Aaltio, Aapo; Buri, Riku; Jokelainen, Antto; Lundberg, Johan

Complementary bidding and cartel detection : Evidence from Nordic asphalt markets

Published in:
International Journal of Industrial Organization

DOI:
10.1016/j.ijindorg.2024.103129

Published: 01/01/2025

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Aaltio, A., Buri, R., Jokelainen, A., & Lundberg, J. (2025). Complementary bidding and cartel detection : Evidence from Nordic asphalt markets. *International Journal of Industrial Organization*, *98*, Article 103129. https://doi.org/10.1016/j.ijindorg.2024.103129

International Journal of Industrial Organization 98 (2025) 103129

Contents lists available at ScienceDirect

International Journal of Industrial Organization

journal homepage: www.elsevier.com/locate/ijio

Complementary bidding and cartel detection: Evidence from Nordic asphalt markets ☆

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

[a] *Finnish Competition and Consumer Authority, Finland*
[b] *Aalto University, Finland*
[c] *Umeå University, Sweden*

ARTICLE INFO

JEL classification:
L22
L74
D44
H57

Keywords:
Procurement
Bidding ring
Collusion
Antitrust
Complementary bidding
Detection

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. (forthcoming) correctly rejects competitive bidding for the cartel period in both markets. The method suggested by Huber and Imhof (2019) achieves a high prediction rate in one of the markets but not in the market where the cartel had a more modest impact on bid distribution. Our results suggest that statistical screening methods with low data requirements can be useful for competition authorities in detecting collusive agreements.

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

☆ We thank the co-editor Matthew L. Gentry, two anonymous referees, Adriano De Leverano, Richard Friberg, Ari Hyytinen, Olli Kauppi, Matias Pietola, Teemu Pekkarinen, Otto Toivanen, the participants of Helsinki GSE Industrial Organization PhD seminar and the participants of EPP 2023 conference for their comments and suggestions. We also thank David Imhof and Martin Huber for assisting with Swiss and Japanese datasets.

* Corresponding author.
E-mail addresses: aapo.aaltio@kkv.fi (A. Aaltio), riku.buri@kkv.fi (R. Buri), antto.jokelainen@aalto.fi (A. Jokelainen), johan.lundberg@umu.se (J. Lundberg).

https://doi.org/10.1016/j.ijindorg.2024.103129
Received 7 March 2023; Received in revised form 14 October 2024; Accepted 13 December 2024

Available online 24 December 2024

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., forthcoming; Chassang et al., 2022), and clustering of losing bids (Lundberg, 2017). Although detailed procurement datasets are still relatively rare, large, economy-wide procurement datasets that contain information on the submitted bids are becoming increasingly available for authorities and researchers.[1] If cartel detection methods based solely on bid distribution data prove to be reliable and accurate, authorities could adopt them for large-scale cartel detection efforts.

In this paper, we study the bidding behavior of two convicted cartels that operated in the Finnish and Swedish asphalt markets during the 1990s and early 2000s. 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 1% 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. Both of these features, the clustered bids and the missing mass of nearly tied bids, have been proposed as markers of collusion in previous literature (Chassang et al., 2022; Clark et al., forthcoming; Imhof et al., 2018). After the start of cartel investigations, both in Finland and Sweden, the distribution of bids becomes unimodal and the share of bids within 10% of the winning bid decreases. The change is considerably larger in Finland. To support the causal interpretation of our results, we conduct several robustness checks and also conduct a difference-in-differences analysis using data from a control market.

The first detection method that we test is a distributional regression approach suggested by Clark et al. (forthcoming). 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. The null hypothesis is that, with competitive bidding, the two distributions should be similar for bid differences 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 close to zero, indicating that the cartel firms avoided leaving bids very close to the winning bid. The null hypothesis is not rejected for the post-investigation period for either of the countries. However, the results for the post-investigation period in Sweden are not as conclusive as they are for Finland.

Finally, the 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. When the predictive model is trained using the data from same country, the model correctly classifies around 90% of the tenders with the Finnish dataset and 74% with the Swedish dataset. We also test the performance of the model when the model is trained with data from cartel operating in another country. For this analysis, we also use data from procurement cartels operating in Switzerland and Japan. When the model is trained using data from another country, the prediction rates decrease for both Finland and Sweden. We also find substantial variation in the prediction rates depending on the model specification and the data used to train the model.

Our results indicate that cartels can result in a 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. We find in particular that the detection methods work well for Finland, where the cartel had a greater impact on the distribution of bids. We propose two potential explanations for the disparity in results between the Finnish and Swedish cartels. First, the Swedish cartel's bidding behavior more closely resembled competitive bidding, leading to a less noticeable difference in bid distributions between competitive and collusive periods. Second, the Swedish cartel did not rig bids in every tender during the cartel period, causing some competitive tenders to be misclassified as collusive. Unlike the Finnish courts, the Swedish Market Court determined that collusion occurred only in specific regions and tenders. Although we are unable to pinpoint the exact tenders considered in the court's decision, we perform multiple robustness checks using different conservative subsamples and find that our results hold across all. However, we cannot entirely dismiss the possibility that the smaller scope of the Swedish cartel contributes to the observed differences between Finland and Sweden.

Both detection methods studied in this paper do have some caveats. The method by Clark et al. (forthcoming) 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. Based on our results, the prediction rate can vary greatly depending on the training data. A second difference between the methods is that the test by Clark et al. (forthcoming) places more emphasis on the difference between the winning and losing bids, while the test

[1] For a description of such datasets see Jääskeläinen and Tukiainen (2019) for Finland, Halonen and Tukiainen (2020) for Sweden, Giuffrida and Rovigatti (2022) 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.

by Huber and Imhof (2019) also uses other features of the bid distribution. One benefit of the more narrow focus is that when a group of firms repeatedly participate in sealed bid first-price auctions, a consistent gap between the winning and losing bids is inconsistent with competitive behavior (see Chassang et al., 2022), while other characteristics, such as low variance of bids, could also result from other features of the market unrelated to collusive behavior. However, an important caveat of the distributional regression test, as pointed out by De Leverano (2023), is that, in settings with complete information, it can flag collusion in a competitive market.

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 information on project costs and simply rely on the distribution of bids. Chassang et al. (2022) document that in Japanese procurement auctions, winning bids tend to be isolated and losing bids are clustered relatively closely to the winner, leading to a bimodal bid distribution. 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. (forthcoming) present similar empirical evidence from a Canadian procurement cartel that also involved both clustering of bids and isolated winning bids. 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 the observation that collusion can alter the distribution of bids, two easily implementable methods have been developed in the literature to detect cartels.[2] Clark et al. (forthcoming) develop a detection method based on a distributional regression, while Huber and Imhof (2019) develop a machine learning model based on statistics calculated from the bid distribution.[3]

Our contribution to the literature is twofold. First, we provide more evidence on the prevalence of isolated winning bids and clustering of losing bids in the context of two large Nordic procurement cartels. Second, we test the performance of two recently developed detection methods. The Finnish cartel has not been previously studied in the cartel detection literature. The Swedish cartel has been studied in previous literature, although with a different focus. Bergman et al. (2020) study the Swedish cartel, and using a spatial econometric model, find a significant positive correlation between complementary cartel bids during the cartel period and no correlation in the period after the cartel. Similarly, Lundberg (2017) illustrates that Moran's I statistic can be used to detect complementary bidding during the cartel period in the Swedish asphalt market. Compared to previous papers using the Swedish data, our paper focuses on showing the existence of isolated winning bids and clustered losing bids, and utilizes detection methods that require only information on the distribution of bids.

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, a cartel operated 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. In both Finland and Sweden, all the largest asphalt paving firms in the market were prosecuted and found guilty of collusion. 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).

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.[4] The employees decided to disclose the cartel to the Competition Authority to escape personal liability. Their new employer applied for, and was later granted, immunity from fines for its involvement in the cartel. Shortly thereafter, the Competition Authority conducted dawn raids, and legal proceedings against eight firms were initiated in March 2003. In its summons

[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.

[3] For follow-up papers on the machine learning model see García Rodríguez et al. (2022) who test the performance of the model in six procurement cartels operating in five different countries, Wallimann et al. (2022) who further develop the model to detect incomplete bidding cartels, Huber et al. (2022) who test the performance of the model when the model is trained in one country and used for prediction in a cartel operating in a different country, and Huber and Imhof (2023) who we propose a detection approach based on deep learning.

[4] 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).

application to the court, the Swedish Competition Authority claimed that the largest firms in the market operated a cartel from 1993 to 2001. Regarding the other firms found guilty, the Competition Authority determined that their involvement in the cartel was limited to specific regions and certain years within the infringement period.[5]

In 2007, the Stockholm District Court ordered seven firms to pay a total of around 50 million euros in fines.[6] Later in 2009 the Swedish Market Court annulled the charges against one of the convicted firms, while increasing the fines for NCC to around 20 million euros. The Market Court found that between 1997 and 2001, the cartel rigged contracts tendered by the Swedish Road Administration in five different regions, while also operating in tenders organized by municipalities. With regard to the largest firms, the court states that the infringement can be seen as a single continuous violation, while for the other firms, the Market Court's decision singles out specific tenders where the firms were found to have participated in the cartel. According to the case handler from the Swedish Competition Authority, the court's conclusion regarding the more limited scope of the cartel can be attributed to the fact that whistleblowers provided information only about the regions where they had been employed, resulting in limited evidence from regions where they had not worked. Later in our analysis, we report results separately for the entire Swedish market and only using firms and areas explicitly mentioned in the court decision. In our main analysis, we use the years 1993-2001 as the cartel period for Sweden, but we also report results focusing on the years specifically mentioned in the Swedish Market Court's decision. In Appendix Section A.8, we provide a more detailed discussion of the legal proceedings of the Swedish case.

According to documentation, three of the four largest Swedish asphalt paving firms started operating the cartel in 1993.[7] In 1995 the fourth big player, an in-house production unit of the Swedish Road Administration, joined the cartel. The four firms met every year to allocate the coming year's contracts and to coordinate their bids. At these meetings, the companies agreed on how they would divide contracts between them and also exchanged information on prices and volumes. Companies not included in the group, and thus potential competitors, were compensated to refrain from bidding or to place such a high bid that they would not win the contract. The losing firms were compensated for not bidding or for bidding high through the use of subcontracting, free services or even direct monetary transfers, so called "p-money". 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 way 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.[8] 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. Unlike the Swedish cartel, the Finnish cartel was found to have operated nationally.

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. 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, 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 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

[5] Note that the severity of cartels is largely determined by the temporal and geographic dimensions of one overall anticompetitive plan. In court proceedings, a crucial question is often to which extent, e.g., proven bid rigging events form one overall infringement and to which extent they should be seen to constitute several smaller ones. To argue for a more prolonged cartel duration, a competition authority is not liable to present proof relating directly to all tenders during the claimed cartel period. The claims of a summons application can therefore be built upon only the procurements with the best evidentiary situation.

[6] In total nine firms were fined, but three of the convicted firms were subsidiaries of the same firm.

[7] Our description of the Swedish cartel is based on the Swedish Market Court's decision (SMC, 2009), and on interviews of Anders Gerde, a case handler at SCA, in Hjalmarsson (2015) and in Kapitalet podcast (2019).

[8] 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).

competitive. Moreover, witness statements from the Swedish cartel revealed that the designated losing bidders were reluctant to submit excessively high bids. Such inflated prices could send negative signals about the firm's competitiveness and pose significant risks if the cartel were to collapse.[9]

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.[10] 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.

Both features discussed above, the clustering of bids and missing bids close to the winner, have also been documented in previous empirical literature. Chassang et al. (2022) find evidence of missing mass of bids close to the winner and bid clustering in suspicious markets in Japan and Clark et al. (forthcoming) provide evidence of similar bidding behavior from a Canadian cartel. Imhof et al. (2018) provides evidence of similar bidding behavior by cartels operating in Switzerland. In summary, based on theory, testimonial evidence, and previous empirical literature, 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 dataset was collected by the Finnish Transport Infrastructure Agency. For years 2010–2019, we have supplemented the original dataset with information from public procurement documents. 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 Finnish data from 1994 to 2009, 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.[11] 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.[12] We also drop tenders where there is more than one winner, information on all bids is not available, the lowest bid did not win, only one bid was submitted, or bids have evident typing errors. In total, these exclude 181 tenders. After cleaning the data, our dataset has information on 4983 bids on 1008 tenders with a total awarded value of 1.9 billion euros.

Table 1 presents summary statistics for both Finland and Sweden before and after the cartel investigations. First, focusing on Finland, we observe 2250 bids on 457 tenders. The average contract value before the cartel investigation was 1.07 million euros and 4.02 million euros after. In the Appendix, we report the development of the project size over time (see Appendix A.6) and also use a model to estimate the cartel overcharge, which based on the results was around 20% (Appendix A.7).[13] Before the start of the investigations, we observe bids from a total of 27 firms and after the investigations from 19 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.[14] 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

9 For a discussion of this see page 33 in Hjalmarsson (2015).

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

11 The bitumen prices are collected from the annual statistics published by Statistics Finland.

12 In Appendix A.2.2 we report some of the results with all years included.

13 Our results are in line with the results of VATT Institute of Economic Research, who estimated that the cartel overcharge in the Finnish asphalt market case was around 15%-25% (VATT, 2011). 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).

14 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. A second potential explanation for the higher number of bidders during the cartel period is that smaller firms might have found it profitable to participate in the market only because the cartel raised prices or because they might have received side payments from participating. In the absence of cartel, these smaller firms might not have found it profitable to compete for the contracts.

Table 1
Summary statistics.

	Finland		Sweden	
	Cartel (1994-2001)	Comp. (2003-2019)	Cartel (1993-2000)	Comp. (2004-2009)
Total awarded (MEUR)	146.4	1286.8	297.4	161.4
Nbr of contracts	137	320	410	141
Avg contract size (MEUR)	1.07	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
Market share of convicted firms	79.5	65.7	77.6	67.2

The total value and the average value of awarded contracts are measured from the price of the winning bid.

period. The increase in the average contract value can be explained by the fact that the size of the projects has increased over time. For observations where paved area is available, the project size is around twice larger in the post-investigation period.[15] 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.

5. Change in the bid distribution

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, therefore, comparable across different-sized projects.

A similar measure has previously been used by Chassang et al. (2022) and Clark et al. (forthcoming). 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. (forthcoming) calculate the bid differences in terms of unit prices. In Appendix A.1, we provide results based on the definition used by Clark et al. (forthcoming). 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 Fig. 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. (forthcoming) and Chassang et al. (2022). Both of these studies find a similar twin-peaked distribution of bid differences from collusive markets.

To test more formally how the distribution of bids changed after the launch of the investigation, we follow Clark et al. (forthcoming) and use a distributional regression approach originally developed by Chernozhukov et al. (2013) and Fortin et al. (2021). 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 estimate the models separately for the Finnish and Swedish datasets using the following linear probability models estimated with OLS:

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

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

[15] For more details, see Appendix A.6.

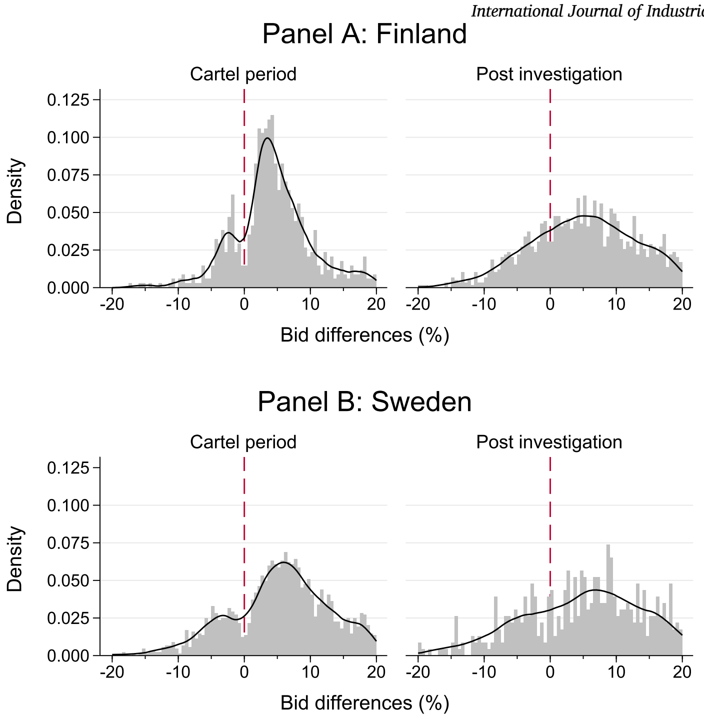

Fig. 1. Distribution of bid differences $\Delta^1_{i,t}$ by country and period.
This figure plots the distribution of differences between a bid and the lowest rival bid for asphalt procurement contracts in Sweden and Finland before and after cartel investigations. The width of the bins is 0.5. The curves correspond to density estimates calculated using an Epanechnikov kernel.

