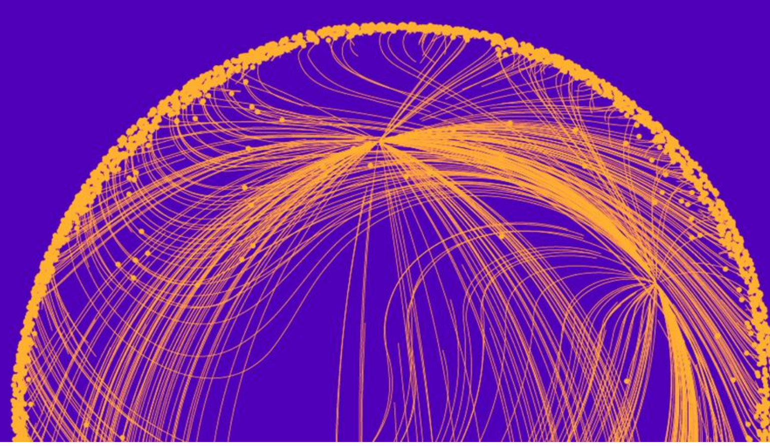
Helsinki GSE

HELSINKI GSE DISCUSSION PAPERS 1 · 2023

Complementary bidding and cartel detection: Evidence from Nordic asphalt markets

Aapo Aaltio
Riku Buri
Antto Jokelainen
Johan Lundberg



HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

HANKEN

Aalto University

Helsinki GSE Discussion Papers

Helsinki GSE Discussion Papers 1 · 2023

Aapo Aaltio, Riku Buri, Antto Jokelainen, and Johan Lundberg:
Complementary bidding and cartel detection: Evidence from Nordic asphalt markets

ISBN 978-952-7543-00-9 (PDF)
ISSN 2954-1492

Helsinki GSE Discussion Papers:
https://www.helsinkigse.fi/discussion-papers

Helsinki Graduate School of Economics
PO BOX 21210
FI-00076 AALTO
FINLAND

Helsinki, January 2023

Complementary bidding and cartel detection: Evidence from Nordic asphalt markets

Aapo Aaltio[a] Riku Buri[b] Antto Jokelainen[c] Johan Lundberg[d]

January 27, 2023

Abstract

A key challenge in cartel enforcement is identifying collusive agreements. We study two major Nordic procurement cartels that operated in the asphalt paving market. We find evidence that during the cartel period bids were clustered and the winning bid was isolated. We implement two cartel detection methods that exploit variation in the distribution of bids. The method developed by Clark et al. (2020) correctly rejects competitive bidding for the cartel period in both markets. The method suggested by Huber and Imhof (2019) predicts a significantly higher probability of collusion for the cartel period in one of the markets. Our results indicate that statistical screening methods with modest data requirements can be useful for competition authorities in detecting collusive agreements.

Keywords: procurement, bidding ring, collusion, antitrust, complementary bidding, detection

JEL Codes: L22, L74, D44, H57

[a]Finnish Competition and Consumer Authority, aapo.aaltio@kkv.fi
[b]Finnish Competition and Consumer Authority and Aalto University, riku.buri@kkv.fi
[c]Aalto University, antto.jokelainen@aalto.fi
[d]Umeå University, johan.lundberg@umu.se

We thank Adriano De Leverano, Olli Kauppi, Matias Pietola, Otto Toivanen, and the participants of Helsinki GSE Industrial Organization PhD seminar for their comments and suggestions.

1 Introduction

Cartels coordinate the actions of their members to increase profits. Comprehensive studies have found that, on average, cartels increase prices by 15 to 30% (Connor and Bolotova, 2006; Boyer and Kotchoni, 2015; Bolotova, 2009; Froeb et al., 1993). Although countries have adopted antitrust laws that prohibit cartels, firms continue to collude. Several cartels have been active in public procurement, which in 2019 represented 12% of the world GDP (Bosio et al., 2022). Therefore, bidding rings potentially impose a significant cost on taxpayers. The key challenge in cartel enforcement is identifying collusive agreements. Previous studies, using data from known cartels, estimate that the probability of a cartel being caught and convicted is only around 10 to 20% per year (Harrington and Wei, 2017; Combe et al., 2008; Bryant and Eckard, 1991).

Although concrete evidence is required for the successful prosecution of cartels, a screening device that flags suspicious behavior in public procurement could potentially help authorities identify collusive agreements at a higher rate and save billions of taxpayers' money. However, finding cartels using statistical methods is complicated by the availability of data. Procurement datasets rarely have detailed project- or firm-specific information, and collecting such data across industries is burdensome. Recently, many studies have focused on detection methods that rely only on bidding information. In this literature, several indicators have been suggested to flag suspicious behavior, for example, low variance of bids (Feinstein et al., 1985; Abrantes-Metz et al., 2006; Imhof et al., 2018), isolated winning bids (Imhof et al., 2018; Clark et al., 2020; Chassang et al., 2022), and clustering of losing bids (Lundberg, 2017). Given that large, economy-wide procurement datasets are becoming increasingly available for authorities and researchers, these methods, if found reliable and precise, could be used by authorities for wide-scale cartel detection.[1]

In this paper, we study the bidding behavior of two convicted cartels that operated in the Finnish and Swedish asphalt markets. The Finnish cartel operated from 1994 to 2002

[1] For a description of such datasets see Jääskeläinen and Tukiainen (2019) for Finland, Halonen and Tukiainen (2020) for Sweden, Giuffrida and Rovigatti (2018) for the U.S., Coviello and Gagliarducci (2017) for Italy, Ferraz et al. (2015) for Brazil, Lee (2022) for Korea, Baránek et al. (2021) for Ukraine, Kawai et al. (2022) for Indonesia, Georgia, Mongolia, Malta, and California, and Baltrunaite (2020) for Lithuania.

and the Swedish cartel from 1993 to 2001. Our paper has two objectives. First, using data before and after the launch of cartel investigations by the competition authorities, we estimate how the distribution of bids changed after the collapse of the cartel. Second, we test the performance of two cartel detection methods, which can be implemented using only information on the distribution of bids.

We find that during the cartel, a large share of bids are within 10% of the winning bid. This clustering of bids is particularly prevalent in the Finnish market. We also observe that during the cartel period, winning bids are isolated, with losing bids typically being at least one percent higher than the winning bid. Together, the clustering of bids and isolated winning bids result in a bimodal distribution of bids during the cartel period. After the start of cartel investigations, the distribution of bids becomes unimodal and the share of bids within 10% of the winning bid decreases. To support the causal interpretation of our results, we conduct a difference-in-differences analysis. As a control market, we use the Californian asphalt market, in which there is no evidence of collusion during the same period. Our results remain largely unchanged in the difference-in-differences analysis.

The first detection method that we test is a distributional regression approach suggested by Clark et al. (2020). The method is based on the observation that while a cartel might find it optimal to leave a gap between the winning bid and the second lowest bid, it does not have similar incentives to manipulate the difference between the losing bids. The method works by comparing two sets of bid differences, where the bid difference is defined as the difference between a bid and the lowest rival bid. The first set of bid differences is calculated from a sample that includes all the bids, whereas the second set is calculated from a sample where the winning bid is excluded. Using the two sets of bid differences, a distributional regression is run with an indicator variable for whether the bid difference is from the original distribution or the one with winning bids excluded. The null hypothesis is that, with competitive bidding, the coefficient of this indicator variable should be statistically insignificant for intervals close to zero. In both Finland and Sweden, the null hypothesis is rejected for the cartel period. In both cases, consistent with the intuition of the test, we find that during the cartel period, the full set of bid differences has a much lower density of bid differences close to zero, indicating that the cartel firms avoided leaving bids very close to the winning bid. In Finland, the null hypothesis is not rejected for the post-investigation period, while in Sweden we find weak evidence of isolated winning bids also in the post-investigation period.

Finally, the second detection method developed by Huber and Imhof (2019) is applied to the same data. It uses machine learning to classify tenders as competitive or collusive. As predictors, the machine learning model uses different statistical screens calculated from the distribution of bids. These include, for example, the standard deviation of the bids and the difference between the winner and the runner-up. The model correctly classifies around 83% of the tenders with the Finnish dataset and 70% with the Swedish dataset. When we evaluate the predictions of the machine learning model over time, we find that in Finland the average collusion probability of tenders starkly decreases after the start of cartel investigation in 2002. However, for Sweden, the model predicts only a modestly higher collusion probability for the cartel period. We also test the performance of the machine learning model when the model is trained with data from one country and then tested with data from the other country. We find that the prediction rates decrease substantially in this cross-market analysis.

Our results indicate that cartels can result in a significant change in the distribution of bids and that statistical cartel detection methods with modest data requirements can be useful for competition authorities in flagging suspicious behavior in public procurement. However, both detection methods studied in this paper do have some caveats. The method by Clark et al. (2020) cannot be used to detect collusion for individual tenders but rather for a group of tenders. This might be an issue if the group has a mix of collusive and competitive tenders. For example, this could be the case if firms collude only in some geographical markets. This can be circumvented by a detailed grouping of tenders, which, however, might be difficult when screening for cartels ex-ante or even unfeasible due to a low number of observations. On the other hand, the method by Huber and Imhof (2019) can predict collusion for each tender individually, but it requires the user to calibrate the predictive model with existing data from both collusive and competitive tenders. Therefore, a clear limitation of the machine learning–based model is the availability of suitable data to train the model and whether the data used to train the model generalizes to the data for which the model is used for prediction. Given that the markets in our setting are very similar, and yet the prediction performance in the cross-market analysis is rather poor, this might be a prevalent constraint in other settings as well.

This paper is related to the literature on cartel detection in auctions. Early papers in the literature often relied on cost data and the estimation of a bidding function (Porter and Zona, 1993; Porter and Zona, 1999; Bajari and Ye, 2003; Aryal and Gabrielli, 2013). More recent papers have focused on methods that do not require the estimation of a bidding function. Imhof et al. (2018) and Imhof (2020) use data from the Swiss road construction sector to document that cartels have resulted in a change in the distribution of bids. Based on these findings Huber and Imhof (2019) develop a machine learning model to detect collusion. Chassang et al. (2022) document that in Japanese procurement auctions, winning bids tend to be isolated. They show that this pattern is inconsistent with competitive behavior in a repeated setting because when the winning bids are persistently isolated, the winners could profitably deviate by increasing their bids. Clark et al. (2020) present empirical evidence from a procurement cartel that bidding involves both clustering and isolated winning bids, and they develop a distributional regression method to detect cartels.[2] Finally, Wachs and Kertész (2019) develops a network-based framework to detect groups of suspicious firms within markets.

Our paper has several novel features. First, instead of proposing a new detection method, we test the performance of two existing methods. Testing the performance of existing methods in new datasets is important for informing policymakers and researchers about the applicability and precision of detection methods in different contexts. This is particularly important for cartel detection because cartels can organize themselves in several different ways, implying that no econometric method or test of collusion is likely applicable in every setting. Second, we study two separate cartels that operated on the same product market in two neighboring countries. This allows us to compare the change in bidding behavior and the performance of the detection methods across two separate but similar cartels. Using testimonial evidence, we discuss how differences in the

[2] There also exist papers that design detection methods with relatively low informational requirements for specific settings, such as Kawai and Nakabayashi (2022) for auctions with a secret reserve price and rebidding, Conley and Decarolis (2016) for auctions where contracts are awarded to the bid closest to a trimmed average bid, and Baránek et al. (2021) for electronic procurement with multiple rounds.

internal organization of the two cartels are potentially linked to the observed differences in cartel bidding behavior. Interestingly, the testimonial evidence suggests that the cartels operated in a similar manner, but we still observe that the distribution of bids during the cartels significantly differs between Finland and Sweden. This finding underlines the notion that any individual statistical screen is unlikely to work in all settings. Finally, unlike most previous papers, we provide evidence of the change in cartel behavior using a difference-in-differences methodology.

Two papers have previously studied the Swedish asphalt cartel. Using a spatial econometric model, Bergman et al. (2020) shows a significant positive correlation between complementary cartel bids during the cartel period, whereas no correlation is shown after the cartel period. Lundberg (2017) illustrates that Moran's I statistic can be used to detect complementary bidding during the cartel period in the Swedish asphalt market. Our paper shows that the Swedish asphalt cartel could have been detected using methods with lower data requirements.

This paper is structured as follows. In the next section, we provide background information on the Finnish and Swedish asphalt cartels. In Section 3, we discuss why cartels might induce a change in the distribution of bids and link it to the testimonial evidence given in the Swedish and Finnish asphalt paving cartel cases. Section 4 describes the data. In Section 5, we analyze how the cartels altered the distribution of bids in Finland and Sweden. In Section 6, we test the performance of the cartel detection methods. Finally, in Section 7, we conclude.

2 Nordic asphalt cartels

Our analyzes are conducted on a dataset that covers publicly procured asphalt paving contracts awarded in Finland and Sweden. In Finland, publicly procured paving contracts can be divided into contracts awarded by municipalities and larger state-level contracts awarded by the Finnish Transport Infrastructure Agency. In terms of contract value, around half of the public demand for paving comes from the state-level contracts. In Sweden, public contracts are procured by the Swedish Road Administration, and the division between state-level contracts and municipality contracts is similar to Finland. In both countries, with a few exceptions, the contracts are allocated using first-price sealed-bid auctions, with the contract awarded to the lowest bidder.

