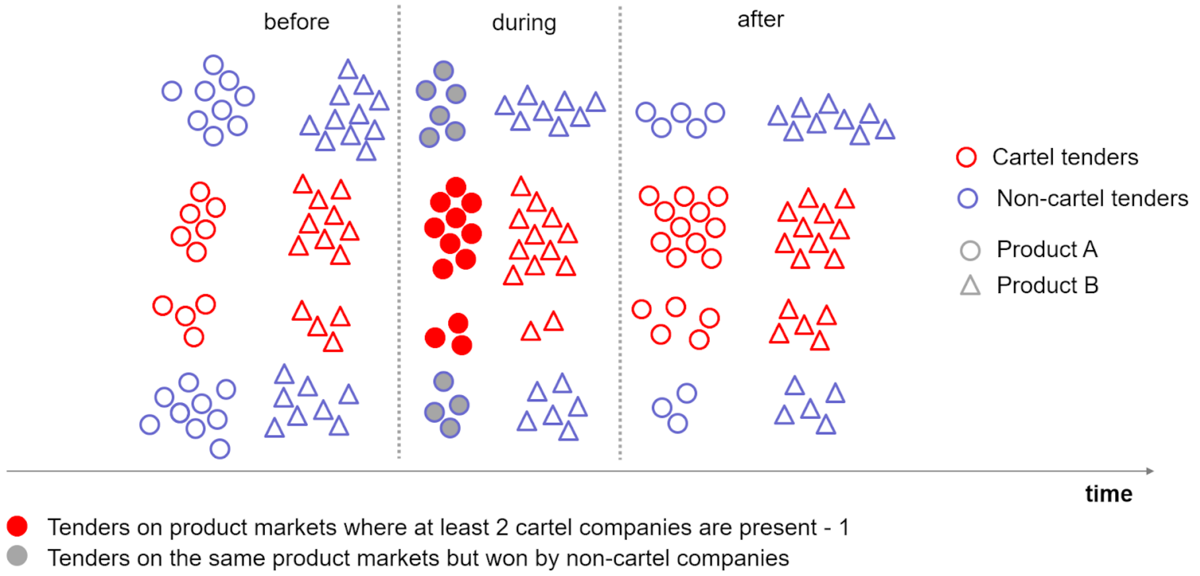
The caveat of matching is that if the algorithm cannot find matches to treated observations, we can lose many cartel contracts. To address this issue, we run a second CEM specification with less restrictive matching dimensions (i.e. and alternative to step 1 above): we use 2 contract size categories instead of 3, 3-digit CPV code categories, and distinguish between during vs. after the cartel period. The looser matching takes place if the restrictive one cannot find any control contracts to at least 10% of the collusive ones. Similarly as for the during-after comparison, we only calculate t-tests for cartels that won at least 6 contracts during the cartel period with at least 6 matching contracts in the comparison group and the indicator value varies.

Note, that cross-sectional matching is ultimately related to the question of how the relevant market (from the cartel point of view) is defined. First, this restriction might underestimate the real extent of the market targeted by the cartel. If companies split contracts by products (and available product codes capture these product variants accurately), then we should indeed pool together markets A and B. Second, this comparison relies on the assumption that the cartel does not cover all contracts of a particular product market and parts of market A are still behave competitively and won by outsider companies (grey circles). For some of the cartels one or neither of these assumptions hold - hence cross-checking the results with for example the during-after comparison, that does not rely on these assumptions is key. We discuss these issues in more detail in the Limitations section.

[24] Test of the following indicators stops here since sample reweighting would be meaningless due to the screens not on contract level: Benford's law, cut-point position, winning probability, concentrated and stable market structure. The algorithm drops out treated observations that do not have control counterpart(s) along match variables.

[25] Practically, the CEM function returns a weight vector and a selected set of observations leading to a weighted t-test or a proportion test. To create the contingency table for the proportion test, every observation worth its weight, hence there can be non-integer values in the table.

FIGURE 2.3. CROSS-SECTIONAL COMPARISON[26]



We compare contract-level indicator averages with simple t-tests or proportion-tests for the during-after and cross-section setup. We applied a 5% and 10% significance level threshold. For example, if the share of single-bidder contracts was 50% during the cartel period but only 10% afterward, and the difference is statistically significant, we take it as a confirmatory test result.

Exceptions

The contract grouping introduced above covers contract-level indicators, we had to use a different logic for those indicators screening for cartel activity at a different level (e.g. company- or market-level).

As Benford's law is defined on a group of contracts, we simply compared the collusive contracts with non-collusive ones either from cartel companies' contracts from after the cartel period or similar contracts during the cartel period. We interpret cases as confirmatory if final prices of collusive contracts do not, while the comparison group do follow Benford's law.

For missing bidder, we cannot apply a statistical test either as we count the unique number of buyer cities companies are bidding at according to their cartel status (i.e. cartel companies during vs after or cartel companies). Therefore, we accept a test as confirming if the unique number of cities companies are bidding at is lower on average based on the collusive contracts vs the control group.

While winning probability would suggest a company-level indicator, we tested it simply by pooling together winning and losing bids of rigged vs. competitive tenders. For example, we calculated the joint winning probability of all cartel companies during the cartel period and compared them with the one they had after the cartel period. We could then simply identify significant differences in winning probability based on proportion tests and the difference-in-difference estimates. Unlike for contract-level tests, we calculate the tests if there is at least 6-6 bids (i.e. not 6-6 awarded contracts) in the treatment and control groups.

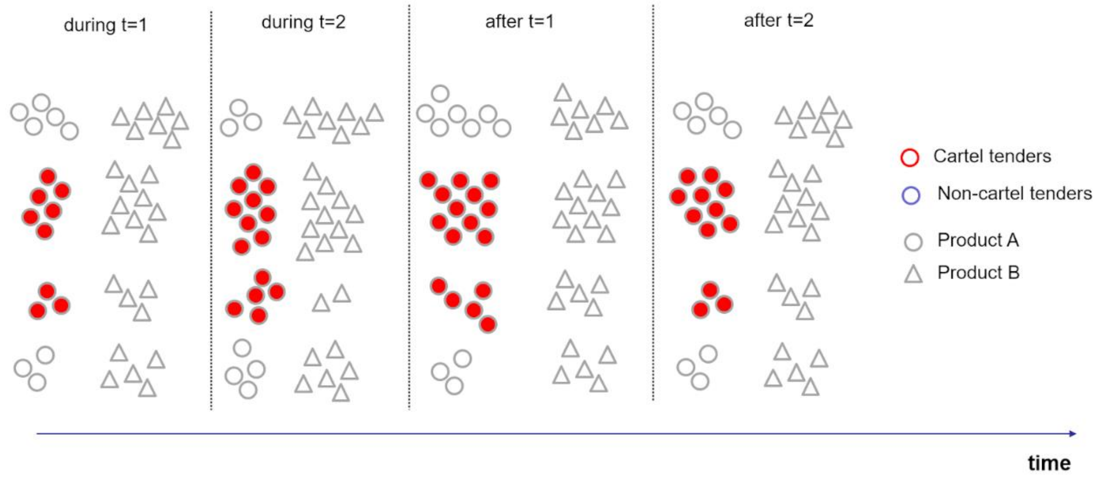
Companies being in a cut-point position is also an exceptional situation, as the indicator is only meaningful if we compare collusive vs. non-collusive groups of contracts. Our indicator is confirmatory

[26] Each symbol represents an awarded contract.

if there is at least one company among the cartel companies during the collusive period in a cut-point position, whereas there are no such company position in the comparison group.

We use two measures to quantify the distribution of contracts among firms in a market - the Hirschman-Herfindal index (HHI) and the average absolute change in market shares - to assess the change in concentration and the stability of market shares, respectively (see section 1.2.3). For these indicators we assume that the cartel covered the full set of relevant contracts on the market during the cartel period (red filled circles of in Figure 2.4). We take a concentration test confirming if the HHI decreases at least by 250 points, while for stability if the stability indicator decreases by at least 1.1%.

FIGURE 2.4. MARKET CONCENTRATION AND STABILITY



2.3 Combining elementary indicators

As cartels are diverse and public procurement datasets are generally noisy (e.g. missing, erroneous data), it is highly likely that elementary collusion risk indicators are highly 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 or screens or into a composite risk score (e.g. Huber and Imhof, 2019). In the literature so far, 3 main approaches to combining elementary collusion risk indicators have been used:

- Naïve equal weights;
- Theory driven indicator combination (sequential or composite score approach); or
- Machine learning-based.

In a naïve equal weight approach, all possible or seemingly valid elementary risk indicators are averaged over and the resulting average incidence of risk factors provide the best estimate for the overall risk of collusion. This approach is generally considered inadequate as it ignores the diversity of cartel strategies which are likely to lead to opposite risk profiles such as high degrees of single bidding with bud suppression or low levels of single bidding with cover bidding strategies.

Theory-driven approaches come in two variants. In the first approach, they combine indicators in sequence, applying them as independent tests where only those cases (markets, companies, etc) are considered risky which pass all tests rather than only some. A typical example of this is demonstrated by Tóth et al (2014) who look at Hungarian public procurement cartels by, first, defining competitive benchmark indicator values then exploring the co-occurrence of risk indicators by cartel type one by one. They look at road construction submarkets defined by regions in Hungary and consider those markets as likely suffering from bid suppression-type cartels which simultaneously have very high HHI, a big increase in HHI over time, and a high relative price of tenders. In addition, some of these markets have a high prevalence of cut-point position bidders too.

In the second theory-based approach, indicators related to the same type of collusion are not considered in sequence rather combined into a score by collusion type (Tóth et al, 2020). In this approach, the risk of each cartel type in Table 2.1 (e.g. type A using withheld bids and subcontracting resulting in a concentrated market structure) is calculated based on as many indicators as can be calculated related to that type. This collusion type by collusion type calculation follows a simple averaging of the corresponding indicators. Then each observation (contract, company, or market) receives the highest score of among the calculated different collusion types.

Supervised machine learning-based approaches, in contrast, make use of a set of known cartel and non-cartel cases, typically by labeling contracts won by cartel and non-cartel members as the outcome variable. An algorithm then learns from a set of collusion risk factors and control variables, serving as predictors or features, how best to predict that label (Huber et al, 2020). Among the various algorithms available for building such predictive models, those are selected which achieve the highest precision on an unseen, test dataset (that is dataset which was not used to fit the model). In the context of fraud, corruption, and collusion, various studies have used the random forest algorithm. We also use this method because of its ability to model a diverse array of different collusive strategies and the markers they leave. Random forest is a supervised machine learning method which predicts the output by constructing multiple decision trees with given features (Breiman, 2001). It is particularly well suited for datasets with many explanatory variables or potential risk indicators and where the same outcome may be the result of multiple different combinations of predictor values (James et al, 2015). In spite of its flexibility and suitability of the complex prediction problem we aim to develop, Random Forest models lead to results which are hard to understand or interpret within our theoretical framework. In other words, in order to achieve high prediction accuracy, we may have to sacrifice some degree of interpretability. We will get back to this caveat in the results section.

2.4 Predicting cartel risks

Once the optimal model is identified using the test-train samples made up of proven cartel and non-cartel cases, it becomes possible to make predictions to the full universe of contracts in a country. 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 below and the use of standard public procurement datasets harnessed from government publication portals, most of these preconditions are largely met. However, the different elementary collusion risk indicators are defined on different levels of observations such as contracts, companies or markets. On higher levels of aggregation such as markets, it is not straightforward how the indicators defined for a particular collusive ring can be calculated for the whole procurement market as it requires a suitable market or sub-market definition. As the authors have discussed elsewhere, it is possible to define suitable market IDs, for example based on procurement classifications such as CPV (Common Procurement Vocabulary) and NUTS (Nomenclature of Territorial Units for Statistics) (Fazekas and Tóth, 2016). We will consider this topic and approaches in the further work section below.

3 Data

In this section we give a bird's eye view of the data we collected for this report. We outline our search for data on proven cartel cases (3.1), give a high-level overview of the procurement data we used (3.2), and then explain how we matched proven cartel cases with public contracting datasets (3.3). Finally, we present summary statistics by countries that show the number of matching contracts by the different tests we implement (3.4).

3.1 Proven cases

Recognizing that there has been little work to date collecting a large number of example cartels in procurement, in this subsection we report on how we screened for proven cartel cases and selected the countries for analysis and provide summary statistics on these bid-rigging cases we found.

3.1.1 Country screening and selection

As a first step, we screened several countries with sufficient data quality for the analysis. We used two main criteria for choosing the most promising ones: availability and quality of public procurement data (especially the availability of bidder information) and the number of proven cartel cases that overlap with procurement data. As a result, we grouped 19 potential countries into three priority groups.

European countries with high quality and detailed procurement data were in priority group 1 regardless of their size (see also next section). These have data on all bidders, as well as proven cases of cartels in public procurement. Larger countries with higher shares of missing data or missing losing bidders were in group 2. Countries with high quality procurement data but no proven public procurement cartel cases and the ones for which we could not clearly confirm the existence of proven cartel cases were in group 3. In addition, we considered a number of non-European countries with high quality procurement data available or a number of proven cases, as a back-up in case that the original list of countries should be expanded. The final country screening and resulting grouping is summarised in the table below.

Based on this screening process and additional data quality checks - e.g. availability of long enough historical data, availability of key variables[27], we shortlisted six countries for the current report, namely: **France, Hungary, Latvia, Portugal, Spain, and Sweden**[28].

[27] For example, we analysed the availability of company names, bidder numbers, dates, that are all key for the analysis.

[28] Note that we identified a couple of key data errors in Lithuanian public procurement data that we plan to fix for an updated report.

TABLE 3.1. COUNTRY SCREENING RESULTS

#	Country	Procurement data collected?	Losing bids available?	Number of proven cartel cases collected	Priority group
1	**Portugal**	yes	yes	2	1
2	**Hungary**	yes	yes	15	1
3	**Lithuania**	Yes (data fixes needed)	yes	13	1
4	**Sweden**	yes	yes	6	1
5	**France**	yes	no	11	2
6	**Spain**	yes	no	17	2
7	**Italy**	yes	no	6	2
8	**Latvia**	yes	no	23	3
9	**Czech**	yes	no	15	3
10	**Slovakia**	yes	no	?	3
11	**Bulgaria**	Yes (low quality)	no	?	3
12	**Poland**	yes	no	?	3
13	**Estonia**	yes	no	4	3
14	**Georgia**	yes	yes	0	3
15	**Mexico**	yes	yes	4	back-up
16	**Chile**	yes	yes	4	back-up
17	**Paraguay**	yes	no	0	back-up
18	**Peru**	no	?	6	back-up
19	**Brazil**	no	?	30	back-up
Total number of collected proven cases:				156	

3.1.2 Cartel case collection

We collected information on the proven cartel cases manually from country specific sources of court rulings by following three search strategies.

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", "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 the process described above, 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 3.2. OVERVIEW OF SOURCES FOR CARTEL CASES[29]

Country	Competition authority	Case repository resource	Number of PP-related cases
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 parsed cartel-level information manually into a data template we developed for storing all relevant information of the case documents[30]. 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[31], information related to the relevant public tender(s) (e.g. tender IDs, product types), and the location. We matched the resulting case dataset to the public procurement data of the respective country as explained in 3.3.

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.

Most proven cartels from the six analysed countries (see above) consist of a relatively small number of firms, see Figure 3.1. The distribution of cartel durations, thought in the literature to be bimodal (Levenstein & Suslow, 2006), is reported in Figure 3.2. We do not observe any bimodal tendency. These results must be interpreted with caution: survivor bias suggests that the true distributions of cartel sizes and duration likely differ from the ones we report. For instance, the chance of a whistleblower emerging may be larger in a cartel with many firms and actors. At the same time, a whistleblower may be more likely to come forward towards the end of a cartel's natural life cycle. In simple terms, we are studying those cartels which have been caught and brought to trial. We also show the relationship between the number of cartel members and cartel length per country (Figure 3.3). Surprisingly, Spanish proven cases seem to be much longer than other countries, and involve significantly more companies as well.

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

[30] 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).

[31] 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.

FIGURE 3.1. DISTRIBUTION OF CARTELS BY CARTEL SIZE

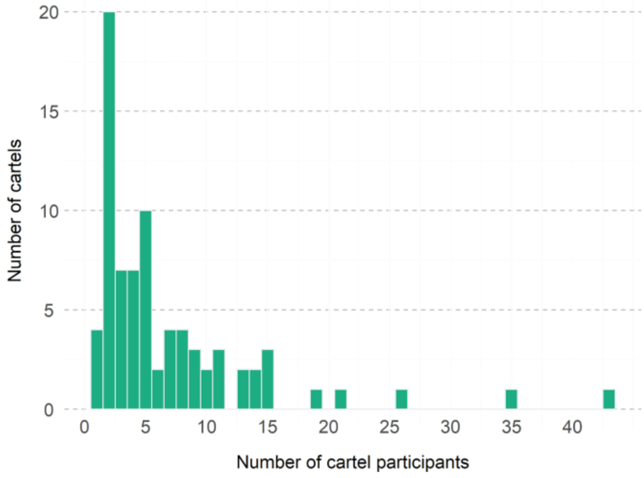


FIGURE 3.2. DISTRIBUTION OF CARTELS BY DURATION

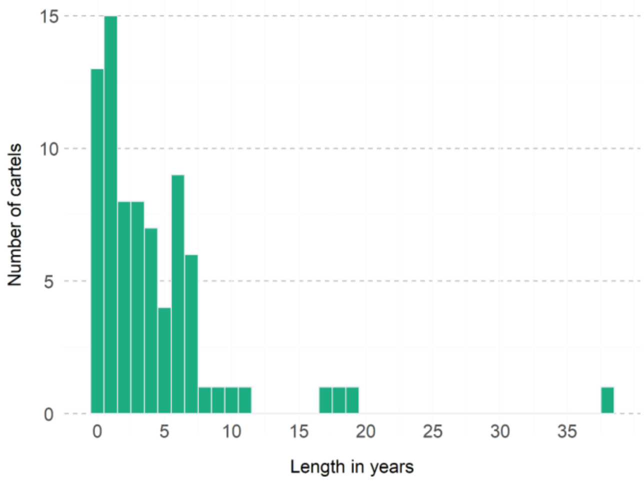
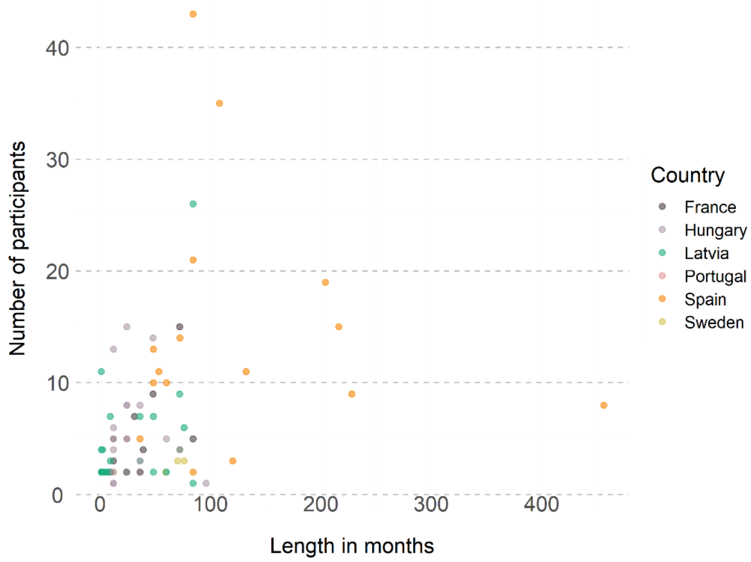


FIGURE 3.3. RELATIONSHIP BETWEEN (PROVEN) LENGTH OF CARTELS AND THE NUMBER OF CARTEL PARTICIPANTS



3.2 Public procurement data

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[32]. 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).

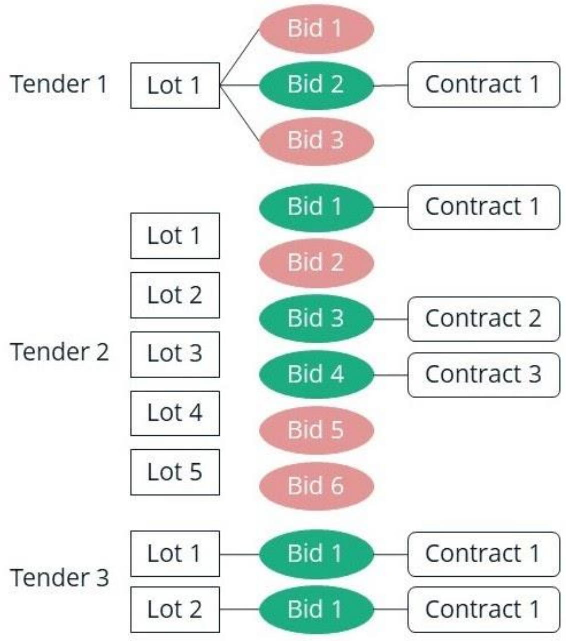
While DIGIWHIST has collected a detailed dataset on public contract across Europe, we have implemented several data quality improvements since the beginning of our project - for example, we have improved several variables that are key for the analysis, such as fixing missing company names, contract values and also connecting the same organization through matching (see section 3.3) and filtered irrelevant data, such as direct contracts where coordination cannot (or hardly can) take place.

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 3.4 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[33]. 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).

[32] 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 analysing the Swedish proven cartel cases.

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

FIGURE 3.4. DATA STRUCTURE



3.3 Data linking

In this section we discuss how we matched cartel case data with procurement data. As we explained in section 3.1, 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[34]. 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 will discuss this issue in more detail in the Methodology section.

TABLE 3.3.MATCHED CARTEL DIMENSION

Matching dimensions	Implemented
Company names	Yes
Cartel period	Yes
Cartel location	No
Cartel buyers	No
Cross bidding	No

[34] 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.

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. The approach was to connect cartel members to procurement records by matching on all available company information - country of operations, company legal name, and address.

This task was complicated by the fact that company names and addresses are not standardized in court documents and public procurement data. Strings containing company names and addresses 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 that was crucial for the matching exercise was to distinguish between company names, their legal forms, and other unnecessary information in this context. Some redundant information such as hyperlinks, 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.

3.4 Final dataset

This section gives a brief overview of the final datasets we used in the analysis. We use two main data sources: a) DIGIWHIST data for France, Hungary, Latvia, Portugal and Spain, and b) an extended version of the historical procurement data provided by Visma Opic that was used by the authors in Fazekas and Tóth (2016). As Table 3.4 shows, we have 77 cartel cases in total from the six analysed countries – out of which we could only test 49 with at least one indicator due to missing data. We have 8520 contracts won by cartel after the collusive period and 5,859 during the cartel period. For the cross-sectional tests we have slightly less cartel contracts (5,731) as some gets filtered out due to lack of similar enough contracts in the control group, that has 187,441 contracts in total – based on the 2-step CEM as explained in the Methodology.[35] In the Appendix, we show the overlap between procurement data and the collected cartel cases; we show the number of contracts by the different matching approaches (company names, cartel time period, and intersecting product markets) that we used for identifying rigged tenders (see Methodology); and also present the (maximum) number of contracts by ‘treatment’ and ‘control’ groups that we can use for the statistical testing.

[35] Note, that CEM weights the control group so that the weighted number of contracts equals the collusive contracts we used for the estimations.

TABLE 3.4. NUMBER OF CONTRACTS INCLUDED IN THE REPORTED TESTS BY COUNTRY

Country	Number of cartels	During vs. after Number of contracts		Cross-sectional Number of contracts	
		during	**after**	**cartel**	**non-cartel***
ES	17	3,991	2,662	3,991	104,160
FR	10	178	1,619	168	27,898
HU	19	848	847	845	20,216
LV	23	528	2,789	415	20,817
PT	2	56	106	56	2,429
SE	6	258	497	256	11,921
TOTAL	**77**	**5,859**	**8,520**	**5,731**	**187,441**

To trace down the signs of bid-rigging on procurement data we define a wide set of elementary indicators that are calculable based on public contracting data.
Table 3.5 summarizes the testable indicators introduced above by our analysed countries.[36]

TABLE 3.5. ELEMENTARY COLLUSION INDICATORS BY COUNTRY

Countries			France	Hungary	Latvia	Portugal	Spain	Sweden
Price based	4	Relative price						
	5	Benford's law						
Bidding patterns	6	Single bidding						
	7	Missing bidders						
	8	Subcontracting						
	9	Consortia						
	10	Cut-point position						
	11	Winning probability						
Market structure	12	Concentrated market structure						
	13	Stable market structure						

[36] Note, that three of the bid price based indicators are only available in Sweden and for most proven cartels bid prices are not available widely enough, we only report those tests in the Appendix and only focus on relative price and Benford's law in the main text.

4 Results Overview

In this section we present our main results by first giving a high-level summary of individual indicator validity and how countries compare in terms of individual indicators picking up the traces of collusion, and an overview of the success rates of all the screens, and second, we discuss the details by each indicator.

Out of the total 77 cartels we collected from court documents, we could test 47 for anti-competitive behaviour with at least one of our screens (Table 4.1). 30 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 47 cartels for which we could run tests, we found 33 that were caught by at least one indicator (70%), 13 were caught by at least three (27%), and 5 were caught by at least five (10%). Very broadly, this suggests that by casting a broad net, we can observe signals of anti-competitive behaviour at scale. We first report these statistics broken down at the country level in

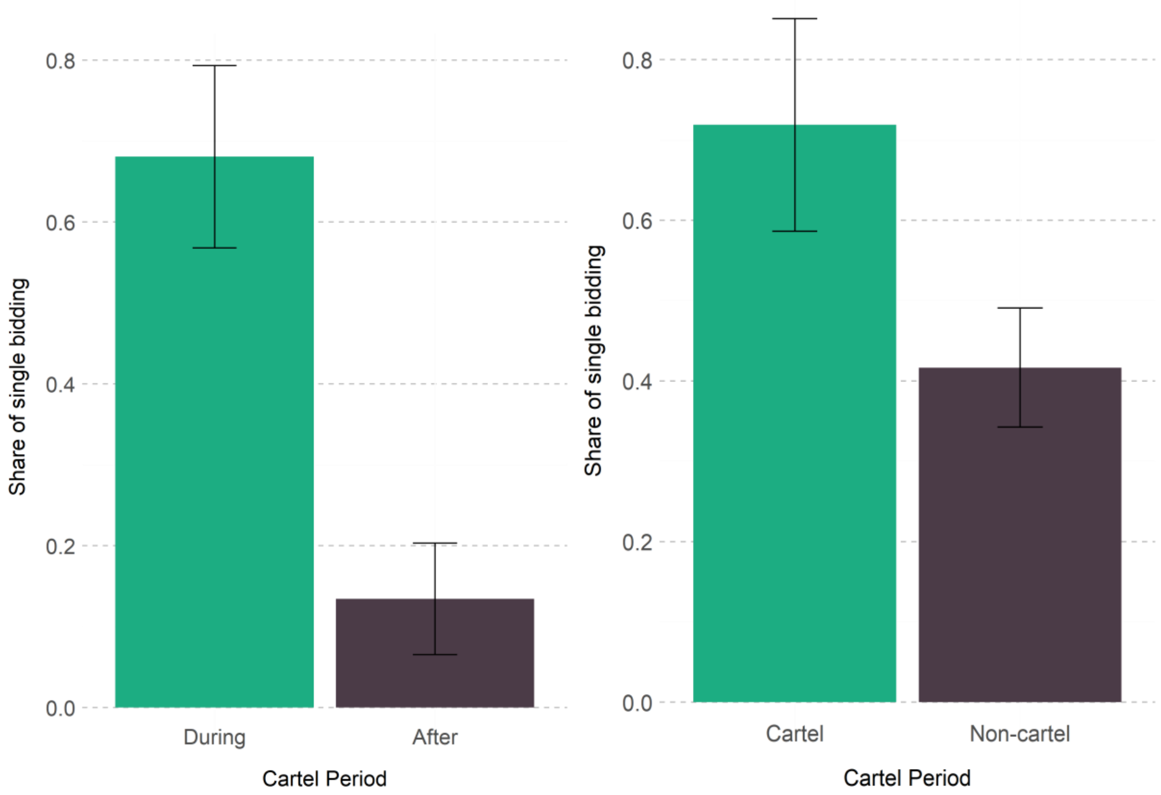
The number of statistically significant indicators varies by country. For example, while 11 out of the 15 analysed cartels are caught by at least one indicator, and 4 caught by at least 5 indicators in Spain, which means that at least 5 tests based on either the during vs. after or the cross-sectional testing logic captured the cartel contracts, only 5 out of the 11 are flagged by one and none are flagged by five indicators in Latvia. This might be explained by the difference in contract numbers – we have 3991 Spanish and only 528 Latvian cartel contracts during the cartel period (Table 1.1).