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.[16]

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%. In Appendix A.2.1, 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.

We perform several additional robustness checks to assess the sensitivity of our findings. The results of these robustness checks are presented in Appendix A.2.2. First, we estimate a model in which 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 of the point-estimates for intervals between 1% and 10% and more than 10%. 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. Third, we estimate the change in the bid distribution separately for the convicted areas in Sweden. The absolute values of the point-estimates increase but still remain smaller than in Finland. Fourth, we estimate the change in the bid distribution focusing on the convicted regions but also limiting the cartel period to years 1997–2001. The results remain similar to when we only use tenders from the convicted regions. Fifth, we estimate the model using tenders in the cartel period where only convicted firms left bids. This increases the absolute value of the point-estimates for Finland and has no notable effect on the point-estimates obtained using data from Sweden. Sixth,

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

Table 2
Distributional effect of the cartel investigations.

Panel A: Finland			
	Within 1%	Between 1% and 10%	More than 10%
Post	-0.018	-0.309***	0.326***
	(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 $p < 0.01$ (***).

we estimate equation (2) for one percent intervals between -20% and 20%. Similarly to our main specification, we see a significant change in the distribution of bids at some of the intervals.

A potential concern is that the results are 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 we use 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.[17] In previous literature, the market has been modeled as competitive and there are no 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. In the difference-in-differences analysis the point-estimates drop for both Finland and Sweden but the main conclusions from the analysis remain similar: in both countries we observe a lower mass of bids close between 1% and 10% of the winning bid, and the point-estimate is larger in Finland. The analysis and results are presented in more detail in Appendix A.2.2.

Overall, we find strong evidence that the cartel altered the distribution of bids in both Finland and Sweden. Interestingly, we find that the change is considerably greater in Finland. In Sections 2 and 3 we point out some differences between the two cartels that could explain these results. One potential key difference is that the Finnish cartel may have been operating more extensively than the Swedish cartel. As discussed above, using data from only the convicted areas in Sweden, we are able to rule out the possibility that a smaller geographical extent of the Swedish cartel alone drives the disparity in the results. The disparity also persists when we focus on the years and firms specified in the Swedish Market Court's decision. However, while we are able to run the test separately for the convicted regions and specific years and firms in Sweden, we are not able to precisely identify all the tenders deemed rigged by the court within our dataset. Therefore, we remain inconclusive as to whether differences in the extent of the cartel can at least partially explain the disparity in results between Finland and Sweden.

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. (forthcoming) and Huber and Imhof (2019). Although 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 results inform whether the two detection methods could have been used to expose the Nordic Asphalt cartels. Our results also inform how the performance of the detection methods differs when the cartel had a large impact on the bid distribution (Finland) and when the cartel had a more modest impact on the distribution of bids (Sweden).

[17] 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.

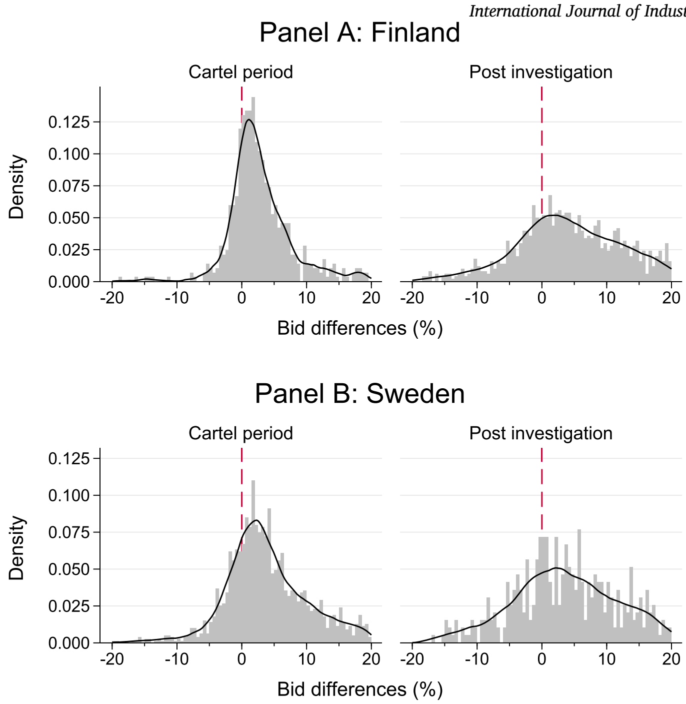

Fig. 2. Distribution of bid difference excluding winning bids $\Delta^2_{i,t}$ by country and period.
This figure plots the difference between a bid and the lowest rival bid when the winning bid is excluded for asphalt procurement contracts in Sweden and Finland before and after the cartel investigations. The width of the bins is 0.5. The curves correspond to density estimates calculated using Epanechnikov kernel.

6.1. Distributional regression test

The cartel detection method suggested by Clark et al. (forthcoming) 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. The detection method involves the comparison of the two distributions in a small interval around zero. Clark et al. (forthcoming) show that if the bidding functions are smooth, under competition, the two distributions can be approximated by the same distribution in a small interval around zero. Bid smoothness is satisfied, for example, in first price auction with independent private values. However, with collusive bidding, the two distributions $\Delta^1_{i,t}$ and $\Delta^2_{i,t}$ can be different for the reasons discussed above.

In Fig. 2, 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.

Clark et al. (forthcoming) present a non-parametric and parametric version of their detection test. To implement both tests, the researcher specifies some relatively small interval $[-H, H]$ close to zero. Then observations within the interval $[-H, H]$ are stacked into one dataset. The non-parametric version then uses Kolmogorov-Smirnov test to determine whether the distribution of $\Delta^1_{i,t}$ and $\Delta^2_{i,t}$ are similar within $[-H, H]$. Small p-values would imply that the empirical distributions are not similar indicating that the market is collusive. In the auction level test auction-level characteristics Z_t can be included as controls by comparing the distribution of the residuals obtained from regressing the bid differences. $\Delta^1_{i,t}$ and $\Delta^2_{i,t}$ on Z_t.[18]

[18] For a step-by-step guide on implementing the non-parametric test with and without controls see section A14 in Clark et al. (forthcoming).

Table 3
Results from Clark et al. (forthcoming) non-parametric cartel detection test.

	Finland P-value	Sweden P-value
Post-investigation - controls	0.569	0.497
Post-investigation - no controls	0.410	0.407
Cartel period - controls	0.000	0.000
Cartel period - no controls	0.000	0.000

This table reports p-values from the non-parametric version of the Clark et al. (forthcoming) test. The controls include region fixed effects and the bitumen index.

To implement the parametric version of the test $[-H, H]$ is split into G equal-sized intervals. Then the following distributional regressions are run separately for each interval g within the G intervals:

$$y_{i,t,g} = \alpha_g + \beta_g \mathbb{1}(f(\Delta^1_{i,t})) + \gamma_g Z_t + \epsilon_{i,t,q} \tag{3}$$

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. With competitive bidding we would expect to see $\beta_g = 0$.

A key step in implementing the test by Clark et al. (forthcoming) is choosing $[-H, H]$. If the interval $[-H, H]$ is set too large, the test could reject the null hypothesis also in a competitive market. To guide our choice of H, we conduct a Monte Carlo simulation. In the Monte Carlo simulation, we first generate bids under competition from a standard first-price auction with private independent values. Then we run the test and check the rejection rates of the test under different choices of intervals. The Monte Carlo simulation is described in more detail in Appendix A.5. Based on the results, we set $[-H, H] = [-5\%, 5\%]$ and $G = 20$. That is, we run distributional regressions for equal-sized intervals g from range -5% to 5%.[19] We also use the same interval for the non-parametric tests.

In Table 3 we present the results from the non-parametric version of the test. The first two rows focus on the post-investigation period, and the last two on the post-investigation period. In the version with controls, we use the same control variables as in the parametric version of the test. For the post-investigation period the null hypothesis is not rejected in neither of the two countries. For the cartel period, the p-value is very small across specification for both countries.

Fig. 3 shows the results from the parametric version of the test 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 -0.078 (s.e. 0.016) lower in the distribution that contains the winning bid. For Sweden, the corresponding point estimate is -0.038 (s.e. 0.009). 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. The results for the post-investigation period are shown on the right-hand side of the Fig. 3. 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 4 of the 20 intervals. However, unlike during the cartel period, no clear visual pattern emerges from the estimates.

Overall, we conclude that the test by Clark et al. (forthcoming) correctly rejects competitive bidding during the cartel period in both Finland and Sweden. The test is able to detect the gap left between the winning and losing bids. Other features, such as the low variance of losing bids, also evident in Fig. 2, are not flagged by the test. In the post-investigation period competitive bidding is not rejected in Finland, while for Sweden the results for the post-investigation period are somewhat mixed. In the parametric version of the test, few of the point-estimates are significant also in the post-investigation period. On the contrary, in the nonparametric version of the test, the null of competition is not rejected during the post-investigation period in Sweden.

6.2. Machine learning test

Next, we examine the performance of a machine learning-based cartel detection method suggested by Huber and Imhof (2019). This method uses machine learning techniques 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 a tender. These screens include, for example, the standard deviation of the bids or the difference between the winning bid and the runner-up. In our application, we use a total of 12 screens. 10 of these are used by Huber and Imhof (2019) while the remaining two are introduced by Huber et al. (2022) and Wallimann et al. (2022). The list of the screens that we use is presented in Table 4. A more comprehensive description of the screens is provided in Appendix A.3. In addition to the 12 screens, we allow the model to have more flexible prediction patterns by also including the second powers and interactions of the 12 screens. This gives us a total of 91 potential predictors.

[19] 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.

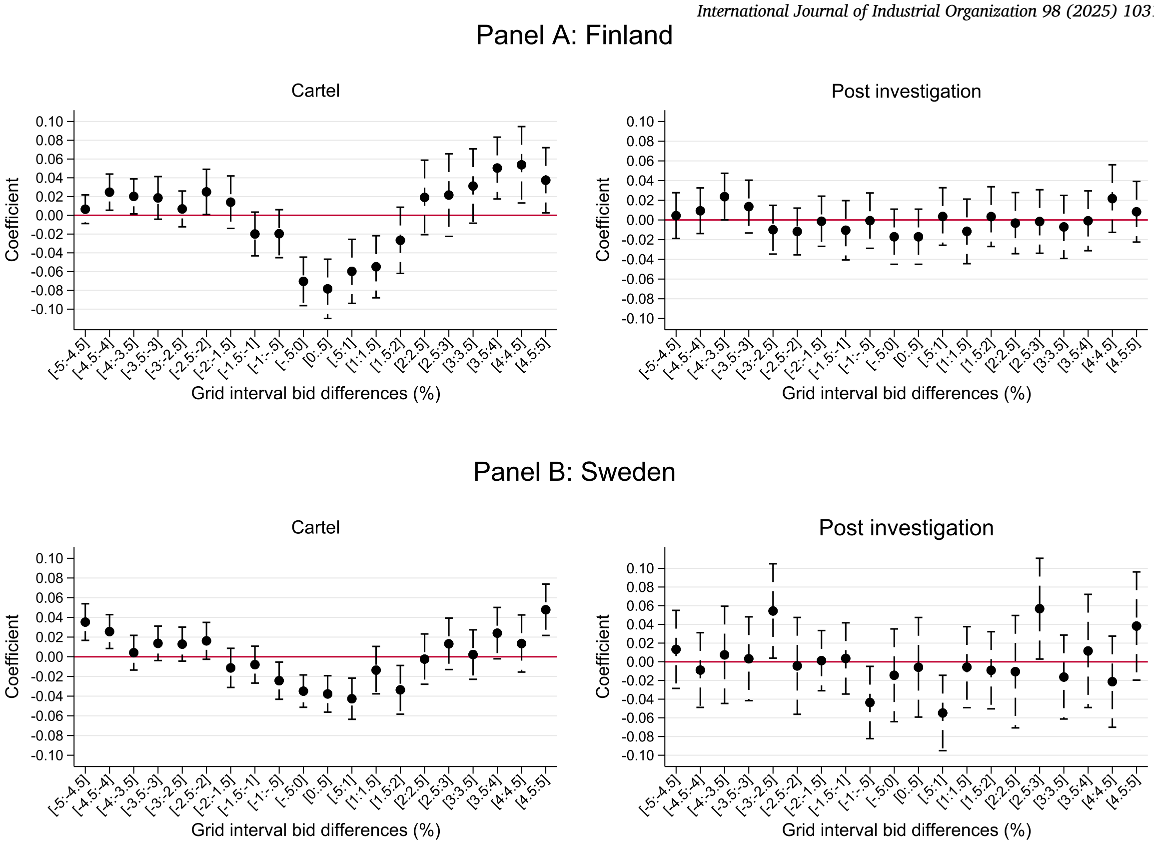

Fig. 3. Results from Clark et al. (forthcoming) parametric cartel detection test.
This figure plots the point estimates and 95% confidence intervals of β_g from equation (3) separately for each country and time period. Confidence intervals are calculated with robust standard errors.

Table 4
Formulas for statistical screens.

Screen	Formula
Standard deviation	σ_t
Spread	$\frac{\max(b_{i,t})-\min(b_{i,t})}{\min(b_{i,t})}$
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
KS-statistic	$KS_t = max(D_t^+, D_t^-)$, $D_t^+ = \max(x_{it} - \frac{rank(x_{it})}{n_t+1})$, $D_t^- = \max(\frac{rank(x_{it})}{n_t+1} - x_{it})$, $x_{it} = \frac{b_{it}-\min(b_{it})}{\max(b_{it})-\min(b_{it})}$

This table presents the screens used in our analysis and their formulas. $b_{i,t}$ 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, respectively. KS-statistic refers to the test statistic of a Kolmogorov-Smirnov test. For details see Appendix A.3.

To calibrate the predictive model, the analyzed 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 main 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:

$$\max_{\delta_0, \boldsymbol{\delta}} \left\{ \sum_{i=1}^{n} \left[y_t(\delta_0 + \sum_{j=1}^{p} \delta_j x_{tj}) - \log(1 + e^{\delta_0 + \sum_{j=1}^{p} \delta_j x_{tj}}) \right] - \lambda \sum_{j=1}^{p} |\delta_j| \right\}. \tag{4}$$

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. To make sure all screens are treated equally by the penalty term, we normalize each predictor to have zero mean and unit variance in the training data.[20] In addition to lasso logit, we report the results for alternative machine learning models in Appendix A.4.3.

We use two different ways to divide the observations into training and test data. First, we use data from the same country for training the model. We refer to this type of analysis as within-country analysis. In the within-country analysis, we use 75% of the observations from the country to train the predictive model and the remaining 25% to test the prediction performance. In addition to within-country specification, we run specifications in which we use data from other countries to train the model. This we refer to as transnational analysis. This relates to a scenario where a competition authority would need to calibrate a predictive model using data from a market that is not the same market from which the prediction is being made.

In training the model for transnational analysis, we supplement the Nordic datasets with datasets from cartels operating in construction sectors in Switzerland and Japan. These datasets have been used in the previous literature (Huber and Imhof, 2019; Huber et al., 2022). For confidentiality reasons, the Swiss data that we use only includes one of the cartels (Ticino cartel) studied in Huber and Imhof (2019). We convert the bids from Switzerland and Japan to euros using exchange rate from year 2021. We predict collusion only from the Finnish and Swedish datasets but use all four to train the model. We present results using each dataset as the training data separately and also when all datasets are pooled into one training dataset. In the transnational analysis, we do not need to split the datasets into training and test samples as with the within-country analysis. However, for the pooled training data we do choose only a subset of tenders from each country such that the number of tenders from each country is equal.[21] This way none of the countries is over-represented in the training sample. Since randomness can affect the split, in both the within-country and transnational analysis, we repeat the split and the prediction 100 times and average the predictive performance over the 100 iterations.

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 by country. 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. Lastly, the table also shows the number of observations (tenders) within each dataset. For both Finland and Sweden, we can see statistically significant differences in the values of the screens between the cartel period and post-investigation period. This applies for both screens capturing bid clustering (e.g., standard deviation, spread, coefficient of variation, or kurtosis) and screens capturing the difference between the winning bid and the losing bids (e.g., absolute difference, skewness, relative difference, or normalized distance). These observations are consistent with the existence of bid clustering and isolated winning bids in the cartel period. Importantly, these statistically significant differences in the screen values across periods suggest that these screens have potential for predicting collusion. This is further supported by the fact that the differences in screen values for the majority of screens are quite similar across all datasets (at least the direction of the change). Hence, also cross-country prediction seems feasible. One notable difference between the countries is that in Sweden the changes in the screen values are smaller than in the other three countries. If the Swedish cartel operated only in some of the tenders – a possibility we've previously discussed – only the collusive tenders would show signs of bid manipulation. As a result, the average screen values during the cartel period would be less pronounced compared to those of a full cartel.