Both in Sweden and Finland, there operated a cartel in the asphalt paving market in the 1990s and early 2000s. The asphalt paving cartels are the largest discovered public procurement cartels in Finland and Sweden. The Swedish cartel operated between 1993 and 2001, and the Finnish cartel between 1994 and 2002. In both Finland and Sweden, all the largest asphalt paving firms in the market were prosecuted and found guilty of collusion. While not all smaller firms were prosecuted, the testimonial evidence suggests that they also participated, voluntarily or by force. At the same time, asphalt paving cartels were also found in other Nordic countries. The Norwegian Competition Authority found that five firms had participated in a nationwide market sharing agreement between 1997 and 2001 (OECD, 2002) while in 1999, the Danish Competition Authority found that several asphalt firms were involved in anticompetitive agreements (OECD, 1999). Many of the largest asphalt paving firms were active in several of the Nordic markets. For example, the market leader of the Finnish asphalt market, Lemminkäinen, was, through its subsidiaries, active in several Nordic markets.

The Swedish Competition Authority started its cartel investigation after receiving information about illegal agreements in the asphalt paving industry in September 2001. The information came from three former employees of the asphalt firm NCC, who had left for a smaller firm in the same industry but had then been caught up in a legal dispute with their former employer over fake invoices related to the cartel.[3] The employees decided to disclose the cartel to the Competition Authority in order to escape personal liability. Their new employer applied for, and was later granted, immunity from fines for its involvement in the cartel. Shortly after, the Competition Authority conducted dawn raids, and legal proceedings against eight firms were initiated in March 2003. In 2009, the Swedish Market Court found the firms guilty of colluding between the years 1993 and 2001 and imposed fines of around 46 million euros.

In Sweden, the four largest firms in the industry operated the cartel. Representatives

[3]Fake invoices were used to make side payments between cartel members. For a discussion on the use of side payments in cartels, see Asker (2009).

of these firms met about twice a year.[4] Guidelines for collaboration in the upcoming year were drawn at a meeting, typically scheduled for late autumn. At the meetings, future projects were allocated among cartel members, and the winning prices for each project were agreed upon. Later, when the deadline for submitting bids for a tender approached, a coordinator working for the designated winner contacted other firms on how to set the losing bids. According to a former manager of one of the major firms, the level of complementary bids was carefully determined to ensure that all bids would seem natural. One of the ways to do this was to set the complementary bids close to the winning bid. (SMC, 2009).

Shortly after the Swedish cartel was discovered, the Finnish Competition Authority began its own in-depth cartel investigation based on the material it received from asphalt market participants.[5] In March 2002, it conducted dawn raids on the premises of several asphalt firms. In 2009, the Supreme Administrative Court of Finland found seven firms guilty of colluding between March 1994 and February 2002 and ordered the convicted firms to pay a total of 83 million euros in fines. While smaller firms were not convicted, witness reports suggest that also smaller firms participated in the cartel.

The court decision and the proposal submitted by the Finnish Competition Authority contain a description of how the Finnish cartel operated. Lemminkäinen, the market leader at the time, was the ringleader of the cartel. The state-level contracts and the contracts offered by municipalities were divided between the cartel participants. Before the deadline of the tender, the firms would coordinate their bids over the phone. According to the manager of one of the convicted firms, before the calls each firm calculated its costs for the project, and then the prices were compared and a predetermined margin was added to the price. The designated winner typically tried to negotiate the price upward, while the others aimed at negotiating the margin lower so that the winning bid would not be unrealistically high and unveil the cartel. The negotiations were managed by the ringleader. Based on witness reports, the complementary bids were set close to the

[4] Our description of the Swedish cartel is based on the Swedish Market Court's decision, the appeal by the Swedish Competition Authority, and on interviews of Anders Gerde, a case handler in the Swedish Competition Authority, in Hjalmarsson (2015) and in Kapitalet podcast, available at `https://podcasts.nu/avsnitt/kapitalet-en-podd-om-ekonomi/asfaltkartellen-del-1-att-bygga-en-kartell`.

[5] The Finnish Competition Authority had been investigating the possibility of a cartel in the asphalt market already in the end of the 1990s based on claims from market participants (Lindberg, 2020).

winning bid, so that the procurer would think that they were getting a correct and fair price. (FCA, 2004; SACF, 2009).

3 Complementary bidding and bid distribution

To develop our empirical hypotheses, in this section we link the testimonial evidence in Finland and Sweden to the theoretical literature on bid rigging. Both the Finnish and the Swedish bid-rigging cartels chose a designated winner for each tender. The designated winner then submitted the lowest bid to the tender, while the other cartel members submitted complementary bids that exceeded the bid of the designated winner. Complementary bids were intended to give the impression of competition to the procurer.

We first discuss why bid rigging can lead to bid clustering. LaCasse (1995) develops a bidding model with endogenous collusion in auction markets, where bidders know that the competition authority can potentially detect collusive behavior. The equilibrium bidding range of the cartel is a subset of the distribution if the firms bid competitively. The cartel truncates the distribution of bids because the winning bid needs to be at least as high as in a competitive market. Furthermore, the cartel cannot set the losing bids too high because unreasonably high bids could lead to antitrust scrutiny. Overall, this leads to a lower variance of bids and clustering of bids under collusion. The testimonial evidence in the Finnish and Swedish asphalt cartels also suggests that complementary bidding led to bid clustering. In both cartels, the firms submitted complementary bids close to the winning bid to make the tender look competitive. Moreover, according to witness statements from the Swedish cartel, the designated losers did not want to submit too high bids to avoid bad advertising. The firms did not want to send negative signals about their competitiveness to private sector clients through their public procurement prices. Additionally, procurement authorities sometimes invited a subset of firms to bid for smaller projects, rather than using an open procedure. Therefore, firms bidding very high on previous projects could potentially be left uninvited for future tenders.

Complementary bidding might also increase the distance between the winner and the runner-up. Chassang et al. (2022) propose two potential reasons for why the winning bids might be isolated under bid rigging. First, nearly identical bids may attract antitrust scrutiny. Cartel members might use tied bids as a randomization device to determine contract allocation.[6] Therefore, many competition authorities list tied, or almost tied, bids as a potential marker for bid rigging. There is also evidence that firms have reacted to this. In the marine hose cartel, which operated between 1986 and 2006 and involved one Swedish firm, one of the internal documents stated that a small difference should be left between the winning bid and the second lowest bid and that identical bids should be avoided (EC, 2009). Second, isolated winning bids may make it easier to ensure the allocation of the contract to the designated winner. This is especially the case in auctions where allocation can be affected by non-price characteristics of the bids, such as quality or completion time, or small trembles can perturb bids.

A caveat with using low variance and bid clustering as an indication of collusion is that it could also be explained by small cost differences between firms. Instead, a consistently high difference between the winner and the runner-up is inconsistent with competitive behavior. Chassang et al. (2022) develop a model with a group of firms that repeatedly participate in first-price sealed-bid procurement auctions. When the winning bid is consistently isolated, the winner can profitably deviate by bidding higher, implying that a persistent gap between the winning bid and the losing bids cannot be a feature of a competitive equilibrium.

In summary, based on both theory and testimonial evidence, we expect that the Finnish and Swedish asphalt cartels potentially led to the clustering of bids and an increase in the difference between the winner and the runner-up. In the following sections, we will test these hypotheses empirically.

4 Data

Our dataset consists of state-level asphalt paving contracts procured by the Finnish Transport Infrastructure Agency in Finland and the Swedish Road Administration in Sweden. For Finland, the dataset covers contracts from 1994 to 2019. For years 1994–2009, the

[6]McAfee and McMillan (1992) show formally that submitting identical bids can be used as a randomization device to determine allocation in bid-rigging cartels.

dataset was collected by the Finnish Transport Infrastructure Agency, and it covers around half of the tenders of that period. For years 2010–2019, we have supplemented the original dataset with information from public procurement documents. For the later period, the coverage of the data is slightly lower. For Sweden, the dataset covers contracts from 1993 to 2009. Both datasets contain information on all submitted bids, the identity of the winner, and the region where the pavement project took place. For the years 1994–2009 in the Finnish data, we also observe detailed information about the project, such as the paving area (m^2), the amount of asphalt (tonnes), and the asphalt quality. For Sweden, we only observe the paving area. We have also collected data on the price of bitumen, as it is one of the main inputs in the production of asphalt.[7] Finally, we have converted Swedish kronor into euros using the exchange rates provided by the central bank of Sweden.

In both datasets, we exclude observations from the year the dawn raids were conducted because it is not clear which of these tenders were still affected by the cartel. Following Bergman et al. (2020) and Lundberg (2017), we also exclude data from Sweden between the dawn raids in 2001 and the first court order in 2003 because the investigated firms might not have immediately understood the seriousness of the charges. For Finland, we also exclude 1994 from the sample because there was a change in the value-added tax in 1994.[8] We also drop tenders where there is more than one winner, information on all bids is not available, the lowest bid did not win, or only one bid was submitted. In total, these exclude 177 tenders. In addition, we drop 4 more tenders from the Swedish data where we suspect typing errors due to unusually high or low bids.[9] After cleaning the data, our dataset has information on 4983 bids on 1008 tenders with a total awarded value of 2.6 billion euros.

Table 1 presents summary statistics for both Finland and Sweden before and after the cartel investigations. For Finland, we observe 2250 bids on 457 tenders. The average contract value before the cartel investigation was 6.35 million euros and 4.02 million euros after the investigation. The observed decrease in the average contract value is explained by both a cartel overcharge and a decrease in the average project size.[10] Before

[7]The bitumen prices are collected from the annual statistics published by Statistics Finland.

[8]In Appendix A3 we report results with all years included.

[9]The largest bid was roughly 10 times larger than the smallest bid.

[10]VATT Institute of Economic Research estimated that the cartel overcharge in the Finnish asphalt

the start of the investigations, we observe bids from a total of 32 firms and after the investigations from 21 firms. Although relatively many firms submitted bids, the market was still fairly concentrated with a Herfindahl–Hirschman index (HHI) of 2302 before the investigation and 2593 after the investigation. The average number of bids was 5.44 before the investigation and 4.70 after. Overall, this increase in market concentration is driven, at least partially, by mergers.[11] This observation is consistent with Dong et al. (2019) who find that introducing a leniency program led to increased merger activity, suggesting that mergers can be a way to replace cartel agreements.

The Swedish dataset consists of 2733 bids on 551 tenders. Of the 551 tenders, 410 come from before the cartel investigation and 141 after. The average contract value was 0.73 million euros in the cartel period and 1.14 million euros in the post-investigation period. The increase in the average contract value can be explained by the fact that the size of the projects has increased over time. Similarly to Finland, we find that market concentration has also increased in the Swedish dataset after the cartel. The number of active firms decreased from 50 to 30, HHI increased from 1942 to 2181, and the average number of bids decreased from 5.25 to 4.11.[12]

market case was around 15% (VATT, 2011). Ultimately, the Helsinki Court of Appeal ordered the asphalt firms to compensate a total of 34 million euros to the State of Finland and several local municipalities (HCA, 2016). For a description of the court proceedings related to the damages cases see Lindberg (2020).

[11] NCC and Destia merged in 2011. The combined market share of the firms in state-level contracts was between 40–50% in 2010 (FCCA, 2011). Other notable mergers were NCC's acquisition of Valtatie Oy in 2008 and, already prior to the ending of the cartel, Skanska's acquisitions of Sata Asfaltti and Savatie in 2000.

[12] In Sweden, there has not been mergers between market leaders after the investigations like in Finland.

Table 1: Summary statistics

	Finland		Sweden	
	Cartel (1994–2001)	Post inv. (2003–2019)	Cartel (1993–2000)	Post inv. (2004–2009)
Total awarded (m EUR)	870.6	1286.8	297.4	161.4
Nbr of contracts	137	320	410	141
Avg contract value (m EUR)	6.35	4.02	0.73	1.14
Nbr of bids per contract	5.44	4.70	5.25	4.11
Nbr of bidding firms	27	19	50	30
HHI	2302	2593	1942	2181

5 Change in the bid distribution

5.1 Descriptive analysis

In Section 3 we developed two hypotheses on how the Finnish and Swedish cartels potentially affected the distribution of bids. In this section, we test these by comparing the distributions of bids during the cartel period and after the launch of cartel investigations. We begin by calculating the difference between a bid and the lowest rival bid in the tender in the following manner:

$$\Delta^1_{i,t} = \frac{b_{i,t} - \Lambda b_{-i,t}}{\Lambda b_{-i,t}} \tag{1}$$

where $b_{i,t}$ refers to the price of bid i in tender t and $\Lambda b_{-i,t}$ denotes the minimum bid of bidders other than i in tender t. A key feature of the measure is that it is scale-invariant and is therefore comparable across different-sized projects.

A similar measure has previously been used by Chassang et al. (2022) and Clark et al. (2020). Our measure is calculated slightly differently compared to these studies.