We report data at the level of indicator and test-type (during vs after, cross-sectional). For each test we report how many total cartels could be screened, and how often a significant signal was observed. The table highlights in particular the issue of data quality heterogeneity across countries and over time. Note for instance that only 6 out of 77 cartels could be tested using the indicators based on the distribution of bid prices in theory and only 1 out of these 6 had enough contracts with non-missing bid prices that allowed statistical testing. Indeed, while quality data on prices is uncommon, data on losing bids is even more rare, limiting the applicability of these indicators. 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.

4.1 Elementary indicators

For example, Figure 4.1 shows a during vs. after and cross-sectional test in practice based on cartel nr 9 from Spain. It shows that single bidding was almost 70% during the collusive period based on contracts won by cartel companies and it dropped to well below 20% after the cartel ended (left). Furthermore, it also shows that during the cartel period, the single bidder share of similar contracts won by companies outside of the cartel was only slightly more than 40%.

FIGURE 4.1. SINGLE BIDDING TEST EXAMPLES (CARTEL 9 – SPAIN)[37] FOR DURING VS. AFTER (LEFT) AND CROSS-SECTIONAL (RIGHT) COMPARISONS. ERROR BARS REPRESENT 95% BOOTSTRAPPED CONFIDENCE INTERVALS.



The number of statistically significant indicators varies by country. For example, while 11 out of the 15 analysed cartels are caught by at least one indicator, and 4 caught by at least 5 indicators in Spain, which means that at least 5 tests based on either the during vs. after or the cross-sectional testing logic captured the cartel contracts, only 5 out of the 11 are flagged by one and none are flagged by five indicators in Latvia. This might be explained by the difference in contract numbers – we have 3991 Spanish and only 528 Latvian cartel contracts during the cartel period (Table 1.1).

TABLE 4.1. INDICATOR TEST SUMMARY BY COUNTRY

Country	All cartels	Cartels tested by at least one test	Number of cartels caught by		
			an indicator	at least 3 indicators	at least 5 indicators
ES	17	15	11	7	4
FR	10	4	4	0	0
HU	19	12	8	2	0
LV	23	10	5	2	0
PT	2	2	2	2	1
SE	6	4	3	0	0
Total	77	47	33	13	5

Table 4.2 summarises the number of potential and confirmatory tests by indicator and test type. As expected, several indicators are not feasible to estimate because of lack of enough contracts or lack or enough available data on specific contracts. Furthermore, even most of the feasible indicators show risks for only a smaller share of cartels. For example, we find only 5 out or the 34 cartels a statistically

[37] Each symbol represents an awarded contract.

lower share of single bidder contracts won by cartel members after the cartel period ended (at a 10% significance level). However, cartel members started to bid in more cities in 13 out of 17 cartel cases once the (estimated) cartel period was over. As we discussed before, only relative price and Benford's law is expected to pick up all cartels in theory. The fact that for example the relative price does not drop more significantly after the cartel period 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.

TABLE 4.2. INDICATOR TEST SUMMARY[38] (*STATISTICAL TESTING IS NOT POSSIBLE – SEE 2.2)

Group	Indicator name	Test type	Tests			Nr. of sign. tests at		Pct. of sign. tests at	
			Potential	Unfeasible	Feasible	5%	10%	5%	10%
Prices	Relative price	during vs. after	61	40	21	3	5	14 %	24 %
		cross-sectional	61	40	21	1	1	5 %	5 %
	Benford's law	during vs. after	77	71	6	0	0	0 %	0 %
		cross-sectional	77	66	11	3	3	27 %	27 %
Bidding patterns	Single bidding	during vs. after	77	43	34	4	5	12 %	15 %
		cross-sectional	77	44	33	5	5	15 %	15 %
	Missing bidders	during vs. after	77	59	18	13	13	72 %	72 %
		cross-sectional	77	40	37	4	4	11 %	11 %
	Subcontracting	during vs. after	50	32	18	4	5	22 %	28 %
		cross-sectional	50	29	21	1	2	5 %	10 %
	Consortium	during vs. after	77	53	24	7	7	29 %	29 %
		cross-sectional	77	39	38	7	7	18 %	18 %
	Cut-point position*	during vs. after	27	13	14	0	0	0 %	0 %
		cross-sectional	27	11	16	5	5	31 %	31 %
	Winning probability	during vs. after	27	13	14	1	1	7 %	7 %
		cross-sectional	27	8	19	6	8	32 %	42 %
Market structure	Concentrated market structure*	during vs. after (cv, narrow)	77	53	24	8	8	33 %	33 %
	Stable market structure*	during vs. after (cv, narrow)	77	66	11	5	5	45 %	45 %

[38] If a test does not have a significance level (Benford's law, cut-point position, and market structure based indicators), the 5% and 10% columns report the same values.
For **Benford's law**, the during vs. after test is considered as a confirming if we find nonconformity during the cartel period and conformity/acceptable conformity/marginally acceptable conformity afterward; the cross-sectional test is considered as confirming if we find nonconformity for cartel companies and conformity/acceptable conformity/marginally acceptable conformity/close conformity for non-cartel companies.
For **cut-point position**, the during vs. after test is valid if we can find at least one cut-point during the cartel period and none afterward; the cross-sectional test is valid if we find at least one cut-point among cartel companies and but none among the non-cartel companies.
For concentration, the during vs. the after test is valid if the Hirschman-Herfindahl Index is higher during the cartel period than afterward.
For **stability**, during and after the cartel period were divided into two-two sub periods. Hence, the average of absolute change in market share is calculated during and after the cartel period. The during vs. after test is valid if the average of absolut market share changes is higher after compared to during the cartel period - i.e. market shares are more fixed during the cartel.
Abbreviations: cv (contract value) and nc (number of contracts) refer to the basis of the market share. In the narrow market definition market participants are only cartel companies; in the wide market definition market participants are based on 4 digits CPV codes.

4.2 Combining indicators

4.2.1 Theory driven

In this section we report the frequency of inferred cartel types by looking at the joint distribution of confirmatory tests by cartel strategies (Table 4.3). In total we found at least 1 confirmatory indicator for 33 cartels based on the total of 74 confirmatory tests. For example, we found at least one valid test[39] of relative bid price, single bidding, subcontracting and consortium for the ES-2 cartel. This suggests that the cartel was following strategy A or D – as having subcontracting and consortium as valid indicators implies that two different strategies were applied for rent reallocation. We emphasize that different indicators highlight signals emitted by different cartel types. This suggests that a screening approach needs to use a variety of indicators to capture different cartel strategies.

TABLE 4.3. INDICATOR TEST SUMMARY BY CARTELS WITH INFERRED CARTEL TYPE

Country	Cartel ID	Benford's law	Relative bid price	Single bidding	Missing bidders	Cut-point position	Subcontracting	Consortium	Winning probability	Concentration	Stability	Cartel type	# indicators
ES	1	No	No	No	No		No	Yes		No	No	C	1
ES	2	No	Yes	Yes	No		Yes	Yes		No	No	A D	4
ES	3		No	Yes	No		No	No		No	Yes	D F	2
ES	4			No	No		No	Yes				C	1
ES	7	No	Yes	No	No		Yes	Yes		No	No	A B C D E	3
ES	9		Yes	Yes	No		No	Yes				A C D F	3
ES	10		Yes	No	No		No	No		No	Yes	D E F G	2
ES	12	No	No	Yes	No		Yes	Yes		Yes	No	A	4
ES	13			Yes	No		No	No				A D F	1
ES	16	No	No	Yes	No		Yes	Yes		Yes	No	A	4
ES	17	Yes	No	No	Yes		Yes	Yes		No		A C D	4
FR	1				Yes			No				A C D F	1
FR	4									Yes		A B C	1
FR	5			No	Yes		No	No		No		A C D F	1
FR	10			Yes	Yes		No	No			Yes	D F	3
HU	1	No	No	No	Yes	Yes		No	Yes	No		A C D F	3
HU	10	Yes	No	No	No	No		No	No	No		A B C D E F G	1
HU	11		Yes	No	Yes	No		No	No	No		A C D F	2
HU	12		No	Yes	Yes	No		No	Yes	No	Yes	D F	4
HU	14				No	No		No	Yes			A C D F	1
HU	15		No	No	No	Yes		No	Yes			A B C D E F G	2
HU	16		No	No	No	No		No	Yes	Yes		A C	2
HU	19		No	No	No	No		Yes	Yes	Yes		C	3
LV	4		No	No	Yes		No					A C D F	1
LV	7		Yes	No	Yes			No		No	Yes	D F	3
LV	14			No	Yes		No	No				A C D F	1
LV	17		No	No	No		Yes	No				A B D E	1
LV	21	No	No	No	Yes		No	No		Yes	No	A C	2
PT	1		No	Yes	Yes	No		No	Yes	Yes		A	4
PT	2		No	No	Yes	Yes		No	Yes	Yes		A C	4
SE	1	Yes		No	No	No		No	Yes			A C D F	2
SE	4			No	No	Yes		No	No			B E G	1
SE	5			No	Yes	Yes		No	No			A B C D E F G	2

39 At least one of the a) during vs. after or b) cross-sectional tests has shown statistically significant changes in the indicator value in the expected direction.

4.2.2 Machine learning

This section introduces the supervised machine learning-based algorithm we fit to predict the likelihood of collusion. It reviews and evaluates selected results from random forest models we applied to the before-after samples pooled from all 6 countries and all cartels with sufficient data. The dataset analysed consists of 13,640 contracts of which 5,476 were awarded to cartel members during the cartel period. We select the best model based on prediction accuracy (correctly classified cases over all cases) in the test sub-sample (30% of the total dataset) which was not used to fit the model. The following elementary collusion risk indicators are used in the model:

- Number of bidders
- Single bidding (yes/no)
- Consortium (yes/no)
- Relative price
- Subcontracted (yes/no)
- Cut-point position (market level)
- HHI (market level)

With winning probability and concentration stability excluded due to low variation and high frequency of missing data. In addition to these risk indicators, two controls were included: country of the cartel and main sector (2-digit CPV code) of the contract.

While we run random forest models which are an ensemble of individual decision trees, we start off by displaying a single decision tree fit to Spanish data (Figure 4.2). It shows for example that contracts awarded in markets with low HHI, but without subcontracting information and with low number of bidders are expected to belong to a collusive market (see red highlights).

FIGURE 4.2. DECISION TREE EXAMPLE, SPAIN, ALL CARTELS, NCONTRACT=6076

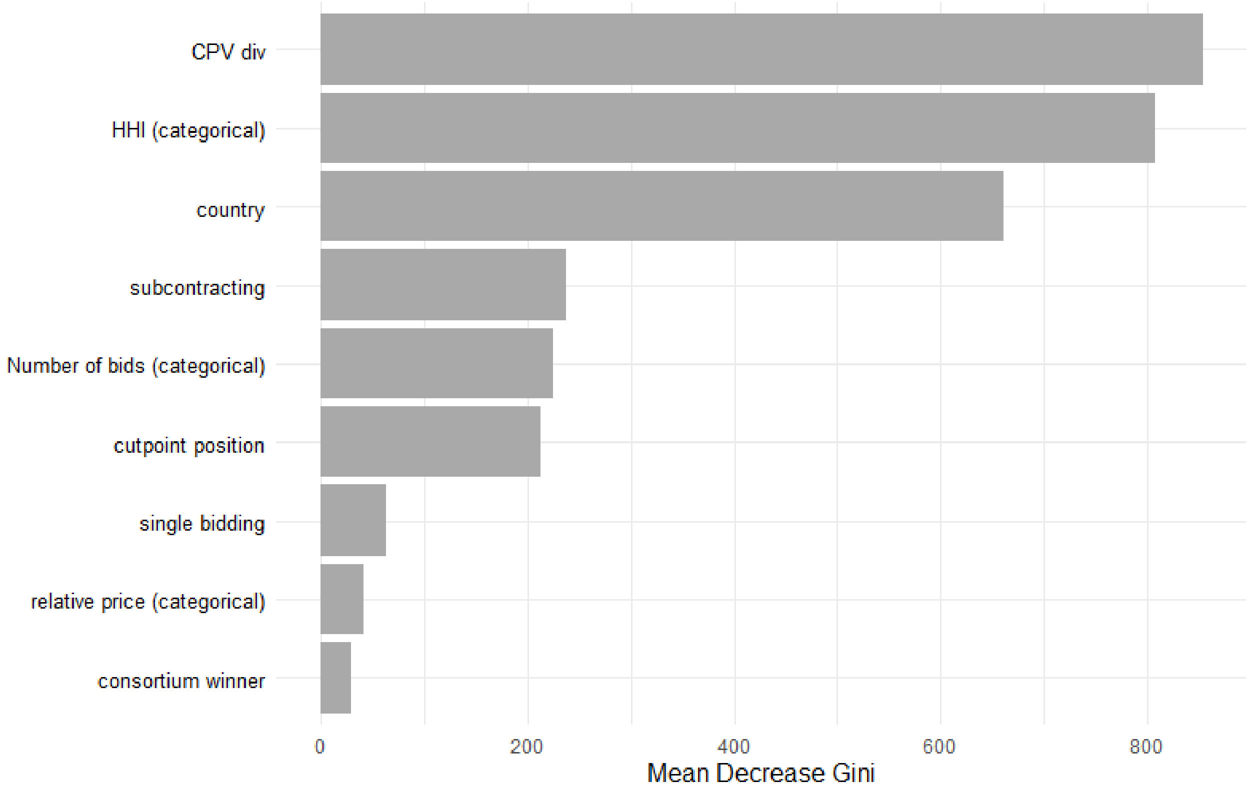
The random forest model maximizing accuracy includes all seven elementary collusion indicators plus country and sector features (we use randomForest library in R, running 100 trees and sampling 3 variables at each run). This model achieves 89.7% prediction accuracy on the test set (precision=84.1%, recall=89.3%) (Table 4.4). Accuracy drops to 88% when country is not used in the model, and further down to 82% when sector is also excluded.

TABLE 4.4. CONFUSION MATRIX, BEST RANDOM FOREST MODEL, ALL COUNTRIES AND CARTELS, BEFORE-AFTER SAMPLE

		Reference	
		No	Yes
Prediction	No	2309	163
	Yes	257	1363

While the overall prediction performance of the model is quite high, we also expect the relationships within the model to correspond to theoretical predictions. First, we establish that while knowing the sector and country are among the most important predictors, some of the 7 elementary collusion risk indicators are similarly important to model accuracy (Figure 4.3). By far, the most important collusion risk indicators is HHI, while subcontracting, number of bids and cut-point are also among the more important predictors.

FIGURE 4.3. VARIABLE IMPORTANCE CHART, FULL MODEL (ALL ELEMENTARY COLLUSION RISK INDICATORS, PLUS SECTOR AND COUNTRY FEATURES INCLUDED)



Looking into the directions and shapes of each predictors' impact on the predicted collusion probability, we find a varied and complex picture (Figure 4.4 and Figure 4.5). The most influential elementary collusion indicator, HHI, is associated with a markedly higher predicted collusion probability when market concentration is low (2086-2693) (Figure 4.5, panel A). While medium levels of HHI (2782-5057) are associated with a close to average predicted probability of collusion. Regarding subcontracting, the absence of subcontractors leads to a markedly lower prediction than their presence (Figure 4.5, panel B). Considering the number of bidders (deciles), The pattern is markedly U-shaped with single bidding leading to a somewhat higher predicted collusion probability (Figure 4.5, panel C).

Then the predicted probability drops starting from two bidders while risk climbs back up at the upper end of the distribution, especially with more than eight (9th decile) or 12 bidders per tender (10th decile).

FIGURE 4.4. SHAPLEY PLOT, FULL MODEL (ALL ELEMENTARY COLLUSION RISK INDICATORS, PLUS SECTOR AND COUNTRY FEATURES INCLUDED)

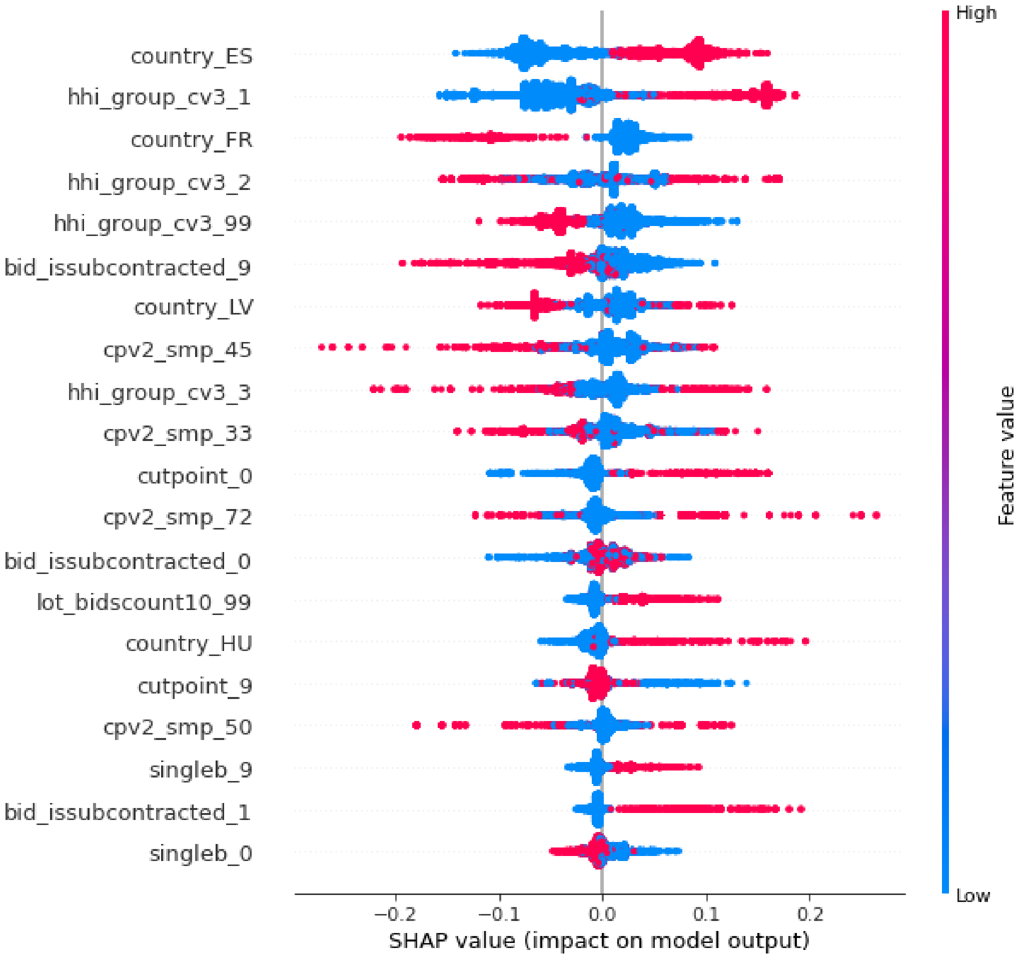
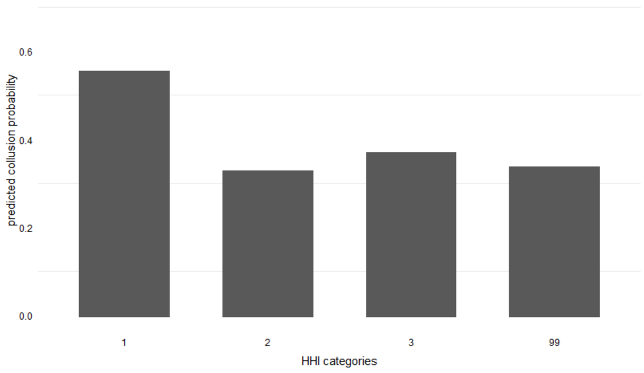
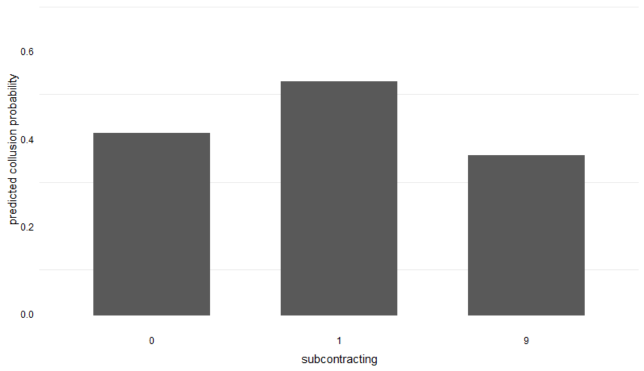


FIGURE 4.5. PARTIAL DEPENDENCE PLOTS FOR SELECTED VARIABLES, FULL MODEL (ALL ELEMENTARY COLLUSION RISK INDICATORS, PLUS SECTOR AND COUNTRY FEATURES INCLUDED)

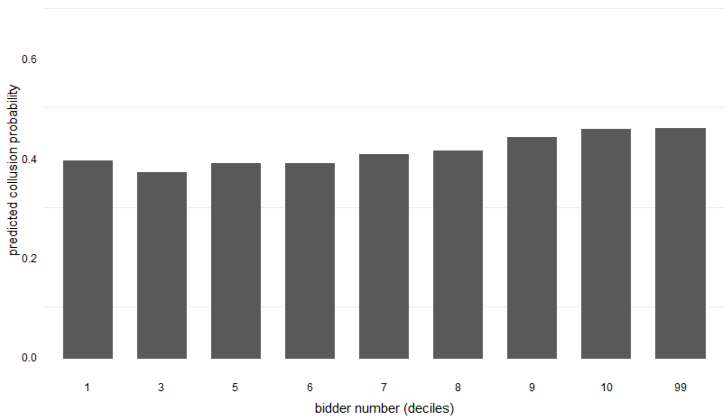
Panel A. HHI categories: low, medium and high



Panel B. subcontracting: yes or no



Panel C. bidder number deciles



4.3 Robustness tests

As we have shown above and as Figure 4.6 also highlights, the number of contracts are closely related to the number of confirming tests that a screening approach can provide. For many cartels, there are not enough data points to detect traces of collusion in a statistically meaningful way. Furthermore, as we already discussed in the Methodology, choosing contracts for the control group – i.e. the non-collusive but comparable contracts – also greatly influences the test results. Therefore, we have implemented a range of robustness checks by loosening or tightening up some of the assumptions behind our testing. For the sake of brevity, we only give the high-level take-aways briefly instead of including all scenarios. These robustness checks can be grouped into four main categories.

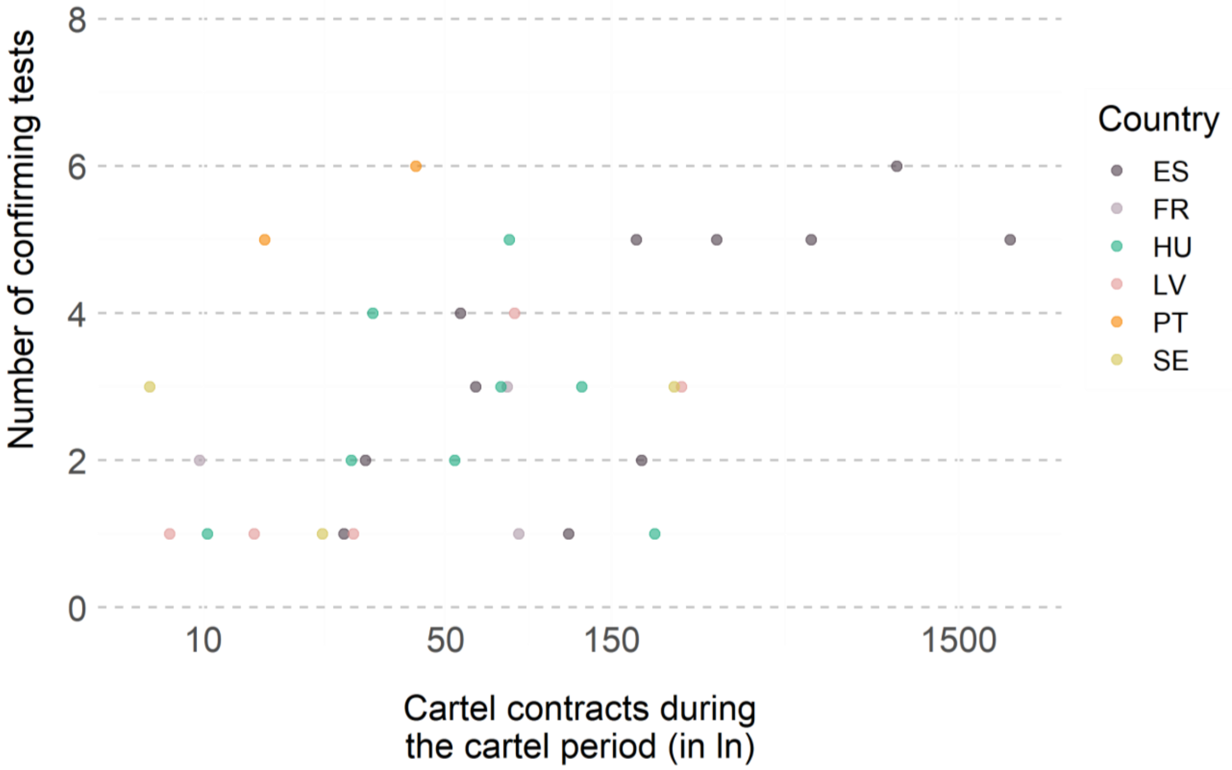
First, as the product market of the cartel activities were not always unambiguously disclosed in the court documents, and finding specific product markets based on CPV codes can be tedious, we did not restrict the cartel contracts based on product codes in the reported tests but simply took all contracts matching the company names and a given time period as collusive tenders. However, it might be the case that cartel companies are colluding in some and competing on other product markets. Therefore, we run tests with a restricted sample as well, where we only keep those contracts that had a product code where at least two cartel members had submitted a bid. Depending on the country and cartel, this ‘intersecting’ product market filter decreased the number of cartel contracts by roughly 0-30% and made some of tests insignificant. However, the overall picture remained largely unchanged.

Second, there are several ways of finding an appropriate control group for a cross-sectional testing. The two-step CEM matching is one of the three different CEM approaches we calculated. One of the unreported is a one-step CEM where unmatched cartel contracts from the first-stage are simply excluded (i.e. this is a more restrictive matching), while the other is more lenient by simply putting back cartel contracts without matching pairs into the treated group with a weight of one, so that we do not lose observations from the already. These estimations lead to only slightly different results – such as having ten fewer or five more significant individual indicator tests.

Third, as we discussed in the Methodology section, we have to rely on assumptions for defining the cartels’ scope in almost all dimensions. We do not only have to accept that all participating companies are convicted, but also that the start and end of the cartel period could be established fairly accurately. Therefore, we also implemented alternative tests with an arbitrary cartel time period, whereby we added extra 3-years to each cartels’ life-span. We assumed that the proven period might have been affected by available evidence and that taking a longer time-period would allow us to better track down collusive behaviour. We also reason that prosecutors are generally happy to win a conviction against a cartel for any of their collusive activity, prioritizing those instances in which they have the clearest evidence of conspiracy. For instance, while a cartel might operate for a decade, a key piece of evidence such as an intercepted phone call or written note between conspirators may only implicate collusion in a single tendering process. Extending the cartel period by three years as a comparison group, the high-level country-by-country results are somewhat better than the ones we got using the conservative time period. We lose fewer cartels that we cannot test due to insufficient data, and the share of cartels detected by one, three or five indicators also increases.

Fourth, in the main text we report all except the market structure based tests with consortia contracts included in both the collusive and competitive contracts. However, finding consortia contracts is based on a probabilistic model (see Appendix), hence we tested all indicators without consortia contracts as well. As adding consortia contracts increases the number of observations, the number of significant tests are higher in the wider sample (in most of the above mentioned different matching methods), however, we only have 5-10 less significant tests for all except market structure base indicators.

FIGURE 4.6. RELATIONSHIP BETWEEN NUMBER OF CARTEL CONTRACTS AND CONFIRMING TESTS (EITHER DURING VS AFTER OR CROSS SECTIONAL). (CORRELATION: 0.51)



5 Discussion

In this study, we analysed whether we can detect public procurement cartels in 6 countries by using publicly available large-scale administrative datasets that combine data on public contracts and proven cartel cases. We tested 10 widely applicable indicators that can be calculated in publicly available datasets in most developed countries.[40]

One of the biggest novelties of our study is its scale. To the best of our knowledge, our study is the first to test cartel screening indicators at this scale on various country datasets that represent well the quality and scope of administrative data available in most developed countries. We have screened 156 proven cartel cases from across 19 countries, and eventually linked case-level data for 77 of them in 6 different country datasets, all in Europe. Finding cartel cases and building a dataset that matches them with administrative contracting data was a significant effort with numerous difficulties. For example, legal documents do not provide information that is easily translatable into a tabular dataset, finding the exact rigged contracts is also hard due to the lack of explicit reference to individual tenders in legal texts.