A concern with statistical screens is that they might capture a change in the average contract size or other market-specific changes instead of collusion. For example, from Table 5 we see that in Switzerland the mean value of bids decreases while in other countries it increases. Similarly, we see that the number of bidders evolves differently in the four countries. Due to this, and following Huber and Imhof (2019), we choose to run the analysis using two different sets of screens. The first set of screens includes all 12 screens reported in Table 5 while the second includes only scale-invariant screens. Scale-invariant screens' values do not depend on the scale of bids, and hence they are not affected by contract size.[22]

[20] The screen values in the test data are scaled using the mean and the variance of the training data.

[21] In the pooled specifications the training data does not include tenders from the country used to test the model.

[22] Scale-invariant screens include the spread, coefficient of variation, kurtosis, percentage difference, skewness, relative difference, normalized distance and KS statistic.

Table 5
Descriptive statistics for statistical screens.

	Finland				Sweden				Switzerland				Japan			
	Cartel mean/sd	Post inv. mean/sd	T-test stat/p-val	K-S test stat/p-val	Cartel mean/sd	Post inv. mean/sd	T-test stat/p-val	K-S test stat/p-val	Cartel mean/sd	Post inv. mean/sd	T-test stat/p-val	K-S test stat/p-val	Cartel mean/sd	Post inv. mean/sd	T-test stat/p-val	K-S test stat/p-val
Standard deviation	0.045	0.378	-10.990	0.869	0.055	0.097	-5.153	0.328	0.028	0.046	-3.269	0.305	0.007	0.031	-10.393	0.595
	(0.037)	(0.326)	(0.000)	(0.000)	(0.065)	(0.088)	(0.000)	(0.000)	(0.024)	(0.040)	(0.001)	(0.010)	(0.018)	(0.031)	(0.000)	(0.000)
Spread	0.144	0.274	-8.023	0.534	0.235	0.256	-0.987	0.222	0.099	0.310	-11.714	0.714	0.035	0.166	-14.410	0.766
	(0.128)	(0.153)	(0.000)	(0.000)	(0.193)	(0.168)	(0.324)	(0.001)	(0.041)	(0.201)	(0.000)	(0.000)	(0.064)	(0.122)	(0.000)	(0.000)
Coefficient of variation	0.045	0.088	-9.293	0.558	0.070	0.081	-1.939	0.250	0.030	0.087	-13.003	0.734	0.011	0.044	-14.397	0.761
	(0.036)	(0.044)	(0.000)	(0.000)	(0.050)	(0.043)	(0.053)	(0.000)	(0.011)	(0.049)	(0.000)	(0.000)	(0.022)	(0.026)	(0.000)	(0.000)
Kurtosis	2.230	2.037	2.938	0.157	2.200	2.007	2.549	0.161	3.208	2.168	4.718	0.459	3.239	3.894	-2.870	0.145
	(0.594)	(0.586)	(0.004)	(0.032)	(0.676)	(0.566)	(0.011)	(0.034)	(1.204)	(0.787)	(0.000)	(0.000)	(1.696)	(3.012)	(0.004)	(0.019)
Absolute difference	0.035	0.179	-8.219	0.605	0.044	0.082	-4.073	0.244	0.045	0.027	2.418	0.378	0.004	0.009	-2.395	0.154
	(0.036)	(0.187)	(0.000)	(0.000)	(0.073)	(0.105)	(0.000)	(0.000)	(0.039)	(0.033)	(0.017)	(0.001)	(0.009)	(0.035)	(0.017)	(0.011)
Percentage difference	0.037	0.049	-2.842	0.227	0.070	0.072	-0.201	0.087	0.051	0.052	-0.132	0.391	0.005	0.020	-3.973	0.235
	(0.034)	(0.042)	(0.005)	(0.000)	(0.085)	(0.065)	(0.841)	(0.579)	(0.020)	(0.049)	(0.895)	(0.000)	(0.008)	(0.056)	(0.000)	(0.000)
Skewness	-0.020	0.313	-4.774	0.247	-0.072	0.080	-1.906	0.188	-0.817	0.203	-7.417	0.583	-0.561	0.474	-10.186	0.451
	(0.671)	(0.608)	(0.000)	(0.000)	(0.702)	(0.633)	(0.057)	(0.008)	(0.730)	(0.615)	(0.000)	(0.000)	(0.842)	(1.273)	(0.000)	(0.000)
Relative difference	2.229	0.905	2.342	0.335	1.910	1.726	0.507	0.169	4.624	0.887	5.549	0.785	1.719	0.803	3.964	0.535
	(8.948)	(1.371)	(0.020)	(0.000)	(3.172)	(2.968)	(0.612)	(0.024)	(3.814)	(1.070)	(0.000)	(0.000)	(1.837)	(2.958)	(0.000)	(0.000)
Normalized distance	1.432	0.809	8.021	0.398	1.385	1.099	2.890	0.191	2.934	1.019	7.843	0.678	3.111	1.667	7.522	0.444
	(0.802)	(0.646)	(0.000)	(0.000)	(0.863)	(0.831)	(0.004)	(0.007)	(1.345)	(0.786)	(0.000)	(0.000)	(1.901)	(2.096)	(0.000)	(0.000)
K-S statistic	0.339	0.370	-2.537	0.146	0.348	0.363	-1.170	0.108	0.429	0.324	4.419	0.466	0.347	0.435	-5.369	0.252
	(0.109)	(0.113)	(0.012)	(0.055)	(0.113)	(0.106)	(0.243)	(0.320)	(0.130)	(0.082)	(0.000)	(0.000)	(0.154)	(0.186)	(0.000)	(0.000)
Mean of bids	1.222	4.468	-13.263	0.739	0.817	1.253	-5.084	0.285	1.044	0.696	1.950	0.355	0.597	0.654	-1.111	0.109
	(0.811)	(2.585)	(0.000)	(0.000)	(0.700)	(0.889)	(0.000)	(0.000)	(0.962)	(0.712)	(0.053)	(0.001)	(0.536)	(0.533)	(0.267)	(0.141)
Number of bids	5.897	5.094	6.510	0.285	5.453	4.755	5.184	0.255	6.537	6.727	-0.452	0.082	11.374	16.188	-19.214	0.658
	(1.349)	(0.984)	(0.000)	(0.000)	(1.225)	(0.883)	(0.000)	(0.000)	(2.147)	(2.300)	(0.652)	(0.985)	(1.821)	(3.335)	(0.000)	(0.000)
Number of tenders	117	265	382	382	380	94	474	474	149	33	182	182	246	192	438	438

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 all countries. Standard errors and p-values are reported in the parentheses. The mean of bids, absolute difference, and standard deviation are reported in millions of euros.

Table 6
Model performance metrics.

Panel A: All screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Acc. (all)	89.7	87.3	76.0	56.7	69.7	68.8	74.1	52.3	73.5	61.7
Acc. (cartel)	89.4	79.4	71.3	97.1	72.6	70.5	81.5	49.0	90.0	61.6
Acc. (post-inv)	89.8	90.7	78.1	38.9	68.4	62.2	44.0	66.0	6.9	62.0
Panel B: Scale-invariant screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Acc. (all)	82.6	74.1	78.5	48.7	75.7	59.7	52.7	51.7	40.6	49.1
Acc. (cartel)	71.3	78.9	71.9	33.2	53.3	56.0	44.3	48.0	34.6	42.9
Acc. (post-inv)	87.6	72.0	81.4	55.6	85.6	74.2	86.5	66.5	65.0	74.0

This table reports the average performance of the predictive model over 100 iterations and measured by accuracy. The row *Test data* refers to the dataset from which the prediction is made from and the row *Training data* refers to the dataset used for training the model. The column *All* refers to a model where training data consists of tenders from all countries except the one being tested on.

Table 6 reports the performance of the predictive model. The performance is measured using accuracy, which represents the share of tenders that were predicted correctly with a probability threshold of 50% for a collusive prediction. In addition to accuracy for the whole sample, we also report accuracy separately for the tenders in the cartel period and tenders in the post-investigation period. Within-county analysis is depicted in the columns where the training data and test data are from the same country (i.e., the first column for Finland and the second column for Sweden) and transnational analysis in the remaining columns. The distribution of predicted collusion probabilities and the predictors chosen by the model are discussed in Appendix A.4.1 and Appendix A.4.2, respectively.

In the within-country analyses, the model was able to classify correctly 89.7% of the tenders in Finland and 74.1% in Sweden. When we use only scale-invariant screens, the prediction rate is 82.6% for Finland and 52.7% for Sweden. The prediction rate for Finland is in line with the prediction rates that have been found from other within-country analyses in the previous literature, while for Sweden the rate is lower. Huber and Imhof (2019) are able to predict 84% of tenders correctly with a dataset covering Swiss cartels, while Huber et al. (2022) predict 88% to 97% of tenders correctly with the Japanese dataset.[23]

In the transnational analysis, the prediction rates decrease on average. When all screens are used and the model is trained with Swedish data and tested with Finnish data, the prediction rate decreases from 89.7% to 87.3%. When the other two countries are used as training data, the prediction rate for Finland is 76.0% with the Swiss data and 56.7% with the Japanese data. When the data are pooled, the prediction rate for Finland is 69.7%. For Sweden, in the transnational analysis when all screens are used to train the model, the prediction rate varies from 52.3% to 73.5%. With only scale-invariant screens, the prediction rates in the transnational analysis range from 48.7% to 78.5% for Finland and from 40.6% to 59.7% for Sweden. In general, the finding that prediction rates decrease in the transnational analysis is consistent with previous literature. For example, Huber et al. (2022) report that prediction rates in their application decrease from 88–97% to 58–90% when moving from within-country analysis to transnational analysis. Although the sign is similar to our study, the prediction rates decrease more substantially in our setting.

We have examined alternative approaches from the machine learning literature to improve the predictive power of our model. First, in Appendix A.4.3 we report results using six alternative machine learning models. Out of the models considered, lasso logit is one of the best performing models. Second, when there is a significant imbalance in the number of collusive and competitive tenders, as is in the case of Sweden (see Table 1), a predictive model can achieve good performance by tilting the baseline collusion probability toward the more frequent type, even when the predictors have poor predictive power. To alleviate this, in Appendix A.4.4, we estimate a version of the model where we weight collusive and competitive predictions inversely to their frequency in the data. With this correction, the baseline collusion probability of the predictions should be roughly 50%. This change has only a minor effect on the predictive performance.

Huber et al. (2022) find that poor transnational performance can be explained by institutional differences across countries. If the screen values are systematically different across countries, the model will have a hard time predicting collusion even if the cartel had a similar effect on the screen values in all countries. From Table 5 we see that such cross-country differences exist in our case. The remedy Huber et al. (2022) suggest for this is to demean the screens by subtracting country-specific means from the screen values. In Appendix A.4.5 we report results when such demeaning is used with our data. With demeaning the prediction rates decrease on average but increase when Japanese data is used to train the model.

As described in Section 2, not all of the firms operating in the Finnish and Swedish asphalt markets were found guilty of collusion by the courts. If some of the firms in the cartel period bid competitively, this could potentially dilute the screens and decrease their predictive power. To account for the possible existence of incomplete cartels, we have run two robustness checks. First, we run a

[23] 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.

model designed to detect incomplete cartels proposed by Wallimann et al. (2022). It uses the same set of screens that we have used so far, but instead of calculating them at the tender level, they are calculated for different bidding coalitions within the tender. The details of the analysis and the results are reported in the Appendix A.4.6. Overall, we do not find significant increases in the prediction rates. In addition to this, we have run the model by excluding the bids submitted by non-convicted Swedish firms. Again, we do not find a significant difference compared to our main specification. In addition to the robustness check with respect to incomplete cartels, we have also conducted robustness check for the existence of a partial cartel. For this we have run a version of the model in which we restrict the Swedish data from the cartel period to only those tenders organized in convicted regions. These results are reported in Appendix A.4.7. Again, we do not find that this substantially increases the performance of the model.

Across specifications, the model performs worse in Sweden compared to Finland. The most straightforward explanation for this is that the screens are not as good predictors in the Swedish data as they are in the Finnish data. This is also evident in Table 5. The T-statistics there show that the changes in the screen values are generally much more pronounced in the Finnish data than they are in the Swedish data. This is in line with our conclusion from Section 5 that the bid distribution was less affected by the cartel in Sweden than in Finland.

The predictive ability of the distributional regression test and the machine learning test differs especially for Sweden. In the machine learning approach, the screens are calculated at the tender level and the prediction is made at the tender level. In contrast, the distributional regression test by Clark et al. (forthcoming) has the unit of observation at bid level, and the test is conducted at the market level. Both reducing the number of observations and making predictions at a more detailed level can make the model unable to predict collusion accurately when the cartel only has a modest effect on the distribution of bids.

One possible explanation for the poorer performance of the model in Sweden is that the cartel may have operated only in specific regions and tenders. When we focus solely on the regions mentioned by the Market Court A.4.7 or the convicted firms A.4.6, there is no clear improvement in the prediction rate. Furthermore, if we assume that models trained with one country's data are suitable for predicting collusion from another country (i.e., transnational analysis works), we would expect the transnational predictions for Swedish data to show a unimodal distribution of low collusion probabilities in the post-investigation period and a bimodal distribution of low and high collusion probabilities during the cartel period. Yet, when we examine the distribution of the predicted collusion probabilities in A.4.1 we do not observe this pattern for Sweden. These findings suggest that the limited scope of the cartel cannot at least fully explain the lower performance of the model in Sweden. Nevertheless, we cannot entirely rule out the possibility that classification issues in the Swedish data during the cartel period may have influenced our results to some extent.

Overall, we find that the results of the machine learning-based model are mixed. For Finland, the model works well across specifications in the within-country analysis. In contrast, in the transnational analysis, the model achieves a high performance only in some of the specifications. To elaborate, the prediction rate for Finland is only 48.7% if Japanese data is used to train the model and we are using only scale-invariant screens, but when the training data is changed to Swedish and all screens are used, the prediction rate raises to 87.3%. For Sweden, the prediction rates are lower than for Finland, but it is not obvious whether this is attributable to poor performance of the model or classification issues with the data.

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 – a bidding pattern observed also in cartels operating in Japan, Switzerland and Canada (Chassang et al., 2022; Clark et al., forthcoming; Imhof et al., 2018).

We find that the cartel had a greater impact on the distribution of bids in Finland compared to Sweden. This may suggest that the Swedish cartel was more effective at mimicking competitive behavior, or alternatively, that the Swedish cartel's influence was more limited in scope. While the Swedish Competition Authority identified the cartel as operating nationwide, the Market Court only found sufficient evidence from specific regions and tenders. We conduct empirical analyses where we restrict the sample only to the convicted firms or regions specified by the court and find that even then the impact remains larger in Finland. However, we cannot precisely identify all the tenders deemed rigged by the court within our dataset, leaving open the possibility that some tenders, during what we define as the cartel period in Sweden, were genuinely competitive. Due to this, we cannot exclude the possibility that the more limited extent of the Swedish cartel at least partially explains the observed disparity in results between Finland and Sweden.

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. (forthcoming), 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. 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 more 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 90% of the tenders in

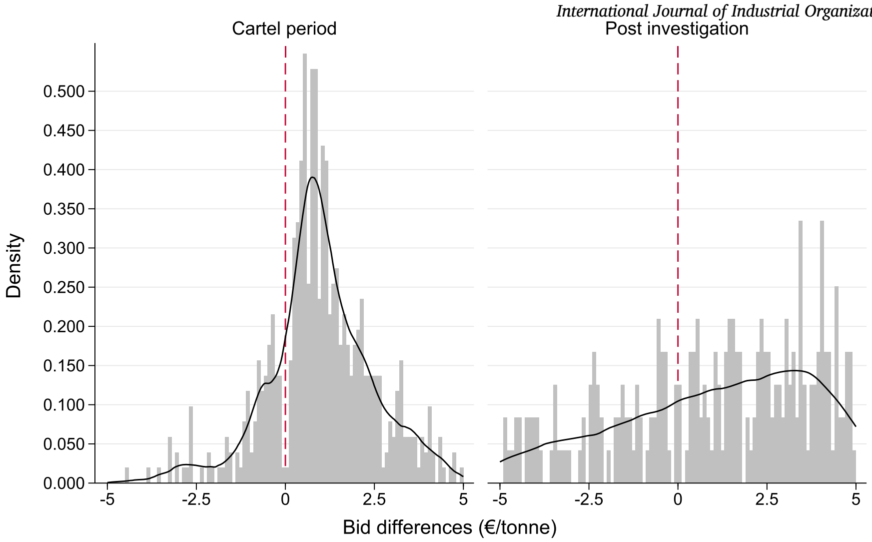

Fig. 4. Distribution of unit price bid differences $\Delta_{i,t}^{1,unit}$.
This figure plots the differences between a bid and the lowest rival bid in terms of unit prices before and after the cartel investigation. The width of the bins is 0.1. The curves correspond to density estimates calculated using an Epanechnikov kernel.

Finland and 74% in Sweden. We also test the performance of the model when the model is trained with data from another country. In this type of analysis, the prediction rates on average decrease substantially for both countries. We also find substantial heterogeneity in the results across model specifications.