Chassang et al. (2022) calculate the bid differences relative to a reserve price, while Clark et al. (2020) calculate the bid differences in terms of unit prices. In Appendix A1, we provide results based on the definition used by Clark et al. (2020). We do not use this as our main measure since we do not observe unit prices for all tenders. Because the asphalt paving auctions in Finland and Sweden do not have reserve prices, we cannot use the definition by Chassang et al. (2022).

We can test our two hypotheses using $\Delta^1_{i,t}$. The first hypothesis regarded the clustering of bids. Clustering would decrease the mass of bid differences $\Delta^1_{i,t}$ at the tails of the distribution and increase the mass of bid differences relatively close to zero. Based on our second hypothesis, collusive bidding might introduce a gap between the winning bid and the losing bids. This would result in a lower mass of bid differences $\Delta^1_{i,t}$ close to zero.

In Figure 1, we have plotted the distribution of $\Delta^1_{i,t}$ during the cartel period and after the investigations for both countries. Panel A shows a histogram and a density estimate of $\Delta^1_{i,t}$ in Finland. There are clearly noticeable differences between the distribution during the cartel and after the cartel investigation. The mass of bid differences within 10% of the winning bid is higher during the cartel period. However, just around zero, the mass of bid differences is somewhat similar before and after the investigation. Additionally, the tails of the distribution taper off more rapidly during the cartel period than after the investigation. Panel B shows a similar pattern for Sweden. We observe a twin-peaked distribution of bid differences during the cartel period with a large mass of bid differences relatively close but not very close to zero. Our findings are in line with the bidding patterns found by Clark et al. (2020) and Chassang et al. (2022). Both of these studies find a similar twin-peaked distribution of bid differences from collusive markets.

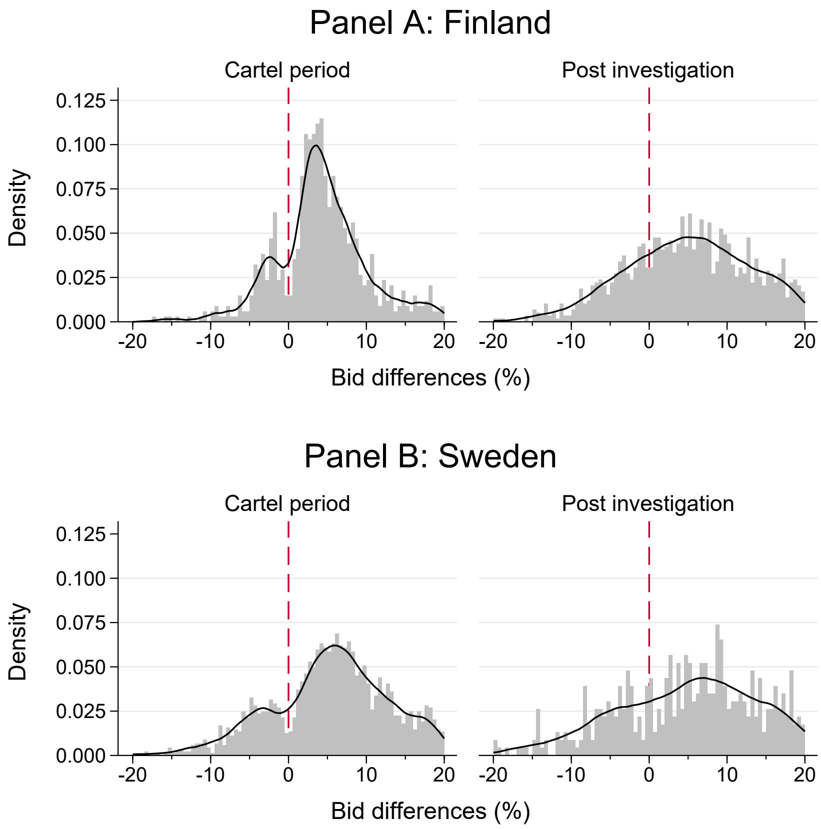


Figure 1: Distribution of bid differences $\Delta^1_{i,t}$ by country and period

This figure plots the distribution of differences between a bid and the lowest rival bid for asphalt procurement contracts in Sweden and Finland before and after cartel investigations. The width of the bins is 0.5. The curves correspond to density estimates calculated using an Epanechnikov kernel.

To test more formally how the distribution of bids changed after the launch of the investigation, we follow Clark et al. (2020) and use a distributional regression approach. We estimate a linear probability model where the outcome variable is a binary variable equal to 1 if the bid difference falls within a given interval of values. The explanatory variable of interest is a binary variable equal to 1 if the observation is from the post-investigation period. The linear probability model is estimated separately for each interval. We also estimate the models separately for the Finnish and Swedish datasets.

More specifically, we estimate the following linear probability model:

$$y_{i,t,g} = \alpha_g + \beta_{1,g}\, post_t + \gamma_g Z_t + \epsilon_{i,t,g} \tag{2}$$

where $y_{i,t,g}$ is a binary variable equal to 1 if the bid difference of bid i in tender t falls within the interval g, $post_t$ is a binary variable equal to 1 if the tender is from the post-investigation period, Z_t is a vector of control variables that include the bitumen index and region fixed effects. Our parameter of interest is $\beta_{1,g}$, which estimates the difference in the share of bid differences within an interval g during the post-investigation period as compared to the cartel period. By construction, the coefficients across intervals sum up to zero. Because of the within-tender correlation in bid differences, we cluster standard errors at the tender level.[13]

The width of the intervals needs to be specified beforehand. Based on our hypotheses in Section 3, we choose three intervals. In the first interval, $y_{i,t,g}$ is equal to 1 if the absolute value of the bid difference falls within 1%. In the second interval, $y_{i,t,g}$ is equal to 1 if the absolute value of the bid difference is within 1–10%. In the third interval, $y_{i,t,g}$ is equal to 1 if the absolute value of the bid difference is larger than 10%. Although these intervals are based on our hypotheses, they are still defined arbitrarily. In Appendix A2, we present results of an alternative specification where we estimate equation (2) for one percent intervals between -20% and 20%.

The results are shown in Table 2. For Finland, we find no statistically significant difference for the share of bid differences near zero (within 1%) between the cartel period and the post-investigation period. However, the share of bid differences at the peaks (between 1% and 10%) is 31 percentage points lower during the post-investigation period while the share of bid differences at the tails (more than 10%) is 33 percentage points higher. Overall, the results indicate that during the cartel period, firms submitted more bids that were relatively close to the winning bid. However, while we observe more bids relatively close to the winning bid, the cartel seems to have avoided leaving the complementary bids very close to the winner. For Sweden, we observe similar results. During the post-investigation period, the share of bid differences within 1% to 10% is lower and the share of bid differences larger than 10% is higher. The statistical significance and the magnitude of the estimates, however, are lower for the Swedish dataset than for the Finnish dataset.

[13] For example, the bid difference of the smallest bid and the bid difference of the second smallest bid are correlated by construction.

We perform two additional robustness checks to assess the sensitivity of our findings to the model specification. First, we estimate a model where we add the size of the project, measured by the paving area, as a control variable. Adding the project size as a control increases the absolute value and statistical significance of β_g for intervals 1–10% and above 10% for both Finland and Sweden. Second, we estimate the model using the full dataset by including tenders from the excluded years (e.g., the investigation years). Including the tenders from the omitted years does not significantly change the results. The results of the robustness checks are reported in Appendix A3.

Table 2: Distributional effect of the cartel investigations

Panel A: Finland

	Within 1%	Between 1% and 10%	More than 10%
Post	-0.018	-0.310***	0.328***
	(0.021)	(0.047)	(0.045)
Observations	2250	2250	2250
Region FE	Yes	Yes	Yes
Bitumen index	Yes	Yes	Yes

Panel B: Sweden

	Within 1%	Between 1% and 10%	More than 10%
Post	-0.027	-0.088*	0.115**
	(0.022)	(0.053)	(0.054)
Observations	2733	2733	2733
Region FE	Yes	Yes	Yes
Bitumen index	Yes	Yes	Yes

The dependent variable is the probability that bid differences fall in a given interval. Post is a dummy equal to 1 if the contract was awarded after the investigation. Panel A shows results for Finland and Panel B for Sweden. Standard errors are clustered at the tender level. Significance at $p < 0.10$ (*), $p < 0.05$ (**), and $p < 0.01$ (***).

5.2 Difference-in-differences analysis

A potential concern in the results presented in the previous section is that they could be driven by something else than the start of cartel investigations and the collapse of the cartels. Given that we find similar results for both countries, we are confident that country-specific changes in procurement practices are not driving our results. However, asphalt paving market–specific changes could potentially have resulted in a similar simultaneous change in bidding behavior both in Finland and Sweden. To test the robustness of our results for asphalt paving market–specific changes, we use data from a control market where there is no evidence of collusion either before or after the start of the cartel investigations in the Nordic asphalt markets.

The control market is the Californian asphalt paving market. Since the prices of the main inputs used in asphalt paving are similar around the world, the Californian asphalt market is exposed to similar cost shocks as the Nordic asphalt markets.[14] In previous literature, the market has been modeled as competitive and there has not been any disclosed cartel investigations in the Californian asphalt paving market during our examination period. Given the above, we believe that the Californian asphalt market provides us with a plausible control market.

The Californian data covers paving contracts procured by the California Department of Transportation from 1999 to 2008.[15] The dataset was originally used by Bajari et al. (2014), and it contains similar information as the Finnish and Swedish datasets: the information on all submitted bids, the identity of the winner, and the region where the pavement project took place. The dataset contains 6914 bids on 1449 contracts.[16]

In Figure 2, we plot a histogram and a density estimate of bid differences $\Delta^1_{i,t}$ for the Californian market before and after the Nordic cartel investigations. We observe that in the Californian market, the distribution of bid differences is similar for both periods.

[14] The main inputs of asphalt are gravel and bitumen. Since the costs of gravel are small compared to bitumen, it is the price of bitumen that mainly determines the production costs. Hence, many countries have also chosen to peg the project prices to the bitumen index. Since bitumen is produced only in few areas, the market for bitumen is global and regional differences in the prices are relatively small.

[15] Contract details from 2001 and the first half of 2003 were not available.

[16] We follow Bajari et al. (2014) in cleaning the dataset. We exclude contracts with only one bidder, contracts with scoring auctions, contracts that were rebid, contracts for which information on all bids was not available, and contracts that were not awarded to the lowest bidder. In total, we drop 113 contracts with 537 bids.

We interpret this as the first indication that there has not been an industry-wide shift in bidding patterns that could explain our findings in the previous section.

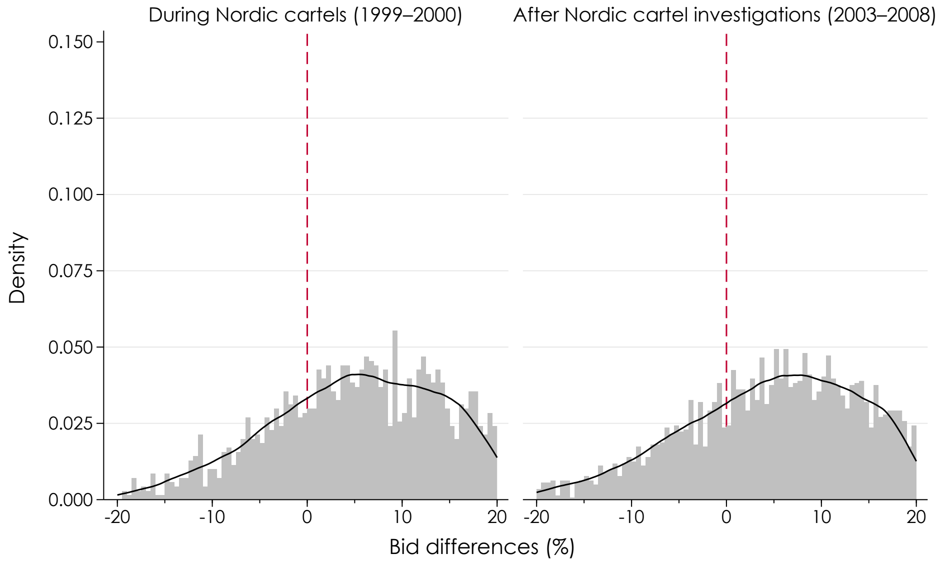


Figure 2: Distribution of bid difference $\Delta^1_{i,t}$ for control market

This figure plots the distribution of differences between a bid and the lowest rival bid for asphalt procurement contracts in California before and after the cartel investigations in Finland (launched in 2002) and Sweden (launched in 2001). The width of the bins is 0.5. The curves correspond to density estimates calculated using an Epanechnikov kernel.

Next, we estimate the following distributional difference-in-differences regression separately for Sweden and Finland:

$$y_{i,t,g} = \alpha_g + \beta_{2,g}\, post_t + \beta_{3,g}\, treat_t + \beta_{4,g}\, treat_t \times post_t + \epsilon_{i,t,g} \tag{3}$$

where, similar to the equation (2), $y_{i,t,g}$ is a binary variable equal to 1 if the bid difference of bid i in tender t falls within the interval g, and $post_t$ is a binary variable equal to 1 if the tender is from the post-investigation period. $treat_t$ is a binary variable equal to 1 for tenders from Sweden and Finland and zero for tenders from California. Our parameter of interest is $\beta_{4,g}$, which will inform how the cartel shifted the distribution of bid differences in Finland and Sweden compared to how the distribution of bids evolved in California. In the regressions, we only include the years available for both markets. Again, we cluster standard errors at the tender level.