Improving data quality is key for precise cartel screening. Three problems in particular made indicator testing difficult: a) missing values of available indicators (e.g. if we do not know the number of bids for 30% of the contracts, statistical testing is more noisy and potentially biased); b) missing key variables

[40] Note, we have also tested three additional bid price distribution based indicators in Sweden that we do not report in the main text.

(e.g. we could not apply price screens in most countries); and c) data structure (i.e. competing bids could not be extracted in many countries for multi-lot tenders).

Our results show that no universal indicator exists that would work across all different cartel strategies. Instead, a strategy-specific combination of indicators can indicate coordination in cases with sufficient number of contracts. A significant share of testable cartels (~65-75%) are picked up by at least one and a non-negligible share (~25-30%) by at least three indicators. While a single indicator might represent only a noisy signal, we consider cases with 3-5 red flags as strong evidence for collusion. It also suggests that cartels rigging many contracts - i.e. the ones causing economic harm for a longer time period - leave strong enough signals that are easier to pick up.

Insignificant results were often caused by the poor quality or complete lack of data and we often could not calculate certain indicators that would corroborate results. For example, we might have a signal in single-bidding but we could not verify suspicion by looking at the prevalence of consortia bidding, which is a trivial co-occurring risk factor. Hence improving data quality and scope along the lines discussed above would allow a significantly more precise testing.

The approach tested does not offer a fully automated tool but guides an analyst and provides a shortlist of indicators and tests to be performed to be able to narrow down the sample of suspicious cases. It can also offer competition authorities a criterion for prioritization of some submarkets and drive allocation of the resources: submarkets with more indicators have to be checked first. For example, whistleblower tips could be quickly followed up by analysing the bidding behaviour of the reported companies with sufficient public contracting data. Furthermore, grouping sets of companies and testing which group has a high-risk bidding patterns and contracting outcomes would help to identify collusion prone company groups. However, this style of analysis will struggle to pick up ad-hoc cases covering only a few contracts. Unlike most previous studies, we do not exploit particular procedural rules to identify collusion risk (e.g. rarely used auction designs), instead our indicators are available and applicable across various regulatory regimes.

A clear advantage of this approach is its broad applicability. As public contracting data is becoming more and more available, which makes these supplementary screens viable across many countries.

6 Limitations

In this section we outline the limitations of our approach. Specifically we describe limitations of our data and some of the assumptions we had to work with. We also reflect on conceptual limitations of our approach, noting the application of ensembles of screens to markets can only be one tool, albeit a powerful one, in the cartel-hunters toolkit.

Limitations of our data can be described in two broad categories. The first pertains to data quality or missing information: different screens depend on having consistent and comparable data on firm identities and behavior. Publicly available databases on procurement generally do not contain accurate information on bid prices and other information relevant to cartel screens such as subcontractors identity. The high incidence of missing data and the common lack of persistent identifiers of market participants provide additional hurdles. The assumption that such data is missing at random may not be true in all cases.

One important added-value of our report is to document the scope and scale of these data quality issues across many markets and their implications for cartel risk estimation. There is a weak if any correlation between national procurement data quality and quality of government. For instance, data from Sweden, traditionally perceived as having excellent quality of government, is no more accessible or clean than data from Portugal or Hungary, both of which have below average quality of government scores among EU member states.

Data quality issues do not only limit the potential of specific cartel screens. They also block the evaluation of the impact of collusive behavior. Estimating the economic harm due to cartels is important for the authorities to prioritize which potential cartels to investigate. They are also important legally, to justify the use of more extensive investigative tools and methods such as surveillance and search warrants.

The second major limitation of our approach is that it assumes that cartels which have been detected are roughly similar to those who evade detection. This is a manifestation of survival bias. It is arguable that cartels who operate or even disband without being uncovered by the authorities are doing something differently. This issue is not new in the anti-collusion literature, in which most studies examine cartels which are revealed by whistleblowers or which arose in highly idiosyncratic circumstances. Indeed, our approach to scanning whole markets could generate novel kinds of examples of cartels for future work. In other words, we do not overcome this traditional limitation of anti-collusion research but offer a new perspective on it.

A related concern is the adaptability of cartels to screens. Indeed, cartels are known to consider the traces they leave behind and often adopt strategies to hide their collusion. This known issue strengthens the case for our approach: it is difficult to adapt to multiple screens at once. If cartels evade screens by decreasing the rents they collect, that represents a major public welfare improvement. If cartels need to coordinate more to evade screens, that increases the risks that they are caught communicating and collaborating. More complex arrangements are likely more difficult to enforce, increasing the risk that cartel members will defect from the collusive agreement and break the cartel from the inside.

As we explained in the Methods section, defining the ‘relevant collusive market’ is not straightforward even in theoretically clear-cut cases, that are based on proven cartel cases. Just to highlight one specific assumption we have worked with: we defined the ‘relevant collusive market’ by taking the intersecting product markets, that are markets where at least two cartel members have submitted a bid. However, this assumption does not fit all cartel cases, especially the ones where the participants split their contracts by product markets. Therefore, an alternative scenario could be also analysed, where we simply take all product codes that cartel members were winning contracts in and re-calculate the cross-sectional and difference-in-differences estimates. However, this approach risks including contracts in the ‘rigged group’ that were indeed unrelated to the collusive activity.

Inferring cartel strategies is only possible with a sufficient number and varying kinds of indicators that enable measuring various aspects of a collusive scheme. As we mentioned before, some key variables are missing or of bad quality in many countries, which makes certain indicators impossible to calculate.

Conclusions and future work

In this report we presented a test of cartel screens on large scale public procurement datasets containing ground-truth cartels. We found that simple screens can provide signals of cartel activity. Though no individual screen works in even a majority of cases, several screens together seem to provide a strong indicator of potential collusion.

We make the following recommendations:

- Data quality needs to be improved, and researchers should be given access to higher quality data.
- Markets can and should be scanned regularly with a broad suite of screens. Groups of firms or submarkets rife with anticompetitive behavior are significantly more likely to set off different kinds of screens, depending on the anti-competitive strategies they adopt.
- The merging of complementary data sources, for instance on firm ownership, can provide additional value.
- Additional screens should be developed, depending on data availability, and added to an ensemble of screens.
- Feedback from competition authority stakeholders will be invaluable in tuning the use of particular screens or groups of screens.

In the future, we will apply our machine learning algorithms to extrapolate to the whole public procurement market and score contracts, groups of firms and sub-markets for the likelihood of collusive behavior. Of course, such algorithms cannot decide for sure whether a cartel is present or not. They work best when used by domain experts in an iterative manner (sometimes referred to as human in the loop or HTL). Algorithms can both widen the scope an individual can investigate and highlight key details. They can also highlight novel emerging cartel strategies, by considering new combinations of screens that a cartel sets off. Our methods to screen markets can vastly narrow down the search space for potential cartels. More work needs to be done in the steps that should come next: within a suspicious market, finding those firms that may be colluding.

To conclude, this chapter represents an initial step towards a big data infrastructure for cartel screening across Europe. We show what is already possible with public data and straightforward screens and suggest what could be done next.

Chapter II

Introduction

In spite of recent scholarship identifying the anti-competitive impact of ownership ties, there is very little we know about the impact of such ties in public procurement. In general, structural links, such as ownership overlaps, can affect the incentive and ability of firms to behave competitively. Yet, while there is substantial theoretical literature on the effects of overlapping ownership on competition, empirical studies on the subject have been limited. At present, structural links are not regulated in a comprehensive or systematic fashion by any competition authority and it is still an open question whether such links may present a competition problem that deserves closer legal scrutiny. To determine the appropriate policy response, there is a need for further in-depth work to show the relationship between structural links and firm behavior.

Assessing the competitive landscape in a market is an important step in the investigations and studies conducted by competition authorities. However, complex corporate connections can make identifying relationships between firms a challenging task. Network approaches are useful in such cases where there is a need to map and understand the structure of relationships between different actors. In this section, we use techniques from network science to conduct an exploratory analysis of corporate ownership in Swedish public procurement markets and investigate its impact on the risk of collusion.

7 Literature

The growing concern of contracting authorities regarding relationships between competitors is evident in the ‘Notice on tools to fight collusion in public procurement and on guidance on how to apply the related exclusion ground’, released by the European Commission in March 2021. Amongst other things, the document considers the issue of affiliated companies submitting bids for the same contract, as well as the matters of joint bidding and subcontracting.

There is substantial literature in economics on structural links and competition (Reynolds & Snapp, 1986; Malueg, 1992; Gilo, 2001; O'Brien & Salop, 2000; Levy et al., 2018). The focus of analysis has largely been on minority ownership, which is defined as shareholdings of less than 50 percent of voting rights. Minority shareholdings can include one-sided ownership by one firm of another, reciprocal direct ownership interests between rivals (cross-ownership), and minority shareholding by third parties (common ownership). According to economic theory, acquisition of minority shares can lead to competitive harm in various ways similar to those arising from acquisitions of control: by reducing competitive pressure between competitors (horizontal unilateral effects); by facilitating coordination among competitors (horizontal coordinated effects)41; or, in case of vertical structural links, by allowing companies to foreclose competitors' access to inputs or customers (vertical effects).

The renewed interest in ownership links was sparked partly by an empirical study by Azar et al. (2018, 2021) showing that common ownership by institutional investors has anti-competitive effects in banking. The authors suggest a generalized Herfindahl-Hirschman Index (GHHI) that accounts for both common ownership and cross-ownership as a measure of market concentration. Conducting an econometric analysis on the prices of bank deposit products, they found that the GHHI was associated with higher fees and deposit thresholds. However, their approach has been criticized by some researchers due to issues related to endogeneity in their model, their definition of control, and the applicability to other industries (OECD, 2017).

41 Using a symmetric Cournot duopoly model, Malueg (1992) however found that, in general, partial cross-ownership has an ambiguous effect on the ability of firms to collude. This ambiguity is due to two counteracting forces: while firms internalize part of the losses that they inflict on competitors when they deviate from a collusive arrangement (thereby weakening the incentive to deviate), cross-ownership also softens competition following a breakdown in the arrangement (thereby strengthening the incentive to deviate).

Structural links are also a subject of scrutiny in other fields. The study of structural links between organizations has a long history in sociology and social network analysis, where, according to (Murray, 2006), there is a strand of literature that "looks at interlocks as structural mechanisms that cement collusion and subsequently help the development of business cartels or monopolies". A survey of work discussing the causes and consequences of interlocking directorates, which refers to the practice of individuals serving on the board of directors of more than one organization, is provided by Mizruchi (1996). According to the author, the reason for the formation of these links can be explicit or inadvertent. Some of those discussed by Mizruchi include collusion, cooptation and monitoring, legitimacy, career advancement, and social cohesion. As a mechanism for collusion, director interlocks are often assumed to facilitate communication among competitors. However, research on the effect of interlocks on actual performance are limited and have differing results. For instance, a few older studies of firms based in the United States found little association between interlocks and profitability (Pennings, 1980; Burt, 1983). On the other hand, in a study of Canadian firms, Carrington (1981) found a positive relationship among industry concentration, interlocking directorates, and profitability.

In network science, the structure of the global ownership network was first extensively studied by Vitali et al. (2011), who investigated the control held by transnational corporations. The authors found that "a large portion of corporate control flows to a small tightly-knit core of financial institutions". Since then, there have been a number of papers studying the topology of corporate networks. In Takes & Heemskerk (2016), the authors conduct a cross-country comparison of interlocking directorates in the global corporate network. They found similar network topologies as Vitali et al. (2011), yet large differences between countries, particularly when it comes to the relation between economic prominence indicators (e.g., revenue) and the network centrality of firms42. These results hint that there is variation in how corporate networks are structured in different countries and support the need to conduct cross-country analysis to compare how these differences might affect firm behavior and competition.

In a more recent paper, van Lidth de Jeude et al. (2019) construct a multiplex network of interactions between companies in Germany and in the United Kingdom, combining ownership links, interlocking directorates, research and development collaborations, and stock correlations. They found low, but non-zero, overlap between the different types of structural links indicating that these complement each other. The implication that corporations may be more closely connected than previously reported highlights the importance of studying different types of links that may facilitate anti-competitive behavior among firms.

Empirical work on the impact of structural links on behavior and outcomes, particularly in bidding markets, is limited. However, interest in the use of new approaches to explore these complicated relationships has grown in recent years. In a working paper, Asai & Charoenwong (2019) study the effects of ownership connections on prices and cost efficiency in public procurement auctions in Singapore. Using identical bidding as an indicator for potential coordination among firms, they found the measure to be strongly correlated to having a shared owner.

8 Methods

8.1 Market Definition

A crucial step when conducting a competitive assessment is defining the "relevant market", which consists of the catalogue of goods or services that are considered substitutable by consumers. If a market is too narrowly defined, non-trivial competition might be excluded from the analysis and the

[42] Centrality measures provide an indication of the relative "importance" of actors in a network. Freeman (1978) provides a discussion of the most common centrality measures used in social network analysis.

competitive constraint imposed on firms consequently underestimated. If it is too broad, the intensity of competition in the market may be overestimated.

Market definition has two dimensions: the product market and the geographic market. Defining a relevant market often calls for detailed economic analyses. The standard framework employed by competition authorities is the test of small but significant and non-transitory increase in price (SSNIP), which seeks to identify the smallest market within which a hypothetical monopolist or cartel could impose a profitable significant increase in price. However, for the purpose of this report, we take a simplified approach and employ the method used by Fazekas and Tóth (2016), which uses product and firm attributes. Specifically, we determine the product groups based on CPV[43] categories and the geographical location based on the NUTS[44] region of the contracting authority to determine the groups of firms that we can consider as competitors.

One advantage of this approach is that we are able to capture both potential and actual competition in a market. Nevertheless, there are obvious drawbacks with using standardized location and product codes for market definition[45]. A suggestion for future work is to combine firm and product attributes with co-bidding information to assign relevant markets. This can be done by applying simple rules or through more sophisticated approaches, such as treating market definition as a multilabel classification problem and using techniques from machine learning.

8.2 Network Construction

For each relevant market, we create a corporate network consisting of 2 layers: ownership (O) and co-bidding (B). A visual representation of a multilayer network is provided in Figure 8.1 (c). The undirected weighted network layer G_B' is a projection of the bipartite graph G_B, which connects firms with tenders, as shown in Figure 8.1 (b). In G_B', each edge has an associated weight $w_{ij}^{[B]}$ indicating the number of tenders where both firms i and j submitted a bid.

The undirected market ownership network G_O' is also a projection; in this case of what we call the elementary ownership network G_O, illustrated in Figure 8.1 (a). In G_O, the vertices (also called nodes) correspond to the economic entities, and the links to the ownership shares connecting them. Building G_O involves creating a network of relationships among firms, tracing back to the level of the ultimate parent entity, if available. An edge $w_{ij}^{[O]}$ in G_O' indicates the existence of an ownership link between competitors i and j.

In this analysis, we make no distinction between the types of ownerships overlaps (e.g., one-sided versus cross-ownership, partial versus full control). $w_{ij}^{[O]}$ is therefore a binary variable that takes on the value 1 when an ownership link is present, and 0 otherwise. Admittedly, this approach disregards valuable information that may have important implications on the competitive environment in a market. Indeed, Gilo et al., (2006) and Shelegia & Spiegel (2015) show how asymmetric stakes in rivals can have differing effects on the behavior of firms. To capture information on the magnitude of ownership linkage, it is possible to associate a weight with each edge $w_{ij}^{[O]}$ that measures the effective stake of

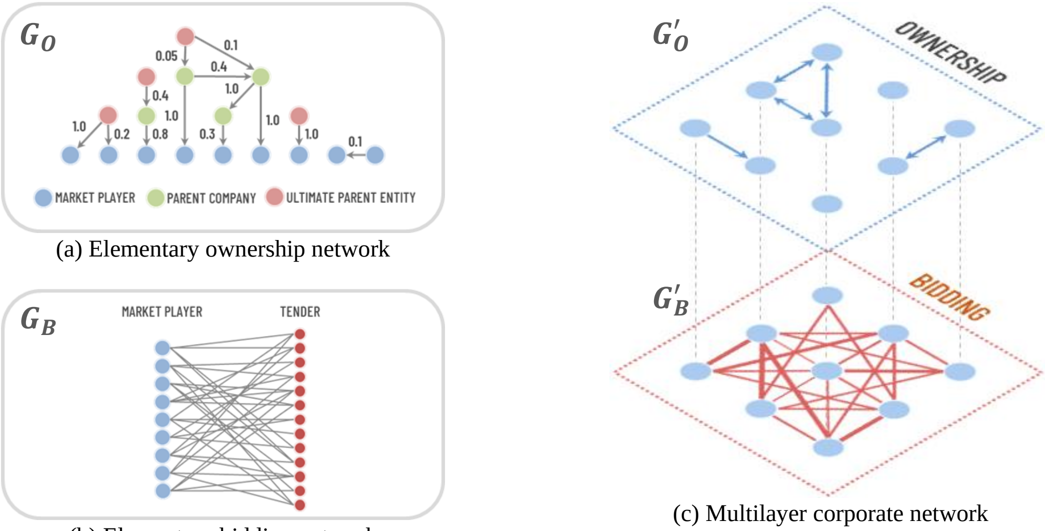
[43] 7 CPV=Common Procurement Vocabulary. For more info see: https://simap.ted.europa.eu/web/simap/cpv

[44] NUTS=Nomenclature of Territorial Units for Statistics. For more info see: https://simap.ted.europa.eu/web/simap/nuts

[45] Mainly, such product classification systems are usually not able to truly capture the demand- and supply-side substitutability of goods and services. For instance, in the United States, the narrowest North American Industry Classification System (NAICS) codes were found to be typically broader than markets defined in actual merger cases (Werden & Froeb, 2018).

firm i in j. This will transform G_O' into a directed weighted network, which we can then analyze and compare against our other network layers. Such an approach is suggested for future work.

FIGURE 8.1. NETWORK CONSTRUCTION



(a) Elementary ownership network

(b) Elementary bidding network

(c) Multilayer corporate network

8.3 Empirical strategy and expected patterns

Our empirical analysis focuses on the relationship between ownership links and competition in public procurement markets. Specifically, we are interested in knowing if certain indicators of structural linkage are correlated with indicators of collusion risk. As discussed previously, structural links may create or enhance the incentive and ability of firms to coordinate their conduct. Collusion requires firms to reach an understanding on the terms of coordination and to able to monitor these terms, allowing deviations to be detected and punished. Ownership links can facilitate this by leading to or increasing information exchange and transparency between competitors (European Commission, 1997).

The relationship between collusion and bidding can be positive or negative, depending on the type of bid-rigging arrangement. For instance, cover bidding and bid rotation can result in more co-bidding between cartel members, while bid suppression and market allocation will lead to less bids submitted by firms. We contend that the first two types of arrangements are more prevalent, particularly in countries with stronger competition enforcement, as they are more difficult to detect. Indeed, cover bidding is generally accepted to be the most recurrent form of agreement in sealed bid auctions and is also used in bid rotation. Our expected results are therefore premised on the assumption that cover bidding is the more likely form of collusive arrangement in Swedish public procurement markets.

We begin our empirical analysis by looking at pair-level data to determine if there is a correlation between the presence of an ownership tie and the frequency of co-bidding. In the presence of bid-covering between affiliated firms, we expect this relationship to be positive. We then further test the hypothesis that structural links weaken competition by relating ownership network measures with collusion risk indicators. Due to the large proportion of observations with missing bid prices in the Swedish dataset, we limit our analysis to bidding pattern and market structure indicators. We study five collusion risk indicators that were found to have at least one valid test in the analysis of existing cartels in Sweden:

- single bidding,
- missing bidders,
- winning probability,

- cut-point position, and
- stability.

Single bidding, missing bidders, and stability are market-level indicators while winning probability and cut-point position are determined at the firm level. For our explanatory variables, we investigate the following ownership network measures:

- density,
- closeness centrality, and
- degree centrality.

Density captures the interconnectivity of firms in the network. It may vary from low density, where a group of firms are loosely connected, to high density, where firms are highly interlinked. It is computed as $\frac{2|E|}{|V|(|V|-1)}$, where $|V|$ and $|E|$ denote the number of vertices and edges, respectively. Markets with higher ownership network density are expected to have less intense competition due to shared knowledge, better information transmission, and higher trust between firms. Therefore, if colluding firms engage in cover bidding, we expect markets with high ownership network density to have less incidence of single bidding and smaller proportion of missing bidders. However, the market structure is expected to be relatively more stable. Using stability of market shares as a collusive marker was proposed by Harrington (2008) under conditions where firms' costs are sufficiently persistent over time relative to their patience.

Harmonic closeness centrality captures the average length of the shortest path[46] between firms in the ownership network. The normalized harmonic centrality of a vertex i is computed as $\frac{1}{|V|-1}\sum_{j\neq i}\frac{1}{dist(i,j)}$. Closeness centrality measures how close a node is to the other nodes in a network; the value is 1 when a firm is directly connected to all the firms in the market, and 0 when it has no ownership link. In social network analysis, firms with higher closeness centrality are often considered important due to their access to greater and quicker information and research exchange. We test this indicator for both our firm-level and market-level regressions, using the average value across firms for the latter. Similar to density, markets with higher harmonic closeness centrality are expected to have less intense competition due to easier coordination and faster communication. The expected relationship with single bidding, missing bidders, and stable market structure is therefore the same. At the firm-level, we expect companies with high closeness centrality to be more likely to be in a cut-position in the co-bidding network. The effect on winning probability is however ambiguous, depending on whether the colluding firm submits more cover bids than genuine bids.

Degree centrality captures the local connections of firms. The degree centrality of a vertex i is defined simply as the number of links incident upon that vertex. Firms with high degree centrality are considered to be important actors due to their direct connections with other firms through ownership links, allowing easier communication and greater possibility to influence or coordinate behavior. Since the measure is a stricter version of closeness centrality, its relationship with our collusion risk indicators should be the same.

A summary of the expected results of our analyses is shown in Table 8.1. We hypothesize that markets with higher ownership structure indicators would have less single bidding and missing bidders, but would also have a more stable market structure. This is based on the presumption that an active cartel in Sweden is more likely to engage in cover bidding or bid rotation, where there is a relatively stable community of bidders that collude between themselves to decide who should win which contract. We

[46] A path is defined as a sequence of nodes where each consecutive pair in the sequence is connected by an edge.

therefore expect more bids beings submitted in markets with relatively high ownership ties, so that colluding firms can maintain the illusion of competition. However, the market structure is expected to be more stable as firms settle on a collusive equilibrium where prices and market shares are fixed.

TABLE 8.1. EXPECTED RELATIONSHIP BETWEEN OWNERSHIP AND COLLUSION RISK INDICATORS.

		Market-level ownership structure indicators		Firm-level ownership structure indicators	
Group	Indicator Name	**Average harmonic closeness centrality**	**Density**	**Harmonic closeness centrality**	**Degree centrality**
Bidding pattern	**Single bidding**	negative	negative		
	Missing bidders	negative	negative		
	Winning probability			indeterminate	indeterminate
	Cut-point position			positive	positive
Market structure	**Stable market structure**	positive	positive		

9 Data

Our two data sources for this analysis are the public procurement data collected by DIGIWHIST[47] and the Orbis Global Database from the Bureau van Dijk company. The latter is the largest cross-country firm-level database encompassing financial statements, production activity, and board members of public and private firms. Importantly, the database also includes information on each company's equity ownership structure, including the names of the owners and their respective shareholdings. This allows for the construction of a network of relationships between firms through ownership. We limited the scope of our analysis to one country (Sweden) and one type of structural link (ownership)[48] for the period 2010-2015.

We use the CPV and NUTS data from DIGIWHIST to categorize tenders into relevant markets. We then identify which firms submitted bids for these tenders to obtain the set of firms that compete in that market. To build the elementary ownership network G_O, we conduct a recursive exploration of the neighborhood of these companies in the Orbis database: first proceeding upstream with a breadth-first search to identify all direct and indirect shareholders of the market participants, and then continuing in a similar way downstream to identify companies that are directly and indirectly owned by the competing firms. Each ownership network G'_O is created from the projection of G_O, reducing the nodes to only those identified as competitors based on the DIGIWHIST data. To illustrate this process, we provide an example with a market consisting of 14 competitors, as shown in Figure 9.1: G_O contains bidders and their shareholders, while G'_O shows only bidders and a link indicating ownership links.

We have chosen to mark competitors i and j as having an ownership link regardless of the number of steps it takes to reach i from j in the elementary ownership network. An alternative approach taken by Asai & Charoenwong (2019) is to only consider connections between firms up to a certain level[49]. However, using such a cut-off may lead us to exclude links that may still be meaningful, especially if the shares held by the common owner is significant.

---

[47] opentender.eu/download

[48] There is however opportunity to expand on this work by including more countries and time periods, as well other types of structural links, such as director interlocks. We discuss this further on the section on next steps and future work.

[49] The authors only consider links up to the fourth degree. They define a fourth-degree connection between firms A and B to exist if A is owned by firm C, B is owned by another firm D, and firm C and D have the same shareholders.

FIGURE 9.1. ILLUSTRATION OF AN ELEMENTARY OWNERSHIP NETWORK AND ITS PROJECTION.

(a) Elementary ownership network G_O (competitors indicated in red)

(b) Network projection G'_O (an edge indicates presence of an ownership link)

Building the co-bidding networks is more straightforward. Using our procurement data, we count the number of times each pair of firms bid on the same tender in a market. We can then analyze these networks together to study their relationship.

Our dataset is restricted by the number of firms in the DIGIWHIST database that we can match in the Orbis database. After removing observations that we are unable to match, we are left with 201,646 data points in our bid-level procurement data, with 6,909 distinct bidders and 64,395 distinct tenders for the period 2010-2015. For each year, we were able to identify between 2,129 to 2,483 markets using our definition for relevant market. More than half of these have less than 10 firms. The distribution of market sizes is shown in Figure 9.2. We also provide the distribution of statistics of interest from the bidding network (Figure 9.3) and ownership network (Figure 9.4).

FIGURE 9.2. FREQUENCY DISTRIBUTION OF MARKET SIZES

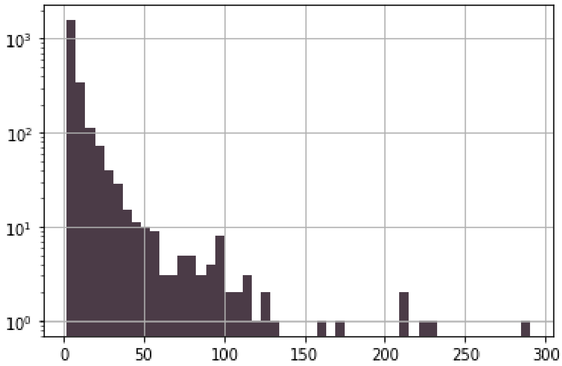


FIGURE 9.3. CO-BIDDING NETWORK DATA, FREQUENCY DISTRIBUTIONS.