Our results suggest that the statistical cartel detection methods we studied do have predictive power in the Nordic markets and hence can be useful for competition authorities in flagging suspicious behavior in public procurement. Consistently, we find that both methods flag collusion with higher certainty in Finland, where the cartel had a larger impact on the bid distribution. In the Swedish cartel, where the impact of the cartel on the bid distribution was more modest, the test by Clark et al. (forthcoming) performs better. Both methods have some limitations. The distributional regression identifies collusive behavior focusing from the gap between the winning and losing bids. In contrast, the machine learning–based method exploits various moments of the bid distribution to detect collusive behavior. However, it requires the user to calibrate the predictive model with existing data on known cartels. Our results indicate that the variation in prediction accuracy is strongly influenced by the selection of training data.

Appendix A

A.1. 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. (forthcoming) define bid differences in unit prices as follows:

$$\Delta_{i,t}^{1,unit} = b_{i,t}^{unit} - \Lambda b_{-i,t}^{unit} \tag{5}$$

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 Fig. 4 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.

We have also replicated the Clark et al. (forthcoming) cartel detection test using $\Delta_{i,t}^{1,unit}$. The results are shown in Fig. 5. Again, we see several statistically significant coefficients during the cartel period. Similarly to our main analysis, for the post-investigation period none of the estimated coefficients is statistically significantly different from zero.

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

$$\Delta_{i,t}^{1,r} = \frac{b_{i,t} - \Lambda b_{-i,t}}{r_t} \tag{6}$$

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

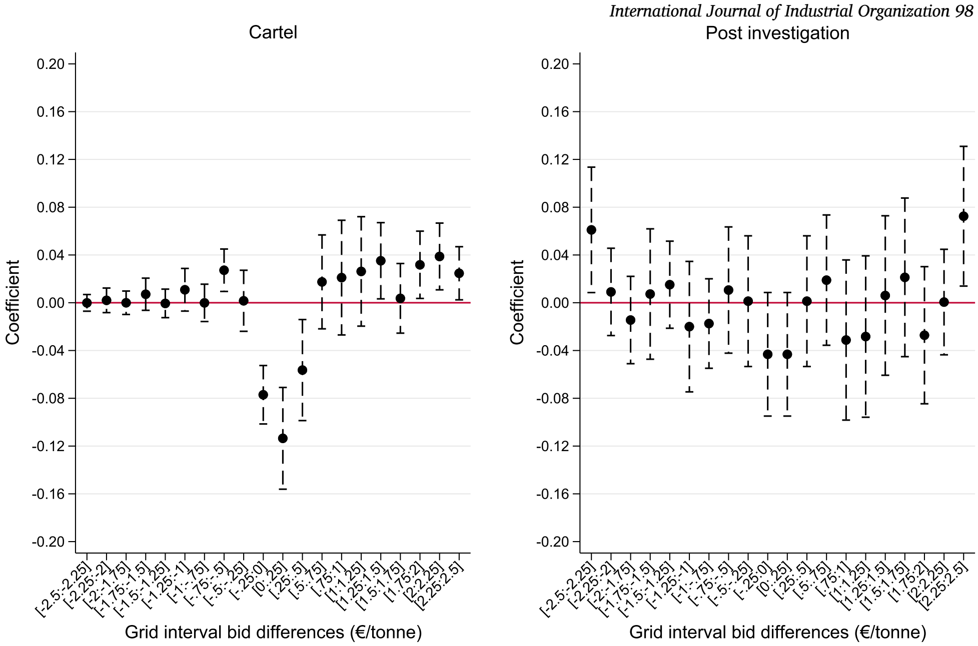

Fig. 5. Results from Clark et al. (forthcoming) cartel detection test when using $\Delta_{i,t}^{1,unit}$.
This figure plots the point estimates and 95% confidence intervals of β_g from equation (3) separately for each country and time period. Confidence intervals are calculated with robust standard errors.

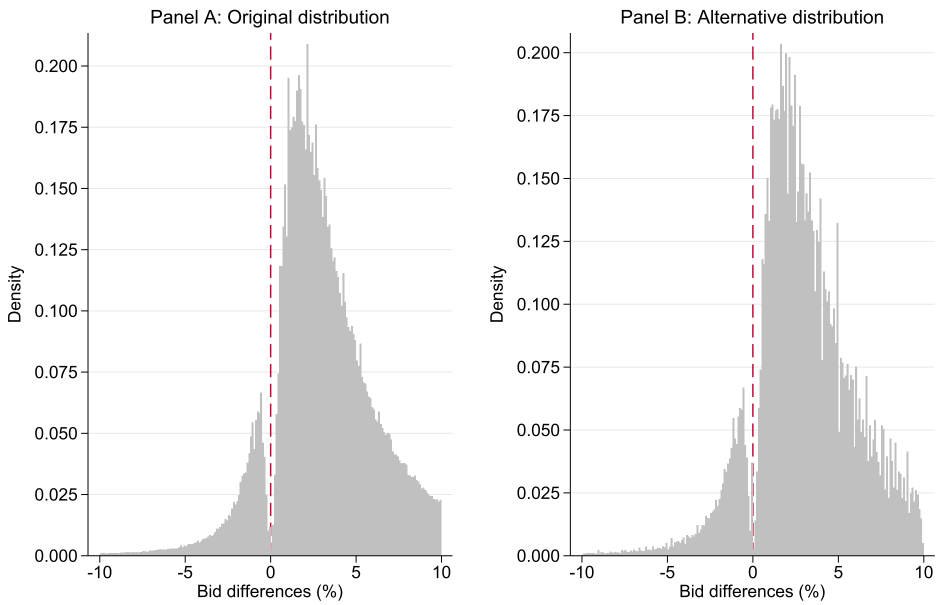

Fig. 6. Distribution of bid differences $\Delta_{i,t}^{1,r}$ (left) and $\Delta_{i,t}^{1}$ (right) with Japanese procurement dataset.
This figure plots the distribution of bid differences for Japanese national procurement data used by Chassang et al. (2022). On the left-hand side, we plot the distribution 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.

distribution as using the original definition. In Fig. 6, 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.

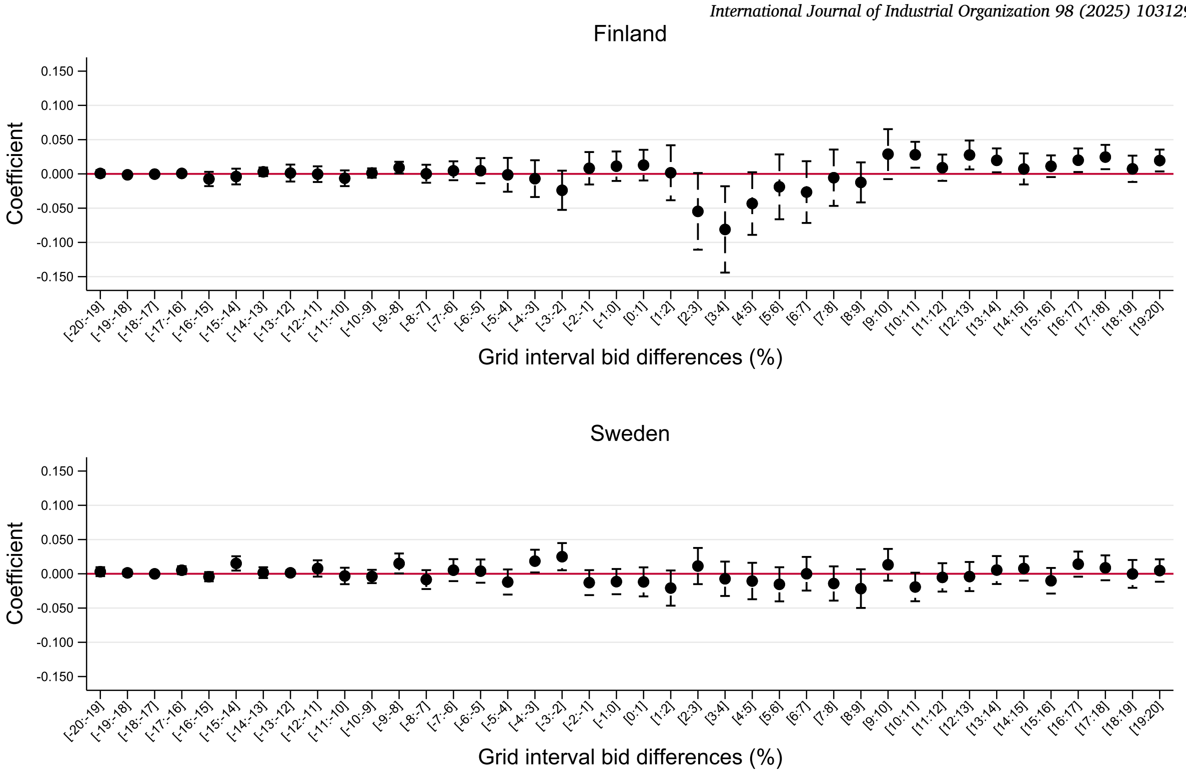

Fig. 7. Distributional regressions using 1% intervals.
This figure plots the results of the distributional regressions with 1% intervals between -20% and 20%. Panel A plots the point estimate of $\beta_{1,g}$ from equation (2) in the pre-post analysis. Panel B plots $\beta_{4,g}$ from equation (7) in the difference-in-differences analysis.

A.2. Robustness checks for change in the bid distribution

A.2.1. Distributional regressions with 1% intervals

In Section 5, we analyze how the distribution of bid differences $\Delta^1_{i,t}$ 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 Fig. 7. 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.

A.2.2. Adding control variables and alternative samples

In Section 5, 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.[24] For Finland, the contract area is available for around 85% of tenders and for Sweden, the area is reported for around 80% of the tenders. 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.

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

A.2.3. Assessing change in bid distribution for only tenders in the convicted areas

In Sweden, the courts found that the cartel operated only in specific regions. In Fig. 8 we plot the distribution of bid differences separately for tenders in convicted areas and other areas. We observe a roughly similar pattern for both area types. Similar to our main analysis there is large mass of bids within 10% of the winning bid but a low mass of bids within 1% of the winning bid. This finding

[24] Asphalt tonnes are 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	-0.014	-0.335***	0.350***
	(0.023)	(0.051)	(0.047)
Observations	1932	1932	1932
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 $p < 0.01$ (***).

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	-0.287***	0.308***
	(0.019)	(0.045)	(0.044)
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 $p < 0.01$ (***).

is consistent with the view of a case handler from the Swedish Competition Authority, Anders Gerde. In an interview, he states the cartel operated nationwide and why only some areas were called out in the decision relates to the fact that the whistleblowers worked in the convicted areas and had knowledge of and documents relating to cartel activities in only those specific regions (Hjalmarsson, 2015).

While both area types show a suspicious bidding pattern during the cartel period, we have still separately analyzed the change in bid distribution focusing only on the convicted areas. The results are shown Table 9. The point-estimates for the change are larger when we restrict the sample to the convicted areas. For example, the point estimate for the probability that a bid falls withing 1% to 10% of the winning bid is -0.184 compared to -0.088 when all areas are included in the regression. The point-estimate, however, still remains lower than the equivalent point-estimate for the change in Finland.

We have also run a version of the model, where we restrict the cartel period to years specifically referred to in the Swedish Market Court's decision. The results are shown in Table 10. The point estimate for the probability that a bid falls within 1% to 10% of the winning bid is -0.164 with a standard error of 0.057. Similarly to the specification reported below, the point estimate for "within 1%" is small and statistically significant.

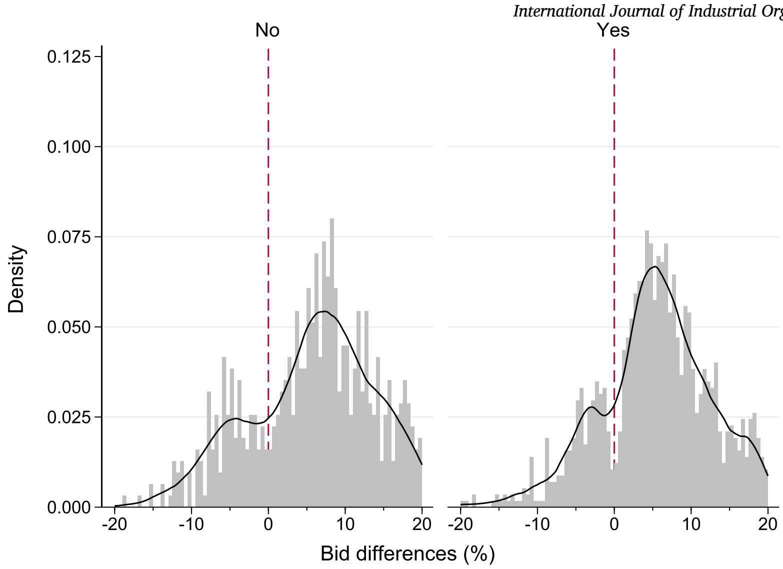

Fig. 8. Distribution of bid difference $\Delta^1_{i,t}$ separately for tenders in convicted areas and other areas in Sweden.
This figure plots the distribution of differences between a bid and the lowest rival bid for asphalt procurement contracts in Sweden during the cartel separately for the convicted areas ("Yes") and other areas ("No"). The width of the bins is 0.5. The curves correspond to density estimates calculated using an Epanechnikov kernel.

Table 9
Distributional effect of the cartel investigations when including only tenders from convicted areas in Sweden.

	Only convicted areas		
	Within 1%	Between 1% and 10%	More than 10%
Post	-0.022	-0.184**	0.207**
	(0.036)	(0.088)	(0.091)
Observations	1514	1514	1514
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. Standard errors are clustered at the tender level. Significance at $p < 0.10$ (*), $p < 0.05$ (**), and $p < 0.01$ (***).

Table 10
Distributional effect of the cartel investigations when including only tenders from convicted areas in Sweden and cartel years from 1997 to 2001.

	Only convicted areas		
	Within 1%	Between 1% and 10%	More than 10%
Post	-0.015	-0.164***	0.179***
	(0.021)	(0.057)	(0.057)
Observations	1054	1054	1054
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. Standard errors are clustered at the tender level. Significance at $p < 0.10$ (*), $p < 0.05$ (**), and $p < 0.01$ (***).

A.2.4. Assessing change in bid distribution using only tenders with convicted firms

In this section, we present results using tenders from the cartel period where only convicted firms left bids. In total, there are 107 tenders for Sweden, where only convicted firms submitted bids, and 73 for Finland. For the post-investigation period, the sample is the same as in our main analysis. The results are presented in Table 11. Compared to our main specification the absolute values of the point-estimates are larger in Finland. For Sweden, there is no notable difference in results. We have also run an analysis where we restrict the sample from the cartel period to tenders where at least 75% of bids were left by convicted firms. In this specification the

Table 11
Distributional effect of the cartel investigations when including tenders with bids submitted by convicted firms.

Panel A: Finland			
	Within 1%	Between 1% and 10%	More than 10%
Post	-0.021	-0.412***	0.433***
	(0.026)	(0.046)	(0.043)
Observations	1803	1803	1803
Region FE	Yes	Yes	Yes
Bitumen index	Yes	Yes	Yes
Panel B: Sweden			
	Within 1%	Between 1% and 10%	More than 10%
Post	0.016	-0.124	0.108
	(0.029)	(0.078)	(0.079)
Observations	865	865	865
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 $p < 0.01$ (***).

number of tenders in the cartel period is 141 for Finland and 216 for Sweden. The results do not notably differ from those reported in Table 11.

A.2.5. Difference-in-differences analysis

A potential concern is that the change in bid distribution is 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.[25] 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.

The Californian data covers paving contracts procured by the California Department of Transportation from 1999 to 2008.[26] 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.[27]

In Fig. 9, 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. 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.

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

$$y_{i,t,g} = \alpha_g + \beta_{2,g}\, post_t + \beta_{3,g}\, treat_t + \beta_{4,g}\, treat_t \times post_t + \epsilon_{i,t,g} \tag{7}$$

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.

[25] 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.

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

[27] 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.

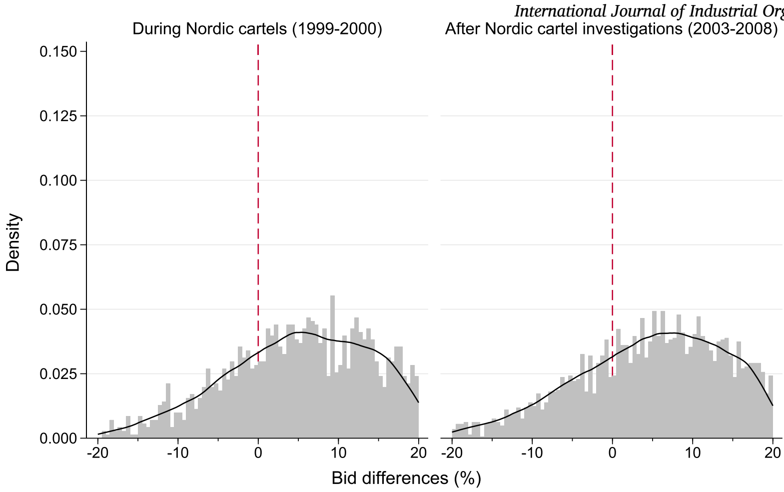

Fig. 9. Distribution of bid difference $\Delta^1_{i,t}$ for control market.
This figure plots the distribution of differences between a bid and the lowest rival bid for asphalt procurement contracts in California before and after the cartel investigations in Finland (launched in 2002) and Sweden (launched in 2001). The width of the bins is 0.5. The curves correspond to density estimates calculated using an Epanechnikov kernel.