The results are shown in Table 3. We again find that in Finland the cartel investigations led to a lower share of bid differences relatively close to the winning bid (between 1% and 10%) and a higher share far from the winning bid (more than 10%). However, the point estimates are slightly smaller in absolute terms when the control market is included. For Sweden, the signs of the coefficient estimates remain the same as in our previous analysis but the magnitude and statistical significance are lower.

The difference-in-differences analysis relies on the common trend assumption. Because we only have two common years of data from the control market and the treatment market before the investigations, we are unable to provide a formal and credible test of the parallel trends assumption. In Appendix A4, we provide descriptive evidence that there existed no clearly observable diverging trends between California and Finland or Sweden in the share of bids at different intervals. We also show that our results are robust to adding country-specific time trends to the difference-in-differences specification.

Overall, the difference-in-differences analysis strengthens our view that in Finland the cartel investigation led to a large and statistically significant shift in the distribution of bid differences, whereas in Sweden the change is less pronounced. Based on the testimonial evidence, the two cartels operated similarly and, in both markets, they aimed at submitting losing bids close to the winning bid. Given this, we find it surprising that the results between the two countries are so different. One apparent difference between the two cartels is that in Finland there was only one ringleader, whereas in Sweden the cartel was operated by four firms. Such differences in the organization of the cartel could potentially explain the observed differences in bidding patterns during the cartel period.

Table 3: Distributional effect of the cartel investigations using a difference-in-differences design

Panel A: Finland			
	Within 1%	Between 1% and 10%	More than 10%
Post × Treat	0.020	-0.238***	0.218***
	(0.023)	(0.075)	(0.073)
Observations	6924	6924	6924
Panel B: Sweden			
	Within 1%	Between 1% and 10%	More than 10%
Post × Treat	-0.021	-0.052	0.073
	(0.020)	(0.046)	(0.047)
Observations	7071	7071	7071

The dependent variable is the probability that bid differences fall in a given interval. Post × Treat is equal to 1 for Finland and Sweden after the launch of cartel investigation. Panel A shows results for Finland and Panel B for Sweden. Standard errors are clustered at the tender level. Significance at $p < 0.10$ (*), $p < 0.05$ (**), and $p < 0.01$ (***).

6 Cartel detection tests

After establishing that the distribution of bids was different during the cartel period, we continue by examining the performance of the detection methods proposed by Clark et al. (2020) and Huber and Imhof (2019). While documenting differences in bidding behavior between collusive and competitive periods is interesting in itself, the key to competition policy is how well these insights can be used to detect future cartels.

Our analysis in the previous section relied on comparing the bidding behavior during a collusive period and a competitive period in the same market. Furthermore, we compared the change in bidding behavior in the suspected market to a control market. Competition authorities interested in detecting collusion most often do not have access to such data. The detection methods tested in this section have more modest data requirements. The distributional regression method by Clark et al. (2020) requires only bidding data from the suspected market, while the machine learning–based method proposed Huber and Imhof (2019) requires, in addition to the data from the suspected market, some data from a related market that has both collusive and competitive tenders.

6.1 Distributional regression test

The cartel detection method suggested by Clark et al. (2020) is based on comparing two different distributions. The first is the distribution of bid differences $\Delta^1_{i,t}$ as defined in equation (1). The second set of bid differences $\Delta^2_{i,t}$ is defined similarly to $\Delta^1_{i,t}$ but excluding winning bids. The intuition of the test is that the difference between losing bids is not similarly affected by bid rigging as the difference between the winning and losing bids is. The cartel does not have similar incentives to manipulate the difference between the losing bids as they have for manipulating the difference between the winner and the runner-up.

In Figure 3, we plot the distribution of $\Delta^2_{i,t}$ for the cartel period and the post-investigation period for both Finland and Sweden. Unlike the original distribution, in this alternative distribution we find no missing mass of bids at zero during the cartel period, indicating that this alternative distribution differs substantially from the original twin-peaked distribution of bid differences $\Delta^1_{i,t}$. However, we do find some differences in the distribution of $\Delta^2_{i,t}$ between the cartel period and the post-investigation period. For both Finland and Sweden, the tails of the distribution taper off more rapidly during the cartel period than in the post-investigation period. This finding is in line with the testimonial evidence where the employees of the firms state that the complementary bids were set close to the winning bid to give an impression of intense competition. However, this finding contradicts LaCasse (1995), who argues that the distribution of losing bids is not informative about the existence of a cartel once the winning bid is known. Our results seem to indicate that imitating competitive behavior can be costly for the cartel, which can result in losing bids, too, conveying information about collusion. This insight is not novel in the literature. For example, Porter and Zona (1999) show that losing bids were not correlated with cost measures during a cartel in Ohio. Another potential reason for losing bids conveying information on collusive behavior is that cartel members were not aware that clustering of bids could be used to detect the existence of a cartel.

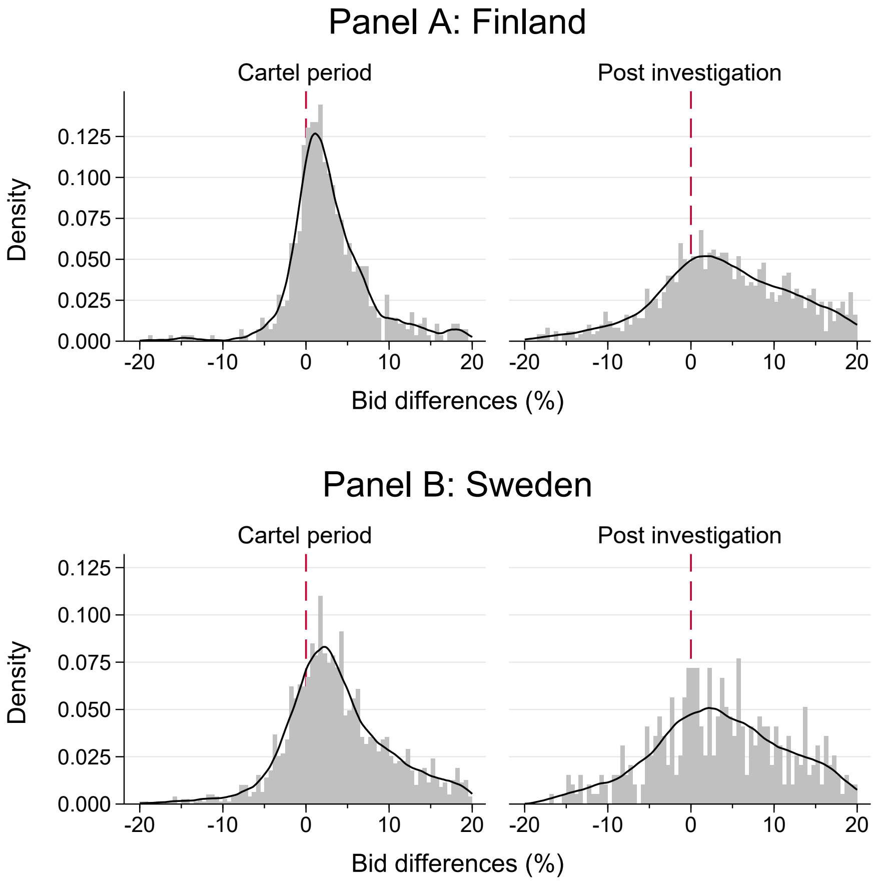


Figure 3: Distribution of bid difference excluding winning bids $\Delta^2_{i,t}$ by country and period

This figure plots the difference between a bid and the lowest rival bid when the winning bid is excluded for asphalt procurement contracts in Sweden and Finland before and after the cartel investigations. The width of the bins is 0.5. The curves correspond to density estimates calculated using Epanechnikov kernel.

To implement the test, the two sets of bid differences, $\Delta^1_{i,t}$ and $\Delta^2_{i,t}$, are stacked, and the following distributional regression is run separately for each interval:

$$y_{i,t,g} = \alpha_g + \beta_g \mathbb{1}(f(\Delta^1_{i,t})) + \gamma_g Z_t + \epsilon_{i,t,q} \tag{4}$$

where $y_{i,t,g}$ is a binary variable equal to 1 if the bid difference of bid i in tender t falls within the interval g. $\mathbb{1}(f(\Delta^1_{i,t}))$ is a binary variable equal to 1 if the observation is from the original distribution where the winning bid is included and zero if it is from the alternative distribution where the winning bid is excluded. Z_t is a vector of control variables. Similarly to the previous section, we include region-fixed effects and the bitumen index as controls and cluster standard errors at the tender level.

The parameter of interest is β_g, which estimates the difference in the density of bids in interval g between the two alternative distributions. Under no collusion, the null hypothesis tested is the following:

$$H_0 : \beta_g = 0 \ \forall \ \in [-H, H] \tag{5}$$

We follow Clark et al. (2020) and implement the test for a close neighborhood of zero. We run the regression for 20 equal-sized intervals between -5% and 5%.[17]

Figure 4 shows the results for both countries during and after the cartel period. In the cartel period, the null is correctly rejected for both Finland (Panel A) and Sweden (Panel B). In both countries, the density of bids close to zero is lower in the distribution with winning bids included. For example, in Finland, the share of bids within 0.0–0.5% of the most competitive rival bid is -0.043 (s.e. 0.009) lower in the distribution that contains the winning bid. For Sweden, the corresponding point estimate is -0.019 (s.e. 0.004). Both countries have several statistically significant negative estimates in a neighborhood around zero. However, the magnitude and statistical significance of the point estimates start to taper off when moving further away from zero. Overall, the test seems to capture well the missing mass of bids close to zero during the cartel period.

The results for the post-investigation period are shown on the right-hand side of the Figure 4. For Finland, the null hypothesis is not rejected in any of the 20 intervals. This indicates that after the cartel period in Finland, there are no notable differences in the two sets of bid differences. For Sweden, we observe significant differences at the $p < 0.05$ level for 2 out of the 20 intervals. However, unlike during the cartel period, no clear visual pattern emerges from the estimates. In addition, having two statistically significant estimates by chance is not unlikely given the number of estimates. Overall,

[17] We have also tested alternative specifications such as using 20 equal-sized intervals between -2% and 2%. The results and conclusions remain similar. These results are available upon request.

we conclude that the test correctly rejects collusion in Finland in the post-investigation period, while for Sweden the results for the post-investigation period are somewhat mixed.

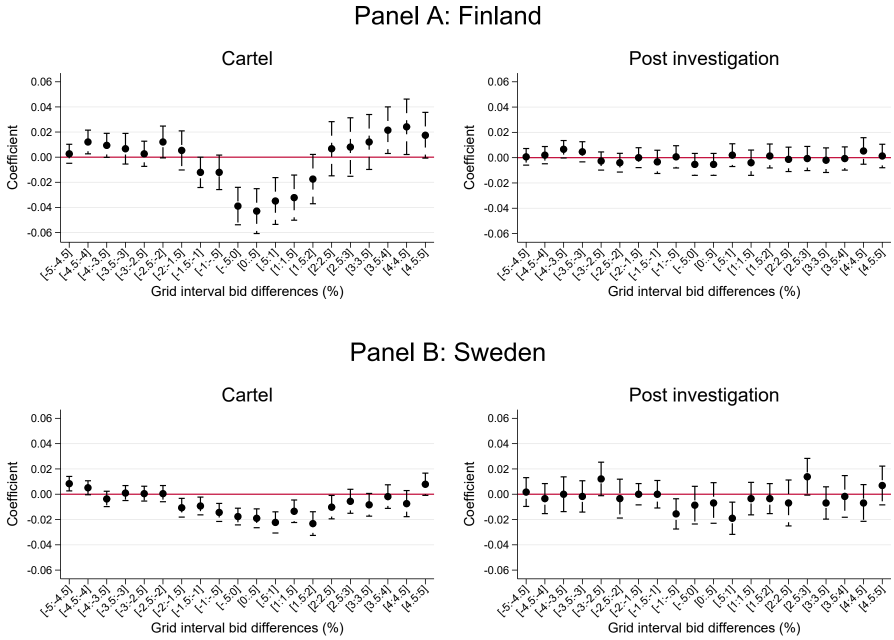


Figure 4: Results from Clark et al. (2020) cartel detection test

This figure plots the point estimates and 95% confidence intervals of β_g from equation (4) separately for each country and time period. Confidence intervals are calculated with robust standard errors.

6.2 Machine learning test

Next, we examine the performance of a cartel detection method suggested by Huber and Imhof (2019). The method uses machine learning to estimate a predictive model that classifies tenders as collusive or competitive. As predictors, the model uses statistical screens computed from the distribution of bids within tenders. These include, for example, the standard deviation of the bids and the difference between the winner and the runner-up. In total, there are 10 statistical screens with each screen capturing a different aspect of the distribution of bids. A list of the statistical screens is presented in Table 4. A detailed description of the screens is provided in Appendix A5. In addition to the 10 screens, the set of predictors also includes second powers of the screens and their interactions.