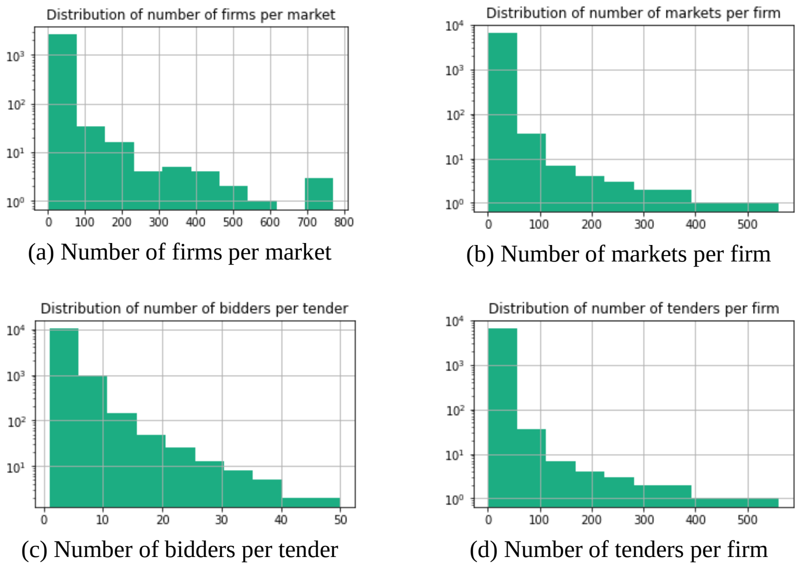
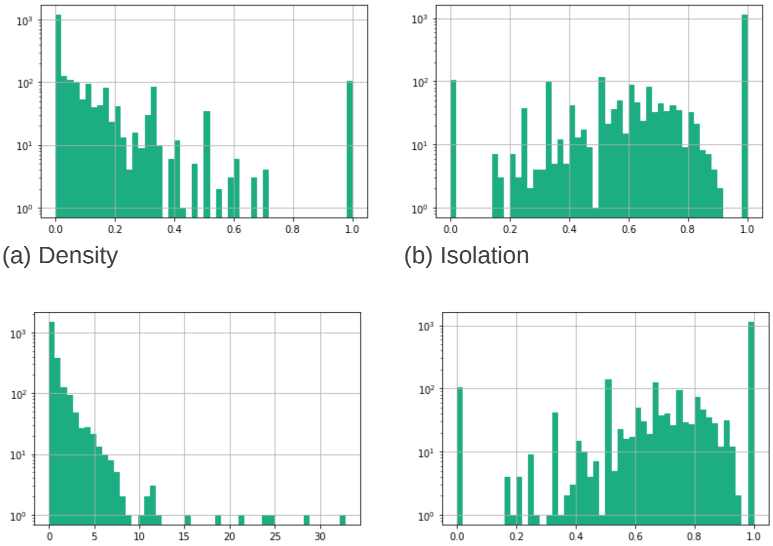


FIGURE 9.4. OWNERSHIP NETWORK DATA, FREQUENCY DISTRIBUTIONS.



(a) Density (b) Isolation

(c) Ave. Harmonic Centrality (d) Disconnection

10 Results

10.1 Correlation Analysis

WE USE PAIR-LEVEL DATA TO INVESTIGATE IF THE PRESENCE OF AN OWNERSHIP TIE RELATES TO THE CO-BIDDING BEHAVIOR OF FIRMS. FOR EACH MARKET M, WE CALCULATE THE POINT BISERIAL CORRELATION COEFFICIENT BETWEEN THE VARIABLES $w_{ij,M}^{[B]}$ (NUMBER OF CO-BIDS) AND $w_{ij,M}^{[O]}$ (OWNERSHIP LINK). THE FREQUENCY DISTRIBUTION OF THESE COEFFICIENTS FOR ALL MARKETS IS SHOWN IN FIGURE 10.1. WE SEE THAT THE CORRELATION BETWEEN OWNERSHIP AND CO-BIDDING DIFFERS ACROSS MARKETS. IN SOME CASES, THE RELATIONSHIP IS NEGATIVE WHILE POSITIVE IN OTHERS. MARKETS IN THE EXTREME SIDES WITH OUTLIER VALUES MAY INDICATE THE PRESENCE OF SUSPICIOUS BEHAVIOR. INDEED, DIVIDING OUR DATASET INTO QUARTILES BASED ON MARKET SIZE (SEE

Figure 10.2), we find that higher negative and positive correlations are more common in smaller markets with less than 9 firms. Based on economic theory, we expect cartels to be more stable when there are fewer firms competing in the market.

FIGURE 10.1. DISTRIBUTION OF PEARSON POINT BISERIAL CORRELATION COEFFICIENTS FOR ALL MARKETS.

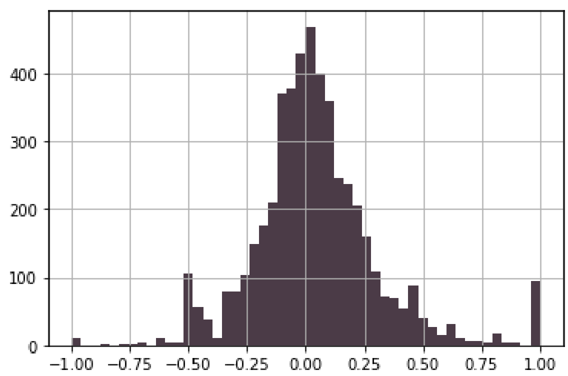
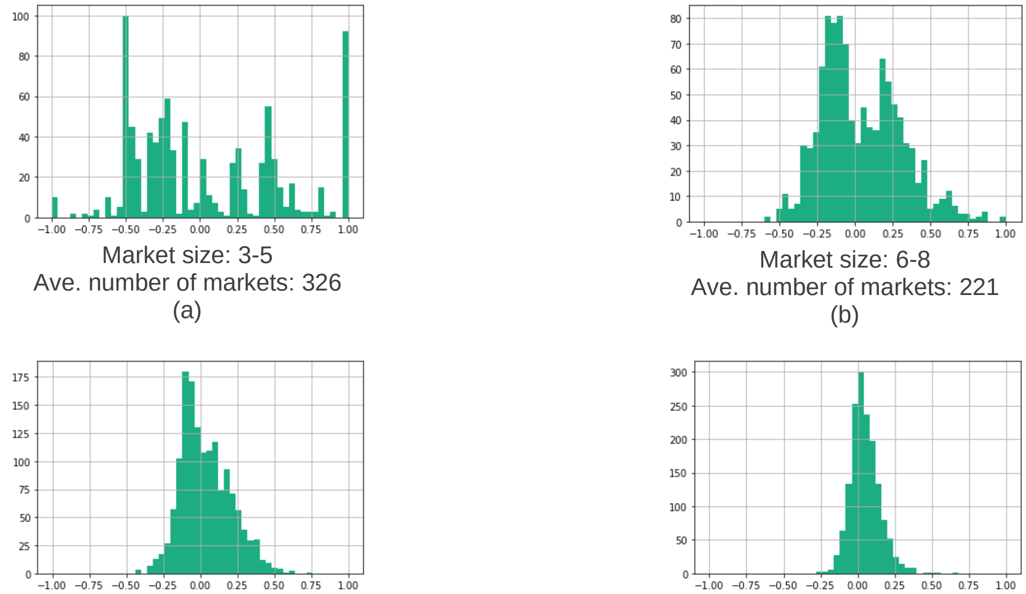


FIGURE 10.2: DISTRIBUTION OF PEARSON POINT BISERIAL CORRELATION COEFFICIENTS FOR DIFFERENT RANGES OF MARKET SIZES.



Market size: 3-5
Ave. number of markets: 326
(a)

Market size: 6-8
Ave. number of markets: 221
(b)

Market size: 9-16
Ave. number of markets: 267
(c)

Market size: 17-320
Ave. number of markets: 262
(d)

We can observe a skew to the right in Figure 10.3, with computed kurtosis of 2.70. This suggests that, often, firms connected by an ownership link are more likely to submit bids for the same contract. We explore the robustness of these results further by looking at the relationship between co-bidding and ownership ties without regard for market definition. This approach has the benefit of also capturing multimarket contact, which can likewise facilitate collusion[50].

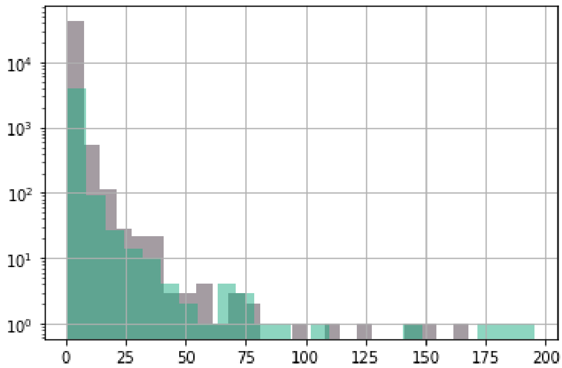


FIGURE 10.3. DISTRIBUTION OF NUMBER OF CO-BIDS.

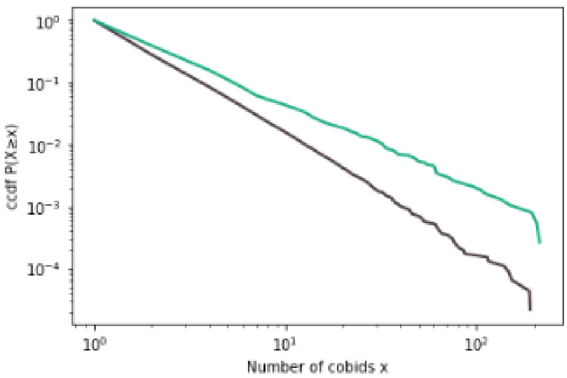


FIGURE 10.4. COMPLEMENTARY CUMULATIVE DISTRIBUTION

Green = with ownership link; Grey = no ownership link

We compare the distribution of the number of co-bids between firms with ownership links versus firms with no ownership link and find a significant difference between the two (see Figure 10.3 and Figure 10.4[51]). We find that similar to the results from the market-level correlations, co-bids appear to be higher between firms with ownership link (see Figure 10.4, for example). The exact reason for such behavior

[50] Such contact may facilitate collusion as punishment for deviation from a collusive equilibrium in one market can also affect other markets.
[51] We also conduct a Mann-Whitney U test to confirm that they are indeed from different distributions.

is uncertain. While such behavior is consistent with bid rigging agreements such as cover bidding and bid rotation, the relationship could also be due to other more innocuous factors. It is, for instance, possible that firms with ownership links are more likely to receive the same information on bidding opportunities. Another possibility is that firms with ownership links are larger and therefore submit more bids, increasing the likelihood of co-bidding with any firm. In fact, we found that on average firms with no ownership link submit bids for only 2 tenders in a year, while the average for those with at least one ownership link is 9.

10.2 Regression Analyses

Our succeeding analyses explores the relationship between ownership network structure and collusion risk. We run the linear regressions on its potential determinants as described in the Methods section. For network order, harmonic closeness centrality, and degree centrality, we use log transformations. Network order, which indicates the number of firms in the relevant market, is used as a control variable in the market-level regressions.

Single bidding

Our first market-level indicator is single bidding, defined as the proportion of tenders that receive only one bid. In the presence of cover bidding, we expect the incidence of single bidding to be smaller.
The results of the simple linear regression using our two proposed network descriptors are shown in Table 10.1. We find that the relationship between proportion of single bidding is negative for both average harmonic closeness centrality and density. Recall that higher values of closeness centrality indicate shorter paths between firms. We therefore find that shorter paths in the ownership network, which signal higher potential for collusion, is associated with a lower proportion of single bidding. This relationship is consistent with cover bidding, as colluding firms submit more bids to maintain the illusion of competition. The effect of network density is the same: markets with more dense ownership structure have lower rate of single bidding.

TABLE 10.1. REGRESSION RESULTS FOR SINGLE BIDDING (MARKET LEVEL, DEPENDENT VARIABLE = PROPORTION OF CONTRACTS WITH SINGLE BIDDING).

	Model 1 b/se	Model 2 b/se
Year		
2010	0.0000 (.)	0.0000 (.)
2011	0.0140* (0.006)	0.0038 (0.005)
2012	0.0064 (0.006)	0.0034 (0.005)
2013	0.0052 (0.006)	0.0035 (0.005)
2014	0.0145* (0.006)	0.0078 (0.005)
2015	0.0262*** (0.006)	0.0175** (0.005)
ln(Order)	-0.0097* (0.005)	-0.0104** (0.003)
ln(Ave. Harmonic Centrality)	-0.0062* (0.003)	
Density		-0.0220* (0.009)
constant	0.0701*** (0.010)	0.0698*** (0.006)
r2	0.0092	0.0036
N	6894	13613

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Missing Bidders

The missing bidders indicator is based on the number of bids submitted. It is a less extreme version of single bidding, indicating the proportion of competitors that did not a submit a bid for a tender. The average is used for this market-level analysis. A high value of the indicator can indicate that a bid suppression scheme is operating in the market; however, a low value may signal cover bidding.

We find that the relationship between proportion of missing bidders and closeness centrality is negative, while that with density is not significant. As discussed above, markets with higher average harmonic centrality are expected to have less intense competition due to easier coordination and faster information flow between companies. Our results indicate that in markets with more firms that are more closely linked through ownership, the average percentage of missing bidders is lower. This finding is consistent with our result from single bidding—in markets where collusion risk due to ownership links is higher, more firms consistently submit bids.

The lack of significance of density suggests an interesting finding: in the case of missing bidders, the topology of the links (who is connected to whom) matters more than the number of links between firms in the ownership network.

TABLE 10.2. REGRESSION RESULTS FOR MISSING BIDDERS (MARKET LEVEL; DEPENDENT VARIABLE = AVERAGE PERCENTAGE OF MISSING BIDDERS).

		Model 1 b/se	Model 2 b/se
Year			
	2010	0.0000	0.0000
		(.)	(.)
	2011	0.0101	0.0093
		(0.008)	(0.006)
	2012	0.0051	0.0094
		(0.008)	(0.006)
	2013	-0.0030	-0.0045
		(0.008)	(0.006)
	2014	-0.0070	-0.0110
		(0.008)	(0.006)
	2015	-0.0133	-0.0159*
		(0.008)	(0.007)
ln(Order)		0.2117***	0.2408***
		(0.006)	(0.004)
ln(Ave. Harmonic Centrality)		-0.0111**	
		(0.004)	
Density			0.0045
			(0.011)
constant		0.0576***	-0.0384***
		(0.012)	(0.008)
r2		0.2922	0.3289
N		6894	13613

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Winning Probability

Winning probability is a firm-level indicator defined as the share of contracts a company has won out of all the bids they submitted for a given relevant market. Our interest lies in the relationship between winning probability and firm-level ownership network measures. Specifically, the degree (i.e., the number of other firms it shares an ownership link with) and the harmonic closeness centrality. In network analysis, closeness centrality is a way of detecting nodes that are able to spread information very efficiently through a graph.

The predicted effect of ownership connections on the winning probability of a firm is ambiguous. A collusive ring engaged in cover bidding can lead to some firms having artificially high winning rates while those participants that submit a larger number of courtesy bids may have artificially low winning rates. Our results show that firms with more ownership connections have lower winning probability. The negative relationship may signal the presence of firms that submit a high number of courtesy bids.

TABLE 10.3. REGRESSION RESULTS FOR WINNING PROBABILITY (FIRM LEVEL; DEPENDENT VARIABLE = WINNING PROBABILITY OF FIRM).

	Model 1 b/se	Model 2 b/se
Year		
2010	0.0000	0.0000
	(.)	(.)
2011	0.0220***	0.0219***
	(0.004)	(0.004)
2012	0.0166***	0.0155***
	(0.003)	(0.003)
2013	0.0049	0.0038
	(0.003)	(0.003)
2014	0.0265***	0.0255***
	(0.003)	(0.004)
2015	0.0366***	0.0357***
	(0.004)	(0.004)
ln(Degree)	-0.0031**	
	(0.001)	
ln(Closeness Centrality)		-0.0033**
		(0.001)
constant	0.3844***	0.3756***
	(0.003)	(0.003)
r2	0.0053	0.0053
N	34806	34806

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Cut-point position

A company is defined to be in a cut-point position if its removal from the co-bidding network increases the number of connected components of the graph. We use a binary indicator to define if a firm is in a cut-point position and use a logistic model for this regression. Consistent with expectation, we find a positive relationship between cut-position and degree and closeness centrality. This indicates that firms with more ownership connections are more likely to be in a cut-point position in the co-bidding network. Firms with high degree or closeness centrality may indicate core actors in a collusive network (Xiao, Ye, Zhou, Ye, & Tekka, 2021) and their importance is emphasized by their position the co-bidding network. These firms in cut-point positions may also indicate the presence of ringleaders or actors taking a central role in the coordination and formation of a cartel.

TABLE 10.4. REGRESSION RESULTS FOR CUT-POINT POSITION (FIRM LEVEL; DEPENDENT VARIABLE = BINARY CUT-POINT POSITION INDICATOR).

	Model 1 b/se	Model 2 b/se
Year		
2010	0.0000	0.0000
	(.)	(.)
2011	-0.0904	-0.0868
	(0.082)	(0.082)
2012	-0.0684	-0.0159
	(0.075)	(0.075)
2013	-0.2181**	-0.1586*
	(0.076)	(0.076)
2014	-0.0547	-0.0061
	(0.076)	(0.076)
2015	-0.0170	0.0367
	(0.076)	(0.076)
ln(Degree)	0.1540***	
	(0.025)	
ln(Closeness Centrality)		0.1324***
		(0.025)
r2		
N	24003.0000	24003.0000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Stability

Market share stability can signal collusive behavior under certain conditions. We define stability as one minus the average absolute change in market shares from the first half to the second half of a year. The lower the value, the less stable the market. Since there is insufficient data on the value of contracts, we used the share of the number of tenders won by each firm. Markets with bid-rigging cartels are expected to have more stable market shares. We therefore expect the relationship between this indicator and our ownership network indicators to be positive. A weakness of the stability indicator is that it is unable to capture partial cartels. If not all active firms in a relevant market participate in the collusive arrangement, then market shares may fluctuate as the number of bids won by these firms changes (Harrington, 2007). Indeed, our results in Table 10.5 show that while there is a positive relationship between stability and our ownership network descriptors, this relationship is not significant when we control for the size of the relevant market.

TABLE 10.5. REGRESSION RESULTS FOR STABILITY (MARKET LEVEL; DEPENDENT VARIABLE = STABILITY).

		Model 1 b/se	Model 2 b/se
Year			
	2010	0.0000	0.0000
		(.)	(.)
	2011	0.0056	0.0061
		(0.009)	(0.008)
	2012	-0.0080	0.0071
		(0.009)	(0.008)
	2013	0.0078	0.0216**
		(0.009)	(0.008)
	2014	0.0138	0.0138
		(0.009)	(0.008)
	2015	-0.0019	0.0032
		(0.009)	(0.008)
ln(Order)		0.0596***	0.0571***
		(0.007)	(0.005)
ln(Ave. Harmonic Centrality)		0.0077	
		(0.004)	
Density			0.0126
			(0.020)
constant		0.6980***	0.6889***
		(0.017)	(0.012)
r2		0.0426	0.0374
N		4292	6120

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A summary of our findings is shown in Table 10.6. Cells highlighted in green are those that are consistent with our expected results. We find that in markets with firms that are more closely related through ownership links, there is smaller incidence of single bidding and less missing bidders. These results also support our findings from the correlation analysis that showed firms with an ownership link are more likely to bid on the same contract. However, the relationship between ownership network density and missing bidders is not significant, suggesting that the structure of the relationships between firms matters more than just the number of links. We also find that ownership indicators have no significant relationship with the stability of market shares. This may be due to the weakness of the indicator as a measure of collusion.

For our firm-level analyses, we find a positive relationship between cut-point position and our ownership structure indicators. This shows that firms with more ownership links are in an important position in the co-bidding network. Winning probability, on the other hand, has a negative relationship. In the presence of cover bidding with a large number of courtesy bids, this result is not unexpected.

TABLE 10.6. ACTUAL RELATIONSHIP BASED ON REGRESSION RESULTS.

		Market-level ownership structure indicators		Firm-level ownership structure indicators	
Group	Indicator Name	**Average harmonic closeness centrality**	**Density**	**Harmonic closeness centrality**	**Degree**
Bidding pattern	**Single bidding**	negative	negative		
	Missing bidders	negative	not significant		
	Winning probability			negative	negative
	Cut-point position			positive	positive
Market structure	**Stable market structure**	not significant	not significant		

11 Next Steps and Future Work

Our study provides groundwork for future research on the competitive effects of structural links in procurement markets. One the one hand, the methodology can be refined using the same data, while on the other hand, the approach can be expanded to incorporate more data and other contexts.

First, our approach using correlations and regressions could only identify the average effects and general relationships. However, it is quite possible that collusion in Sweden is more of an outlier behavior rather than the norm. So further work, could depart from the regression results and instead look at extreme or outlier cases where the hypothesized relationships are the strongest. This would mean to flag cases with the densest ownership ties and collate them with the highest value cartel screens (i.e., highest cartel risks). In addition, convergence among different ownership and competitive behavior-based indicators should not be bivariate because cartels are most likely to leave multiple markers at the same time. This implies that on top of looking at bivariate relationships such as ownership density and single bidding, it could further strengthen measurement validity to look at sets of indicators corresponding to the same cartel type. This could be done, for example, by looking at the highest risk markets indicator by indicator until a subset of markets of companies are identified where are relevant cartel screens point at cartel presence.

Second, aside from these, there is a need to explore other methods for market definition. Since our analysis uses data aggregated at the market level, our results are highly sensitive to the way markets are defined. As previously mentioned, combining firm and product attributes with co-bidding information to assign relevant markets is a better approach. Recall that our elementary bidding network is a bipartite graph $G = (U, V, E)$, whose nodes can be partitioned into two disjoint and independent sets U and V, such that there is no edge between nodes belonging to the same set. In this case V is the set of firms and U is the set of tenders. Two firms $i, j \in V$ can be considered competitors and belonging to the same relevant market if either (a) there exists a path $\{i, u, j\}$, $u \in U$, or (b) j is reachable[52] from i subject to certain constraints related to the attributes of the nodes on the path from i to j. Using this approach should allow us to identify both the actual and potential competitors in a market. Determining the appropriate constraints should be part of the research.

Third, a natural next step is to expand the scope of the investigation by including data from other countries and conducting cross-country comparisons. This would also allow us to test other collusion risk indicators that use price data, which is available for certain countries. Another valuable addition to this work would be to include a measure of ownership stakes between firms. With this information we can distinguish among the types of ownership links, enabling more sophisticated analysis.

Lastly, we suggest studying other types of structural links. Our focus so far has been on ownership; however, other connections between firms—such as director interlocks, joint bidding agreements, and

[52] A node j is defined to be reachable from i if there exists a path which starts with i and ends with j.

subcontracting—may likewise have anti-competitive effects. These links also provide the possibility for firms to control or influence the behavior of their competitors, aside from providing a channel to transfer information that can enhance the stability of a collusive equilibrium. Such research, identifying and characterizing the complex interactions between firms and investigating their impact on competition, can help guide discussions of policy responses to structural links.

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.

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

Asai, K., & Charoenwong, B. (2019). Ownership Networks and Bid Rigging.

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

Athey, Susan, and Kyle Bagwell. (2001). Optimal Collusion with Private Information. The RAND Journal of Economics 32, no. 3 428–65.

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

Azar, J., Raina, S., & Schmalz, M. (2021). Ultimate Ownership and Bank Competition. Financial Management. 2022; 51: 227– 269.

Azar, J., Schmalz, M. C., & Tecu, I. (2018). Anticompetitive Effects of Common Ownership. The Journal of Finance, 73(4), 1513-1565.

Backus, M., Conlon, C., & Sinkinson, M. (2019). The Common Ownership Hypothesis: Theory and Evidence. Brookings Economic Studies.

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. University of Kentucky Doctoral Dissertations. 191. https://uknowledge.uky.edu/gradschool_diss/191

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

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, Mihály, Luciana Cingolani, & Bence Tóth (2018), Innovations in Objectively Measuring Corruption in Public Procurement. In Helmut K. Anheier, Matthias Haber, and Mark A. Kayser (eds.) Governance Indicators. Approaches, Progress, Promise. Oxford University Press, Oxford. Ch. 7. Pp. 154-185.

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

Freeman, L. C. (1978). Centrality in Social Networks: Conceptual Clarification. Social Networks, 1(3), 215-239.

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

Gilo, D. (2001). The Anticompetitive Effect of Passive Investment. Michigan Law Review, 99(1), 1-47.

Gilo, D., Moshe, Y., & Spiegel, Y. (2006). Partial Cross Ownership and Tacit Collusion. Source: The RAND Journal of Economics RAND Journal of Economics, 37(1), 81-99.

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.

http://www.oecd.org/competition/cartels/42594486.pdf

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. (2020). Transnational machine learning with screens for flagging bid-rigging cartels. Université de Fribourg.

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, K., & Nakabayashi, J. (2014). Detecting large-scale collusion in procurement auctions. Available at SSRN 2467175.

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.

Levy, N., Spiegel, Y., & Gilo, D. (2018). Partial Vertical Integration, Ownership Structure, and Foreclosure. American Economic Journal: Microeconomics, 10(1), 132-180.

Malueg, D. (1992). Collusive behavior and partial ownership of rivals. International Journal of Industrial Organization.

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.

Mizruchi, M. S. (1996). What Do Interlocks Do? An Analysis, Critique, and Assessment of Research on Interlocking Directorates. Annual Review of Sociology, 22, 271-298.

Morozov, I. & Podkolzina E. (2013) Collusion detection in procurement auctions. Higher School of Economics Research Paper No. WP BPR 25/EC/2013

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

Murray, G. (2006). Capitalist Networks and Social Power in Australia and New Zealand. Aldershot: Ashgate Publishing Limited.

O'Brien, D., & Salop, S. (2000). Competitive Effects of Partial Ownership: Financial Interest and Corporate Control. Antitrust Law Journal, 67(3), 559-614.

O'Brien, D., & Waehrer, K. (2017). The Competitive Effects of Common Ownership: We Know Less than We Think. SSRN Electronic Journal.

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. (1993). Detection of bid rigging in procurement auctions. Journal of political economy, 101(3), 518-538.

Porter, R. H., & Zona, J. D. (1997). Ohio school milk markets: An analysis of bidding (No. w6037). National Bureau of Economic Research.

Reynolds, R., & Snapp, B. (1986). The Competitive Effects of Partial Equity Interests and Joint Ventures. International Journal of Industrial Organization, 4(2), 141-153.

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

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.

Shelegia, S., & Spiegel, Y. (2015). Cost-reducing investments under partial cross-ownership. 114(1).

Takes, F. W., & Heemskerk, E. M. (2016). Centrality in the global network of corporate control. Social Network Analysis and Mining, 6(97).

Takes, F. W., Kosters, W. A., Witte, B., & Heemskerk, E. M. (2018). Multiplex network motifs as building blocks of corporate networks. Applied Network Science, 3(39).

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.

van Lidth de Jeude, J., Aste, T., & Caldarelli, G. (2019). The Multilayer Structure of Corporate Networks. New Journal of Physics, 21.

Vitali, S., Glattfelder, J. B., & Battiston, S. (2011). The Network of Global Corporate Control. PLoS ONE, 6(10).

Vives, X. (2020). Common ownership, market power, and innovation. International Journal of Industrial Organization, 70.

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.

Werden, G. J., & Froeb, L. M. (2018). Don’t Panic: A Guide to Claims of Increasing Concentration. Antitrust Magazine, 34(2).

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

WP BRP 25/EC/2013. National Research University Higher School of Economics.

12 Appendix A

12.1 Finding consortia

To find consortium bids, company names were cleaned by standard text cleaning steps (for example, removing special characters, setting them lowering case etc.).
After that consortium companies were classified by a 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

12.2 Country datasets

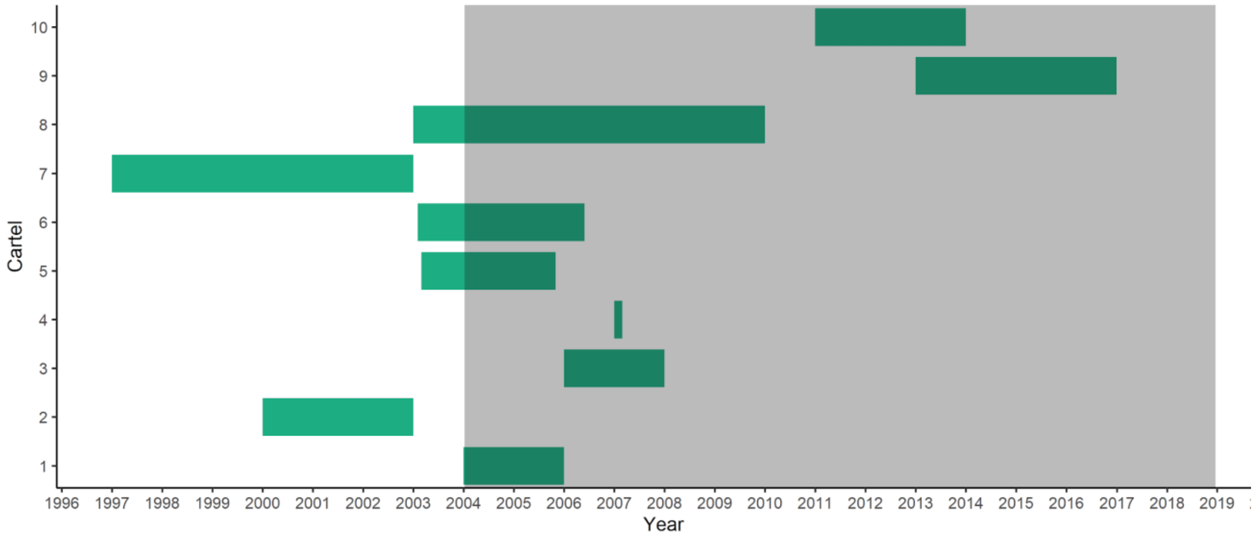
12.2.1 France

The available data for analysis of anti-competitive behavior of firms in France include information on ten cartels that operated in the country and public procurement records. The cartels selected for this report operated in France from 1997 (the oldest available cartel case #7) to the end of 2016 when the most recent cartel #9 was discovered. The available procurement data provide contract-level tendering information collected by DIGIWHIST coming from both Europe-wide TED as well as national data sources. Our dataset only includes contract level information (i.e. no bid details), with a similar level of tender details as in Spain or Latvia.