Table 12
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 $p < 0.01$ (***).

The results are shown in Table 12. Similar to results from our main specification, we 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.

To test the robustness of the difference-in-differences specification, we have added several control variables to the regression. To control for changes in demand, we have added the aggregate gross domestic product and the gross domestic product of the construction sector to the regression.[28] We also added several variables related to market structure. These include the number of bidders in the tender, HHI, and number of active bidders on the market in a given year.[29] The results are shown in Table 13.

[28] For Finland and Sweden the source of GDP data is Eurostat and for California from the United States Regional Economic analysis project.

[29] The variables related to market structure are potentially bad controls because, as discussed in Section 4, the ending of the cartel potentially affected the market structure. We also run a version of the regression where we only add the control variables related to demand. The results are very similar to when both sets of control variables are added to the regression. These results are available on request.

Table 13
Distributional effect of the cartel investigations using a difference-in-differences design with controls.

Panel A: Finland			
	Within 1%	Between 1% and 10%	More than 10%
Post * Treat	0.006	-0.255***	0.249***
	(0.029)	(0.081)	(0.079)
Observations	6924	6924	6924
Panel B: Sweden			
	Within 1%	Between 1% and 10%	More than 10%
Post * Treat	-0.030	-0.002	0.033
	(0.025)	(0.057)	(0.057)
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 $p < 0.01$ (***).

When Finland is used as the treatment market, the point estimates are slightly larger in absolute value, and their statistical significance increases. In contrast, when Sweden is used as the treatment market, the point estimates and their statistical significance are lower than without the control variables.[30]

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.[31] 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 Fig. 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.

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:

$$y_{i,t,g} = \alpha_g + \beta_{2,g}\, post_t + \beta_{3,g}\, treat_t + \beta_{4,g}\, treat_t \times post_t + \beta_{5,g} t + \epsilon_{i,t,g} \tag{8}$$

where everything is as in equation (7) 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 14. 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 (7) 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 (7) 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%.[32] 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.

[30] We have also run a version of the difference-in-differences specification where we have added both country-specific time and the additional control variables to the specification. For Finland this increases the point estimates further. For Sweden, the point estimates are larger, but they remain statistically insignificant. These are available on request.

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

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

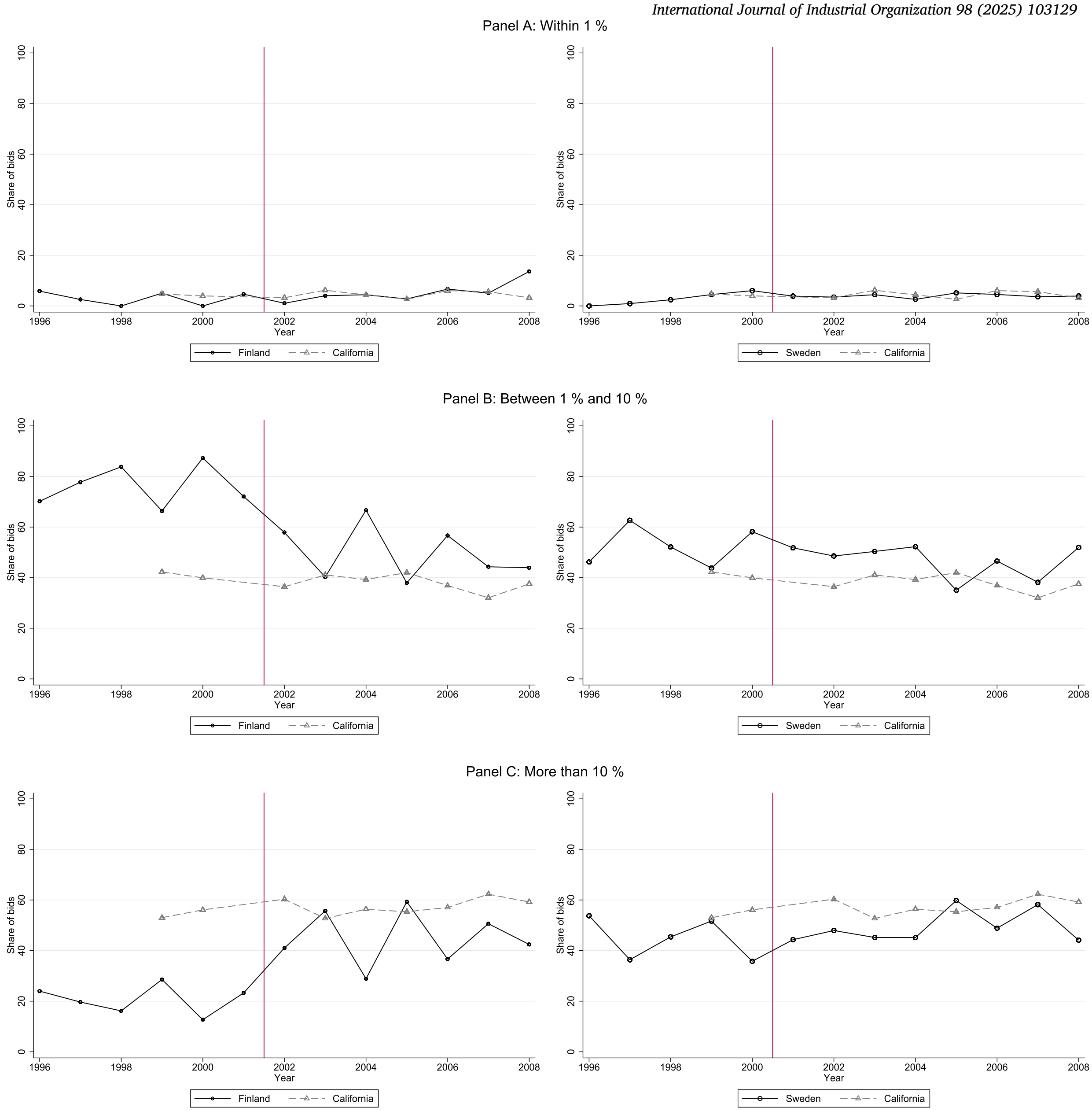

Fig. 10. Pre-treatment trends.
This figure plots the share of bids in different intervals over time. Panel A focuses on intervals 0–1%, Panel B on intervals 1–10%, and Panel C on intervals >10%. The gray line depicts California. The black line depicts Finland in the graphs on the left-hand side and Sweden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

A.3. Statistical screens of the machine learning model

In Section 6.2, we use 12 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), and the subsequent papers by Huber et al. (2022) and Wallimann et al. (2022). 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 12 statistical screens in more detail.

As discussed in Section 3, bid rigging may affect the dispersion of bids. Standard deviation, spread and coefficient of variation 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

Table 14
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 $p < 0.01$ (***).

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. The number of bids and contract value are used to control for different procurement types. Finally, Kolmogorov-Smirnov statistic (KS-statistic) measures non-parametrically the likeliness of the empirical cumulative distribution function (cdf) of the bids to the cdf of some baseline distribution (in our case uniform distribution). As the cdf of the bids under collusive bidding is likely different from the cdf of competitive bids, it will show up as difference in the KS-statistic. Note that in our application of KS test we also normalize the bids to an interval of [0, 1]. This is to match support of the empirical cdf with the support of the uniform distribution. The formulas to calculate each of the 12 screens are presented in Table 4.

A.4. Additional results for machine learning

In this section we present additional analyses that we have conducted for the machine learning model discussed in Section 6.2.

A.4.1. Distributions of machine learning predictions

In this section we show in more detail the distribution of predicted collusion probabilities by the machine learning model described in Section 6.2. In Fig. 11 and Fig. 12, we plot the share of tenders over the predicted collusion probability for Finland and Sweden, respectively. The predicted collusion probability is calculated as the average over the iterations in which the tender was in the test sample. The distributions are shown separately by period (columns) and by training data (rows). We show the results from the specification where only scale-invariant screens are used.

Based on these distributions, we can see that there are large differences in the predicted collusion probabilities over different training data. When focusing on the distributions resulting from a model trained with Swedish data, the predicted probabilities are clustered close to 50%. On the other hand, the models trained with Swiss or Japanese data seem to give predictions that are closer to the end points (0% or 100%). Therefore, even though the accuracies with these models are quite similar (as with Swedish and Swiss data), the distribution of the predicted probabilities may differ significantly. We believe that this result can be explained by the different scales in the screen values across countries and how that affects the magnitude of the coefficients chosen for the logit model.

A.4.2. 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.

In our 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 15 we report five predictors with the highest average value of the absolute coefficient for the different specifications. From Panel A, which shows the predictors when all screens are considered, the mean of bids and the number of bids seem to be among the most frequently used predictors. Other important

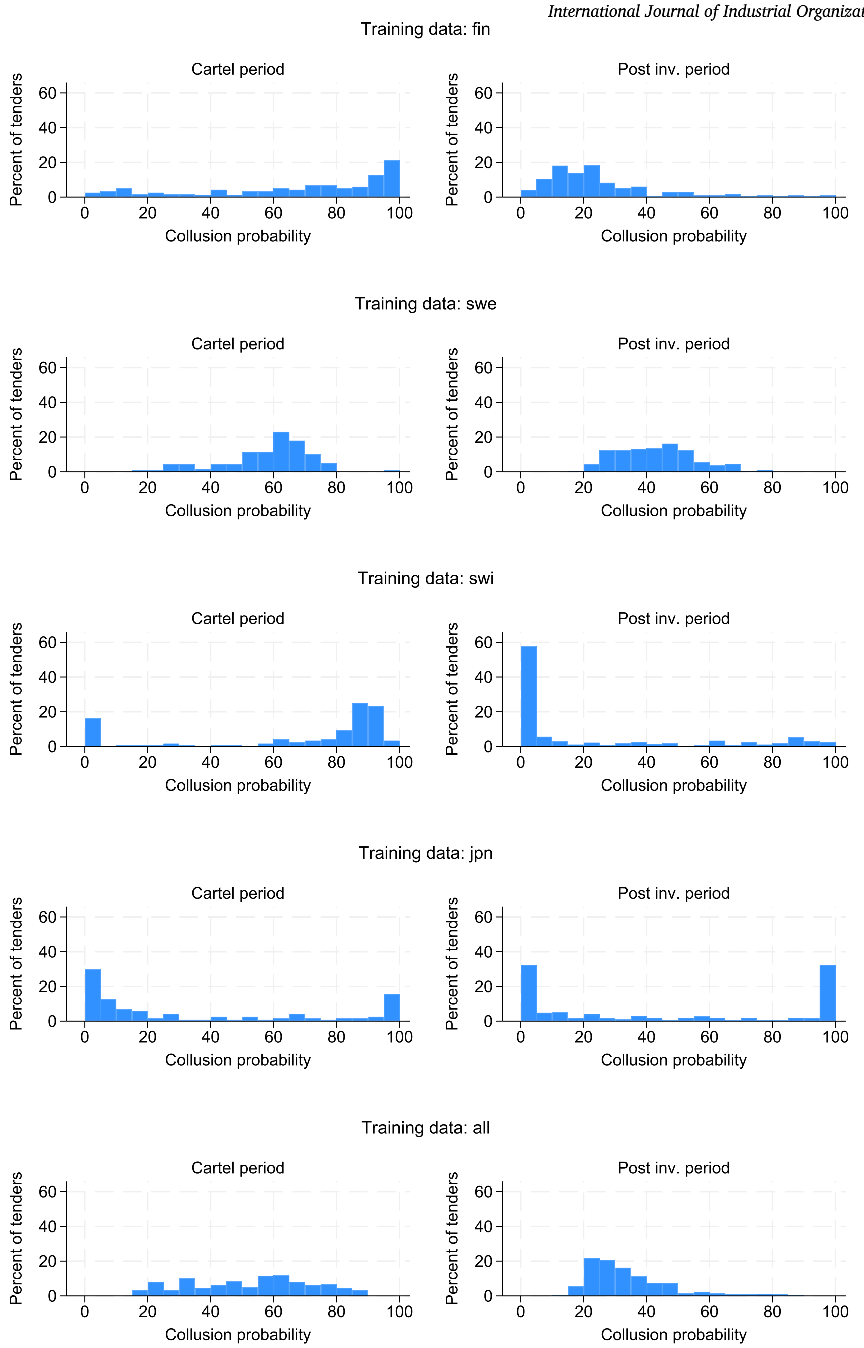

Fig. 11. Distribution of collusion probabilities – Finland.

predictors are the relative distance and variance-based screens such as standard deviation and coefficient of variation. In Panel B, which reports the most important predictors when we exclude the scale-dependent screens and number of bids, the model seems to shift toward using screens that focus on the difference between bids (relative distance, percentage difference, normalized distance, coefficient of variation).

A.4.3. Alternative machine learning models

In Section 6.2 we use lasso logit as our baseline model for predicting collusion. The choice of lasso logit is motivated by the fact that it is a model that is relatively easy to understand, it has been used in the existing literature on cartel screening, and that it has relatively good prediction performance in the Nordic cases. In this section, we present results for alternative machine learning models.

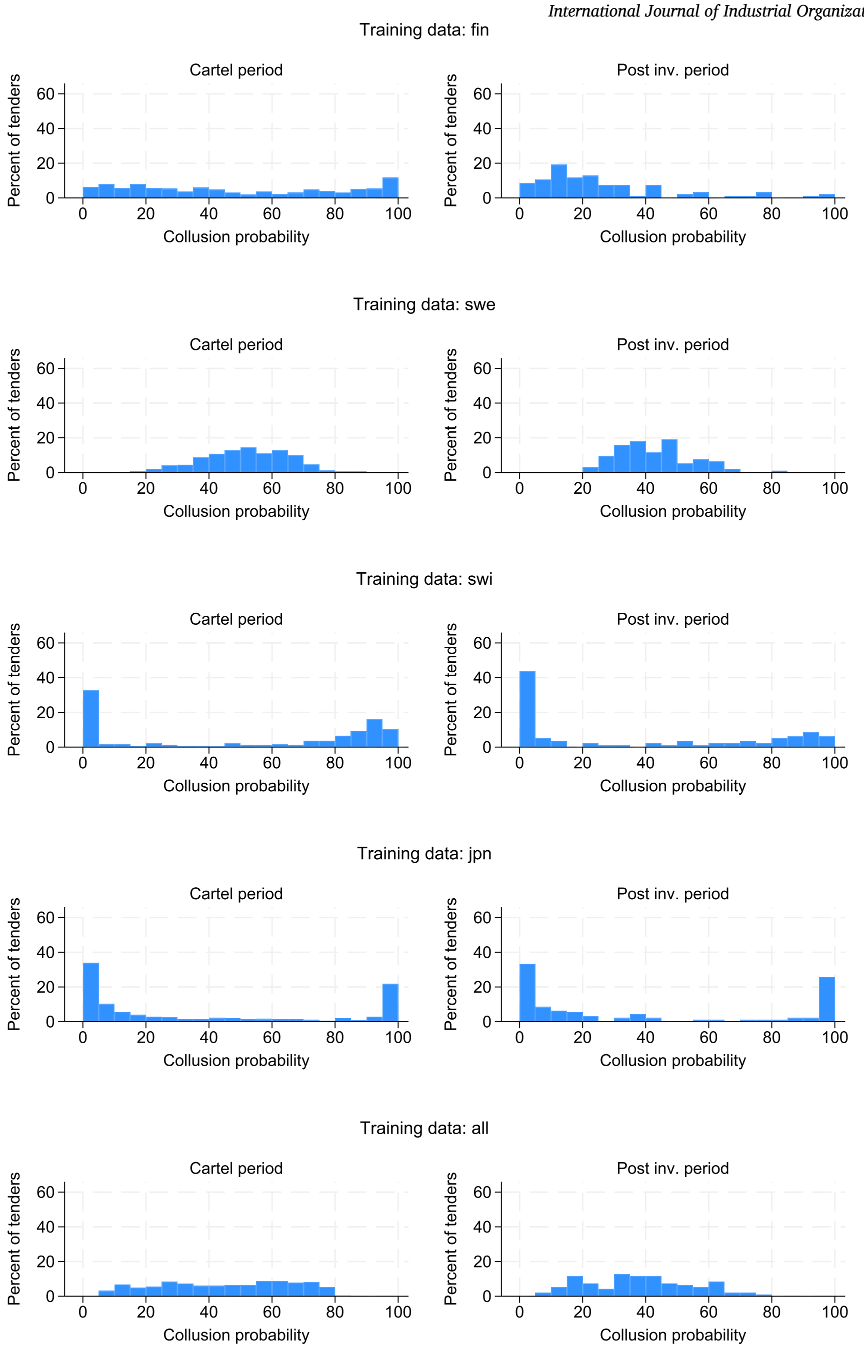

Fig. 12. Distribution of collusion probabilities – Sweden.