Table 4: Formulas for statistical screens

Screen	Formula
Standard deviation	σ_t
Coefficient of variation	$\frac{\sigma_t}{\mu_t}$
Kurtosis	$\frac{n_t(n_t+1)}{(n_t-1)(n_t-2)(n_t-3)} \sum_{i=1}^{n} (\frac{b_{i,t}-\mu_t}{\sigma_t})^4 - \frac{3(n_t-1)^2}{(n_t-2)(n_t-3)}$
Absolute difference	$b_{2,t} - b_{1,t}$
Percentage difference	$\frac{b_{2,t}-b_{1,t}}{b_{1,t}}$
Skewness	$\frac{n_t}{(n_t-1)(n_t-2)} \sum_{i=1}^{n} (\frac{b_{it}-\mu_t}{\sigma_t})^3$
Relative distance	$\frac{b_{2,t}-b_{1t}}{\sigma_{-1,t}}$
Normalized distance	$\frac{b_{2,t}-b_{1,t}}{\frac{\sum_{i=1,j=i+1} b_{j,t}-b_{i,t}}{n_t-1}}$
Number of bids	n_t
Mean of bids	μ_t

This table presents the formulas used to calculate the statistical screens used in the predictive model. $b_{i,t}$ refers to the i'th lowest bid in tender t. μ_t, σ_t, and n_t refer to the mean, standard deviation and number of bids in tender t.

To calibrate the predictive model, the dataset is divided into two parts: a training set and a test set. The training set is used for estimating the model parameters, and the test set is used for evaluating the predictive performance of the model given the estimated parameters. This requires prior knowledge of the true values (i.e., whether a given tender was competitive or collusive) of the tenders in the dataset. After the model has been estimated and the predictive performance is considered adequate, the predictive model can be used to predict collusion from datasets where collusion is not known ex-ante.

We follow Huber and Imhof (2019) and use lasso logit regression as our predictive model. It is a logistic regression where the number of predictors is restricted with a so-called penalty term. By limiting the number of predictors, the model chooses only the best predictors and hence avoids overfitting. The estimation of the model's parameters is based on the following optimization problem:

$$\max_{\delta_0,\,\boldsymbol{\delta}} \left\{ \sum_{i=1}^{n} \left[y_t(\delta_0 + \sum_{j=1}^{p} \delta_j x_{tj}) - \log(1 + e^{\delta_0 + \sum_{j=1}^{p} \delta_j x_{tj}}) \right] - \lambda \sum_{j=1}^{p} |\delta_j| \right\}. \tag{6}$$

where t indexes a tender and j a predictor in our data, y_t is the cartel indicator, δ_0 and $\boldsymbol{\delta}$ denote the intercept and slope of the predictors, $\boldsymbol{x}$ is the vector of predictors (i.e., the statistical screens), and λ is the coefficient of the penalty term. The penalty term limits the number of predictors of the model based on the sum of their coefficients' absolute values. Since the coefficient of the penalty term cannot be simultaneously estimated with the predictor coefficients, it is estimated with 15-fold cross-validation within the training data. We do this by first splitting the training data into 15 sections (called folds). Then a candidate penalty term coefficient is assigned for each fold, and the rest of the parameters are estimated for each fold with the penalty term coefficient as given. Finally, the performance of each model is tested with the other 14 folds, and the penalty term coefficient from the best-performing model is chosen for the final model. In addition to lasso logit, we also report the results for alternative machine learning models in Appendix A7.

The analysis is performed using two different ways to divide the observations into training and test data. In the first analysis, the training and test data are from the same country. The training sample contains 75% of observations and the test sample 25% of observations. The sampling is repeated 100 times, and the performance measures are averaged over the 100 repetitions. In the second analysis, we use data from one country as training data and evaluate the performance using data from the other country.

Before reporting the results of the predictive model, in Table 5 we report the means and standard errors of the statistical screens, separately for the cartel period and the post-investigation period. We also report the test statistics and p-values of a Welch's t-test and Kolmogorov–Smirnov test. The former tests whether the two samples have the same mean whereas the latter tests whether the two samples are from the same probability distribution, hence also considering the shape of the distribution. Consistent with our earlier findings, the screens capturing bid clustering (standard deviation, coefficient of variation, and kurtosis) show statistically significant differences between the cartel period and the post-investigation period. Similarly, the observed differences in the normalized distance and skewness are in line with the existence of an isolated winning bid during the cartel period. In 8 of the 10 statistical screens, the change in the mean has a similar sign in both Finland and Sweden. Overall, the signs are also similar to those reported by Huber and Imhof (2019) on the Swiss dataset.

Table 5: Descriptive statistics for statistical screens

	Finland				Sweden			
	Cartel	Post inv.	T-test	K–S test	Cartel	Post inv.	T-test	K–S test
	mean/sd	mean/sd	stat/p-val	stat/p-val	mean/sd	mean/sd	stat/p-val	stat/p-val
Standard deviation	0.419	0.286	4.271	0.267	0.110	0.060	4.768	0.277
	(0.399)	(0.255)	(0.000)	(0.000)	(0.116)	(0.071)	(0.000)	(0.000)
Coefficient of variation	0.097	0.054	9.048	0.507	0.089	0.078	1.926	0.239
	(0.050)	(0.044)	(0.000)	(0.000)	(0.052)	(0.058)	(0.055)	(0.000)
Kurtosis	2.037	2.230	-2.921	0.157	2.007	2.200	-2.829	0.161
	(0.587)	(0.597)	(0.004)	(0.037)	(0.569)	(0.677)	(0.005)	(0.039)
Absolute difference	0.203	0.208	-0.243	0.149	0.090	0.043	4.391	0.268
	(0.235)	(0.206)	(0.808)	(0.028)	(0.119)	(0.071)	(0.000)	(0.000)
Percentage difference	0.056	0.046	1.736	0.204	0.077	0.071	0.866	0.122
	(0.049)	(0.061)	(0.084)	(0.001)	(0.068)	(0.086)	(0.387)	(0.090)
Skewness	0.269	-0.026	4.428	0.222	0.094	-0.066	2.620	0.199
	(0.609)	(0.655)	(0.000)	(0.000)	(0.580)	(0.692)	(0.009)	(0.001)
Relative distance	1.259	2.408	-1.556	0.319	1.530	1.704	-0.696	0.129
	(3.155)	(8.225)	(0.122)	(0.000)	(2.357)	(2.816)	(0.487)	(0.075)
Normalized distance	0.839	1.377	-7.147	0.353	1.032	1.354	-4.370	0.219
	(0.634)	(0.777)	(0.000)	(0.000)	(0.721)	(0.847)	(0.000)	(0.000)
Number of bids	4.703	5.438	-4.592	0.249	4.106	5.254	-9.442	0.367
	(1.253)	(1.684)	(0.000)	(0.000)	(1.193)	(1.382)	(0.000)	(0.000)
Mean of bids	4.441	6.723	-5.223	0.270	1.264	0.800	5.414	0.252
	(2.770)	(4.783)	(0.000)	(0.000)	(0.933)	(0.687)	(0.000)	(0.000)

In this table, we report the period-specific means and standard errors of the statistical screens as well as the test statistics and p-values of a Welch's t-test and a Kolmogorov–Smirnov test between the two periods. We do this separately for Finland and Sweden. Standard errors and p-values are reported in the parentheses.

Table 6 reports the performance of the predictive model.[18] The performance is measured using accuracy, balanced accuracy, and mean squared error (MSE). Accuracy represents the share of all tenders that were predicted correctly with a probability threshold of 50% for a collusive prediction. Balanced accuracy is defined as the arithmetic mean of the accuracy for collusive tenders and the accuracy for competitive tenders. MSE is calculated based on the mean squared error between the predicted collusion probability and the true value of the tender.

In Panel A, we report the main results. The first two columns focus on the within-country analysis where both training and test data are from the same country. For Finland, the model was able to predict 83% of the tenders correctly. For Sweden, the prediction rate is 70%. The prediction rate for Finland is close to what has been found in the previous literature, but for Sweden the rate is significantly lower. Using the same method and the same set of predictors, Huber and Imhof (2019) are able to predict 84% of tenders correctly with a Swiss road construction procurement dataset. Using a slightly different set of predictors, Huber et al. (2022) predict 88% to 97% of tenders correctly with a Japanese procurement dataset. Silveira et al. (2022) achieve an even higher prediction rate when they apply a similar predictive model to a non-procurement setting in the Brazilian retail gasoline market.

The remaining two columns focus on the transnational analysis. When the model is trained with the Swedish data and tested on the Finnish data, the prediction rate drops from 83% to 79%. When the model is trained with the Finnish data and tested on the Swedish data, the prediction rate drops considerably more from 70% to 36%. Our finding is consistent with the results in the previous literature. Huber et al. (2022) report that prediction rates drop from 88–97% to 58–90% when moving from within-country analysis to transnational analysis.

Because the bids in our dataset are not in unit prices, some of the statistical screens might capture changes in project sizes instead of cartel behavior. To take this into account, in Panel B of Table 6, we report the performance of a model where we have excluded scale-dependent statistical screens from the analysis. These include the mean of bids,

[18] In Appendix A6 we report which statistical screens were picked by our model and discuss the importance of different statistical screens as predictors.

the absolute difference, and the standard deviation of the bids. In addition, we have excluded the number of bidders, which also potentially captures factors other than cartel behavior. In the within-country analysis, dropping the scale-dependent screens keeps the prediction rate for Finland nearly unchanged, whereas for Sweden the prediction rate drops significantly. In the transnational analysis, dropping the scale-dependent screens slightly decreases the prediction rate when the model is trained with the Swedish data and tested on the Finnish data. Inversely, the prediction rate increases when the model is trained with the Finnish data and tested on the Swedish data. The results suggest that using scale-dependent screens can be problematic if the market exhibits some type of structural changes other than the cartel, such as a trend in project sizes.

Table 6: Model performance metrics

Panel A: All screens				
	Within-country		Transnational	
	Finland	Sweden	Swe–Fin	Fin–Swe
Accuracy	83.40	69.77	78.91	35.87
Balanced accuracy	78.11	67.24	71.85	56.03
MSE	0.128	0.216	0.164	0.380
Panel B: Scale invariant screens				
	Within-country		Transnational	
	Finland	Sweden	Swe–Fin	Fin–Swe
Accuracy	82.12	59.64	74.98	53.24
Balanced accuracy	78.61	65.11	75.43	66.64
MSE	0.143	0.246	0.197	0.317

In this table we report performance measures of our main predictive model by country and type of analysis.

In Figure 5, we plot the yearly average collusion probability predicted by our model. In the Finnish data, there is a clear drop in collusion probability when the cartel ends. With the Swedish data, the collusion probability remains almost the same before and after the investigation. This provides further indication that the machine learning model predicts collusion poorly from the Swedish data. The fact that the model performs better in Finland is consistent with our earlier findings that the cartel had a more significant impact on the distribution of bids in Finland than in Sweden.

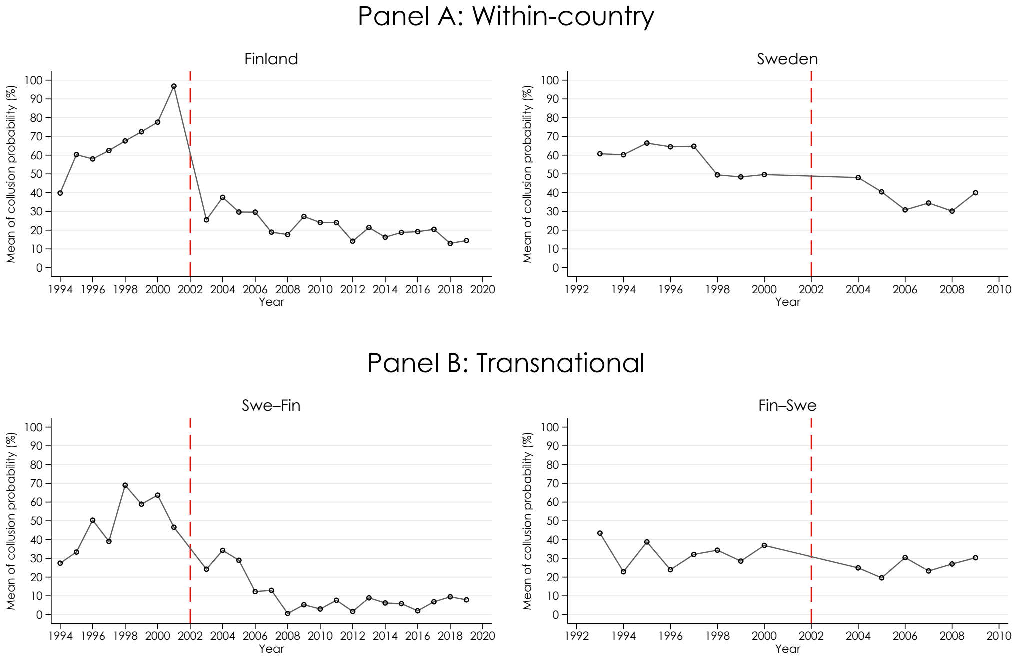


Figure 5: Average collusion probability by year

In this figure we have plotted the predicted average collusion probability of tenders by year. All screens are included.