The dataset covers the whole contract value range of procurement records above the national minimum threshold for reporting between 2004 to 2019[53]. The number of contracts is evenly distributed across years. From the analyzed sample we exclude outright awards given that in such contracts companies do not have control over whether they are contracted directly by procuring bodies. The cleaned and filtered dataset used for the analysis includes 1.83 million awarded contracts.

Figure 12.1 shows the overlap between the collusive period and our data coverage by each cartel. Cartels #2 and #7 operated over 1997-2003 and 2000-2003, respectively, which are out of the scope of contracting data. For cartels #1, #5, #6, and #8 we have contracts only from during and after the cartel period, while for more recent cartels #3, #4, #9, and #10, we also have contracts awarded to the cartel members before the proven start of the collusive period. Of note, the lifespan of cartel #4 is significantly shorter than those of other available cartels, which might pose challenges with identifying contracts awarded to cartel members during the period of cartel activities.

FIGURE 12.1.OVERLAP BETWEEN PUBLIC CONTRACTING DATA COVERAGE (GREY SPACE) AND CARTEL (GREEN BARS) TIME PERIODS

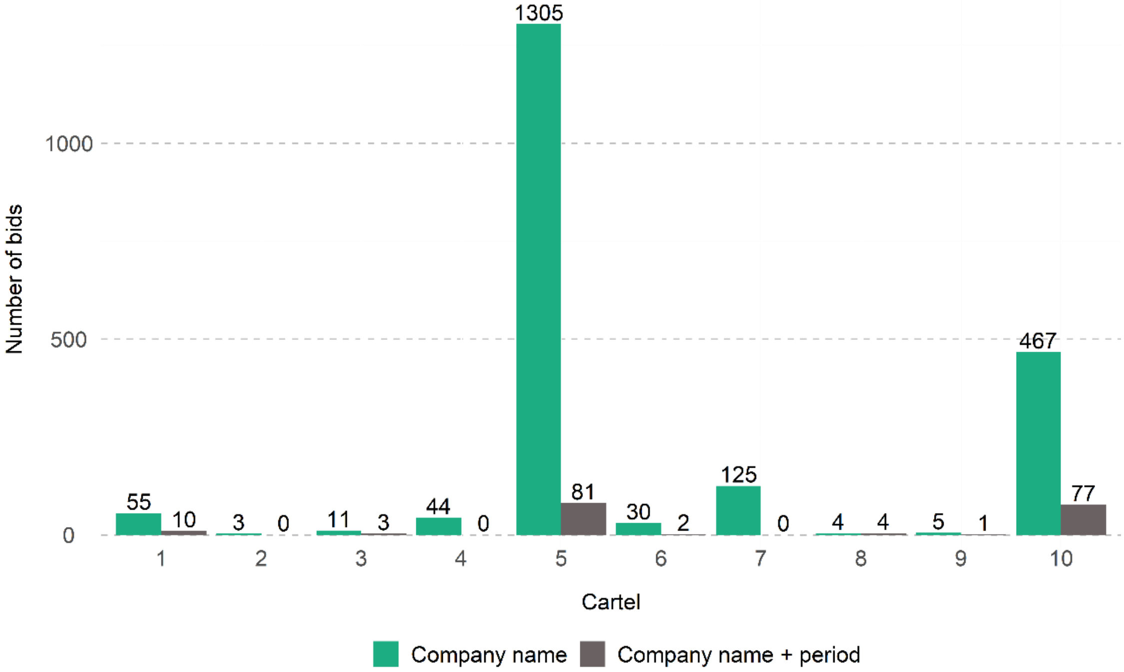


The number of contracts varies significantly by the different sample definitions and cartels (

Figure 12.2). We found a relatively high number of contracts for cartels #1, #5, and #10 for all three sample variants. At the same time, although cartel members #4 and #7 also won contracts, all of these were outside of the proven cartel period or were out of the scope of our dataset entirely.

[53] Note that high-value tenders, that are published on TED, are only available since 2006.

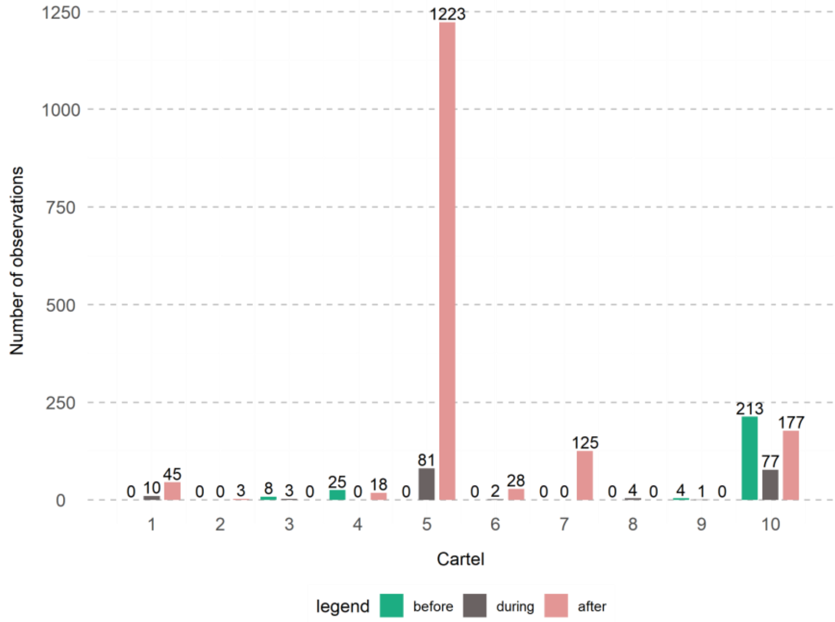
FIGURE 12.2. NUMBER OF MATCHED CONTRACTS BY SAMPLE DEFINITIONS



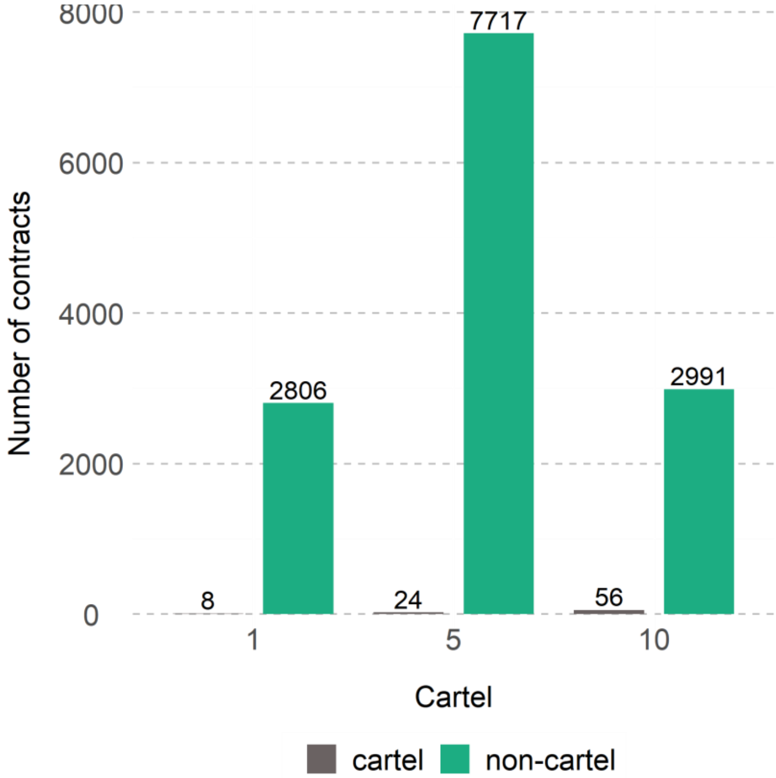
As we explained in the Methods and Data sections, we group contracts by the two implemented tests (a) during vs. after, b) cross-sectional, hence we show the number of contracts in each contract group that we use for statistical testing (Figure 12.3 – Panel A) shows the number of contracts used in during vs. after the cartel comparison. Out of ten available cartels, only four (#1, #5, and #10) have enough contracts both during and after the cartel period for statistical testing. The rest of the cartels are excluded from during-after testing due to the lack of enough data. For cross-sectional comparison Panel B, cartels potentially suitable for testing are #1, #5, and #10. Note that we show the number of potentially available contracts for risk calculations, hence the final number of contracts with calculable individual indicators can be lower due to missing data.

FIGURE 12.3. NUMBER OF CONTRACTS USED IN STATISTICAL TESTING

A: During-after comparison



B: Cross-sectional comparison

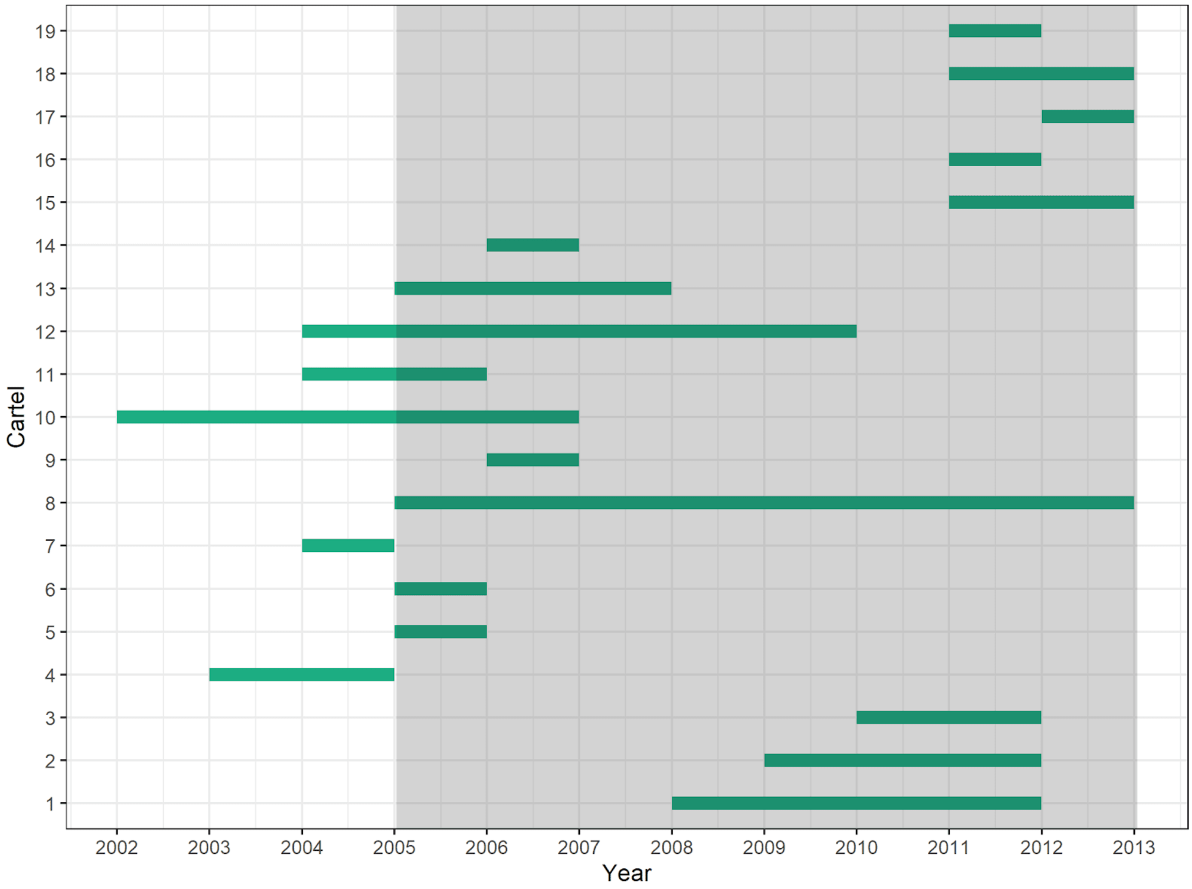


12.2.2 Hungary

We use bid-level tendering data between 2005 and 2013 to test cartel indicators in Hungary. We restricted the dataset to these years as information on losing bidders were only published in this time period. The number of observations varies greatly across years: the number of observations significantly lower - varying around 4000 - until 2009, and it increases for later years (possibly due to the change in publication rules and natural fluctuation in public spending). 2010 is an outlier with almost 15000 observations.

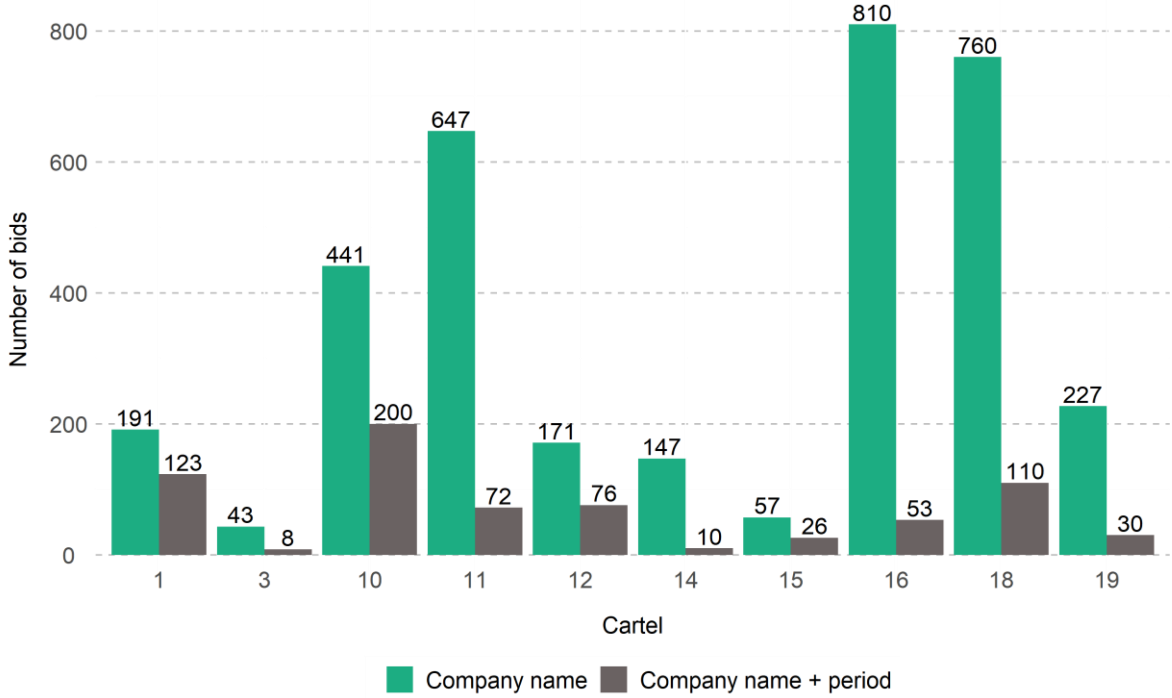
Public contracting data does not cover the proven collusion period of Cartels 4 and 7 (Figure 12.4). Furthermore, we do not have data with losing bidders for the after-cartel period for cartels 8, 15, 17 and 18, hence the during vs. after cartel period comparisons cannot be estimated. However, we can still estimate cross-sectional differences due to their long enough lifespans.

FIGURE 12.4. OVERLAP BETWEEN PUBLIC CONTRACTING DATA COVERAGE (GREY) AND CARTEL TIME PERIODS



As in all other countries, the number of matching contracts varies greatly by the applied filters (Figure 12.5). For example, while we could identify 810 contracts won by cartel 16, only 53 were advertised during the cartel period. The number of matches is often less than 5 for other cartels - that are omitted from Figure 12.5. While we found a number of contracts won by the members of cartels 11, 12 16 and 19 most of these tenders were outside of the proven cartel period.

FIGURE 12.5. NUMBER OF CONTRACTS AWARDED BY SAMPLE DEFINITIONS[54]

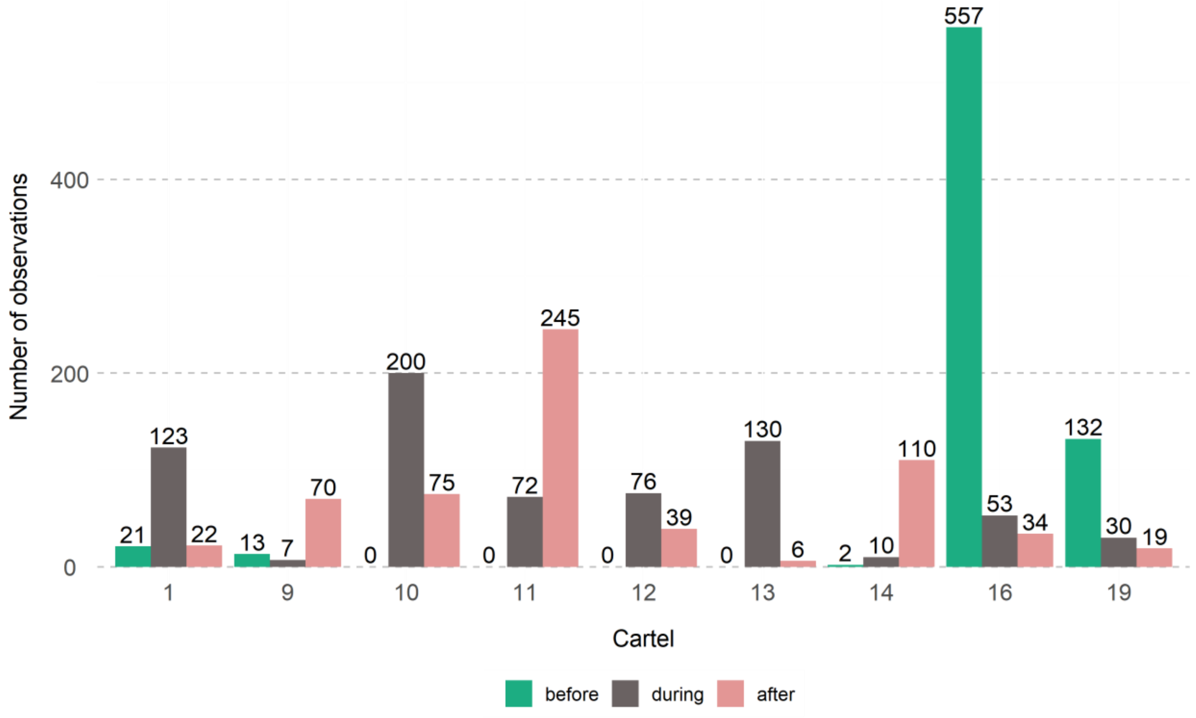


The next three figures show the number of contracts that we use for statistical testing a) during and after the cartel period, b) for cross-sectional. Several cartels are out of scope for the analysis due to lack of enough observations (omitted cases from Figure 12.6). No cartels have at least 50 contracts both during and after the cartel period. For instance, cartel 1 won a relatively high number of contracts during the cartel period (83) but it won only 20 contracts after the proven cartel period (Panel A). Note, that Figure 12.6 shows the number of potentially available contracts for risk calculations and the number of contracts with calculable individual indicators can be less.

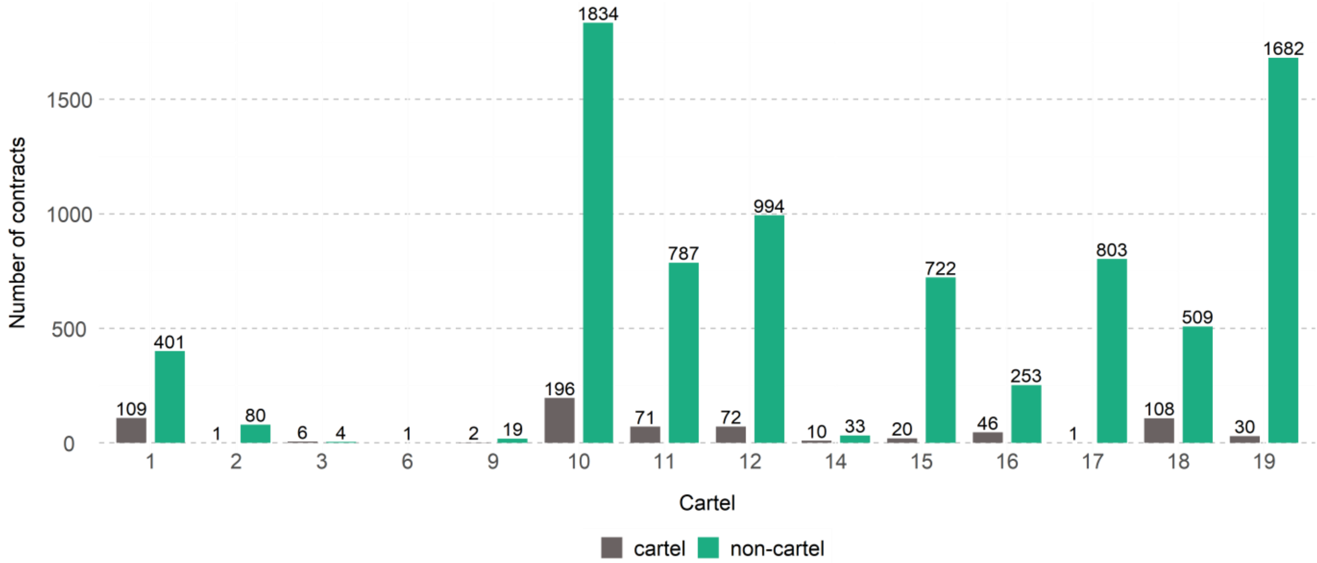
Panel B shows the potential number of observations during the cartel split into rigged and competitive contracts. Most often we have enough observations for non-cartel tenders but the low number of rigged contracts decreases the chance for stable estimations.

[54] Cartel cases are omitted where we have less than 5 observations in any of the three dimensions.

FIGURE 12.6. NUMBER OF CONTRACTS USED IN THE THREE TESTS
Panel A: During-after[55]



Panel B: Cross-section[56]

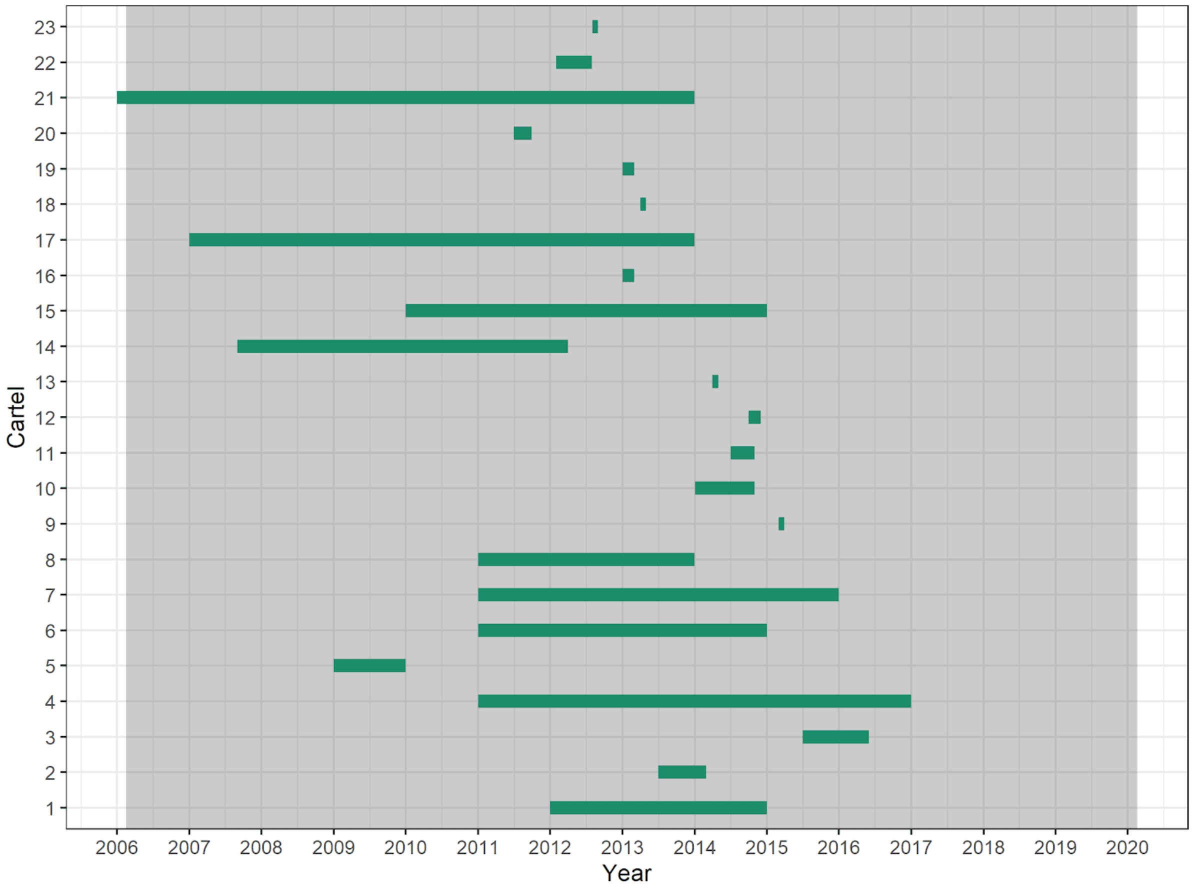


55 Cartels with less than 5 observations either during or after the collusive period are excluded.
56 Cartels with less than 5 contracts in the cartel or non-cartel group are excluded.

12.2.3 Latvia

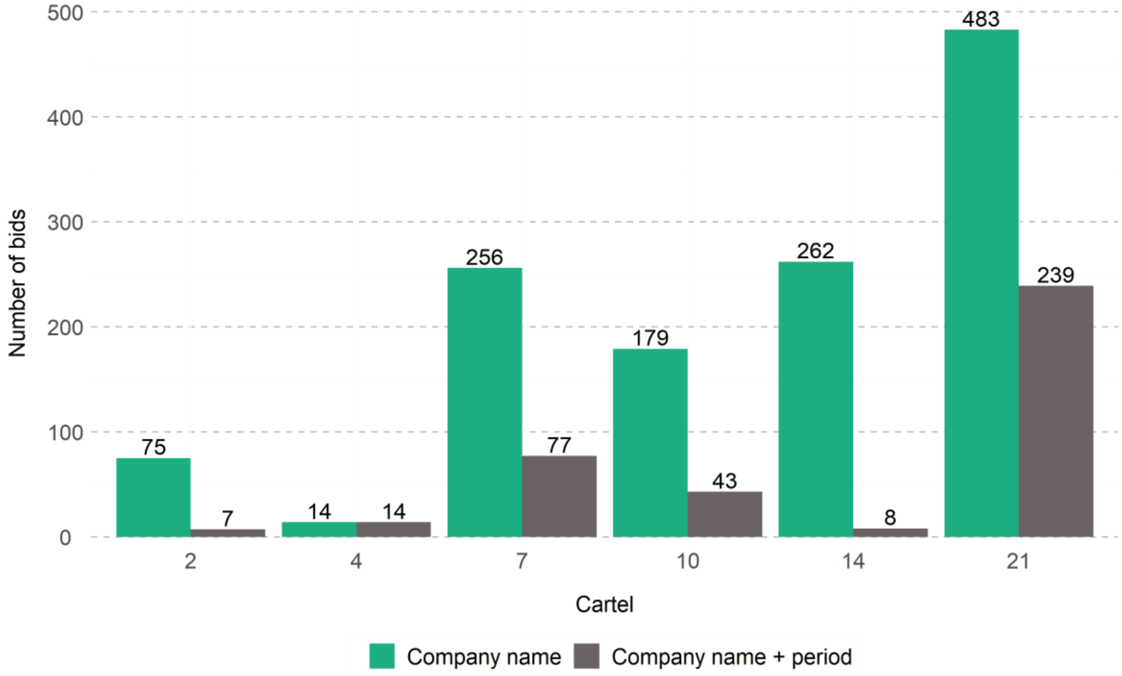
We use contract level tendering data that can be split into three periods between 2006 and 2020 based on the number of observations. Only a few thousand contracts were published up to 2010. In 2011 and 2012 the number of recorded contracts were less than 10 thousand. Between 2013 and 2020 there were approximately 20 thousand contracts in each year. Thus, the later the cartel period is, the higher the chance to have large enough sample size for performing tests. As for other countries, we removed outright awards and innovation partnerships from the analysed set of tenders as companies do not have control over whether they are contracted directly (they can only reject it). The Gantt chart shows that the proven cartel period is extremely short in many cases such as cartels 9, 11, 12, 13, 16,18, 19, 20 and 23. But we have a long enough time period after the cartels' dissolution in each case.