The first model we consider is a decision tree. Decision tree is an algorithm that classifies tenders by making consecutive binary decisions, and where each binary choice is made based on the value of a screen. When these binary choices are put together, a decision tree is formed. The second classifier is a nearest neighbor algorithm. The nearest neighbor algorithm (also known as K-neighbors) chooses the classification of a tender based on the classification of the K nearest neighbors of the tender within feature space (i.e., the tenders that have similar values for the screens). The third algorithm, Naive Bayes, bases classification on the posterior probability of Bayes theorem (while assuming conditional independence between features). For our fourth algorithm we have Neural network which uses a layered structure with a large number of small activation functions that convert features into signals and which then are used to determine the classification. This method is often used especially in image processing. The fifth algorithm, random forest, is a so-called ensemble method that takes the average of the predictions by multiple decision trees. This type of algorithm has been often used in the previous literature on cartel detection, notably by Huber et al. (2022) and Wallimann et al. (2022). The last algorithm

Table 15
The most important predictors.

Panel A: All screens

Finland										Sweden									
Fin		Swe		Swi		Jpn		All		Fin		Swe		Swi		Jpn		All	
Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)
normd × mean	3.32	std × nbids	3.11	cv^2	4.04	$nbids^2$	2.61	$nbids^2$	2.98	std × nbids	3.10	normd × mean	4.06	cv^2	4.05	$nbids^2$	2.61	$nbids^2$	3.03
rd^2	2.19	cv^2	1.79	spread × cv	3.32	nbids	1.71	normd	1.82	cv^2	1.62	rd^2	3.12	spread × cv	3.32	nbids	1.70	normd	1.96
mean × nbids	1.84	std	1.46	$spread^2$	2.66	cv	1.66	mean × nbids	1.26	std	1.35	rd × ks	2.15	$spread^2$	2.66	cv	1.64	cv	1.94
normd × nbids	1.84	cv × mean	1.46	perc_diff	1.69	spread	1.58	nbids	1.10	cv × mean	1.35	normd × nbids	2.09	perc_diff	1.68	spread	1.57	mean × nbids	1.70
rd × normd	1.49	skew × normd	1.27	std × kurt	0.60	mean	1.36	cv	1.04	skew × normd	1.26	rd × nbids	2.03	std × kurt	0.60	mean	1.36	cv × ks	1.68
penalty term	49.79	penalty term	284.41	penalty term	20.71	penalty term	118.34	penalty term	79.67	penalty term	232.74	penalty term	64.77	penalty term	6.04	penalty term	14.90	penalty term	148.59

Panel B: Scale-invariant screens

Finland										Sweden									
Fin		Swe		Swi		Jpn		All		Fin		Swe		Swi		Jpn		All	
Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)	Screen	Mean (abs.)
rd^2	5.20	cv^2	2.09	cv^2	5.23	spread	5.78	normd	1.13	cv^2	1.70	rd^2	7.89	cv^2	5.09	spread	5.77	kurt × rd	1.01
kurt × normd	2.36	skew × normd	1.67	spread × cv	3.87	cv	4.77	kurt × rd	1.13	skew × normd	1.41	perc_diff	2.41	spread × cv	3.72	cv	4.77	$normd^2$	0.90
perc_diff	2.31	$spread^2$	1.56	$spread^2$	2.68	cv^2	4.53	$normd^2$	1.01	perc_diff	1.31	skew × ks	2.34	$spread^2$	2.51	cv^2	4.53	cv	0.89
$perc\_diff^2$	2.29	perc_diff	1.47	perc_diff	1.39	rd	4.10	kurt × perc_diff	0.76	$spread^2$	1.23	$perc\_diff^2$	2.30	perc_diff	1.32	rd	4.09	normd	0.77
skew × ks	2.03	$skew^2$	1.25	kurt	1.24	rd^2	3.08	cv × ks	0.67	$skew^2$	1.07	kurt × normd	2.14	kurt	1.13	rd^2	3.07	spread	0.68
penalty term	145.17	penalty term	305.18	penalty term	7.91	penalty term	211.90	penalty term	55.92	penalty term	106.45	penalty term	145.22	penalty term	9.19	penalty term	142.59	penalty term	26.60

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 A.3.

Table 16
Model performance metrics for alternative machine learning models.

Panel A: Scale-invariant screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Lasso logit										
Acc. (all)	82.6	74.1	78.6	48.7	75.5	59.3	52.7	51.7	40.7	49.2
Acc. (cartel)	71.3	78.9	72.0	33.2	51.3	55.5	44.3	48.0	34.6	42.8
Acc. (post-inv)	87.6	72.0	81.5	55.5	86.2	74.2	86.5	66.6	65.0	75.0
Decision tree										
Acc. (all)	77.9	30.7	72.8	61.1	75.0	79.6	43.1	55.8	25.8	48.8
Acc. (cartel)	56.9	99.9	61.5	8.5	55.6	99.7	32.3	54.7	9.6	41.9
Acc. (post-inv)	87.0	0.1	77.8	84.3	83.5	0.3	86.8	60.5	91.3	76.9
Nearest neighbor										
Acc. (all)	80.8	35.7	70.3	65.1	76.0	78.2	43.8	59.3	27.9	50.4
Acc. (cartel)	57.4	98.7	75.8	26.6	56.2	96.5	31.7	60.3	12.0	44.6
Acc. (post-inv)	91.0	7.8	67.9	82.1	84.7	5.7	92.4	55.2	92.5	73.9
Naive Bayes										
Acc. (all)	68.6	64.9	73.3	72.5	64.6	33.0	34.3	35.7	43.3	51.5
Acc. (cartel)	35.9	12.8	24.8	66.8	56.7	21.3	21.7	23.2	34.7	48.3
Acc. (post-inv)	82.7	87.9	94.7	75.0	68.1	79.3	85.1	86.2	78.4	64.3
Neural network										
Acc. (all)	76.8	45.7	72.0	54.4	65.8	73.2	45.9	55.9	53.3	50.3
Acc. (cartel)	60.7	91.6	78.5	63.4	60.6	83.8	35.5	57.4	54.7	45.2
Acc. (post-inv)	83.7	25.4	69.1	50.5	68.1	31.5	87.7	50.2	47.9	70.6
Random forest										
Acc. (all)	80.1	30.8	75.2	64.1	77.5	79.5	40.9	55.8	23.2	47.9
Acc. (cartel)	56.7	99.8	71.2	6.0	57.7	99.5	28.3	53.7	5.4	40.3
Acc. (post-inv)	90.3	0.3	76.9	89.7	86.2	0.4	91.8	64.6	94.8	78.6
AdaBoost										
Acc. (all)	80.2	31.1	74.5	65.7	73.4	79.1	41.7	56.6	25.4	48.9
Acc. (cartel)	57.4	99.6	70.5	11.0	54.7	98.7	30.4	55.6	9.6	42.2
Acc. (post-inv)	90.0	0.8	76.3	89.8	81.7	1.1	87.5	60.8	88.9	75.9

that we consider, AdaBoost (also known as adaptive boosting), is another ensemble method, which in this case uses subsequent simple prediction models (usually small decision trees), re-weights the sample between the predictions to put more emphasis on the harder-to-predict instances, and then takes the weighted average of the predictions. All the models we consider here are implemented using the Scikit-learn package in Python. For more detailed information on the models, we suggest the reader to have a look at the excellent documentation available at Scikit-learn website.

When testing the performance of these alternative models, we only include scale-invariant screens to the model. For finding the optimal hyperparameters for each model, we have used a randomized search over a specified parameter grid and took the best performing parameter values with cross-validation. Table 16 reports the results of alternative learners. The first specification in the table is lasso logit for reference. The small differences between the lasso logit results in Table 6 and Table 16 are explained by randomness in the composition of training and test samples.

The results from all models are reported in Table 16. In both Sweden and Finland, the best within-country performance is obtained with lasso logit. For transnational analyses the performance varies quite substantially. For Finland lasso logit is the best performing model for all transnational analyses except for the case where Japanese data is used for training the model. For Japanese training set, Naive Bayes seems to perform the best. For Sweden, there is much more variance in transnational prediction rates compared to Finland. For a specification where Finland is the training set, other models perform better than lasso logit, but they are strongly biased toward one type of prediction. For the Swiss-trained models, nearest neighbor seems to perform best, while for the Japanese-trained model the best model is neural network. However, the prediction rates with these models still remain very low: the best learner with the Swiss data obtains a prediction rate of 59.3% and with the Japanese data 53.3%.

Overall, lasso logit seems to consistently be among the best performing models. Based on these results, there is no single model that would dominate all the other models – it depends on the specification which one is the best model to use. However, it seems that the performance across the models is around the same if the prediction rate is overall relatively good (as is the case with Finland). On the other hand, if there is significant variance in the prediction rate among the models, then none of the models seems to obtain a good prediction rate (as is the case with Sweden). Our conclusion from this result is that the prediction performance is mainly influenced by how well the data that is used for training the model matches with the data from which the prediction being made. If these data match, the choice of learner has little effect on the performance – they all seem to work well. On the other hand, if the data do not match, choosing an alternative learner doesn't seem to fix the performance.

A.4.4. Balanced objective function

In Section 6.2 we discussed the potential issue of unbalanced data in our sample. When the parameters of a machine learning model are estimated, the parameter values are set to maximize the out-of-sample prediction rate within the training sample (using cross-validation). This means that if the share of classes (collusive vs competitive) are unbalanced in the training sample, the estimated

Table 17
Model performance metrics with balanced objective function.

Panel A: All screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Acc. (all)	89.2	81.2	63.5	59.7	69.7	61.8	75.4	45.9	72.8	61.7
Acc. (cartel)	92.4	77.9	56.5	96.4	72.6	59.0	86.6	39.4	88.7	61.6
Acc. (post-inv)	87.9	82.6	66.6	43.6	68.4	73.1	30.4	72.3	8.4	62.0
Panel B: Scale-invariant screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Acc. (all)	79.7	74.8	76.6	50.6	75.7	57.6	58.5	45.9	39.9	49.1
Acc. (cartel)	76.2	75.4	56.4	33.6	53.3	52.8	53.0	39.8	33.5	42.9
Acc. (post-inv)	81.2	74.5	85.5	58.1	85.6	76.5	80.8	70.3	66.0	74.0

Table 18
Model performance metrics with country-wise demeaning.

Panel A: All screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Acc. (all)	89.3	53.3	47.9	47.2	57.6	69.3	42.7	57.4	54.5	58.1
Acc. (cartel)	89.9	24.4	61.7	7.6	73.1	71.9	29.8	62.6	53.8	56.9
Acc. (post-inv)	89.0	66.1	41.8	64.7	50.8	59.2	94.8	36.2	57.2	62.7
Panel B: Scale-invariant screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Acc. (all)	82.3	65.2	45.8	69.9	56.5	59.2	57.6	56.0	58.9	60.2
Acc. (cartel)	69.9	58.9	67.2	74.5	85.5	55.7	52.6	60.2	57.4	61.7
Acc. (post-inv)	87.7	68.0	36.3	67.9	43.6	73.0	78.0	38.9	64.8	54.0

parameters will be such that the baseline prediction leans towards the more common class. One remedy for this is to weight the objective function in the training phase so that the cost of misprediction is higher for the underrepresented class. If the weights are set inversely to their proportion in the dataset, the baseline predicted collusion probability should be roughly 50%. In Table 17, we report the results when the objective function is weighted.

In expectation, the main benefit of balancing is that the accuracy for the cartel period and for the post-investigation period should become closer to each other. Overall, we find that balancing the objective function does not uniformly affect the results across specifications. In some cases, the prediction rates between the two periods converge while in other cases they remain the same or even diverge.

A.4.5. Demeaned screens

Huber et al. (2022) suggest that when conducting transnational screening, one should demean the screens by subtracting country-specific means from the screen values. The reason for this is that institutional differences between countries may result in the screens having systematically different values. For example, if the training country has a smaller coefficient of variation for both collusive and competitive tenders than the test country, this would cause the model to always predict tenders from the test country to be collusive. Therefore, the suggestion is to demean the screen values to avoid this. In Table 18 we report results when the screens are demeaned before training the model.

The expected benefit of the demeaning would be that the transnational prediction performance improves. The results here show that the prediction rates for all transnational specifications actually drop significantly or remain the same except for the case where Japanese data is used for training the model and only scale-invariant screens are used. Based on the results, demeaning of screens – at least in our application – does not generally seem to improve the transnational prediction performance.

A.4.6. Incomplete cartels

In Section 6.2 we discuss how the cartel being incomplete can affect the performance of the model. For checking the robustness of our results for incomplete cartels, we perform the machine learning analysis using screens suggested by Wallimann et al. (2022). These screens are intended to predict collusion even in cases where there is a mix of collusive and competitive bids within a tender. In our implementation of the method, we use the same set of screens than in 4, but now calculated at a coalition level within each tender. We specify the coalitions to be of size 3 and 4. The aggregation from coalition level to tender level is done using mean,

Table 19
Model performance metrics with method by Wallimann et al. (2022).

Panel A: All screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Acc. (all)	88.0	88.8	81.7	75.8	73.8	69.2	73.0	48.7	35.8	55.6
Acc. (cartel)	90.5	86.2	64.7	21.4	68.3	71.3	78.6	41.6	22.2	53.1
Acc. (post-inv)	87.0	89.9	89.1	99.4	76.1	61.0	50.3	77.5	90.8	65.6
Panel B: Scale-invariant screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Acc. (all)	77.9	64.7	78.7	52.0	71.5	64.9	48.3	50.5	50.5	51.3
Acc. (cartel)	67.1	78.5	64.0	73.3	67.2	64.8	39.9	45.2	50.4	46.6
Acc. (post-inv)	82.6	58.7	85.1	42.8	73.3	65.4	82.4	71.8	51.1	70.4

Table 20
Model performance metrics with only convicted firms in Sweden.

Panel A: All screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Acc. (all)	89.7	71.6	76.0	56.7	79.9	70.5	61.5	57.7	61.7	62.1
Acc. (cartel)	89.4	21.6	71.3	97.1	73.4	77.6	71.3	53.0	92.5	62.2
Acc. (post-inv)	89.8	93.6	78.1	38.9	82.7	58.2	44.0	66.0	6.9	62.0
Panel B: Scale-invariant screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Acc. (all)	82.6	67.7	78.5	48.7	75.8	65.5	49.9	57.5	46.9	56.5
Acc. (cartel)	71.3	57.2	71.9	33.2	61.0	72.1	29.2	52.4	36.6	46.6
Acc. (post-inv)	87.6	72.4	81.4	55.6	82.4	54.0	86.5	66.5	65.0	74.0

median, min, and max of the coalition level screen values. We do not include the second powers nor the interactions of the screens in this analysis. The results for the Nordic data are shown in Table 19.

The predictive performance of the incomplete collusion test is quite similar to that reported in the main specification in Table 6. There are some differences between this model and the baseline model, but these differences are mixed depending on the dataset used for training and whether all or only scale-invariant screens were used. Using screens for incomplete cartels doesn't seem to uniformly increase or decrease the prediction performance.

In addition to the robust method by Wallimann et al. (2022), we have estimated a specification of the baseline model from section 6.2 where we have excluded the bids from non-convicted firms in Sweden. These results are shown in Table 20. Overall, the results don't seem to significantly change from the baseline model from Table 6.

Lastly, we have also tested prediction at the coalition-level rather than tender-level as suggested by Imhof and Wallimann (2021). That is, the cartel prediction is being made for each bidding coalition rather than for each tender. This prediction is fundamentally different from the other models since the object of prediction is a coalition instead of a tender, but it still might be a useful tool for competition authorities. However, in our data the number of unique bidding coalitions in very low in some periods. This means that the machine learning model has too few observations to reasonably estimate the model. Since, this approach was not well suited for our context, these results are not included in the paper but are available on request.

Overall, since neither the method by Wallimann et al. (2022) nor the exclusion of non-convicted firms seems to improve the performance for Sweden, we find it unlikely that incompleteness of the Swedish cartel could explain the poor prediction performance.

A.4.7. Partial cartels

In section 6.2 we mention the concern that the poor performance for Sweden may be explained if some of the tenders in the cartel period are actually competitive. With this in mind, we have estimated the machine learning model with only the tenders that we are confident that are collusive. These are the tenders from the four regions in Sweden that were the focus of the court proceedings and from which the court found evidence of collusion. In Table 21 we show these results. Compared to the results of the main specification, the performance hardly changes.