7 Conclusions

A key challenge for competition authorities is to identify illegal agreements. Statistical methods that flag suspicious behavior could potentially help authorities to identify collusive agreements at a higher rate. In this paper, we studied the bidding behavior of two convicted cartels that operated in the Nordic asphalt paving markets. We began our analysis by estimating how the distribution of bids changed after the competition authorities launched their cartel investigations. We find that during the cartel, the variance of bids was lower both in Finland and Sweden, with a higher mass of bids clustered relatively close to the winner. Our second finding is that the cartels avoided leaving bids very close to the designated winner. Together, the clustering of bids and the isolated winning bid generated a bimodal, twin-peaked bid distribution during the cartels. Overall, we find that the distribution of bids was substantially more altered by the cartel in Finland. Our findings remain similar when we perform a difference-in-differences analysis using the Californian asphalt paving market as a control group.

We link some of our findings to the testimonial evidence provided in the court decisions. Consistent with the clustering of bids, both cartels aimed to leave complementary bids close to the winning bid to give an impression of intense competition. Overall, based on the testimonial evidence, the cartels operated somewhat similarly in the two countries. However, we do point out some differences between the two cartels, such as the fact that the Finnish cartel was run by one firm, while the Swedish cartel had four ringleaders. However, we are unable to relate these differences directly to the observed difference in bidding behavior between the two cartels.

After presenting evidence that the distribution of bids was altered by the cartels in both Finland and Sweden, we examine the performance of two cartel screening methods suggested in the previous literature. The first screening method, introduced by Clark et al. (2020), compares two distributions of bid differences. The first distribution contains the difference between a bid and the lowest rival bid. The second set of bid differences are defined similarly but excluding winning bids. Using a distributional regression approach, we find that the detection method correctly rejects competitive behavior for the cartel period in both Finland and Sweden. The method does not reject competitive bidding for the period after the investigations in Finland. For Sweden, the results of the method for the post-investigation period are inconclusive. The second detection method that we examine is a machine learning–based method introduced by Huber and Imhof (2019). This method predicts tenders as collusive or competitive by using predictors calculated from the distribution of bids within a tender. When the model is trained with data from the same country, the model correctly classifies 83% of the tenders in Finland and 70% in Sweden. In a transnational analysis, where the training data and test data are from different countries, the prediction rates are substantially lower.

Our results suggest that statistical cartel detection methods with modest data requirements can be useful for competition authorities in flagging suspicious behavior in public procurement. The two methods have different limitations. The distributional regression method cannot be used to detect collusion for individual projects but rather for a group of tenders. The machine learning–based method can be used to predict collusion for individual tenders but requires the user to calibrate the predictive model with existing data on known cartels. Based on our results, finding suitable training data can be difficult.

The two cartels studied in this paper operated in the same product market in two neighboring countries during the same time period. Still, we find that the bidding behavior of the cartels and the performance of the detection methods differ between the two countries. This suggests that details on how cartels operate can have a significant effect on the performance of different detection methods and that no single detection method is likely to be able to detect all kinds of conspiracies. Given that the details on how a potential cartel operates are unobservable by the competition authority screening for collusion, we suggest using several complementary detection methods instead of relying on the results of only one particular method.

References

Abrantes-Metz, Rosa, Luke Froeb, John Geweke, and Christopher T. Taylor (2006) "A variance screen for collusion," *International Journal of Industrial Organization*, 24 (3), 467–486.

Angrist, Joshua D. and Jörn-Steffen Pischke (2009) "Mostly Harmless Econometrics: An Empiricist's Companion," *Princenton University Press.*

Aryal, Gaurab and Maria F. Gabrielli (2013) "Testing for collusion in asymmetric first-price auctions," *International Journal of Industrial Organization*, 31 (1), 26–35.

Asker, John (2009) "A Study of the Internal Organization of a Bidding Cartel," *American Economic Review*, 100 (3), 724–762.

Bajari, Patrick, Stephanie Houghton, and Steven Tadelis (2014) "Bidding for Incomplete Contracts: An Empirical Analysis of Adaptation Costs," *American Economic Review*, 104 (4), 1228–1319.

Bajari, Patrick and Lixin Ye (2003) "Deciding Between Competition and Collusion," *Review of Economics and Statistics*, 85 (4), 971–989.

Baltrunaite, Audinga (2020) "Political Contributions and Public Procurement: Evidence from Lithuania," *Journal of the European Economic Association*, 18 (2), 541–582.

Baránek, Bruno, Leon Musolff, and Vítězslav Titl (2021) "Detection of Collusive Networks in E-procurement," *Working paper.*

Bergman, Mats A., Johan Lundberg, Sofia Lundberg, and Johan Y. Stake (2020) "Interactions Across Firms and Bid Rigging," *Review of Industrial Organization*, 56, 107–130.

Bolotova, Yuliya V. (2009) "Cartel overcharges: An empirical analysis," *Journal of Economic Behavior & Organization*, 70, 321–341.

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

Boyer, Marcel and Rachidi Kotchoni (2015) "How Much Do Cartel Overcharge?" *Review of Industrial Organization*, 47, 119–153.

Bryant, Peter G. and E. Woodrow Eckard (1991) "Price Fixing: The Probability of Getting Caught," *Review of Economics and Statistics*, 73 (3), 531–536.

Chassang, Sylvain, Kei Kawai, Jun Nakabayashi, and Juan Ortner (2022) "Robust Screens for Noncompetitive Bidding in Procurement Auctions," *Econometrica*, 90, 315–346.

Clark, Robert, Decio Coviello, and Adriano de Leverano (2020) "Complementary bidding and the collusive arrangement: Evidence from an antitrust investigation," *ZEW Discussion Papers*, 20-052.

Combe, Emmanuel, Constance Monnier, and Renaud Legal (2008) "Cartels: The probability of getting caught in the European union," *Bruges European Economic Research papers*, 12.

Conley, Timothy G. and Francesco Decarolis (2016) "Detecting Bidders Groups in Collusive Auctions," *American Economic Journal: Microeconomics*, 8 (2), 1–38.

Connor, John M. and Yuliya Bolotova (2006) "Cartel overcharges: Survey and meta-analysis," *International Journal of Industrial Organization*, 24, 1109–1137.

Coviello, Decio and Stefano Gagliarducci (2017) "Tenure in Office and Public Procurement," *American Economic Journal: Economic Policy*, 9 (3), 59–105.

Dong, Ailin, Massimo Massa, and Alminas Zaldokas (2019) "The effects of global leniency programs on margins and mergers," *RAND Journal of Economics*, 50 (4), 883–915.

European Commission (EC) (2009) "COMMISSION DECISION of 28.1.2009 relating to a proceeding under Article 81 of the Treaty and Article 53 of the EEA Agreement, Case COMP/39406 – Marine Hoses."

Feinstein, Jonathan S., Michael K. Block, and Frederick C. Nold (1985) "Asymmetric Information and Collusive Behavior in Auction Markets," *American Economic Review*, 75 (3), 441–460.

Ferraz, Claudio, Frederico Finan, and Dimitri Szerman (2015) "Procuring Firm Growth: The Effects of Government Purchases on Firm Dynamics," *Working paper*.

Finnish Competition and Consumer Authority (FCCA) (2004) "Esitys kilpailunrikkomismaksun määräämisestä eräille asfalttialan yrityksille kilpailunrajoituslaissa kielletyistä hinta- ja tarjousyhteistyöstä, markkinoiden jakamisesta sekä tietojenvaihdosta (in Finnish)."

——— (2011) "Esitys markkinaoikeudelle asiassa NCC Roads Oy / Destia Oy:n sekä Destia Kalusto Oy:n asfalttipäällysteliiketoiminta (in Finnish)."

Froeb, Luke M., E. Robert A. Koyak, and Gregory J. Werden (1993) "What is the effect of bid-rigging on prices?" *Economics Letters*, 42, 419–423.

Giuffrida, Leonardo M. and Gabriele Rovigatti (2018) "Can the private sector ensure the public interest? Evidence from federal procurement," *ZEW Discussion Papers*, 18-045.

Halonen, Kirsi-Maria and Janne Tukiainen (2020) "Competition and Litigation in Swedish Public Procurement," *Konkurrensverket uppdragsforskningsrapport*, 2020:1.

Harrington, Joseph E. and Yanhao Wei (2017) "What Can the Duration of Discovered Cartels Tell Us About the Duration of All Cartels?" *Economic Journal*, 127, 1977–2017.

Helsinki Court of Appeal (HCA) (2016) "Tiedote 20.10.2016 Helsingin Hovioikeuden tuomiot asfalttikartelli -asioissa (in Finnish)."

Hjalmarsson, Erik (2015) "Asfaltdjungeln. En ANT-analys av en svensk kartell (in Swedish)," *Stockholm School of Economics. Masters thesis.*

Huber, Martin and David Imhof (2019) "Machine learning with screens for detecting bid-rigging cartels," *International Journal of Industrial Organization*, 65, 277–301.

Huber, Martin, David Imhof, and Rieko Ishii (2022) "Transnational machine learning with screens for flagging bid-rigging cartels," *Journal of the Royal Statistical Society Series A*, 185 (3), 1074–1114.

Imhof, David (2020) "Detecting bid-rigging cartels with descriptive statistics," *Journal of Competition Law and Economics*, 15 (4), 427–467.

Imhof, David, Yavuz Karagök, and Samuel Rutz (2018) "Screening For Bid Rigging—Does It Work?" *Journal of Competition Law and Economics*, 14 (2), 235–261.

Jääskeläinen, Jan and Janne Tukiainen (2019) "Anatomy of Public Procurement," *VATT Working Papers*, 18.

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

Kawai, Kei, Jun Nakabayashi, and Daichi Shimamoto (2022) "A Study of Bid-rigging in Procurement Auctions: Evidence from Indonesia, Georgia, Mongolia, Malta, and State of California," *Working paper.*

LaCasse, Chantale (1995) "Bid Rigging and the Threat of Government Prosecution," *RAND Journal of Economics*, 26 (3), 398–417.

Lee, Munseob (2022) "Government Purchases and Firm Growth," *Working paper.*

Lindberg, Leena (2020) "Kaksikymmentä vuotta asfalttikartelliprosesseja (in Finnish)," *Defensor Legis*, 4, 632–649.

Lundberg, Johan (2017) "On cartel detection and Moran's I," *Letters in Spatial and Resource Sciences*, 10, 129–139.

McAfee, Preston and John McMillan (1992) "Bidding Rings," *American Economic Review*, 82 (3), 579–599.

OECD (1999) "Annual Report on Competition Policy Developments in Denmark."

——— (2002) "Competition law and policy in Norway. Norway's report to the OECD for the year 2002."

Porter, Robert and John Douglas Zona (1993) "Detection of Bid Rigging in Procurement Auctions," *Journal of Political Economy*, 101 (3), 518–538.

——— (1999) "Ohio School Milk Markets: An Analysis of Bidding," *RAND Journal of Economics*, 30 (2), 263–288.

Silveira, Douglas, Silvinha Vasconcelos, Marcelo Resende, and Daniel O. Cajueiro (2022) "Won't Get Fooled Again: A supervised machine learning approach for screening gasoline cartels," *Energy Economics*, 105, 1–18.

The Supreme Administrative Court of Finland (SACF) (2009) "Korkeimman hallinto-oikeuden päätös 29.9.2009 (in Finnish)."

The Swedish Market Court (SMC) (2009) "Dom, DNR A 2/07, (in Swedish)."

VATT Institute for Economic Research (2011) "Asfalttikartellin valtiolle aiheuttamien vahinkojen arviointi – Loppuraportti (in Finnish)."

Wachs, Johannes and János Kertész (2019) "A Network Approach to Cartel Detection in Public Auction Markets," *Scientific Reports*, 9, 1–10.

Appendix

A1. Alternative definitions of bid differences

We calculate bid differences by dividing the difference between a bid and the lowest rival bid by the lowest rival bid (see equation (1)). In this section, we discuss two alternative definitions used in the previous literature.

Clark et al. (2020) define bid differences in unit prices as follows:

$$\Delta_{i,t}^{1,unit} = b_{i,t}^{unit} - \Lambda b_{-i,t}^{unit} \tag{7}$$

where $b_{i,t}^{unit}$ refers to the per tonne price of bid i in tender t and $\Lambda b_{-i,t}^{unit}$ refers to the per tonne price of the smallest competing bid in the tender t. In our datasets, asphalt tonnes are available only for 189 tenders in Finland between 1994 and 2009. For the bids in these tenders, we calculate $\Delta_{i,t}^{1,unit}$ and plot them in Figure 6 before and after the cartel investigations. During the cartel period, we observe a similar twin-peaked distribution of bid differences using the unit price–based distribution as we do when using our main definition of bid differences.