FIGURE 12.7. OVERLAP BETWEEN PUBLIC CONTRACTING DATA COVERAGE (GREY) AND CARTEL TIME PERIODS (GREEN)



The number of matching contracts varies greatly by the filters we apply (Figure 12.8). For example, while we could identify 483 contracts won by cartel 21, only 239 of them were advertised during the cartel period. While we found many contracts won by the members of cartel 2 and 14, most of these tenders were outside of the proven cartel period. The number of matching contracts by these dimensions is often zero for other cartels - as the often extremely short proven cartel periods have already suggested.

FIGURE 12.8. NUMBER OF CONTRACTS AWARDED BY SAMPLE DEFINITIONS[57]



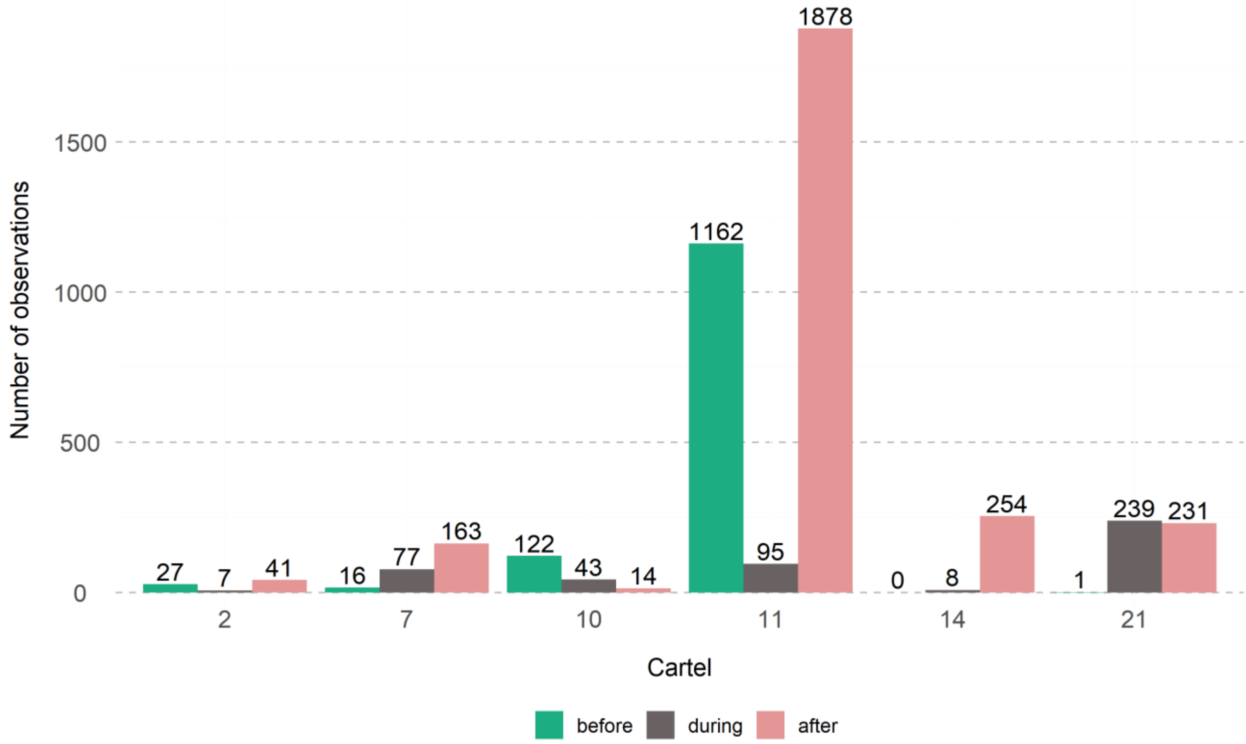
The next two figures show the number of contracts that we use for statistical testing a) during and after the cartel period, and b) cross-sectional.

Several cartels get out of scope from the analysis due to lack of enough observations (Figure 12.9). For instance, only cartels 7, 11 and 21 have potentially enough contracts both during and after the cartel period to compare risk indicator levels before and after the cartel period (Panel A). Note, that Figure 12.9 shows the number of potentially available contracts for risk calculations and the number of contracts with calculable individual indicators can be less - depending on whether the specific indicator values are available.

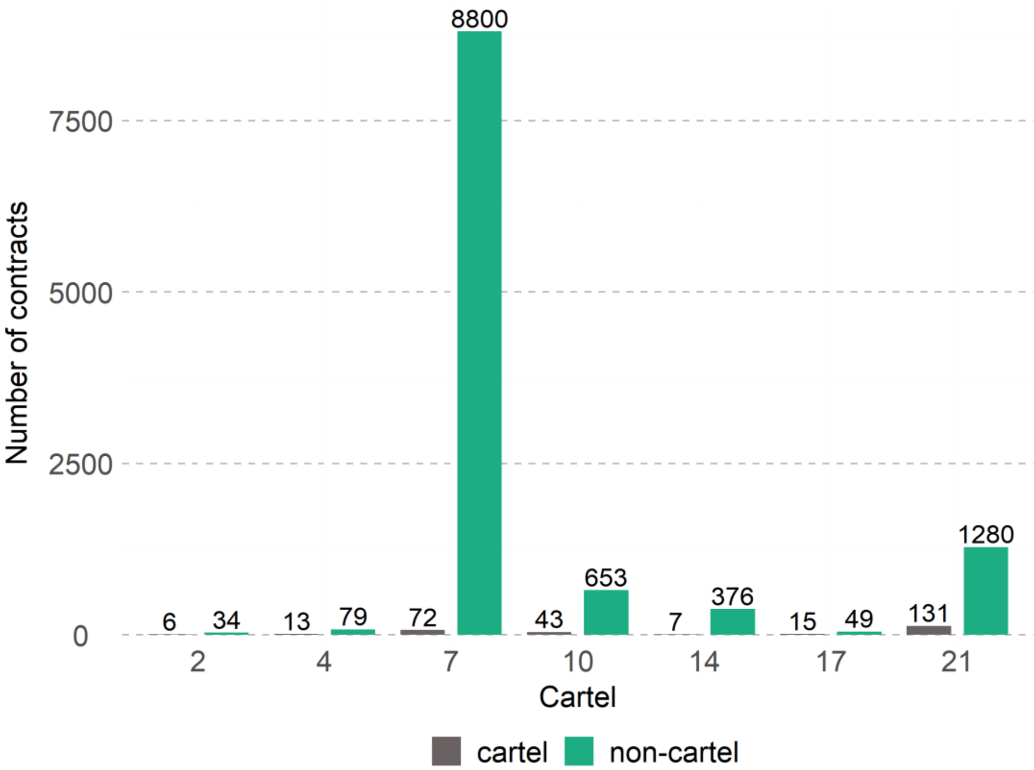
Panel B shows the potential number of observations during the cartel split into cartel and non-cartel contracts. We often have enough contracts from the competitive comparison group but the lack of enough cartel contracts makes statistical testing impossible for many cartels. The most promising cases are cartels 7 and 21 again.

[57] Cartels with less than 5 observations in any of the three dimensions are excluded.

FIGURE 12.9. NUMBER OF CONTRACTS USED IN THE THREE TESTS
Panel A: During-after[58]



Panel B: Cross-section[59]

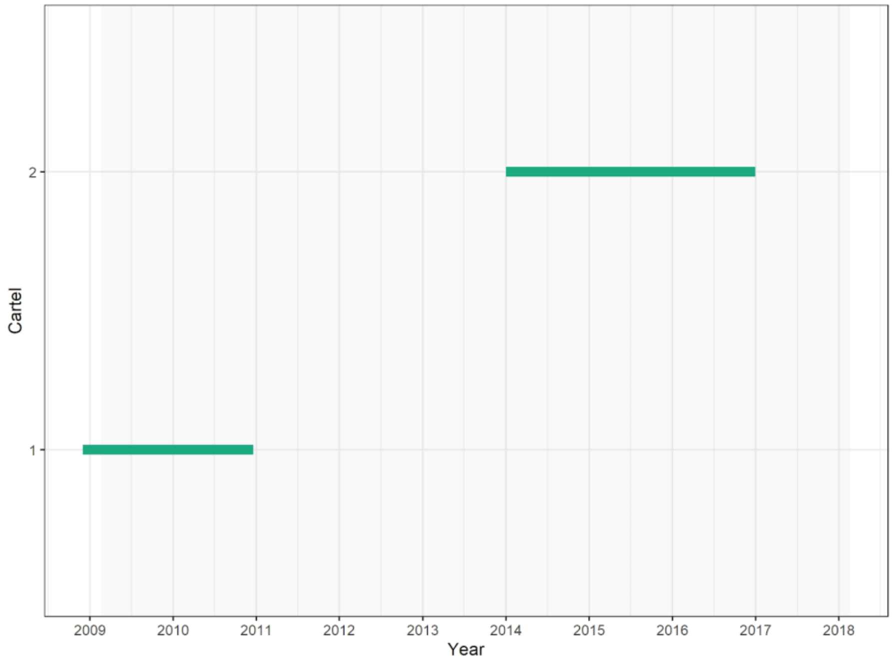


[58] Cartels with less than 5 contracts during or the after the cartel periods are excluded.
[59] Cartels with less than 5 contracts in the cartel or non-cartel group are excluded.

12.2.4 Portugal

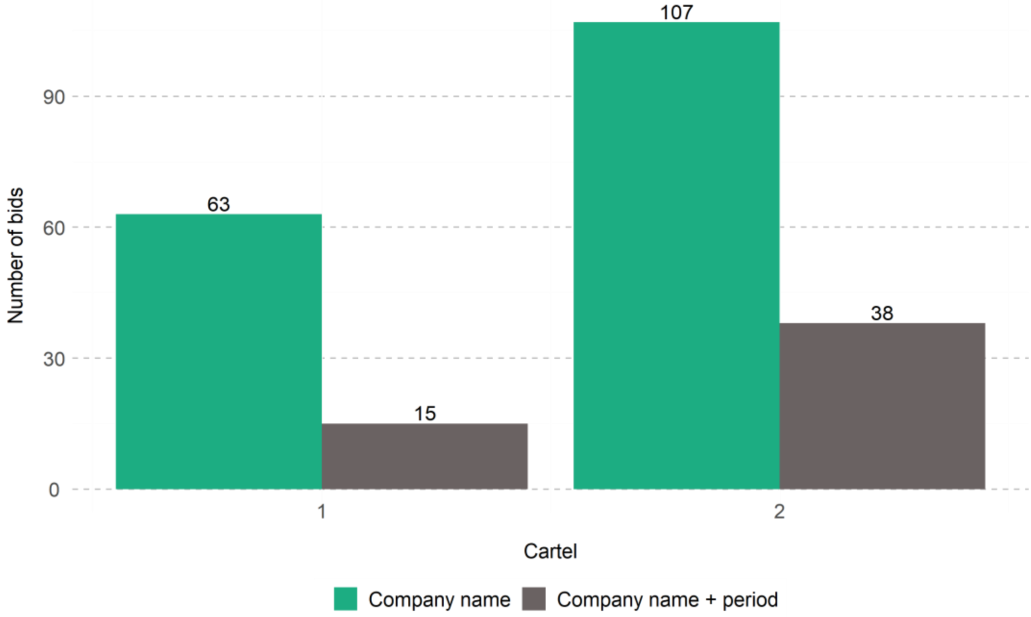
We use bid-level tendering data that contains losing and winning bids as well, that was collected from the central Portuguese website. Due to the way losing bids are published at the Portuguese public procurement website, losing bidders are listed by each tender and are not grouped by the lots they are competing for. Therefore, we can only use tenders in the analysis that awarded one lot (or contract). The data consists of almost 45,000 tenders and 38,000 have only one lot supporting the use of a great share of losing bids. Before 2014 we have less than 6000 contracts per year, and even less than 2500 in 2009 and 2010, which can have an effect on the cartel 1 related tests as the Gantt chart suggests (Figure 12.10). Contract numbers jump up to the ten thousands from 2014 onward, however, the relatively short time period after the end of the 2nd cartel could affect its statistical tests.

FIGURE 12.10. OVERLAP BETWEEN PUBLIC CONTRACTING DATA COVERAGE (GREY) AND CARTEL TIME PERIODS (GREEN)



We find only a couple of dozen relevant contracts for the two cartels that fit our filters (Figure 12.11). For example, while we could identify 107 contracts won by cartel 2, only 38 of them were advertised during the cartel period.

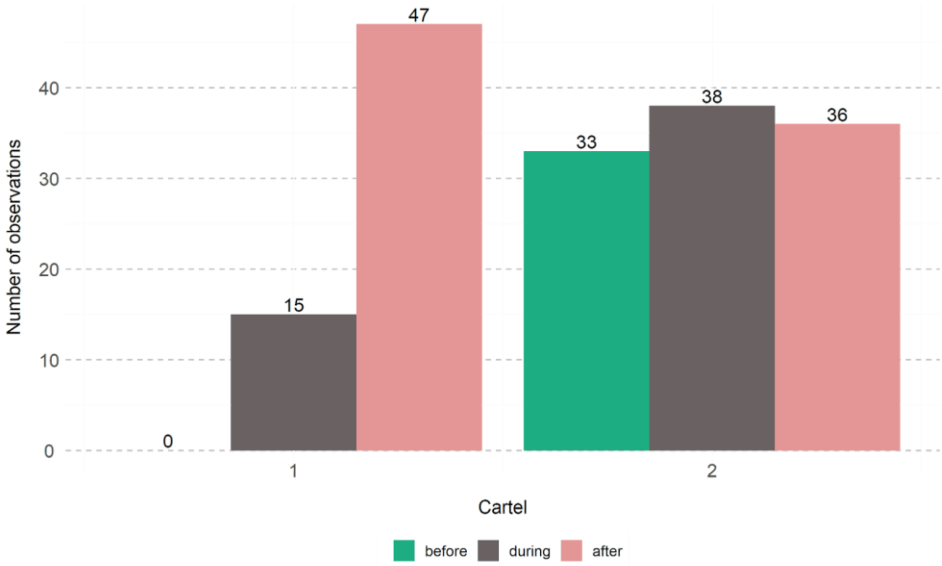
FIGURE 12.11. NUMBER OF CONTRACTS AWARDED BY SAMPLE DEFINITIONS



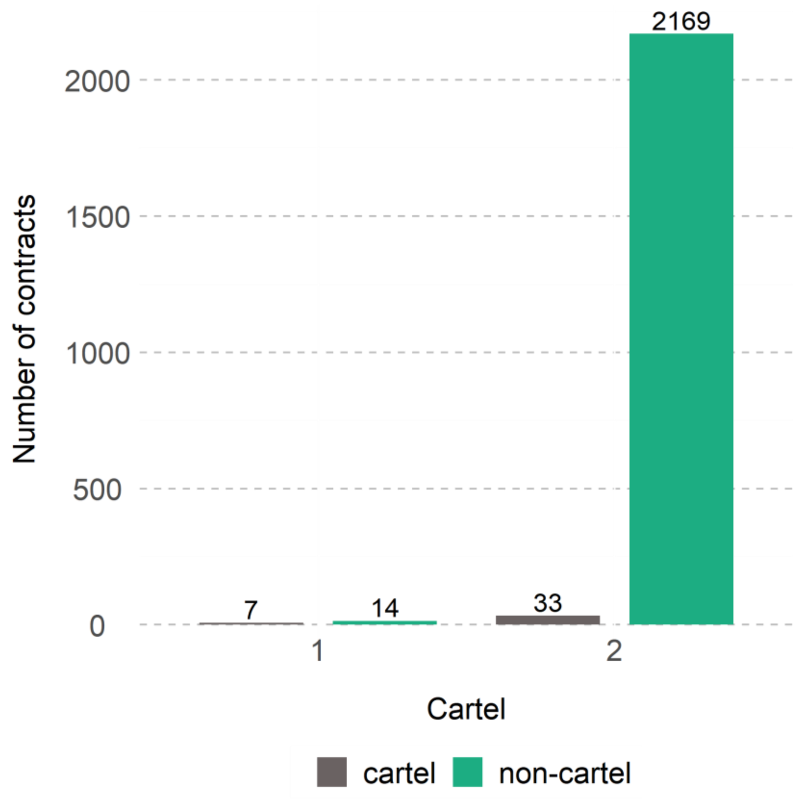
The next two figures show the number of contracts that we use for statistical testing a) during and after the cartel period, and b) for cross-sectional tests. Both cartels have low numbers of matching contracts leading to small sample estimations. Note that all panels of Figure 12.12 show the potentially available contracts for risk calculations and the number of contracts with calculable individual indicators can be less. Panel B shows the number of contracts during the cartel period split into cartel and non-cartel contracts.

FIGURE 12.12. NUMBER OF CONTRACTS USED IN THE THREE TESTS

Panel A: During-after



Panel B: Cross-section



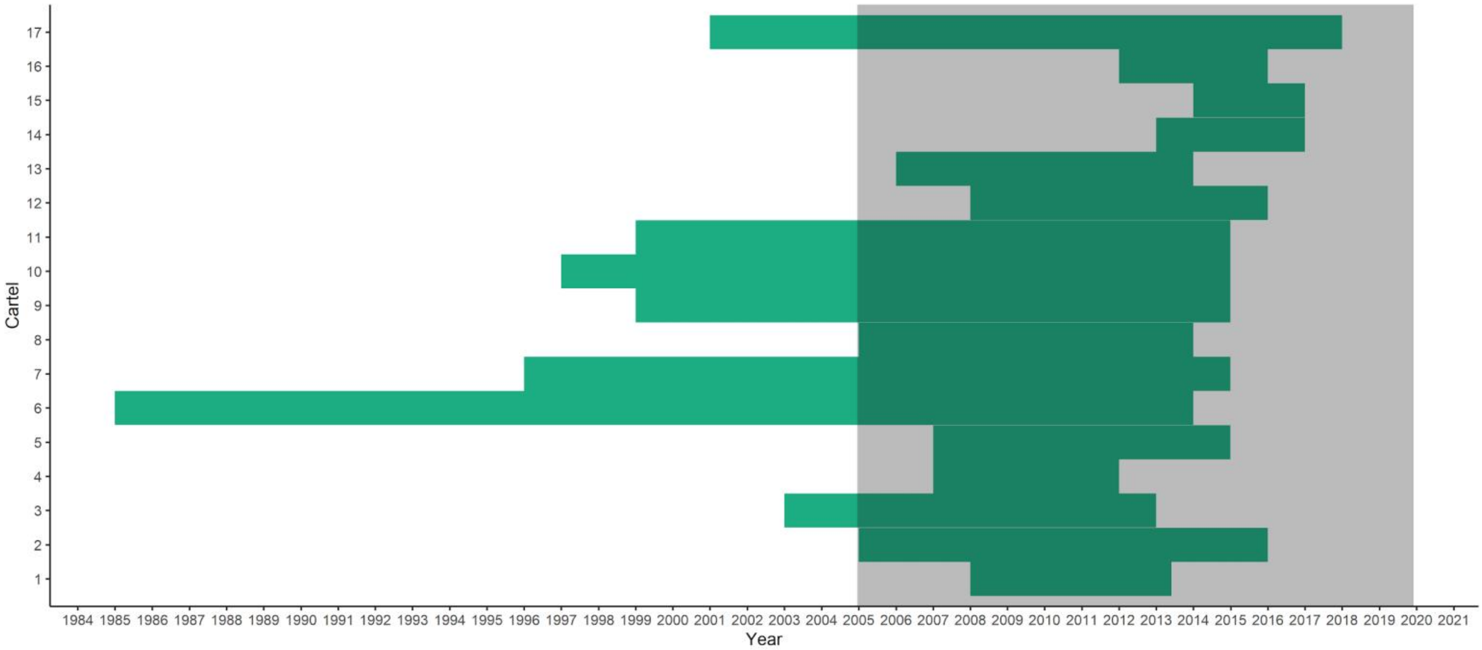
12.2.5 Spain

The public procurement cartel analysis in Spain is based on 17 detected cartels. They were active from 1985 (the oldest available cartel case #6) to the end of 2016 when the most recent cartels #1 and #17 were discovered. The available procurement data provide contract-level tendering information collected by DIGIWHIST coming from both Europe-wide TED as well as national data sources. The dataset has information only on winning companies (i.e. no bid details), with a similar level of tender details as in France or Latvia.

The dataset covers the whole contract value range of procurement records above the national minimum threshold for publication. Procurement records are available from 2006 to 2020. The number of contracts represented in the sample is evenly distributed across years with a slightly increasing trend over time. From the analysed sample we exclude outright awards given that companies do not have control over whether they are contracted directly by procuring bodies in such contracts. The cleaned and filtered dataset used for the analysis includes roughly 341 thousand observations.

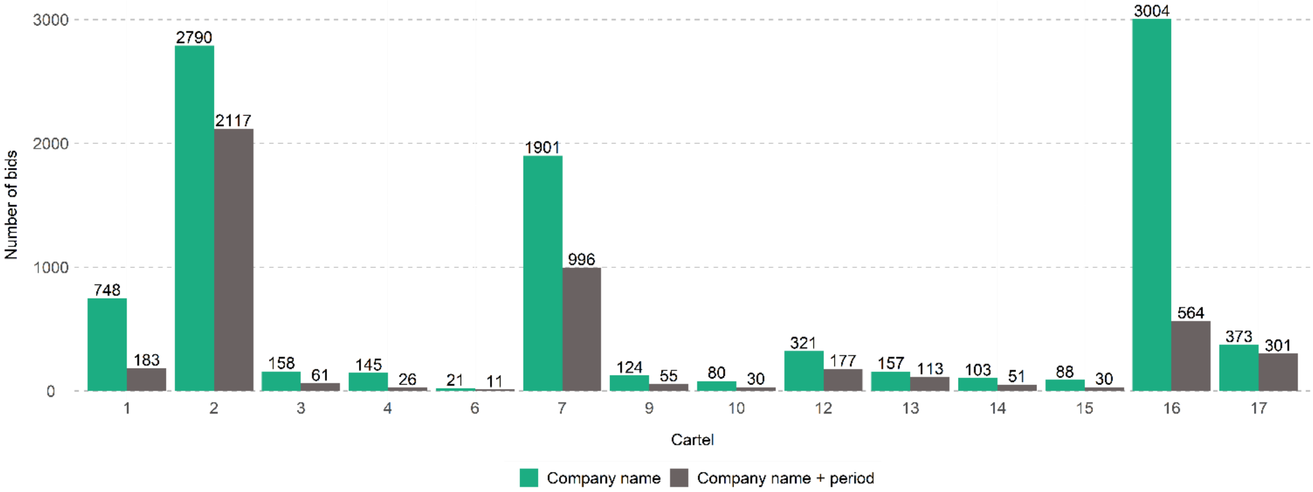
Figure 12.13 shows the overlap between the collusive period and our data coverage by each cartel. Many cartels have relatively long lifespans ranging from a minimum of three years to a maximum of 29 years. Some of these cartels have emerged years before the starting date of available procurement records (for instance, cartels #6, #7, #9, #10, and #11), hence we do not have data before the start of the cartel. Nevertheless, our data covers a few extra years after the cartel period for all proven cartels.

FIGURE 12.13. OVERLAP BETWEEN PUBLIC CONTRACTING DATA COVERAGE (GREY SPACE) AND CARTEL (GREEN BARS) TIME PERIODS



As in all other countries, the number of contracts varies significantly by the different sample definitions and cartels (Figure 12.14). However, we found at least 5 contracts for all cartels and all sample definitions except for cartels #5 and #8.

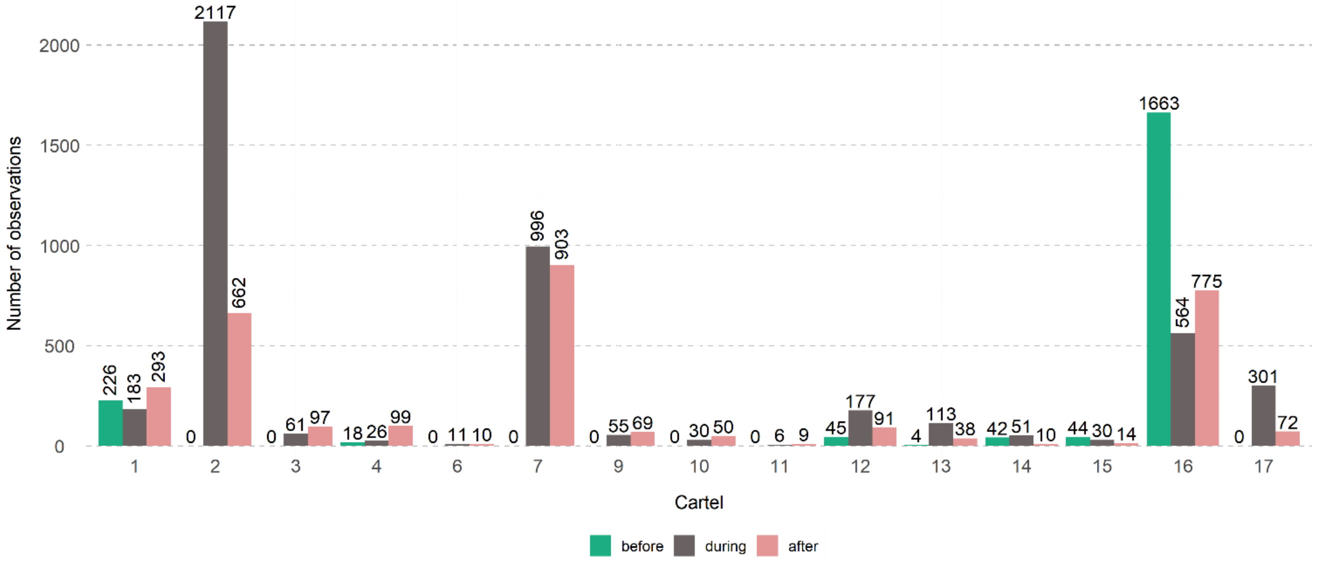
FIGURE 12.14. NUMBER OF MATCHED CONTRACTS BY SAMPLE DEFINITIONS



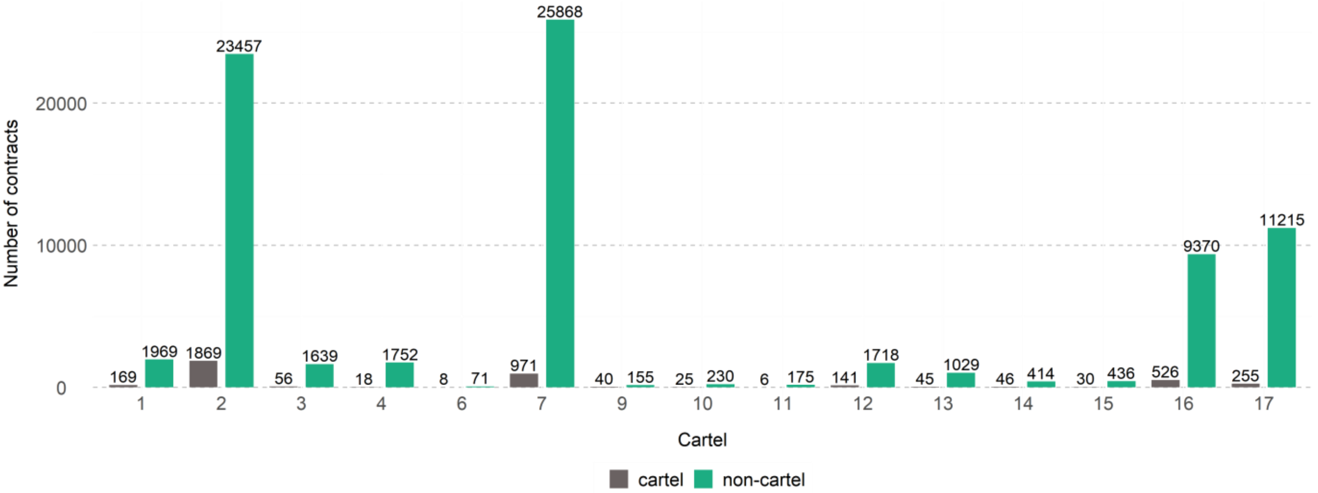
As we explained in the Methodology and Data sections, we group contracts by the three implemented tests (a) during vs. after, b) cross-sectional, and c) difference-in-differences), hence we show the number of contracts in each contract group that we use for statistical testing (Figure 12.15 – Panel A) shows the number of contracts used in during vs. after the cartel comparison. All 17 cartels (except for cartels #5 and #8) have at least 10 observations for both during and after the cartel period. For cross-sectional comparison (Panel B) all of the cartels except #5 and #8 are potentially suitable for testing. Note that we show the number of potentially available contracts for risk calculations, hence the final number of contracts with calculable individual indicators can be lower due to missing data.