In addition to this specification, we have run other alternative specifications where we excluded different subsets of tenders from the Swedish data. In one specification, we exclude Swedish tenders from before 1995. The intuition here is that during these years the cartel might still have been forming and hence the bidding rigging scheme might not have been fully developed. In another specification, we exclude Swedish tenders that were very small in size. These tenders seemed to be potential outliers when looking

Table 21
Model performance metrics with only convicted areas in Sweden.

Panel A: All screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Acc. (all)	89.7	86.1	76.0	56.7	62.6	70.7	72.2	56.8	65.5	64.8
Acc. (cartel)	89.4	83.2	71.3	97.1	71.4	73.1	82.9	53.4	87.9	65.9
Acc. (post-inv)	89.8	87.4	78.1	38.9	58.7	64.6	44.0	66.0	6.9	62.0
Panel B: Scale-invariant screens										
Test data	Finland					Sweden				
Training data	Fin	Swe	Swi	Jpn	All	Fin	Swe	Swi	Jpn	All
Acc. (all)	82.6	75.4	78.5	48.7	76.4	66.3	61.6	56.6	43.9	55.5
Acc. (cartel)	71.3	77.3	71.9	33.2	52.3	62.1	52.1	52.8	35.9	48.4
Acc. (post-inv)	87.6	74.6	81.4	55.6	87.0	76.9	86.5	66.5	65.0	74.0

at the screen values. However, none of these alternative specifications made distinctive differences to the prediction results. These specifications are not reported in the paper but are available on request. Overall, based on the result here, the results of our main specification in Section 6.2 seem to be robust to the choice of tenders from the Swedish data.

A.5. Monte Carlo simulation results for distributional regression method

In Section 6.1 we discuss the choice of interval [-H, H]. As mentioned, a too large interval would cause the distributional regression test to reject the null hypothesis also in a competitive market. In their paper, Clark et al. (forthcoming) provide a proof regarding the false rejection of competitive bidding. They show that, if under competition the distribution of bids is smooth, then in a sufficiently small interval Δ^1_{ij} and Δ^2_{ij} can be approximated by the same distribution. The smoothness of bids implies the smoothness of each of the n-order statistics, which in turn implies that the distribution of the differences, $\Delta_k (k = 1, 2)$ are also smooth. For smooth Δ^1 and Δ^2 and for a sufficiently small interval H, the conditional distributions $\Delta^1_{ij} | \Delta^1_{ij} \in [-H, +H]$ and $\Delta^2_{ij} | \Delta^2_{ij} \in [-H, +H]$ can be both approximated by the same distribution. This is because in a small interval any distribution can be approximated by a rectangular bin. Bid smoothness is satisfied, for example, in sealed-bid first-price auctions with independent private values. For more details of the proof, we refer the reader to Clark et al. (forthcoming).

Hence, if the interval is set small enough, the test does not reject competitive bidding under competition. To investigate the appropriate interval width in our setting, we conduct a Monte Carlo simulation exercise. We simulate datasets from independent private value procurement auctions where n risk neutral bidders are bidding for a particular contract. Bidders are symmetric and the cost of carrying out the contract is drawn i.i.d. from a differentiable density f with a bounded support $[\underline{c}, \overline{c}]$. The bidder who submits the lowest bid is selected as the winner and awarded the contract.

For each simulation, we have 200 tenders. This is roughly the average of the number of tenders in the competitive period in Sweden and Finland. The number of bidders is derived from a discrete uniform distribution between three and five. For each bidder, we draw the costs using two different distributions. The first is a uniform distribution between 0.5 and 1 and the second is normal distribution with a mean of 0.75 and standard deviation of 0.15. The two distributions have the same mean and roughly the same standard deviation. The standard deviation was chosen such that the standard deviation of the bids roughly matches our dataset. For each bid, we solve for the equilibrium bids. With uniform distribution, the optimal bid function can be solved analytically as a function of the cost and the number of bidders. With a normal distribution, we solve the optimal bids numerically using value function iteration. In total, we have 200 simulated datasets.

For each simulated dataset, we conduct the distributional regression test using 10 different bin widths from 0.5 to 5.0. For each width, we estimate the model using 20 intervals. With the smallest bin the intervals vary from -5.0 and -4.5 to 4.5 and 5.0 and with the largest width they vary from -50 to -45 and 45 to 50. For each simulated dataset and interval width, we calculate the share of statistically significant estimates at the $p < 0.05$ level and then average over the simulations.

The results are shown in Fig. 13 separately for each width and distribution type. The horizontal line depicts 5%, which is the expected share of significant point estimates given our significance level and that the distributional regression analysis is performed 20 times. With the smallest width, the share of significant coefficients is almost exactly 5%. Once the bin width is increased, the share significant results increase markedly. With 2.5 bin width, the share of significant results with uniform distribution is already above 10%. The simulation results are in line with the proof. Given a sufficiently small interval, the test does not reject competitive behavior when data is generated by competitive bidding behavior in a standard first-price procurement auction with independent private values. We have also examined the magnitude of the point estimates. The largest point estimates are obtained when the largest bin width is used. With the largest bin width (5 percentage points) for the centermost intervals the point-estimates are close to .04. This is considerably smaller than the largest point estimates observed in the Finnish data and reported in Fig. 3. However, they are quite close to those estimated using the Swedish data.

We have also tested how the sample size affects the choice of bin width. In Fig. 14 we plot the rejection rate of the null hypothesis for three different sample sizes and five different bin widths. The sample size varies from 1000 to 1 million tenders, while the width

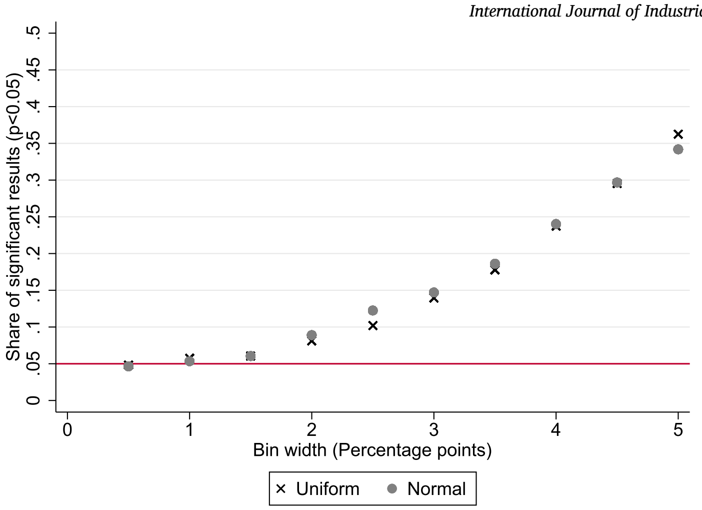

Fig. 13. Monte Carlo Simulation results with different cost distributions.

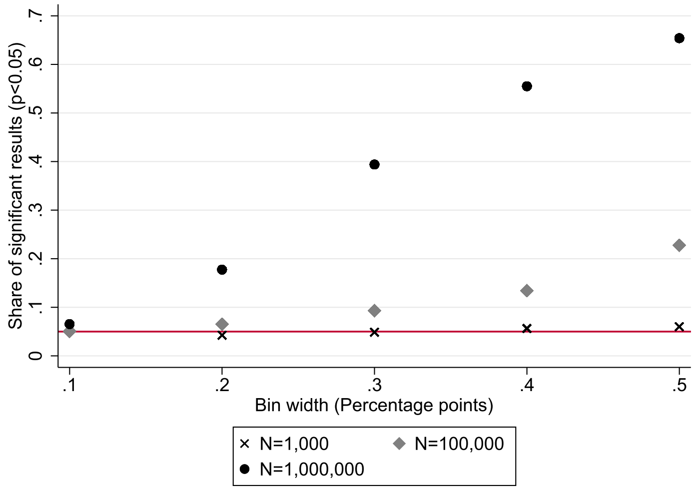

Fig. 14. Monte Carlo Simulation results with different sample sizes.

of the bin varies from 0.1% to 0.5%. With the smallest sample size, the rejection rate is close to 5% with all bin widths. With a sample size of 10,000 and with the largest bin width, the null hypothesis is rejected in more than 20% cases. However, with the smallest bin width, the rejection rate has converged to almost 5%. With the largest sample size, the null hypothesis was rejected in majority of the cases with the two largest bin widths. The rejection rate decreases sharply when the bin width is decreased and with the smallest bin width the rejection is just above 6%.

We have also tested the non-parametric version of the test. In Fig. 15, we plot the average p-value obtained from the Kolmogorov–Smirnov test with different interval sizes. The costs are from uniform distribution and the number of tenders is set to 200. With smaller interval sizes the p-values are higher indicating that the null of competition is correctly not rejected. When the interval size is 25% or more then the average p-value is below .1 indicating that in many cases the null hypothesis is incorrectly rejected.

The results of the Monte Carlo simulation demonstrate the importance of choosing a small enough bin width. With 5% bin width, the largest we use, the null hypothesis is rejected at $p < 0.05$ in around 40% of the time. Once the width of the bin is decreased, the share of significant results starts to converge to the correct rate of 5%. With the chosen parameters, the correct rejection rate is achieved with a bin width of around 0.5%. The Monte Carlo simulation also demonstrates that if the sample size is increased, the bin width required to reject the null under competition decreases. In very large samples this could mean that the required bin width approaches zero, but in smaller samples, the required bin width is reasonable. Given that, in practice, most of the screened markets do not have millions of tenders, this is unlikely to be a major issue.

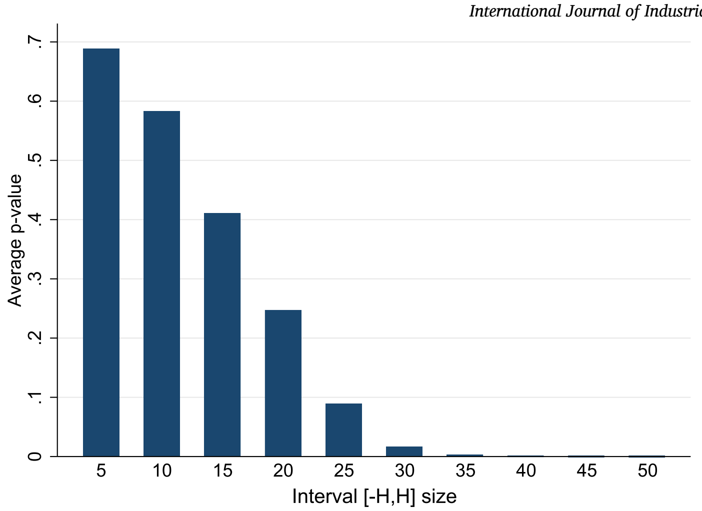

Fig. 15. Monte Carlo Simulation results for non-parametric test.

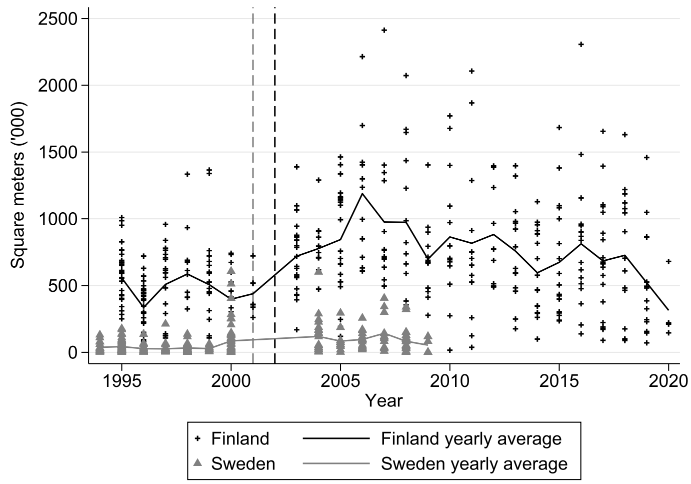

Fig. 16. Development of project size over time.
This figure depicts the development of project size over time. Each individual dot represents a single tender and the line presents the development of yearly average prices.

A.6. Development of project size over time

We observe the area paved for around 85% of the tenders for the Finnish data and around 80% for the Swedish data. In Fig. 16 we plot the development of project size over time. Two interesting observations arise from the Figure. First, there is substantial heterogeneity in project size within a year. The largest projects can be over ten times larger than the smaller projects. The second interesting observation is that in both countries the average project size increases over time. In both countries the project size is around twice larger after the cartel than during the cartel.

A.7. Cartel overcharge

In this Section, we briefly discuss the price development before and after the cartel in Finland and Sweden. As asphalt projects differ in size, as well as other details of the project, a simple comparison of prices before and after the cartel is unlikely to be informative on the actual price effects of the cartel.

For the Swedish data, the information on project characteristics is scarce. We only observe the size of the project measured in square meters. Using the same data as we, Bergman et al. (2020) report that the price per square meter is around 30% higher during the cartel period. We have also run a simple regression of the logarithm of the winning bid on the bitumen index, area and a cartel indicator. The point-estimate for the cartel indicator is .29 indicating a price increase of around 30% during the cartel. However, caution should be taken in interpreting this causally. There is substantial heterogeneity in the price per square meter across projects.

Fig. 17. Price development before and after the cartel in Finland.
This figure depicts the development of the bitumen index and the average per square meter price in the Finnish data. Bitumen index originates from the monthly level data merged to projects and averaged across the projects in a given year.

Table 22
Cartel overcharge estimate for Finland.

	Log price	Log price
Cartel	0.195*	0.254***
	(0.105)	(0.093)
Area / 10000	0.028***	0.023***
	(0.003)	(0.002)
Area^2 / 10000	-0.000***	-0.000***
	(0.000)	(0.000)
Bitumen	0.000	-0.001
	(0.001)	(0.001)
Consumer Price Index	0.025**	0.032***
	(0.011)	(0.009)
Bitumen compensated	0.798***	0.716***
	(0.082)	(0.075)
Asphalt mass / 10000		0.065***
		(0.008)
r2	0.82	0.88
N	256	225

The dependent variable is the logarithm of price. Significance at $p < 0.10$ (*), $p < 0.05$ (**), and $p < 0.01$ (***).

Hence, changes in the unobserved project characteristics can explain some of the observed difference in average price over the two periods.[33]

For the Finnish data we have more detailed information on project characteristics which allows us to conduct a more thorough analysis of the price effects of the cartel. We start by plotting the average price per square meter and the bitumen index in Fig. 17.

As seen in the figure, before the dawn raids in 2002 the average prices and bitumen index follow each other closely. However, after 2002 we see a clear drop in the average price with no visible change in the bitumen index. This provides us with the first indication that, as expected, the breakdown of the cartel resulted in lower prices in the Finnish asphalt market.

In Table 22 we provide results on the estimated cartel overcharge using several project characteristics as controls. These include the size of the project measured in area and asphalt mass, the bitumen price index, and an indicator variable which equals one if the price of the contract is tied to the development of the bitumen price (i.e., the buyer carries the risk of bitumen price changes). Our choice of control variables is guided by the report of the VATT Institute for Economic Research, which was used in court proceedings related to cartel damages. Following the analysis by VATT, we also include tenders from year 2002 to the sample. The tenders organized before March we categorize as collusive and the rest as competitive.

Based on the results, the price level was around 20% to 30% higher during the cartel period. Our results differ slightly from those reported in VATT Institute for Economic Research (2011). This is due to a slightly smaller sample, because in our paper we exclude the tenders with only one bidder or where the price information is not available for all bids.

[33] The variation is particularly large in smaller projects.

A.8. Legal proceedings in the Swedish asphalt cartel

The Market Court's judgments, the summons applications by the Swedish Competition Authority, and the Swedish Competition Authority's own summary of the case are all in Swedish. To facilitate the English-reading audience to get a better understanding of the legal process, we have chosen to provide a English translation of parts of the documentation below.

A.8.1. Summons application (Stockholm District Court (2003) Dnr 341/2003, pp. 5–8.)

The companies NCC AB, Vägverket, Skanska Sverige AB, Peab Asfalt AB, Peab Sverige AB, Peab Asfalt Syd AB, Sandahls Grus & Asfalt AB, Kvalitetsasfalt i Mellansverige AB, Svenska Väg AB, ODEN Entreprenad AB, and Bygg och Miljö i Östergötland AB have intentionally or negligently violated the prohibition in Section 6 of the Competition Act (KL) by entering into agreements or applying coordinated practices aimed at significantly hindering, restricting, or distorting competition, or resulting in such an outcome. The violations are specified in points i–xiii and are not minor.