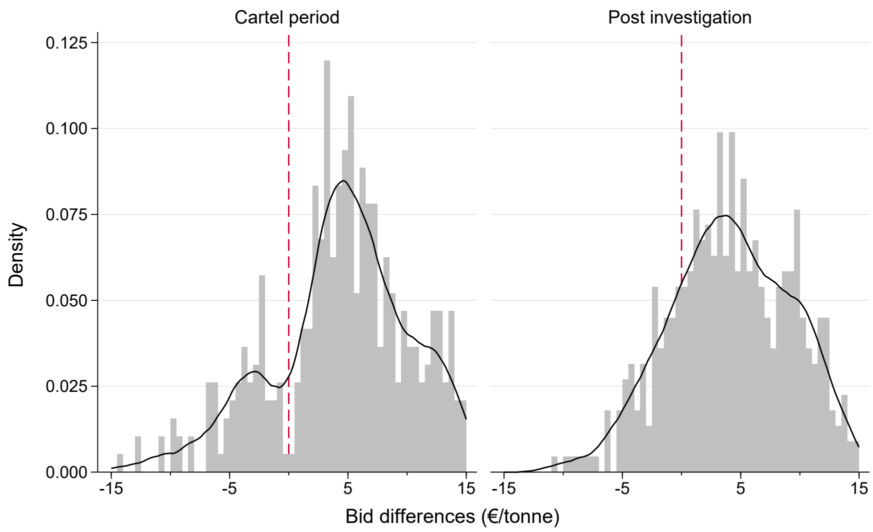


Figure 6: Distribution of unit price bid differences $\Delta_{i,t}^{1,unit}$

This figure plots the differences between a bid and the lowest rival bid in terms of unit prices before and after the cartel investigation. The width of the bins is 0.5. The curves correspond to density estimates calculated using an Epanechnikov kernel.

We have also replicated the Clark et al. (2020) cartel detection test using $\Delta_{i,t}^{1,unit}$. The results are shown in Figure 7. Again, we see several statistically significant coefficients during the cartel period. However, the magnitude and statistical significance are slightly lower than in our main analysis. Similarly to our main analysis, for the post-investigation period none of the estimated coefficients is statistically significantly different from zero.

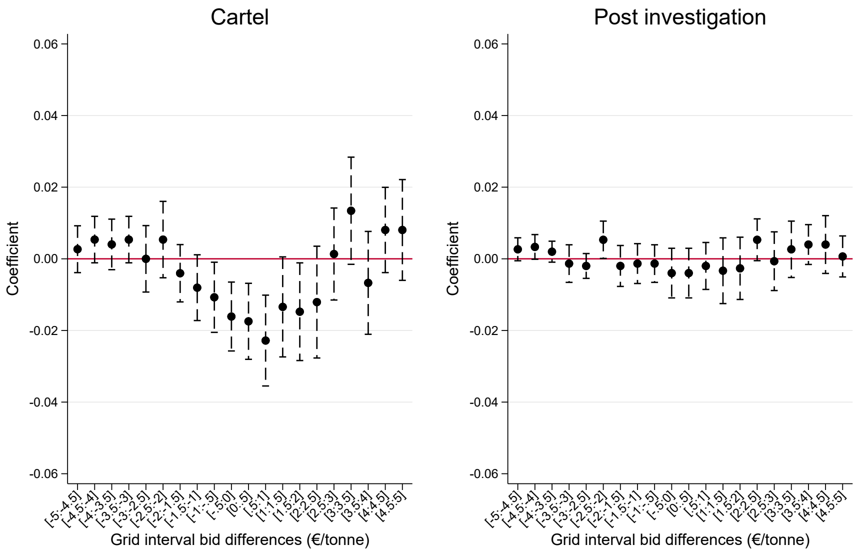


Figure 7: Results from Clark et al. (2020) cartel detection test when using $\Delta_{i,t}^{1,unit}$

This figure plots the point estimates and 95% confidence intervals of β_g from equation (4) separately for each country and time period. Confidence intervals are calculated with robust standard errors.

Chassang et al. (2022) define bid differences as follows:

$$\Delta_{i,t}^{1,r} = \frac{b_{i,t} - \Lambda b_{-i,t}}{r_t} \tag{8}$$

where r_t refers to the reserve price in tender t.

We cannot use this definition even for a subset of our data because the tenders covered in our data do not have a reserve price. However, we have tested whether using our definition of bid differences in the Japanese procurement data produces a similar distribution as using the original definition. In Figure 8, we plot both $\Delta_{i,t}^{1}$ and $\Delta_{i,t}^{1,r}$ in the Japanese data. Both distributions have a missing mass of bid differences around zero.

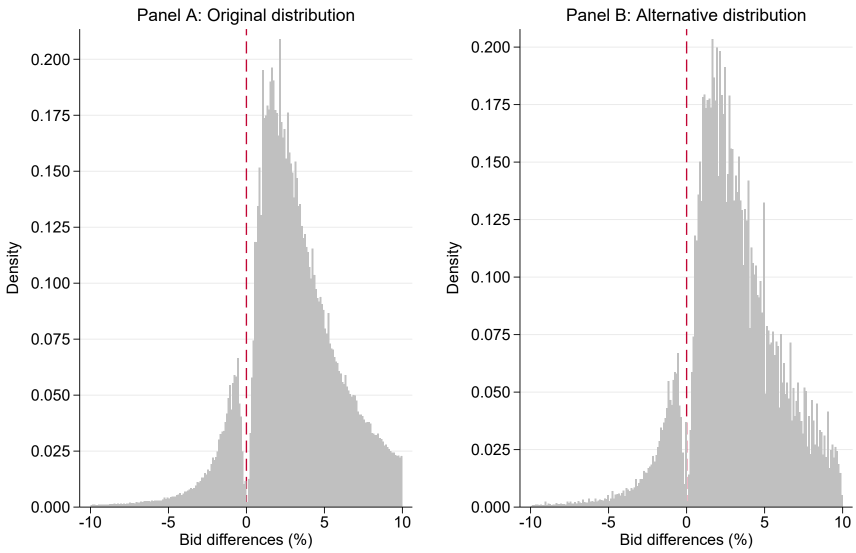


Figure 8: Distribution of bid differences $\Delta_{i,t}^{1,r}$ (left) and $\Delta_{i,t}^{1}$ (right) with Japanese procurement dataset

This figure plots the distribution of bid differences for Japanese national procurement data used by Chassang et al. (2022). On the left-hand side, we plot the distribution of bid differences divided by the reserve price (Chassang et al., 2022, replicated Figure 1(b)). On the right-hand side, we plot the distribution of bid differences divided by the lowest rival bid.

A2. Distributional regressions with 1% intervals

In Section 5, we analyze how the distribution of bid differences $\Delta_{i,t}^{1}$ changes in Finland and Sweden after the launch of cartel investigations. In our main analysis, we run distributional regressions with three intervals: within 1%, between 1% and 10%, and over 10%. As a robustness check in this section, we provide results when using 1% intervals between -20% and 20%.

The results are shown in Figure 9. Panel A shows the pre-post analysis conducted in Section 5.1 and Panel B shows the difference-in-differences analysis conducted in Section 5.2. For Finland, both analyses indicate that the share of bids between 2% and 10 % of the winner decreases after the investigation. In contrast, the share of bid differences over 10% increases after the investigation. For Sweden, the signs of the point estimates are similar but the magnitude and significance are considerably lower.

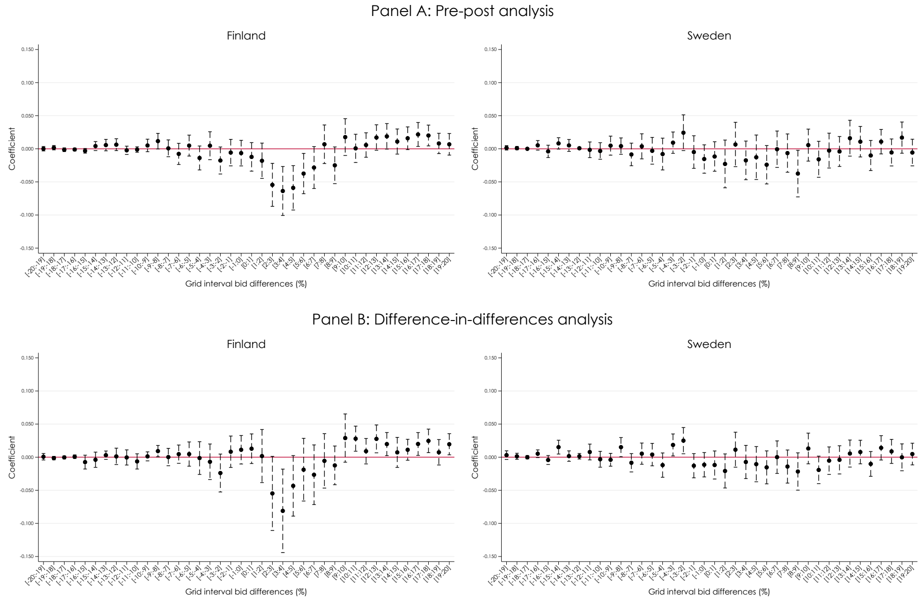


Figure 9: Distributional regressions using 1% intervals

This figure plots the results of the distributional regressions with 1% intervals between -20 % and 20 %. Panel A plots the point estimate of $\beta_{1,g}$ from equation (2) in the pre-post analysis. Panel B plots $\beta_{4,g}$ from equation (3) in the difference-in-differences analysis.

A3. Robustness checks for distributional regressions

In Section 5.1, we estimate the change in the bid distribution after the cartel investigations in Finland and Sweden. In this section, we provide results from two robustness checks. First, we add project size to the regression equation (2) as a control variable. We have decided to use the paving area as a measure of project size because it is available for both countries.[19] For Finland, the contract area is available for 103 tenders (32%) in the post-investigation period and for all tenders in the cartel period. For Sweden, the area is reported for 335 tenders (82%) in the cartel period and for 105 tenders (74%) in the post-investigation period. Both in Finland and Sweden, the average contract area is around twice larger in the post-investigation period. The results are shown in Table 7. In this specification, the estimated change in the bid distribution after the cartel investigations is larger than in our main specification. Also, the statistical significance increases compared to our main specification.

[19] Asphalt tonnes is only available for a small subset of Finnish tenders.

Table 7: Distributional effect of the cartel investigations when including project size controls

Panel A: Finland

	Within 1%	Between 1% and 10%	More than 10%
Post	-0.019	-0.312***	0.331***
	(0.026)	(0.068)	(0.064)
Observations	1297	1297	1297
Region FE	Yes	Yes	Yes
Bitumen index	Yes	Yes	Yes

Panel B: Sweden

	Within 1%	Between 1% and 10%	More than 10%
Post	-0.025	-0.133**	0.158***
	(0.026)	(0.058)	(0.059)
Observations	2190	2190	2190
Region FE	Yes	Yes	Yes
Bitumen index	Yes	Yes	Yes

The dependent variable is the probability that bid differences fall in a given interval. Post is a dummy equal to 1 if the contract was awarded after the investigation. Panel A shows results for Finland and Panel B for Sweden. Standard errors are clustered at the tender level. Significance at $p < 0.10$ (*), $p < 0.05$ (**), and $p < 0.01$ (***).

In our main specification, we exclude the tenders that were organized during the investigation years or for Sweden, the years 2002 and 2003. In our second robustness check, we run the distributional regressions with these tenders included in the sample. We classify these tenders competitive. The results are shown in Table 8. The results, particularly for Finland, remain largely unchanged compared to our main specification.

Table 8: Distributional effect of the cartel investigations when including all years

Panel A: Finland

	Within 1%	Between 1% and 10%	More than 10%
Post	-0.022	-0.288***	0.310***
	(0.019)	(0.045)	(0.043)
Observations	2345	2345	2345
Region FE	Yes	Yes	Yes
Bitumen index	Yes	Yes	Yes

Panel B: Sweden

	Within 1%	Between 1% and 10%	More than 10%
Post	-0.013	-0.069*	0.082**
	(0.014)	(0.037)	(0.037)
Observations	3799	3799	3799
Region FE	Yes	Yes	Yes
Bitumen index	Yes	Yes	Yes

The dependent variable is the probability that bid differences fall in a given interval. Post is a dummy equal to 1 if the contract was awarded after the investigation. Panel A shows results for Finland and Panel B for Sweden. Standard errors are clustered at the tender level. Significance at $p < 0.10$ (*), $p < 0.05$ (**), and $p < 0.01$ (***).

A4. Pre-trends in difference-in-differences analysis

In Section 5.2, we perform a difference-in-differences analysis where we compare the development of bid distributions in Finland and Sweden before and after the cartel investigations with the development of asphalt paving procurement auctions in California. The difference-in-differences analysis relies on the parallel trend assumption, which requires that the outcomes for the control and treatment groups would develop similarly in the absence of treatment. Typically, the parallel trends assumption is evaluated by examining whether the treatment and control groups exhibit a similar trend in the pre-treatment period.

A commonly used strategy to assess whether the control and treatment groups followed a similar trend is to estimate an event-study specification which includes interaction terms between a treatment group dummy and a time variable. We only observe two common pre-investigation years for both the Californian asphalt market and the Finnish and Swedish asphalt markets.[20] The event-study specification also requires to omit one period before the treatment, leaving us with only one pre-treatment period. Because of this, we only conduct a descriptive analysis of the pre-treatment trends. In Figure 10, we plot the share of bids at different intervals before and after the cartel investigations in the treatment and control groups. Based on a visual inspection, it seems that the share of bid differences in different intervals is fairly stable over time, with no clear trends in any of the three countries. However, we do observe that there is a clear level change in the share of bids at two intervals (1–10% and over 10%) for Finland after the cartel investigations. The absence of trends suggests that it is a reasonable assumption that without the cartel investigations in Finland and Sweden, the share of bid differences in different intervals would have remained close to their pre-investigation levels.