FIGURE 12.15. NUMBER OF CONTRACTS USED IN STATISTICAL TESTING

Panel A: During-after comparison



Panel B: Cross-sectional comparison



12.2.6 Sweden

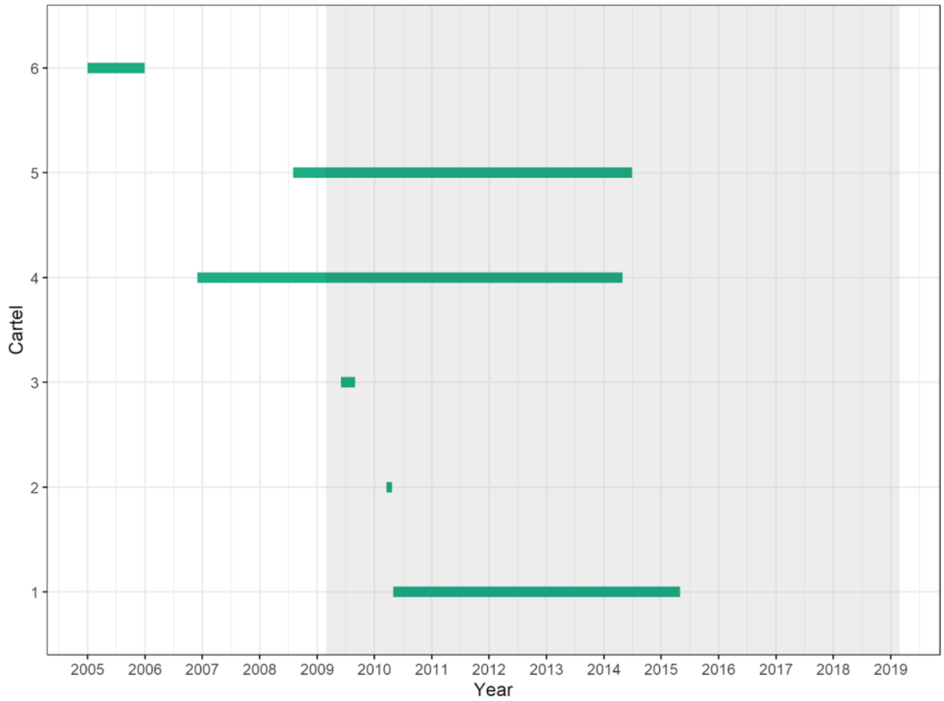
Given that no central repository of procurement announcements exists in Sweden, we use bid-level tendering data provided by Visma Opic. A thorough description of the dataset can be found in Fazekas and Toth (2016), here we only mention its main features relevant for the purposes of our analysis[60]. It contains bid-level data between 2009 and 2018, the number of contracts is evenly distributed across years. We remove direct contracts from the analysed set of tenders as in theory, bid rigging companies do not have control over whether they are contracted directly or not. However, we keep cancelled contracts as they might contain useful information on the rigging scheme61. The involved companies

[60] Although we use an updated version in this study, that includes years 2017 and 2018, the data cleaning procedures remained the same as the ones applied to the version used in Fazekas and Toth (2017).

[61] During bid submission, companies cannot know whether the tender would be cancelled, hence their bidding behaviour should still be affected by their agreement.

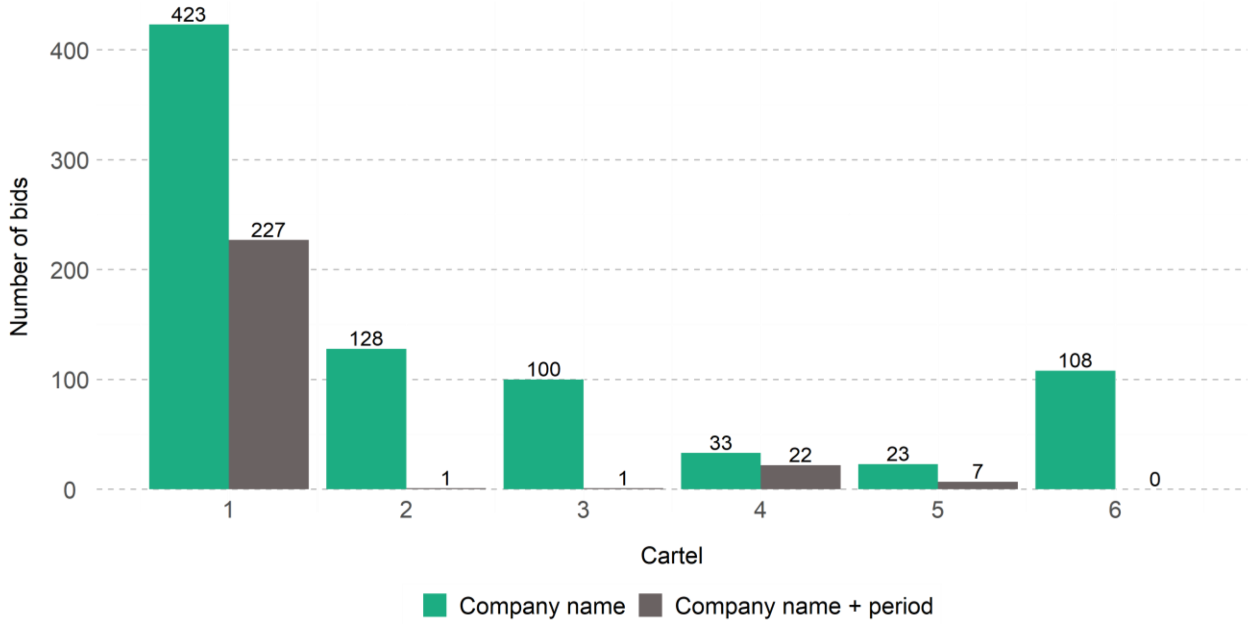
are expected to submit or withhold bids on cancelled tenders as they are unaware of cancellation at the time of bid submission.

FIGURE 12.16. OVERLAP BETWEEN PUBLIC CONTRACTING DATA COVERAGE (GREY) AND CARTEL TIME PERIODS



The number of matching contracts varies greatly by the applied filters (Figure 12.17). For example, while we could identify 227 contracts won by the two companies of cartel one that were advertised during the cartel period, the number of matches is often zero for other cartels. While we found a number of contracts won by the members of cartel two, three and six, most of these tenders were outside of the proven cartel period or were out of the scope of our dataset.

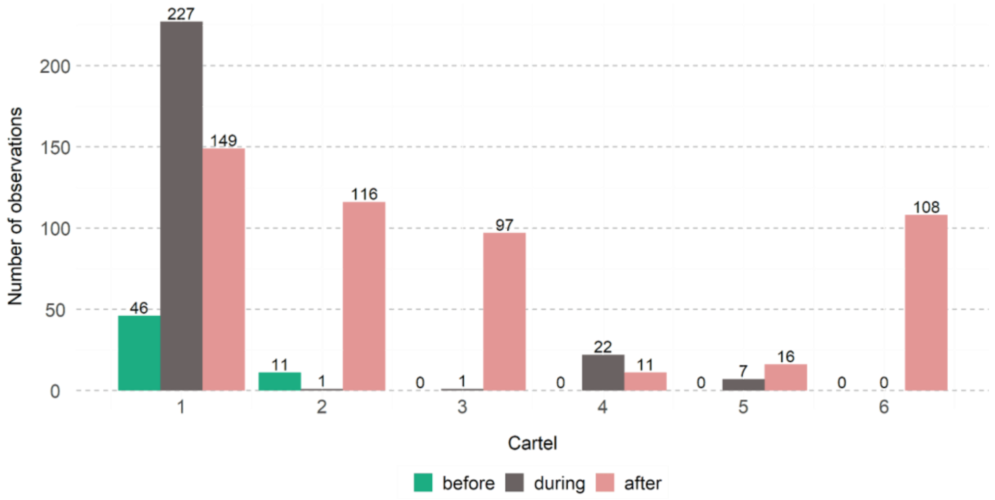
FIGURE 12.17. NUMBER OF CONTRACTS AWARDED BY SAMPLE DEFINITIONS



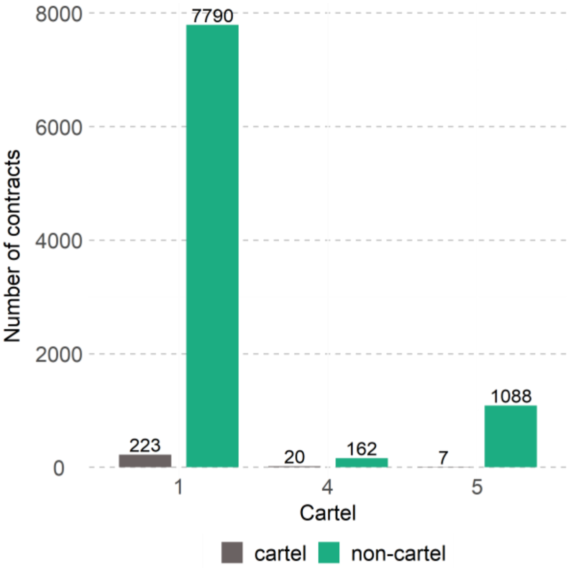
The next two figures and the tables show the number of contracts that we use for statistical testing a) during and after the cartel period, and b) for cross-sectional tests. Several cartels get out of scope from the analysis due to lack of enough observations (Figure 12.18). For example, only cartel one, four and five have potentially enough contracts both during and after the cartel period to compare risk indicator levels before and after the cartel period (Panel A). Note, that Figure 12.18 shows the number of potentially available contracts for risk calculations, hence the number of contracts with calculable individual indicators can be less.

FIGURE 12.18. NUMBER OF CONTRACTS USED IN THE THREE TESTS

Panel A: During-after



Panel B: Cross-section



13 Appendix B

13.1 Price-based indicators

13.1.1 Bid price distribution-based indicators

In this section we report the results based on the bid price distribution-based indicators. As bid prices for losing bids are only available for Swedish contracts and the share of missing prices is significant, we could only test them for one cartel. The first indicator is the relative difference between the first and second lowest bid price. We defined it as a binary variable which has value 1 if the relative price difference is close to 5%, 10%, 15%, 20%, 25% or 30%, and 0 otherwise – hence it picks up suspicious offer price differences. The sample is restricted to cartel companies and we have excluded all cartels with less than five awarded contracts both during and after the cartel period.

Table 13.1 Panel A show the during vs. after tests for cartel 1 in Sweden. It shows 56 winning bids during the cartel period, 11% of which had suspicious losing bids. In contrast, we have 34 cartel contracts after the cartel period with a 6% artificial bid share. Although the difference between the two periods has a negative sign, it is not statistically significant at the 10% level.

For the cross-sectional test (Panel B), the sample is narrowed down to the contracts won during the cartel period. With these restrictions cartel 1 submitted 56 winning bids, 14% of which have suspicious losing bids, during the cartel period. In contrast, we have far more non-cartel contract observations, of which 14% are suggestive of an artificially set bid price. Thus, the difference is the opposite of what the theoretical framework suggests.

TABLE 13.1. RELATIVE DIFFERENCE INDICATOR DURING VS. AFTER ESTIMATIONS

A: During vs. after

Country	Cartel ID	Contracts		Average		Diff	P-value
		During	After	During	After		
SE	1	56	34	0.11	0.06	-0.05	0.345

B: Cross-sectional

Country	Cartel ID	Contracts		Average		Diff	P-value
		Non-cartel	Cartel	Non-cartel	Cartel		
SE	1	52	1,835	0.1	0.13	0.03	0.692

Relative standard deviation, sometimes called the coefficient of variation, is defined at contract level as the standard deviation of bid prices divided by their average. We defined the indicator as a binary variable that captures extremely low and high values. We set the indicator to 1 when the relative standard deviation is less or greater than the 10% and 90% quantiles of the distribution of bid prices.

First, we test the indicator among cartel company contracts during vs. after the proven cartel period (Table 13.2 Panel A). The sample is restricted to cartel companies and we have excluded all cartels with less than five awarded contracts both during and after the cartel period. For example, Cartel 1 submitted 59 winning bids, with an extreme bid distribution rate of 22% during the cartel period. In contrast, we have 38 cartel contract observations after the cartel period, 11% of which have an extreme distribution. Thus, the difference of the two periods has a negative sign in line with our expectation, but the test is not significant at the 10% level.

For the cross-sectional test (Panel B), the sample is narrowed down to tenders won during the cartel period. We compare those cartel and non-cartel companies who won a bid where at least two cartel companies submitted a bid. Cartel 1 submitted 59 winning bids of which 21% had extreme distributions during the cartel period. In contrast, we have far more non-cartel contracts containing a share of 22% extreme values. Thus, the difference of the two periods has a positive sign in line with the theoretical expectation, however, it is not significant at 10% level.

TABLE 13.2. RELATIVE STANDARD DEVIATION INDICATOR DURING VS. AFTER ESTIMATIONS

Panel A: During vs after

Country	Cartel ID	Contracts		Average		Diff	P-value
		During	After	During	After		
SE	1	59	38	0.22	0.11	-0.12	0.119

Panel B: Cross-sectional

Country	Cartel ID	Contracts		Average		Diff	P-value
		Non-cartel	Cartel	Non-cartel	Cartel		
SE	1	53	1,595	0.23	0.25	0.02	0.587

The **relative price range** of bids on a contract is defined as the difference between the highest and lowest bid price divided by the mean of bid prices. Hence, we define the relative price range indicator as a binary variable, that has a value of 1 when the range is less than 10% or greater than the 90% quantiles of the distribution. Our expectation is that collusive behaviour manifests as extreme values for this statistic.

First, the shares of the indicator among cartel contracts are tested during vs. after the proven cartel period as summarized by Table 13.3. The sample is restricted for cartel companies and we have excluded all cartels with less than five awarded contracts both during and after the cartel period. Cartel 1 won 59 contracts, of which 20% had an extreme range, during the cartel period. In contrast, we have 38 cartel contract observations after the cartel period with an extreme bid range of 11%. Thus, the difference of the two periods has a negative sign supporting our expectation, but the test is not significant at the 10% level (Table 13.3).

For the cross-sectional test, the sample is narrowed down to the tenders that were won during the cartel period. We compare those cartel and non-cartel companies who won a bid where at least two cartel companies submitted a bid. We have excluded all cartels with less than five awarded contracts for the cartel and non-cartel groups. Cartel 1 won 49 contracts during the cartel period and 19% had an extreme relative price range. In contrast, we have far more non-cartel contracts with an 20% extreme relative price range share.

TABLE 13.3. RELATIVE DIFFERENCE INDICATOR DURING VS. AFTER ESTIMATIONS
Panel A: during vs after

Country	Cartel ID	Contracts		Average		Diff	P-value
		During	After	During	After		
SE	1	59	38	0.2	0.11	-0.1	0.161

Panel B: cross-sectional

Country	Cartel ID	Contracts		Average		Diff	P-value
		Non-cartel	Cartel	Non-cartel	Cartel		
SE	1	54	1,923	0.22	0.16	-0.06	0.162

13.1.2 Relative price

Our results vary between test types and countries. First, the simple during vs. after comparison shows that 5 out of or the total 21 potential tests show significant differences that correspond to the collusive logic - i.e. relative prices are significantly higher compared to the after collusion period (Table 13.4). We find the similar results for the cross-sectional comparison: 3 out of the total 24 cartels have higher relative prices for the rigged contracts compared to the competitive control group (Table 13.5).

The results also suggest that the robustness of these indicators varies across countries. While for example it works 3 out of 11 cartels in Spain using the during-after tests (Table 13.4), that is more than 50%, it only captured 1 out of the 5 Hungarian cartels where calculation was even feasible.

TABLE 13.4. RELATIVE PRICE INDICATOR DURING VS. AFTER ESTIMATIONS

Country	Cartel ID	Contracts		Average		Diff	P-value
		During	After	During	After		
LV	2	7	28	1.00	1.00	0.00	0.648
LV	7	35	50	1.00	0.97	-0.03	0.014
LV	10	20	14	0.99	1.00	0.01	0.919
LV	11	28	644	0.98	0.99	0.01	0.765
LV	21	70	120	0.97	0.98	0.01	0.769
HU	1	72	10	0.95	0.92	-0.03	0.237
HU	10	101	39	1.00	1.00	0.00	0.716
HU	11	24	153	1.00	0.99	-0.01	0.010
HU	12	37	31	0.98	0.97	-0.01	0.284
HU	19	29	15	0.95	1.00	0.05	0.983
ES	1	29	42	0.80	0.82	0.02	0.659
ES	2	325	146	0.88	0.85	-0.03	0.055
ES	3	7	11	0.80	0.72	-0.09	0.196
ES	7	11	8	0.88	0.64	-0.24	0.001
ES	9	24	39	0.97	0.94	-0.02	0.065
ES	10	12	6	0.94	0.92	-0.03	0.362
ES	12	19	6	0.85	0.92	0.07	0.802
ES	16	86	82	0.83	0.81	-0.02	0.259
ES	17	58	8	0.84	0.85	0.01	0.549
PT	1	11	36	0.74	0.86	0.12	0.979
PT	2	29	34	0.82	0.87	0.05	0.948

TABLE 13.5. RELATIVE PRICE INDICATOR CROSS-SECTIONAL ESTIMATIONS

Country	Cartel ID	Contracts		Average		Diff	P-value
		Non-cartel	Cartel	Non-cartel	Cartel		
LV	4	7	39	1.00	1.00	0.00	0.933
LV	7	31	2,984	1.00	0.99	0.01	0.321
LV	10	20	254	0.99	0.99	0.00	0.998
LV	17	9	29	0.95	0.99	-0.03	0.341
LV	21	32	428	0.97	0.97	-0.01	0.694
HU	1	62	185	0.99	0.99	0.00	0.800
HU	10	97	870	0.99	1.00	0.00	0.021
HU	11	24	233	1.00	1.00	0.00	0.735
HU	12	36	643	0.98	0.99	-0.02	0.011
HU	15	15	604	1.00	0.99	0.01	0.397
HU	16	12	62	0.98	0.99	-0.01	0.631
HU	19	27	1,441	0.98	0.98	0.00	0.998
ES	1	17	200	0.86	0.87	-0.01	0.731
ES	2	216	1,333	0.92	0.92	0.00	0.536
ES	7	8	57	0.94	0.92	0.01	0.780
ES	9	23	48	0.97	0.97	0.00	0.980
ES	10	9	12	1.02	0.88	0.14	0.001
ES	12	10	63	0.91	0.92	-0.01	0.785
ES	16	51	221	0.89	0.92	-0.03	0.094
ES	17	37	1,116	0.89	0.91	-0.01	0.412
PT	2	17	1,144	0.85	0.87	-0.02	0.364

13.1.3 Benford's law

First, we tested whether first digits of the bid prices follow the theoretical distribution during vs. after the proven cartel period. The sample is restricted for cartel companies and we have excluded all cartels with less than 100 awarded contracts both during and after the cartel period. For example, cartel 21 in Latvia won 149 and 218 contracts during and after the cartel period, respectively. Its mean absolute deviation[62] (MAD) is approximately equal in the two periods leading to significant deviation from theoretical distribution (nonconformity). But conformity would be expected after the cartel period. Thus, none of the indicator tests signal collusive behavior (Table 13.6).

TABLE 13.6. BENFORD'S LAW TEST DURING AND AFTER THE CARTEL PERIOD

Country	Cartel ID	Contracts		MAD		Conformity	
		During	After	During	After	During	After
LV	21	149	218	0.020	0.019	Nonconformity	Nonconformity
SE	1	211	126	0.019	0.024	Nonconformity	Nonconformity
ES	1	173	238	0.026	0.018	Nonconformity	Nonconformity
ES	2	1,835	627	0.013	0.013	Marginally acceptable conformity	Marginally acceptable conformity
ES	7	729	764	0.009	0.019	Acceptable conformity	Nonconformity
ES	16	511	756	0.012	0.010	Acceptable conformity	Acceptable conformity

For the cross-sectional test, the sample is narrowed down to the contracts that were won during the cartel period. We compare the theoretical and empirical digit distribution separately among cartel and non-cartel companies who won a contract where at least two cartel companies submitted a bid (Table 13.7). We have excluded all cartels with less than 100 awarded contracts for the cartel and non-cartel groups. On the intersecting product market of cartel 17 in Spain, non-cartel and cartel companies submitted 255 and 11,215 bids respectively. The mean absolute deviation was 0.014 among non-cartel contracts, which corresponds to marginally acceptable conformity. In contrast, cartels' price digit distribution deviates far more from the theoretical one with 0.028 MAD value which results in nonconformity supporting our expectation. Cross sectional Benford's law test indicates cartel 1 in Sweden, cartel 10 from Hungary, and cartel 17 in Spain.

[62] It indicates the magnitude of difference between the theoretical and empirical distribution of digits. The higher value it has, the lower chance the distribution follows the particular pattern.

TABLE 13.7. BENFORD'S LAW TEST FOR CARTEL AND NON-CARTEL BIDS

Country	Cartel ID	Contracts		MAD		Conformity	
		Cartel	Non-cartel	Cartel	Non-cartel	Cartel	Non-cartel
LV	21	131	1,280	0.034	0.021	Nonconformity	Nonconformity
SE	1	223	7,790	0.031	0.003	Nonconformity	Close conformity
HU	1	109	401	0.026	0.024	Nonconformity	Nonconformity
HU	10	196	1,834	0.031	0.012	Nonconformity	Acceptable conformity
HU	18	108	509	0.028	0.021	Nonconformity	Nonconformity
ES	1	169	1,969	0.027	0.036	Nonconformity	Nonconformity
ES	2	1,869	23,457	0.012	0.013	Marginally acceptable conformity	Marginally acceptable conformity
ES	7	971	25,868	0.008	0.002	Acceptable conformity	Close conformity
ES	12	141	1,718	0.027	0.015	Nonconformity	Nonconformity
ES	16	526	9,370	0.012	0.006	Acceptable conformity	Close conformity
ES	17	255	11,215	0.028	0.014	Nonconformity	Marginally acceptable conformity

13.2 Bidding pattern indicators

13.2.1 Single bidding

Out of the 34 cartels where the number of contracts were sufficient for t-tests for the during and after cartel period, only four showed a significant drop in the probability of single bidder contracts at 10% significance levels (Table 13.8). Five out of the 34 possible tests showed that the share of single-bidder contracts are significantly higher for collusive tenders based on the cross-sectional tests (Table 13.9). The results also show that this indicator is not robust across different testing approaches: only one cartel is flagged by all two tests (ES-9) - the 9th Spanish cartel shows a 27-55 % points decrease in the share of single-bidding in those contracts that are won by cartel members after the collusive period.

TABLE 13.8. SINGLE-BIDDING INDICATOR DURING VS. AFTER ESTIMATIONS[63]

Country	Cartel ID	Contracts		Average		Diff	P-value
		During	After	During	After		
LV	2	7	41	0.00	0.29	0.29	0.881
LV	7	77	163	0.09	0.29	0.20	1.000
LV	10	43	14	0.05	0.00	-0.05	0.500
LV	11	95	1,874	0.36	0.39	0.03	0.702
LV	14	7	253	0.00	0.71	0.71	1.000
LV	21	233	225	0.18	0.19	0.01	0.522
SE	1	212	145	0.13	0.14	0.01	0.501
SE	4	19	11	0.00	0.09	0.09	0.611
SE	5	7	15	0.00	0.27	0.27	0.820
HU	1	106	19	0.19	0.11	-0.08	0.290
HU	10	134	47	0.22	0.15	-0.07	0.216
HU	11	37	175	0.19	0.28	0.09	0.825
HU	**12**	**46**	**36**	**0.41**	**0.25**	**-0.16**	**0.095**
HU	13	90	6	0.23	0.00	-0.23	0.204
HU	16	43	17	0.33	0.53	0.20	0.879
HU	19	29	19	0.17	0.11	-0.07	0.410
FR	5	22	636	0.00	0.07	0.07	0.806
FR	10	62	144	0.44	0.49	0.06	0.728
ES	1	103	231	0.04	0.14	0.10	0.994
ES	2	1,659	635	0.34	0.40	0.05	0.990
ES	3	43	91	0.42	0.38	-0.03	0.426
ES	4	20	90	0.00	0.07	0.07	0.740
ES	6	11	10	0.27	0.00	-0.27	0.123
ES	7	538	710	0.07	0.07	0.00	0.526
ES	**9**	**47**	**67**	**0.68**	**0.13**	**-0.55**	**0.000**
ES	10	17	48	0.12	0.46	0.34	0.986
ES	12	133	86	0.18	0.23	0.05	0.779
ES	**13**	**56**	**38**	**0.21**	**0.03**	**-0.19**	**0.011**
ES	14	17	10	0.24	0.50	0.26	0.838
ES	15	28	14	0.04	0.00	-0.04	0.500
ES	**16**	**419**	**750**	**0.20**	**0.13**	**-0.07**	**0.001**
ES	17	248	72	0.05	0.10	0.04	0.866
PT	**1**	**15**	**47**	**1.00**	**0.66**	**-0.34**	**0.011**
PT	2	38	36	0.29	0.19	-0.10	0.248

[63] In Sweden and Portugal, we kept unique awarded contracts to calculate the share of single bidder contracts.

TABLE 13.9. SINGLE-BIDDING INDICATOR CROSS-SECTIONAL ESTIMATIONS[64]

Country	Cartel ID	Contracts		Average		Diff	P-value
		Non-cartel	Cartel	Non-cartel	Cartel		
LV	4	13	79	0.15	0.04	-0.11	0.166
LV	7	72	8,800	0.08	0.33	0.24	1.000
LV	10	43	653	0.05	0.04	-0.01	0.500
LV	17	15	50	0.60	0.86	0.26	0.967
LV	21	127	1,277	0.20	0.19	-0.01	0.443
SE	1	208	7,297	0.13	0.15	0.01	0.653
SE	4	17	143	0.00	0.08	0.08	0.749
SE	5	7	826	0.00	0.12	0.12	0.661
HU	1	99	308	0.18	0.25	0.07	0.888
HU	10	130	1,078	0.21	0.24	0.03	0.751
HU	11	36	275	0.19	0.15	-0.04	0.336
HU	12	42	781	0.43	0.33	-0.10	0.115
HU	15	19	719	0.16	0.20	0.04	0.558
HU	16	37	204	0.30	0.50	0.20	0.980
HU	18	76	396	0.17	0.30	0.13	0.983
HU	19	29	1,677	0.17	0.18	0.01	0.500
FR	5	10	515	0.00	0.00	0.00	0.500
FR	**10**	**44**	**1,481**	**0.50**	**0.20**	**-0.30**	**0.000**
ES	1	63	739	0.05	0.14	0.09	0.964
ES	**2**	**1,368**	**12,689**	**0.34**	**0.24**	**-0.10**	**0.000**
ES	**3**	**38**	**476**	**0.47**	**0.27**	**-0.20**	**0.006**
ES	4	7	275	0.00	0.13	0.13	0.682
ES	6	8	54	0.38	0.54	0.16	0.680
ES	7	308	6,507	0.11	0.16	0.05	0.987
ES	**9**	**32**	**122**	**0.72**	**0.45**	**-0.27**	**0.006**
ES	10	13	155	0.15	0.26	0.10	0.686
ES	**12**	**97**	**1,067**	**0.23**	**0.14**	**-0.08**	**0.023**
ES	13	32	407	0.31	0.46	0.14	0.919
ES	14	15	109	0.27	0.20	-0.06	0.413
ES	15	24	352	0.04	0.22	0.18	0.965
ES	16	371	5,300	0.21	0.19	-0.02	0.155
ES	17	185	6,824	0.05	0.09	0.04	0.959
PT	2	30	1,974	0.30	0.39	0.09	0.801

[64] In Sweden and Portugal, we kept unique awarded contracts to calculate the share of single bidder contracts.

