

Unclassified

English - Or. English

3 November 2022

**DIRECTORATE FOR FINANCIAL AND ENTERPRISE AFFAIRS
COMPETITION COMMITTEE**

Working Party No. 3 on Co-operation and Enforcement

Data Screening Tools for Competition Investigations – Note by Denmark

28 November 2022

This document reproduces a written contribution from Denmark submitted for Item 3 of the 136th OECD Working Party 3 meeting on 28 November 2022.

More documents related to this discussion can be found at
www.oecd.org/daf/competition/data-screening-tools-for-competition-investigations.htm

Ms Despina PACHNOU
[Email: Despina.PACHNOU@oecd.org]

JT03506563

Denmark

1. The Danish Competition and Consumer Authority (DCCA) uses digital screening tools and methods to detect anticompetitive conduct, through the development and utilization of a software (“Bid Viewer”) to identify potential collusion in public tenders. Bid Viewer utilizes computational screening methods, including machine learning and artificial neural networks, and is designed to uncover suspicious patterns in large public procurement datasets. The use of this tool and other enforcement efforts at the DCCA seek to ensure that competition for public contracts is fair and that public funds are used efficiently.

2. An article about the tool Bid Viewer has been published by the DCCA¹ where methodologies and use of screens are generically described and explained. Additionally, the tool has been presented at several international conferences, for example the CMA Data, Technology and Analytics Conference 2022² and the ICN AEWG Economists webinar³.

3. Bid Viewer is developed in-house to ensure full control over the methods, enable code sharing with other national competition authorities, and longevity of know-how. In brief, in the case of coordination between competitors, systematic patterns can be observed. It is possible to develop computational methods (“screens”) that flag suspicious bids and potentially coordinating companies. Both theoretical and empirical models, published by academic researchers, have inspired the development of Bid Viewer.

4. The screens implemented and utilized are based on such empirical evidence of cartel behavior and the models are trained and validated on datasets with known collusive and non-collusive tenders. Data required for such screens include tender and company identifiers, all winning and losing bid sums, tender date, geographical location, and market and submarket identifiers for each tender. Three complementary methods are used as screens to flag tenders and companies with potentially suspicious collusive bidding patterns:

1. Statistical indicators derived from all the bids of a tender, such as the normalized relative distance, percent difference, coefficient of variation, skewness, and kurtosis. The value of each indicator is either associated with competitive or non-competitive bidding behavior. These indicators are described in detail in the article¹.
2. Statistical indicators are used as input into machine learning screens, which combine multiple indicators into a model. It is possible to apply such models with the aim of identifying price fixing¹. In brief, the models are trained and evaluated on tender data including known cartels, and because the model assigns a higher probability score to tenders with known collusion it is possible to apply the model

¹ Kultima Roat, J. (2022), “Collusion detection in public procurement using computational methods”, Danish Competition and Consumer Authority: Competitive Markets and Consumer Welfare 56, https://www.en.kfst.dk/media/cnldn11q/bid-viewer_56_seneste.pdf

² Collusion detection in public procurement using computational methods (June 2022) (2:38:33) <https://www.youtube.com/watch?v=Iovsp5aHcuU>

³ AEWG Economists webinar (March 2022) (18:55) <https://www.youtube.com/watch?t=1134&v=cJ1Nsn82vS8&feature=youtu.be>

to national procurement data from other sources and assume that tenders with a higher score may have a higher risk of collusion.

3. Company bidding pattern analysis, of both individual businesses and groups of companies. Typically, this analysis will be performed by comparing companies' bids and participation in tendering processes over time, geographic regions, and markets. Illustrated examples are presented in the article¹. They include identification of bidders' price fixing through alternating winners, collusion through geographic and non-geographic market sharing, and when companies coordinate and agree to participate in tenders without the intention of winning ("sham" bidding). Patterns of alternating winners can be identified by observing if certain companies take turn having the lowest bid in shared tenders. Market sharing identifies when companies are expected to bid in the same markets but only appear to bid separately in submarkets. "Sham" bidding is identified by comparing the number of expected and actual wins and losses a company has, and if these substantially differs, this may be indicative of potential coordination between companies.
5. The screens help differentiate between non-collusive bidding patterns and potentially unlawful, collusive bidding patterns. Suspicious tenders and companies are flagged and are further investigated prior to opening and proceeding with an investigation. Subsequent investigatory steps include determining whether the observed bidding patterns can be explained by external factors not included in the computational analysis.
6. The screens are promising and correctly flags suspicious behavior in previously published datasets with known cartels. Several such flags have also been raised in national procurement datasets. DCCA continues to evaluate the flags' accuracy and usefulness as starting points for a traditional case investigation. So far, no cartels have been identified with the use of Bid Viewer.
7. The development and use of Bid Viewer has multifaceted benefits beyond identifying potentially suspicious behavior using the screens locally in Denmark. The existing international collaborations with code and data sharing benefits both authorities from countries where *ex officio* screening methodologies already are an integrated part of the authority's work through a shared tool to co-develop and utilize nationally, as well as providing a head start to authorities in countries without prior experience but ready to explore screening methods nationally.
8. There are many lessons learnt from this on-going project. First, it is not possible for a single national authority to stay up to date on newly published methods and continuously develop and implement new screens. The DCCA engages actively with academic researchers as well as other national competition authorities to ensure the accuracy, longevity, and validation of implemented screens. Second, it is important to utilize complementary screening methods. It is unlikely that any two cartels operate identically and therefore it is important to use a variety of methods developed and tested on data from different markets and countries. Third, a screening tool must be versatile to also allow exploratory analysis of the data. Flagging tenders and companies by one or multiple screens is generally only the first step in an analysis. A screening tool should also be able to further, in-depth explore and arrive at potential explanations of why and how these companies and tenders were flagged.
9. The DCCA uses screening tools to support many kinds of investigations. Specifically Bid Viewer is used to identify bid rigging, i.e., horizontal agreements.
10. Data about public tenders are collected from public authorities manually and at different occasions, compared to, for example, a centralized system containing the tender

data. The collected information generally includes information about each tender, the winning and losing bids, information about the companies and, if needed, more detailed information about each bid for a subset of tenders.

11. Often the data must be pre-processed by extracting it from the public authority's journalizing system in formats such as PDF, Word and Excel documents, into a structured table-like format with information about each bid. This process is generally lengthy and non-trivial. At times web scraping has been used to gather data from public sources.

12. Generally, non-public procurement data required for screening include both winning and losing bids and information about the participating companies. This data is commonly collected through voluntary collaboration with the public authorities. Currently in Denmark there is no single data source with procurement data available.

13. Our screening team consists of several staff members with varying backgrounds such as data scientists, investigators, lawyers, and economists.

14. Our tool, Bid Viewer, is developed in close collaboration with other national competition authorities inside and outside the EU. With the collaborators, we share code, data, know-how and screening experiences.

15. The greatest benefits of using a screening tool are that it is a proactive method to identify collusive behavior of selected markets, enabling detection of otherwise undetectable collusion and targeted investigations. A sophisticated screening tool also conserve agency resources because multiple markets and companies can be screened simultaneously. The development of these screens and methods also provides in-house know-how that can be applied to other internal data science projects.

16. Currently, easy access to large quantities of structured data with winning and losing bids from national public authorities remain the greatest challenge. Generally, some data exists where only the winning bid is known, but such data poses a great challenge because most relevant screens also require the losing bids.

17. A second challenge is access to appropriate training data from historical procurements over time periods with and without known collusion. Such data is used to e.g., train machine learning models.

18. The challenge of acquiring the in-house expertise has currently been mitigated through international collaboration and an active international network.