The companies have entered into agreements or coordinated as follows:

i. NCC AB, Vägverket, Skanska Sverige AB, and Peab Asfalt AB have participated in market division, directly or indirectly set prices, exchanged information on prices and volumes, and decided that they would exchange information on bids and decide which company would win the contract for state, municipal, and private procurements regarding asphalt paving. The violation, as far as NCC AB, Skanska Sverige AB, and Peab Asfalt AB are concerned, covered the period from July 1, 1993, to October 24, 2001, and for Vägverket from May 1, 1995, to October 24, 2001.

ii. II Sandahls Grus & Asfalt AB, Kvalitetsasfalt i Mellansverige AB, Svenska Väg AB, PNB Asfalt AB, Bygg och Anläggning P-A Berglund AB, and ODEN Asfalt AB have – to the extent specified under each procurement below (a–f) – participated in market division, directly or indirectly set prices, exchanged information on bids, and accepted and acted in accordance with the improper conduct described in point I, regarding which company would win the contract in the following procurements:

   (a) Sandahls Grus & Asfalt AB, Bygg och Anläggning P-A Berglund AB, Svenska Väg AB, Kvalitetsasfalt i Mellansverige AB, and ODEN Asfalt AB regarding asphalt paving at the Swedish National Road Administration, Southeast region, in 1998.
   (b) Bygg och Anläggning P-A Berglund AB regarding asphalt paving at the Swedish National Road Administration, Southeast region, in 1999, 2000, and 2001.
   (c) Sandahls Grus & Asfalt AB, Svenska Väg AB, PNB Asfalt AB, and Kvalitetsasfalt i Mellansverige AB regarding asphalt paving at the Swedish National Road Administration, West region, in 2001.
   (d) Sandahls Grus & Asfalt AB, Svenska Väg AB, Kvalitetsasfalt i Mellansverige AB, and ODEN Asfalt AB regarding asphalt paving at the Swedish National Road Administration, Mälardalen region, in 2000.
   (e) Sandahls Grus & Asfalt AB, Svenska Väg AB, and Kvalitetsasfalt i Mellansverige AB regarding asphalt paving at the Swedish National Road Administration, Mälardalen region, in 2001.
   (f) Sandahls Grus & Asfalt AB and Kvalitetsasfalt i Mellansverige AB regarding asphalt paving at the Swedish National Road Administration, Central region, in 2001.

iii. NCC AB, Skanska Sverige AB, Vägverket, and Peab Mellersta AB, prior to the Swedish National Road Administration's procurement for the reconstruction of an interchange on the E4 road at Linköping in Östergötland County (Ullevirondellen) in 1998, participated in market division, directly or indirectly set prices, exchanged information on bids, and decided which company would win the contract. NCC AB, Skanska Sverige AB, Vägverket, and Peab Mellersta AB also exchanged information with and persuaded another company to participate in market division and price setting for bids.

iv. Peab Mellersta AB, NCC AB, Skanska Sverige AB, and Vägverket, prior to Åtvidaberg Municipality's procurement in 1998 for the construction of a new pedestrian and bicycle bridge, participated in market division, directly or indirectly set prices, exchanged information on bids, and decided which company would win the contract. Peab Mellersta AB, NCC AB, Skanska Sverige AB, and Vägverket also exchanged information with and persuaded another company to participate in market division and price setting for bids.

v. Peab Sverige AB, NCC AB, Skanska Sverige AB, and Vägverket, prior to Linköping Municipality's procurement in 2000 for maintenance paving and reconstruction of bus stops, participated in market division, directly or indirectly set prices, exchanged information on bids, and decided which company would win the contracts.

vi. Bygg och Miljö i Östergötland AB participated in market division, directly or indirectly set prices, exchanged information on bids, and accepted and acted in accordance with the improper conduct described in point V regarding which company would win the contracts.

vii. Peab Mellersta AB, NCC AB, and Skanska Sverige AB, prior to Kuwait Petroleum Svenska AB's procurement in 1998 for a project named "Q8 Östgötaporten Mark," participated in market division, directly or indirectly set prices, exchanged information on bids, and decided which company would win the contract. Peab Mellersta AB, NCC AB, and Skanska Sverige AB also exchanged information with and persuaded another company to participate in market division and price setting for bids.

viii. Peab Mellersta AB, NCC AB, and Skanska Sverige AB, prior to Tekniska Verken i Linköping AB's procurement in 1998 for a ramp at Gärstadverken in Linköping, participated in market division, directly or indirectly set prices, exchanged information on bids, and decided which company would win the contract. Peab Mellersta AB, NCC AB, and Skanska Sverige AB also exchanged information with and persuaded another company to participate in market division and price setting for bids.

ix. NCC AB and Bygg och Anläggning P-A Berglund AB entered into agreements or coordinated on market division regarding Emmaboda Municipality during the period 1997–2000, Tingsryd and Karlshamn Municipalities during the period 1998–2000, and Värnamo, Vetlanda, and Vaggeryd Municipalities during the period 1997–2000.
x. NCC AB and Bygg och Anläggning P-A Berglund AB entered into agreements or coordinated on limiting the production and sale of asphalt during the period January 1, 2000, to October 24, 2001.
xi. NCC AB and Bygg och Anläggning P-A Berglund AB, in addition to the violation under point X, have entered into agreements or coordinated on market division and the limitation of production and sale of asphalt in 2000.
xii. NCC AB and Skanska Sverige AB, in December 1999, entered into agreements or coordinated to make an agreement with another company regarding the limitation of production and sale of asphalt. 12. XIII NCC AB, Skanska Sverige AB, and Vägverket, in January 2001, entered into agreements or coordinated on market division regarding tank coating operations.

A.8.2. Market Court's decision (Market Court's judgment, pages 99–102)

The Market Court finds that the companies were guilty of violations of Section 6 of the Competition Act (KL) and Article 81.1 of the EC Treaty as follows:

- NCC and Peab Asfalt, together with other companies, divided the Swedish National Road Administration contracts for paving works among themselves and, in most cases, agreed on the bid price that the company winning the respective contract would state in its bid in the Southeast region in 1997, 1998, and 2000, in the West region in 2000 and 2001, in the Mälardalen region in 2000 and 2001, and in part of the Central region in 2001.
- NCC and Peab Asfalt, together with other companies, agreed on the division of the six municipal procurements for paving works in the municipalities of Landskrona, Ystad, Klippan, Kristianstad/Bromölla, Hässleholm, and Helsingborg/Höganäs in 2000, and that the other companies, which would not win the procurement, would refrain from competing by submitting "taken prices".
- NCC and Peab Asfalt, together with other companies in the Southeast region in 1997, agreed with [...] and [...] that the latter companies would refrain from competing for the Swedish National Road Administration procurements in exchange for compensation.
- In the Southeast region in 1998, agreements were made between, on the one hand, NCC and Peab Asfalt and other companies, and on the other hand, [...] and [...] (now Peab Asfalt Syd), that [...] and [...], would refrain from competing for the Swedish National Road Administration procurements in exchange for compensation.
- NCC and Peab Asfalt, together with other companies in the Southeast region in 2000, agreed with a company that it would refrain from competing in the Swedish National Road Administration procurements in exchange for compensation. NCC and Peab Asfalt also, together with other companies, entered into a three-year agreement with [...] that [...] would refrain from competing for the Swedish National Road Administration procurements and limit its production of asphalt in the region in exchange for compensation, and that [...] would win the procurements made by the municipalities of Emmaboda, Tingsryd, and Karlshamn, and that the three-year agreement would be renegotiated.
- NCC and Peab Asfalt, together with other companies, made agreements with PNB Asfalt (now Peab Asfalt Syd), [...] and other companies that these companies would not compete for the Swedish National Road Administration procurements in the West region in 2001. Furthermore, NCC and Peab Asfalt, together with other companies, made a special agreement with [...], that [...] would win a group.
- NCC and Peab Asfalt, together with other companies, agreed with [...] that [...] would not compete for the Swedish National Road Administration procurements in the Mälardalen region in 2000 and that [...] would receive so-called p-money for this.
- The individual violations reported above, as far as NCC and Peab Asfalt are concerned, should be seen as a single continuous violation in breach of Section 6 of the Competition Act (KL) during the period 1997 to February/March 2001 and also in breach of Article 81.1 of the EC Treaty during the period January 1 to February/March 2001.
- NCC, together with another company, in connection with Götene Municipality's procurement in 1999 and Ulricehamn Municipality's and Ekås Road Association's procurements in 2000 for asphalt paving works, exchanged information on bid prices, which is considered improper coordinated practices in breach of Section 6 of the Competition Act.
- NCC and [...] in 2000, in breach of Section 6 of the Competition Act, made an agreement that [...] would not establish an asphalt plant in Östergötland.
- NCC and two other companies, in 2001, in breach of Section 6 of the Competition Act, agreed that NCC would dispose of certain equipment in exchange for being allowed to participate in the distribution of paving contracts within this type of operation.
- NCC and Peab Sverige, together with other companies, in breach of Section 6 of the Competition Act, agreed on which procurement each company would win and the bid prices the companies would submit for five procurements (packages) for maintenance paving and reconstruction of bus stops that Linköping Municipality made in 2000.

References

Abrantes-Metz, Rosa, Froeb, Luke, Geweke, John, Taylor, Christopher T., 2006. A variance screen for collusion. Int. J. Ind. Organ. 24 (3), 467–486.
Angrist, Joshua D., Pischke, Jörn-Steffen, 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.
Aryal, Gaurab, Gabrielli, Maria F., 2013. Testing for collusion in asymmetric first-price auctions. Int. J. Ind. Organ. 31 (1), 26–35.
Asker, John, 2009. A study of the internal organization of a bidding cartel. Am. Econ. Rev. 100 (3), 724–762.
Bajari, Patrick, Ye, Lixin, 2003. Deciding between competition and collusion. Rev. Econ. Stat. 85 (4), 971–989.
Bajari, Patrick, Houghton, Stephanie, Tadelis, Steven, 2014. Bidding for incomplete contracts: an empirical analysis of adaptation costs. Am. Econ. Rev. 104 (4), 1228–1319.

Baltrunaite, Audinga, 2020. Political contributions and public procurement: evidence from Lithuania. J. Eur. Econ. Assoc. 18 (2), 541–582.
Baránek, Bruno, Musolff, Leon, Titl, Vítězslav, 2021. Detection of Collusive Networks in E-procurement. Working paper.
Bergman, Mats A., Lundberg, Johan, Lundberg, Sofia, Stake, Johan Y., 2020. Interactions across firms and bid rigging. Rev. Ind. Organ. 56, 107–130.
Bolotova, Yuliya V., 2009. Cartel overcharges: an empirical analysis. J. Econ. Behav. Organ. 70 (1–2), 321–341.
Bosio, Erica, Djankov, Simeon, Glaeser, Edward, Shleifer, Andrei, 2022. Public procurement in law and practice. Am. Econ. Rev. 112 (4), 1091–1117.
Boyer, Marcel, Kotchoni, Rachidi, 2015. How much do cartel overcharge? Rev. Ind. Organ. 47, 119–153.
Bryant, Peter G., Eckard, E. Woodrow, 1991. Price fixing: the probability of getting caught. Rev. Econ. Stat. 73 (3), 531–536.
Chassang, Sylvain, Kawai, Kei, Nakabayashi, Jun, Ortner, Juan, 2022. Robust screens for noncompetitive bidding in procurement auctions. Econometrica 90 (1), 315–346.
Chernozhukov, Victor, Fernández-Val, Iván, Blaise, Melly, 2013. Inference on counterfactual distributions. Econometrica 81 (6), 2205–2268.
Clark, Robert, Coviello, Decio, De Leverano, Adriano, forthcoming. Complementary bidding and the collusive arrangement: evidence from an antitrust investigation. Am. Econ. J. Microecon.
Combe, Emmanuel, Monnier, Constance, Legal, Renaud, 2008. Cartels: the probability of getting caught in the European Union. Bruges Eur. Econ. Res. Pap. 12.
Conley, Timothy G., Decarolis, Francesco, 2016. Detecting bidders groups in collusive auctions. Am. Econ. J. Microecon. 8 (2), 1–38.
Connor, John M., Bolotova, Yuliya, 2006. Cartel overcharges: survey and meta-analysis. Int. J. Ind. Organ. 24 (6), 1109–1137.
Coviello, Decio, Gagliarducci, Stefano, 2017. Tenure in office and public procurement. Am. Econ. J.: Econ. Policy 9 (3), 59–105.
De Leverano, Adriano, 2023. Clustered bids in first-price auctions: collusion or competition? Econ. Lett. 233.
Dong, Ailin, Massa, Massimo, Zaldokas, Alminas, 2019. The effects of global leniency programs on margins and mergers. Rand J. Econ. 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., Block, Michael K., Nold, Frederick C., 1985. Asymmetric information and collusive behavior in auction markets. Am. Econ. Rev. 75 (3), 441–460.
Ferraz, Claudio, Finan, Frederico, Szerman, Dimitri, 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).
Finnish Competition and Consumer Authority (FCCA), 2011. Esitys markkinaoikeudelle asiassa NCC Roads Oy / Destia Oy: n sekä Destia Kalusto Oy: n asfalttipäällysteliiketoiminta (in Finnish).
Fortin, Nicole M., Lemieux, Thomas, Lloyd, Neil, 2021. Labor market institutions and the distribution of wages: the role of spillover effects. J. Labor Econ. 39 (2), 369–412.
Froeb, Luke M., Koyak, E. Robert A., Werden, Gregory J., 1993. What is the effect of bid-rigging on prices? Econ. Lett. 42 (4), 419–423.
García Rodríguez, Manuel J., Rodríguez-Montequín, Vicente, Ballesteros-Perez, Pablo, Love, Peter E.D., Signor, Regis, 2022. Collusion detection in public procurement auctions with machine learning algorithms. Autom. Constr. 133.
Giuffrida, Leonardo M., Rovigatti, Gabriele, 2022. Supplier selection and contract enforcement: evidence from performance bonding. J. Econ. Manag. Strategy 31 (4), 980–1019.
Halonen, Kirsi-Maria, Tukiainen, Janne, 2020. Competition and Litigation in Swedish Public Procurement. Uppdragsforskningsrapport, vol. 1. Konkurrensverket.
Harrington, Joseph E., Wei, Yanhao, 2017. What can the duration of discovered cartels tell us about the duration of all cartels? Econ. J. 127 (604), 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. Masters thesis. Stockholm School of Economics (in Swedish).
Huber, Martin, Imhof, David, 2019. Machine learning with screens for detecting bid-rigging cartels. Int. J. Ind. Organ. 65, 277–301.
Huber, Martin, Imhof, David, 2023. Flagging cartel participants with deep learning based on convolutional neural networks. Int. J. Ind. Organ. 89, 29–46.
Huber, Martin, Imhof, David, Ishii, Rieko, 2022. Transnational machine learning with screens for flagging bid-rigging cartels. J. R. Stat. Soc., Ser. A 185 (3), 1074–1114.
Imhof, David, 2020. Detecting bid-rigging cartels with descriptive statistics. J. Compet. Law Econ. 15 (4), 427–467.
Imhof, David, Wallimann, Hannes, 2021. Detecting bid-rigging coalitions in different countries and auction formats. Int. Rev. Law Econ. 68, 106016.
Imhof, David, Karagök, Yavuz, Rutz, Samuel, 2018. Screening for bid rigging—does it work? J. Compet. Law Econ. 14 (2), 235–261.
Jääskeläinen, Jan, Tukiainen, Janne, 2019. Anatomy of Public Procurement. VATT Working Papers, 18.
Kapitalet podcast, 2019. 117: Asfaltkartellen, del 1 – Att bygga en kartell. Podcasts.nu, Internet Archive. https://web.archive.org/web/20241008145546/. https://podcasts.nu/avsnitt/kapitalet-en-podd-om-ekonomi/asfaltkartellen-del-1-att-bygga-en-kartell, 2019.
Kawai, Kei, Nakabayashi, Jun, 2022. Detecting large-scale collusion in procurement auctions. J. Polit. Econ. 130 (5), 1364–1411.
Kawai, Kei, Nakabayashi, Jun, Shimamoto, Daichi, 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 J. Econ. 26 (3), 398–417.
Lee, Munseob, 2022. Government Purchases and Firm Growth. Working paper.
Lindberg, Leena, 2020. Kaksikymmentä vuotta asfalttikartelliprosesseja (in Finnish). Def. Legis 4, 632–649.
Lundberg, Johan, 2017. On cartel detection and Moran's I. Lett. Spat. Resour. Sci. 10, 129–139.
McAfee, Preston, McMillan, John, 1992. Bidding rings. Am. Econ. Rev. 82 (3), 579–599.
OECD, 1999. Annual Report on Competition Policy Developments in Denmark.
OECD, 2002. Competition law and policy in Norway. Norway's report to the OECD for the year 2002.
Porter, Robert, Zona, John Douglas, 1993. Detection of bid rigging in procurement auctions. J. Polit. Econ. 101 (3), 518–538.
Porter, Robert, Zona, John Douglas, 1999. Ohio school milk markets: an analysis of bidding. Rand J. Econ. 30 (2), 263–288.
Silveira, Douglas, Vasconcelos, Silvinha, Resende, Marcelo, Cajueiro, Daniel O., 2022. Won't Get Fooled Again: a supervised machine learning approach for screening gasoline cartels. Energy Econ. 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).
Wallimann, Hannes, Imhof, David, Huber, Martin, 2022. A machine learning approach for flagging incomplete bid-rigging cartels. Comput. Econ.