[20] Note that in the Californian data year 2001 is omitted.

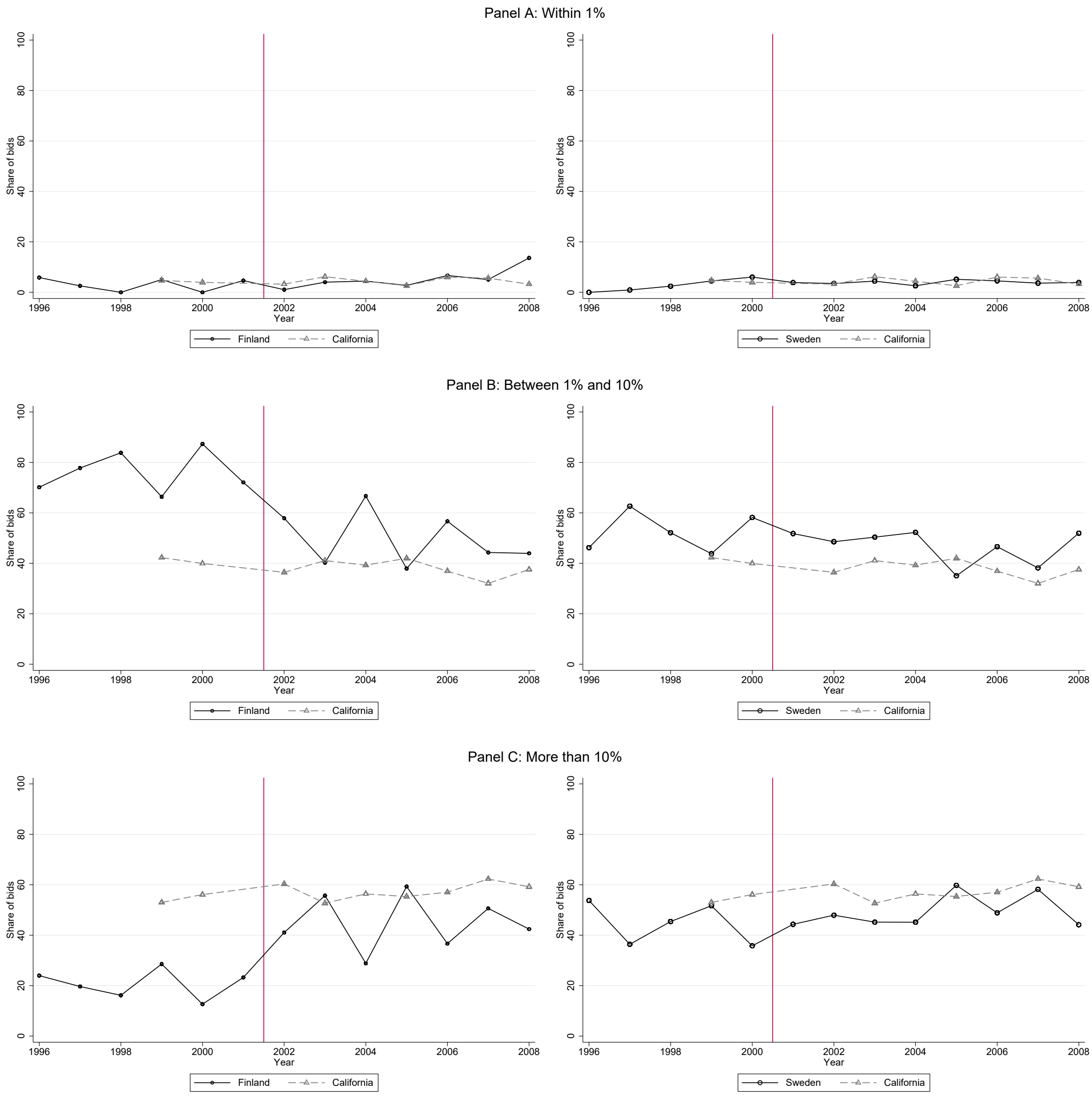


Figure 10: Pre-treatment trends

This figure plots the share of bids in different intervals over time. Panel A focuses on intervals 0–1%, Panel B on intervals 1–10%, and Panel C on intervals >10%. The gray line depicts California. The black line depicts Finland in the graphs on the left-hand side and Sweden on the right-hand side. Note that the years when the investigations started are omitted in our main analysis.

As an alternative check on the difference-in-differences analysis, we have added treatment- and control-group-specific time trends to the list of controls. Specifically, we estimate the following:

$$y_{i,t,g} = \alpha_g + \beta_{2,g}\, post_t + \beta_{3,g}\, treat_t + \beta_{4,g}\, treat_t \times post_t + \beta_{5,g} t + \epsilon_{i,t,g} \tag{9}$$

where everything is as in equation (3) and $\beta_{5,g}$ is a group-specific coefficient for a linear time trend t. As noted by Angrist and Pischke (2009), this allows the treatment and control units to follow different trends in a limited but potentially revealing way.

The results are shown in Table 9. In both markets, adding the group-specific time trends increases both the magnitude and statistical significance of our key explanatory variable. We have also estimated a version of equation (3) where we replace the $post_t$ term with year fixed-effects. $\beta_{4,g}$ remains largely unchanged in this alternative specification. Finally, we have also estimated equation (3) using all years in the estimation sample. In this specification, the point estimates and statistical significance are smaller. However, for Finland, $\beta_{4,g}$ remains statistically significant for the intervals between 1–10% and over 10%.[21] Overall, given that the share of bids in different intervals does not exhibit a general time trend in any of the three markets and that our results are robust to including group-specific time trends, we believe that the parallel trends assumption is satisfied in our setting.

[21] The results from these two alternative specifications are available upon request.

Table 9: Distributional effect of the cartel investigations using difference-in-differences with group-specific time-trends

Panel A: Finland

	Within 1%	Between 1% and 10%	More than 10%
Post × Treat	-0.058	-0.347***	0.405***
	(0.053)	(0.123)	(0.122)
Observations	6924	6924	6924

Panel B: Sweden

	Within 1%	Between 1% and 10%	More than 10%
Post × Treat	-0.045	-0.197*	0.243*
	(0.050)	(0.120)	(0.124)
Observations	7071	7071	7071

The dependent variable is the probability that bid differences fall in a given interval. Post × Treat is equal to 1 for Finland and Sweden after the launch of cartel investigation. Panel A shows results for Finland and Panel B for Sweden. Standard errors are clustered at the tender level. Significance at $p < 0.10$ (*), $p < 0.05$ (**), and $p < 0.01$ (***).

A5. Statistical screens of the machine learning model

In Section 6.2, we use 10 statistical screens calculated from the distribution of bids (and their second powers and interactions) as predictors in the predictive model. The set of statistical screens is chosen following Huber and Imhof (2019) and Imhof (2020). The statistical screens aim to capture the clustering of bids and the manipulated difference between the winning bid and the losing bids. In this section, we discuss each of the 10 statistical screens in more detail.

As discussed in Section 3, bid rigging may affect the dispersion of bids. Coefficient of variation and standard deviation capture this directly by measuring the variation of bids within a tender. Kurtosis, which measures the tailedness of a distribution, is also used to detect changes in the dispersion of bids. Bid rigging may also affect the difference between the lowest and second lowest bid. Absolute difference and percentage difference capture this, with the former measuring the difference between the winner and the runner-up in monetary terms and the latter in percentages. Skewness, which measures the symmetry of the bid distribution, is also used to capture the isolation of the winning bid. Bid rigging might simultaneously affect the difference between the lowest and second lowest bid as well as the difference among losing bids. To capture this, relative distance and normalized distance are used. Relative distance is calculated by dividing the difference between the lowest and second lowest bids by the standard deviation of the losing bids. In normalized distance, the difference between the lowest and second lowest bid is divided by the average distance between all adjacent bids. Finally, the number of bids and contract value are used to control for different procurement types. The formulas to calculate each of the 10 screens are presented in Table 4.

A6. Predictors of the machine learning model

In Section 6.2, we use machine learning to predict collusion in the Finnish and Swedish asphalt markets. One of the main benefits of using machine learning is that it chooses the best predictors from a large set of possible predictors. In this section, we discuss the relative importance of different predictors in predicting collusion. These results could be of interest to researchers and competition authorities who plan to apply the machine learning–based screening method in other settings.

In our within-country analysis, we run the predictive model for 100 iterations and take averages over these iterations. Since each iteration has a different training sample, the chosen predictors also vary over iterations. To measure the importance of different predictors, we take an average of the absolute values of the predictor coefficients. The screens are normalized to zero mean and unit variance, so the magnitude of the coefficients is an estimate of the predictors' importance.

In Table 10 we report five predictors with the highest average value of the absolute coefficient for the different specifications. In Panel A in all four columns, the mean of bids is one of the important predictors. Other important predictors are the number of bids for both Finland and Sweden, the normalized distance and relative distance for Finland, and the standard deviation, coefficient of variation, kurtosis, and absolute difference for Sweden. In Panel B we report the most important predictors when we exclude the scale-invariant predictors and number of bids. When these screens are excluded, several screens that focus on the difference between bids (relative distance, percentage difference, normalized distance) are listed as the most important predictors.

Table 10: The most important predictors

Panel A: All screens

Finland		Sweden		Fin–Swe		Swe–Fin	
Screen	Abs value	Screen	Abs value	Screen	Abs value	Screen	Abs value
mean	3.59	std × nbids	3.62	mean	3.71	std × nbids	3.48
normd × mean	3.51	std	1.62	normd × mean	3.60	std	1.75
mean^2	2.86	cv × mean	1.62	mean^2	3.00	cv × mean	1.75
nbids × mean	2.65	mean	1.28	nbids × mean	2.90	kurt × diff_abs	1.27
rd^2	2.03	diff_abs^2	1.25	rd^2	2.34	mean	1.15

Panel B: Scale invariant screens

Finland		Sweden		Fin–Swe		Swe–Fin	
Screen	Abs value	Screen	Abs value	Screen	Abs value	Screen	Abs value
rd^2	6.55	diff_perc	1.83	rd^2	9.35	diff_perc	1.91
diff_perc	2.35	skew × normd	1.38	normd	2.23	cv × normd	1.39
normd	2.10	cv × normd	1.37	diff_perc	2.17	skew × normd	1.33
kurt × normd	1.98	skew × rd	1.36	cv × rd	1.85	skew	1.08
diff_perc^2	1.57	skew	1.03	cv × diff_perc	1.64	skew × rd	1.08

This table reports the five predictors with the highest average absolute value of the model coefficients. ”mean” refers to the mean of bids in a tender, ”normd” to the normalized distance, ”nbids” to the number of bids, ”rd” to the relative distance, ”std” to the standard deviation, ”cv” to the coefficient of variation, ”diff_abs” to the absolute difference, ”diff_perc” to the percentage difference, ”kurt” to the kurtosis, and ”skew” to the skewness. For formulas of the screens see Table 4 and for a discussion of individual screens see Section A5.

A7. Results of alternative machine learning models

In Section 6.2 we use lasso logit (regularized multinomial) to predict collusion. In this section, we present results for six alternative machine learning models. Nearest neighbor (K-neighbors) chooses the classification based on the labels of K nearest neighbors in the feature space. Naive Bayes bases the classification on the posterior probability of the

Bayes theorem when assuming conditional independence between features. Decision tree classifies by making consecutive binary decisions based on feature values, thus forming a decision tree. Random forest is an ensemble method that takes the average of multiple deep decision trees with bootstrapping. AdaBoost (adaptive boosting) is another ensemble learning method. It uses subsequent simple prediction models (usually small decision trees), reweighs the sample between the predictions to put more emphasis on the harder-to-predict instances, and finally takes the weighted average of the predictions. Neural network uses a layered structure with a large number of activation functions that convert features into signals that determine classification. All models are implemented using the Scikit-learn package in Python. All statistical screens, including the scale-dependent screens, are included in the models.

In Table 11, we show the mean squared error for all alternative machine learning models. For the Finnish dataset in within-country analysis, the smallest mean squared error is achieved with the Regularized multinomial (i.e., the lasso logit) and the Random forest models. For Sweden, the models that achieve the lowest MSE are the Neural network and the Random forest. For transnational analysis, the best performing models are the Regularized multinomial when Sweden is used for training the models and Finland for testing, and AdaBoost when Finland is used for training the models and Sweden for testing.

Table 11: Mean squared error for alternative machine learning models

	Within-country		Transnational	
	Finland	Sweden	Swe–Fin	Fin–Swe
Regularized multinomial	0.128	0.216	0.164	0.380
Nearest neighbor	0.138	0.149	0.306	0.419
Naive Bayes	0.225	0.266	0.306	0.676
Decision tree	0.164	0.160	0.287	0.447
Random forest	0.127	0.148	0.251	0.374
AdaBoost	0.198	0.177	0.261	0.271
Neural network	0.132	0.139	0.228	0.434

In this table we report the mean squared error (MSE) of alternative predictive models by country and type of analysis.