13.2.2 Missing bidders

TABLE 13.10. MISSING BIDDERS INDICATOR DURING VS. AFTER ESTIMATIONS

Country	Cartel ID	Contracts		Company		Average		Diff
	Cartel	After	During	After	During	After	During	
LV	7	2	2	163	64	16.50	15.50	1.00
LV	14	4	3	152	8	19.75	2.33	17.42
LV	21	14	14	231	218	8.07	7.29	0.79
SE	1	2	2	299	382	107.00	112.00	-5.00
SE	4	3	3	23	39	6.33	12.33	-6.00
SE	5	2	2	21	18	10.00	6.00	4.00
HU	1	5	8	37	44	6.20	4.25	1.95
HU	10	4	4	299	739	44.00	106.50	-62.50
HU	11	11	7	544	246	28.18	24.71	3.47
HU	12	3	3	92	87	22.33	19.00	3.33
HU	16	4	4	51	62	9.50	10.50	-1.00
HU	19	5	5	52	62	9.20	10.20	-1.00
FR	1	3	3	13	10	4.00	2.33	1.67
FR	5	7	4	298	81	26.86	16.50	10.36
FR	10	2	6	95	77	32.50	11.17	21.33
ES	17	13	13	72	68	3.54	3.38	0.15
PT	1	4	3	21	15	3.75	2.67	1.08
PT	2	5	5	166	123	11.60	10.60	1.00

TABLE 13.11. MISSING BIDDERS INDICATOR CROSS-SECTIONAL ESTIMATIONS

Country	Cartel ID	Contracts		Company		Average		Diff
	Cartel	Company	Cartel	Company	Cartel	Company	Cartel	
LV	4	4	2	35	13	4.25	3.50	0.75
LV	7	111	2	2,908	72	6.01	17.00	-10.99
LV	14	249	3	330	8	1.23	2.00	-0.77
LV	17	22	2	49	15	1.64	5.00	-3.36
LV	21	432	14	1,280	131	1.74	5.14	-3.40
SE	1	125	2	9,317	496	33.59	137.00	-103.41
SE	4	75	3	306	44	3.48	13.00	-9.52
SE	5	69	2	720	18	8.36	6.00	2.36
HU	1	720	10	2,504	222	2.85	13.50	-10.65
HU	3	9	3	14	10	1.56	3.00	-1.44
HU	10	10	4	894	733	44.30	105.50	-61.20
HU	11	831	7	2,197	244	2.06	24.57	-22.51
HU	12	111	3	2,472	139	14.88	28.67	-13.78
HU	14	174	10	218	14	1.14	1.20	-0.06
HU	15	2,595	7	6,684	57	2.19	6.57	-4.38
HU	16	30	4	313	70	7.00	11.75	-4.75
HU	18	62	5	899	192	8.68	16.20	-7.52
HU	19	546	5	4,546	66	6.21	10.60	-4.39
FR	1	253	3	593	7	1.34	1.33	0.01
FR	5	4,783	2	7,600	22	1.13	8.00	-6.87
FR	10	98	2	506	22	2.55	9.50	-6.95
ES	1	1,301	55	1,624	73	1.05	1.25	-0.20
ES	2	11,146	635	14,566	1,354	1.13	1.58	-0.44
ES	3	294	37	344	51	1.03	1.27	-0.24
ES	4	940	9	998	10	1.02	1.11	-0.09
ES	7	20,447	682	25,456	883	1.10	1.19	-0.09
ES	9	102	13	112	26	1.02	1.23	-0.21
ES	10	61	15	63	19	1.00	1.20	-0.20
ES	11	138	5	149	6	1.00	1.20	-0.20
ES	12	1,252	64	1,675	127	1.18	1.47	-0.29

Country	Cartel ID	Contracts		Company		Average		Diff
	Cartel	Company	Cartel	Company	Cartel	Company	Cartel	
ES	13	1,019	39	1,029	45	1.00	1.00	0.00
ES	14	401	41	414	46	1.00	1.10	-0.10
ES	15	395	28	410	29	1.02	1.04	-0.02
ES	16	4,885	157	6,690	422	1.15	1.94	-0.79
ES	17	5,662	81	7,002	139	1.07	1.35	-0.28
PT	1	12	3	14	7	1.08	2.33	-1.25
PT	2	403	5	12,624	198	13.10	12.60	0.50

13.2.3 Winning probability

Out of all cartels from Portugal, Sweden, and Hungary, 14 had sufficient number of contracts to perform t-tests for during vs. after the cartel period. Out of these 14 cartels, only one had a significant drop in winning probability at a 90% confidence interval (Table 13.12). The cross-sectional estimates were based on 13 cartels, out of which eight had a significantly higher winning probability than the respective control groups (Table 13.13).

TABLE 13.12. WINNING PROBABILITY INDICATOR DURING VS. AFTER ESTIMATIONS

Country	Cartel ID	Contracts		Average		Diff	P-value
		During	After	During	After		
SE	1	509	296	0.42	0.49	0.07	0.976
SE	4	45	23	0.42	0.48	0.06	0.572
SE	5	20	20	0.35	0.75	0.40	0.987
HU	1	229	35	0.50	0.57	0.07	0.720
HU	9	12	72	0.50	0.78	0.28	0.953
HU	10	732	296	0.26	0.23	-0.03	0.174
HU	11	244	816	0.28	0.27	-0.01	0.402
HU	12	133	92	0.50	0.42	-0.07	0.175
HU	13	464	16	0.27	0.38	0.11	0.741
HU	14	14	67	0.71	0.88	0.17	0.881
HU	16	75	49	0.67	0.65	-0.01	0.500
HU	19	65	52	0.45	0.37	-0.08	0.244
PT	**1**	**15**	**85**	**1.00**	**0.55**	**-0.45**	**0.001**
PT	2	206	133	0.18	0.27	0.09	0.959

TABLE 13.13. WINNING PROBABILITY INDICATOR

Country	Cartel ID	Contracts		Average		Diff	P-value
		Non-cartel	Cartel	Non-cartel	Cartel		
SE	**1**	**423**	**23,327**	**0.49**	**0.36**	**-0.14**	**0.000**
SE	4	32	394	0.53	0.47	-0.06	0.303
SE	5	12	2,256	0.58	0.56	-0.02	0.500
HU	**1**	**182**	**1,322**	**0.57**	**0.39**	**-0.18**	**0.000**
HU	10	690	5,157	0.27	0.37	0.09	1.000
HU	11	214	1,804	0.32	0.37	0.05	0.905
HU	**12**	**112**	**2,849**	**0.56**	**0.35**	**-0.21**	**0.000**
HU	**14**	**12**	**66**	**0.83**	**0.53**	**-0.30**	**0.051**
HU	**15**	**33**	**2,269**	**0.61**	**0.35**	**-0.26**	**0.002**
HU	**16**	**66**	**711**	**0.67**	**0.48**	**-0.18**	**0.003**
HU	18	179	1,028	0.57	0.55	-0.02	0.382
HU	**19**	**61**	**5,961**	**0.48**	**0.31**	**-0.17**	**0.004**
PT	**2**	**141**	**12,473**	**0.21**	**0.16**	**-0.05**	**0.050**

13.2.4 Subcontracting

Out of the 18 cartels with the number of observations sufficient for t-tests for the during and after cartel period, five cartels show a significant decrease in the share of subcontracting at a 10% significance level (Table 13.14). Interestingly, all five cartels originate from Spain; while cartels from Latvia and France did not show any significant results. In comparison to during vs after t-tests, cross-sectional estimations were based on observations of 21 cartels (Table 13.15). Out of those, five cartels confirm the theoretical expectation.

TABLE 13.14. SUBCONTRACTING INDICATOR DURING VS. AFTER ESTIMATIONS

Country	Cartel ID	Contracts		Average		Diff	P-value
		During	After	During	After		
LV	2	7	40	0.00	0.03	0.03	0.500
LV	10	43	14	0.09	0.00	-0.09	0.281
LV	11	95	1,869	0.00	0.00	0.00	0.500
LV	14	8	252	0.00	0.01	0.01	0.500
LV	21	239	226	0.09	0.12	0.03	0.831
FR	5	12	502	0.33	0.17	-0.16	0.136
FR	10	62	120	0.03	0.08	0.05	0.842
ES	1	161	170	0.22	0.18	-0.04	0.255
ES	**2**	**1,996**	**612**	**0.19**	**0.09**	**-0.10**	**0.000**
ES	3	55	52	0.18	0.15	-0.03	0.449
ES	4	21	22	0.29	0.14	-0.15	0.204
ES	**7**	**984**	**822**	**0.04**	**0.00**	**-0.03**	**0.000**
ES	9	54	63	0.09	0.05	-0.04	0.276
ES	**12**	**151**	**77**	**0.13**	**0.00**	**-0.13**	**0.001**
ES	13	113	17	0.27	0.12	-0.16	0.139
ES	15	28	13	0.14	0.08	-0.07	0.465
ES	**16**	**364**	**293**	**0.18**	**0.14**	**-0.04**	**0.079**
ES	**17**	**265**	**47**	**0.23**	**0.06**	**-0.16**	**0.009**

TABLE 13.15. SUBCONTRACTING INDICATOR CROSS-SECTIONAL ESTIMATIONS

Country	Cartel ID	Contracts		Average		Diff	P-value
		Non-cartel	Cartel	Non-cartel	Cartel		
LV	4	13	79	0.92	0.95	-0.03	0.661
LV	10	43	653	0.91	0.94	-0.04	0.305
LV	14	7	376	1.00	0.94	0.06	0.501
LV	**17**	**15**	**49**	**0.87**	**0.98**	**-0.12**	**0.058**
LV	21	131	1,280	0.85	0.87	-0.02	0.521
FR	5	7	159	1.00	0.91	0.09	0.402
FR	**10**	**49**	**1,983**	**1.00**	**0.91**	**0.09**	**0.030**
ES	**1**	**148**	**1,635**	**0.76**	**0.65**	**0.12**	**0.004**
ES	2	1,753	20,827	0.80	0.81	-0.01	0.390
ES	3	50	1,585	0.82	0.81	0.01	0.869
ES	4	14	1,619	0.71	0.65	0.06	0.640
ES	6	6	61	1.00	0.72	0.28	0.138
ES	**7**	**959**	**25,609**	**0.96**	**0.97**	**-0.01**	**0.038**
ES	9	40	149	0.92	0.96	-0.04	0.346
ES	10	19	156	1.00	0.91	0.09	0.173
ES	12	109	1,377	0.90	0.89	0.01	0.855
ES	**13**	**45**	**1,028**	**0.87**	**0.63**	**0.24**	**0.001**
ES	14	40	358	0.98	0.98	-0.01	0.776
ES	15	28	365	0.86	0.91	-0.05	0.406
ES	16	328	5,078	0.80	0.81	-0.01	0.642
ES	17	229	8,335	0.76	0.74	0.03	0.324

13.2.5 Consortia

TABLE 13.16. CONSORTIA INDICATOR DURING VS. AFTER ESTIMATIONS

Country	Cartel ID	Contracts		Average		Diff	P-value
		During	After	During	After		
LV	21	239	231	0.00	0.00	0.00	0.507
SE	5	20	21	0.05	0.00	-0.05	0.490
HU	1	237	37	0.01	0.00	-0.01	0.500
HU	10	739	302	0.00	0.00	0.00	0.678
HU	11	247	839	0.00	0.00	0.00	0.500
HU	13	469	16	0.00	0.06	0.06	0.996
HU	19	66	52	0.21	0.21	0.00	0.500
FR	1	10	45	0.00	0.07	0.07	0.528
FR	5	81	1,223	0.01	0.02	0.01	0.575
FR	10	77	177	0.05	0.17	0.12	0.990
ES	**1**	**183**	**293**	**0.52**	**0.20**	**-0.32**	**0.000**
ES	**2**	**2,117**	**662**	**0.19**	**0.15**	**-0.04**	**0.006**
ES	3	61	97	0.03	0.02	-0.01	0.500
ES	**4**	**26**	**99**	**0.50**	**0.15**	**-0.35**	**0.000**
ES	7	996	903	0.06	0.39	0.33	1.000
ES	**9**	**55**	**69**	**0.29**	**0.12**	**-0.17**	**0.013**
ES	10	30	50	0.10	0.04	-0.06	0.275
ES	11	6	9	0.17	0.00	-0.17	0.416
ES	**12**	**177**	**91**	**0.15**	**0.02**	**-0.12**	**0.002**
ES	13	113	38	0.00	0.03	0.03	0.717
ES	14	51	10	0.00	0.10	0.10	0.820
ES	15	30	14	0.03	0.14	0.11	0.758
ES	**16**	**564**	**775**	**0.14**	**0.06**	**-0.09**	**0.000**
ES	**17**	**301**	**72**	**0.34**	**0.14**	**-0.20**	**0.001**

TABLE 13.17. CONSORTIA INDICATOR CROSS-SECTIONAL ESTIMATIONS

Country	Cartel ID	Contracts		Average		Diff	P-value
		Non-cartel	Cartel	Non-cartel	Cartel		
LV	7	72	8,800	0.00	0.00	0.00	0.500
LV	10	43	653	0.00	0.01	0.01	0.500
LV	14	7	376	0.00	0.09	0.09	0.565
LV	17	15	49	0.07	0.09	0.03	0.500
LV	21	131	1,280	0.00	0.03	0.03	0.938
SE	1	496	25,407	0.00	0.01	0.01	0.982
SE	4	44	495	0.00	0.03	0.03	0.716
SE	5	18	2,741	0.06	0.02	-0.03	0.420
HU	1	222	2,622	0.01	0.04	0.04	0.991
HU	9	7	236	0.00	0.04	0.04	0.500
HU	10	733	5,585	0.00	0.06	0.06	1.000
HU	11	245	2,201	0.00	0.06	0.06	1.000
HU	12	139	5,640	0.00	0.05	0.05	0.993
HU	14	14	260	0.00	0.02	0.02	0.500
HU	15	57	7,039	0.00	0.03	0.03	0.823
HU	16	70	558	0.00	0.04	0.04	0.920
HU	18	192	1,280	0.00	0.06	0.06	1.000
HU	**19**	**66**	**7,841**	**0.21**	**0.05**	**-0.16**	**0.000**
FR	1	8	2,806	0.00	0.06	0.06	0.500
FR	5	24	7,717	0.00	0.03	0.03	0.554
FR	10	56	2,991	0.05	0.05	0.00	0.500
ES	**1**	**169**	**1,969**	**0.55**	**0.44**	**-0.11**	**0.003**
ES	**2**	**1,869**	**23,457**	**0.20**	**0.15**	**-0.06**	**0.000**
ES	3	56	1,639	0.04	0.08	0.04	0.823
ES	4	18	1,752	0.50	0.41	-0.09	0.286
ES	6	8	71	0.00	0.13	0.13	0.690
ES	**7**	**971**	**25,868**	**0.06**	**0.03**	**-0.03**	**0.000**
ES	9	40	155	0.35	0.38	0.03	0.563
ES	10	25	230	0.08	0.14	0.06	0.689
ES	11	6	175	0.17	0.02	-0.14	0.196
ES	**12**	**141**	**1,718**	**0.13**	**0.07**	**-0.05**	**0.015**
ES	13	45	1,029	0.00	0.01	0.01	0.562
ES	14	46	414	0.00	0.08	0.08	0.947
ES	15	30	436	0.03	0.07	0.03	0.633
ES	**16**	**526**	**9,370**	**0.15**	**0.12**	**-0.03**	**0.036**
ES	**17**	**255**	**11,215**	**0.35**	**0.24**	**-0.11**	**0.000**
PT	1	7	14	0.00	0.07	0.07	0.500
PT	2	197	14,148	0.00	0.00	0.00	0.539

13.2.6 Market cut-points

Table 13.18 presents results of network cut-point detection for during vs. after comparison. Based on the 17 cartels we have tested, none of the cartels fit fully the theoretical expectation. However, the number of cut-point position companies decreased for cartel 1 and 19 in Hungary after the cartel period.

TABLE 13.18. CUT-POINT POSITION INDICATOR DURING VS. AFTER ESTIMATIONS

Country	Cartel ID	Cut points		Contracts		Cartels		Entities	
		During	After	During	After	During	After	During	After
SE	1	0	0	227	149	2	2	2	2
SE	4	0	0	20	11	3	3	3	3

Country	Cartel ID	Cut points		Contracts		Cartels		Entities	
		During	After	During	After	During	After	During	After
SE	5	0	0	7	16	2	2	2	2
HU	1	2	1	121	21	10	4	10	4
HU	9	0	0	7	70	1	4	1	4
HU	10	0	0	181	69	4	4	4	4
HU	11	0	2	64	236	6	11	6	11
HU	12	0	1	67	39	3	3	3	3
HU	13	0	0	127	6	1	3	1	3
HU	14	0	0	1	19	10	11	10	11
HU	16	0	0	38	22	4	4	4	4
HU	19	3	1	30	19	5	5	5	5
PT	1	0	1	15	48	3	5	3	5
PT	2	0	0	40	47	5	5	5	5

Cross-sectional comparison has revealed more cartel cases fitting the theoretical expectation. To reiterate, for the cross-sectional comparison, we select contracts won by cartel members during the cartel period on a relevant product market (a product code on which at least two cartel members have submitted a bid). As a control group, we select contracts won by ordinary firms during the cartel period on the same product market. Out of 16 cartels we were able to test (Table 13.19), four cartels (cartel #2 from Portugal and cartels #4 and #5 from Sweden, and #1 from Hungary) have their members placed as cut-points in a network of rigged contracts while none of the companies in the control group had such position.

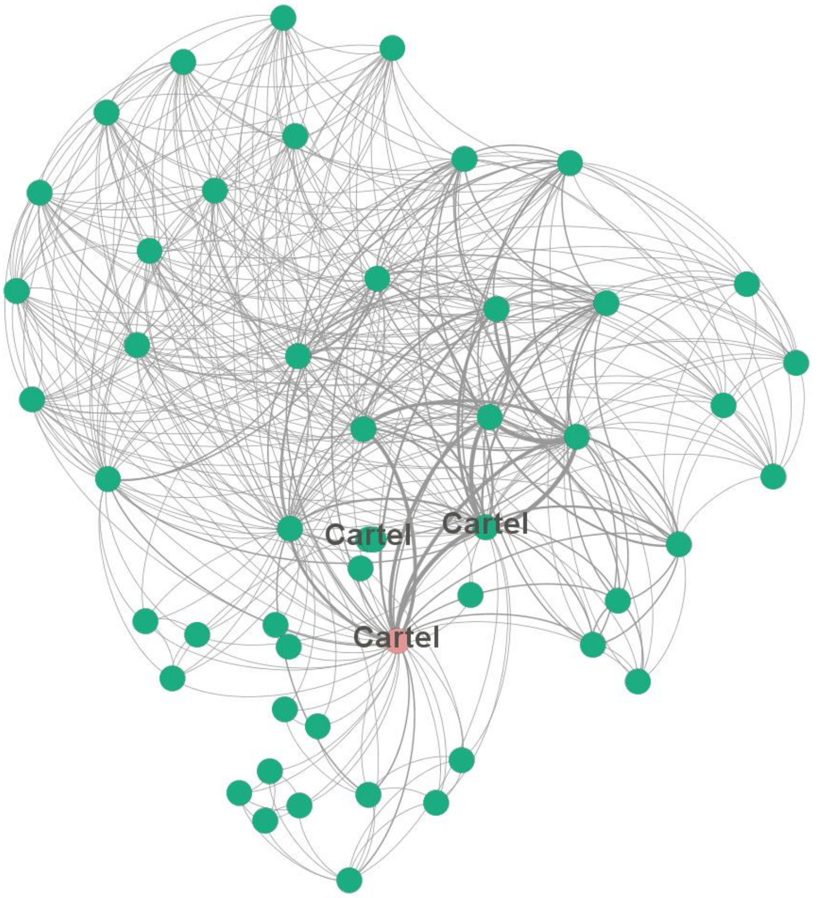
TABLE 13.19. CUT-POINT POSITION INDICATOR CROSS-SECTIONAL ESTIMATIONS

Country	Cartel ID	Cut points		Contracts		Companies		Entities	
		Cartel	Non-cartel	Cartel	Non-cartel	Cartel	Non-cartel	Cartel	Non-cartel
SE	1	2	1	223	6,237	2	2	216	4,650
SE	2	0	0	1	310	1	0	4	648
SE	**4**	**1**	**0**	**18**	**91**	**3**	**2**	**56**	**259**
SE	**5**	**1**	**0**	**7**	**505**	**2**	**2**	**16**	**1,024**
HU	**1**	**4**	**0**	**107**	**294**	**10**	**8**	**167**	**403**
HU	3	0	0	4	4	3	1	3	5
HU	10	2	2	177	1,354	4	4	172	1,392
HU	11	3	1	63	469	6	6	104	888
HU	12	2	2	63	852	3	3	72	1,019
HU	14	0	0	1	29	10	0	14	29
HU	**15**	**3**	**0**	**20**	**698**	**6**	**1**	**45**	**1,383**
HU	16	3	1	32	127	4	2	28	236
HU	18	0	0	70	145	5	3	106	206
HU	19	5	1	30	1,541	5	4	59	2,194
PT	1	0	0	7	13	3	0	3	11
PT	**2**	**2**	**0**	**30**	**1,974**	**5**	**5**	**81**	**1,321**

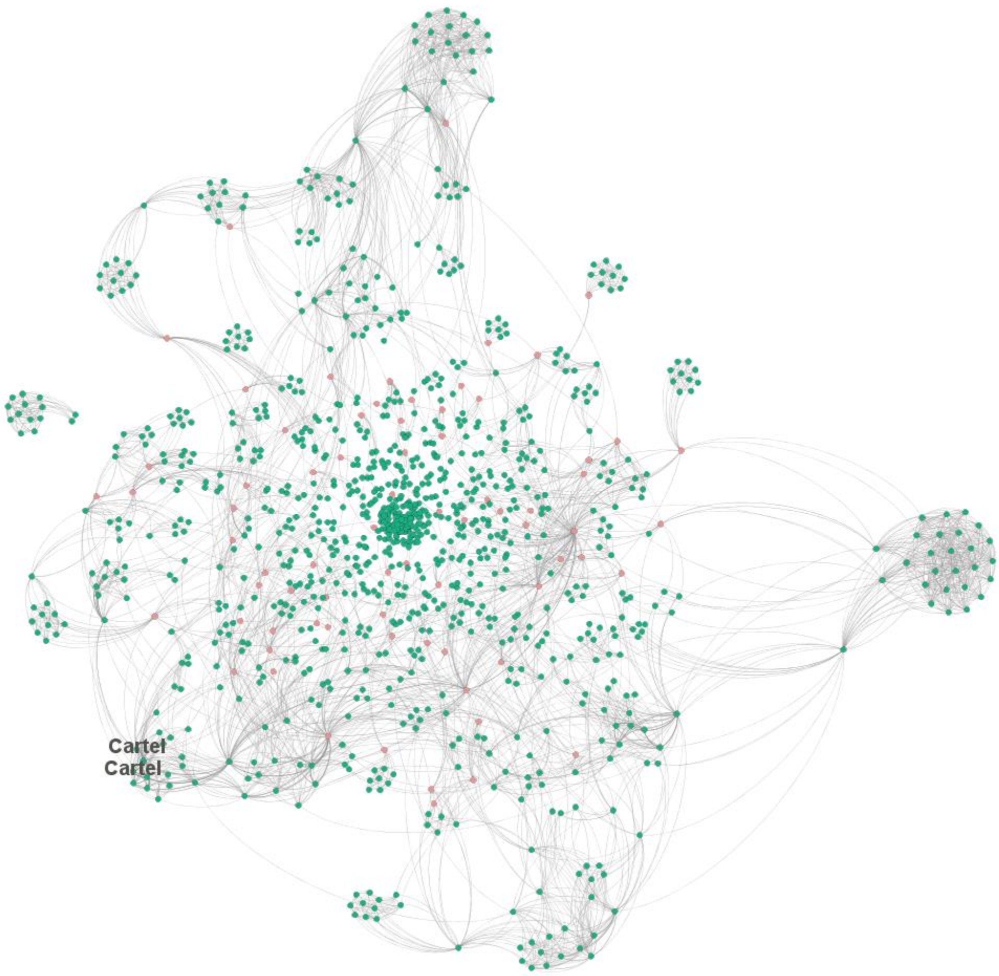
One of the three cartels' network supporting theoretical expectations is presented below (Figure 13.1 – Panel A) illustrates a wide co-bidding network defined by the sample of rigged contracts for cartel #4 in Sweden. Interestingly, these contracts won by cartel members during the cartel period on the relevant product market are characterized by high competition and frequent co-bidding. Despite the high level of connectivity, this network has a cartel company in a cut-point position. In the control group (Panel B), an extensive co-bidding network also has cut-points, but none of these firms are cartel members.

FIGURE 13.1. CUT-POINT POSITIONS IN A CO-BIDDING PROCUREMENT NETWORK - SWEDEN, CARTEL #4, CROSS-SECTIONAL COMPARISON

Panel A: Rigged contracts



Panel B: Control group



13.3 Market structure

13.3.1 Concentration

Market concentration is compared during vs. after the proven cartel periods, as reported in Table 13.20. We assume for all cartels that both periods have the same length (2 years) – i.e. we calculate concentration based on a +-2 year time window around the end of the collusive period. The unbalanced length of periods would potentially distort our calculations since the longer the examined period is, the more likely a market is to have additional participants. We do not calculate the indicator for a cartels without at least one cartel company winning at least three contracts in both periods. For this indicator we apply a narrow market definition, which consists of only cartel members, that most probably underestimates the market, however it tracks down the relative share of the relevant companies – i.e. the cartel members – whose contract allocation practices are in the focus. Future work should consider alternative methods to define markets.

TABLE 13.20. CONCENTRATION INDICATOR DURING VS. AFTER ESTIMATIONS. MARKET SHARES ARE CALCULATED BASED ON CONTRACT VALUES USING NARROW MARKET DEFINITION.

Country	Cartel ID	Companies		HHI		Diff
		During	After	During	After	
LV	2	3	3	3,617	4,021	404
LV	7	2	2	9,286	9,050	-237
LV	10	7	2	4,378	5,312	934
LV	13	11	5	2,116	5,235	3,119
LV	**21**	**12**	**9**	**5,594**	**3,980**	**-1,615**
HU	1	8	2	3,337	8,166	4,829
HU	10	4	4	3,264	5,474	2,210
HU	11	6	9	2,903	3,048	144
HU	12	3	2	3,720	5,057	1,338
HU	**16**	**4**	**3**	**4,657**	**4,044**	**-613**
HU	**19**	**5**	**4**	**4,596**	**4,279**	**-317**
FR	**4**	**2**	**2**	**10,000**	**8,992**	**-1,008**
FR	5	2	3	10,000	9,890	-110
ES	1	4	5	5,443	9,020	3,577
ES	2	10	10	2,162	3,271	1,109
ES	3	3	3	7,079	9,309	2,230
ES	7	7	6	2,086	3,267	1,181
ES	10	5	3	7,022	7,591	570
ES	**12**	**3**	**3**	**6,404**	**5,094**	**-1,309**
ES	15	4	4	5,011	8,866	3,855
ES	**16**	**13**	**10**	**2,949**	**2,693**	**-256**
ES	17	7	4	4,519	8,620	4,101
PT	**1**	**3**	**3**	**7,875**	**4,440**	**-3,435**
PT	**2**	**4**	**5**	**4,756**	**2,675**	**-2,080**

13.3.2 Stability

Stability is tested by looking at the difference in the absolute market share change during vs. after the proven cartel period (Table 13.21).

TABLE 13.21. STABILITY INDICATOR DURING VS. AFTER ESTIMATIONS. MARKET SHARES ARE CALCULATED BASED ON CONTRACT VALUES.

Country	Cartel ID	All firms	Cartels		Abs. change in market share		Diff
			Trans. 1	Trans. 2	Trans. 1	Trans. 2	
LV	**7**	**2**	**2**	**1**	**100 %**	**95 %**	**-5 %**
LV	21	12	11	4	95 %	95 %	-1 %
HU	**12**	**3**	**3**	**2**	**89 %**	**63 %**	**-25 %**
FR	**10**	**2**	**2**	**2**	**100 %**	**94 %**	**-6 %**
ES	1	5	3	4	91 %	97 %	7 %
ES	2	9	8	7	93 %	96 %	3 %
ES	**3**	**4**	**3**	**1**	**94 %**	**2 %**	**-92 %**
ES	7	7	7	5	90 %	95 %	5 %
ES	**10**	**2**	**1**	**1**	**83 %**	**50 %**	**-33 %**
ES	12	3	3	2	54 %	87 %	33 %
ES	16	12	9	9	96 %	96 %	0 %