Organisation for Economic Co-operation and Development

DAF/COMP/WP3/WD(2022)40

Unclassified

English - Or. English

28 November 2022

DIRECTORATE FOR FINANCIAL AND ENTERPRISE AFFAIRS
COMPETITION COMMITTEE

Cancels & replaces the same document of 25 November 2022

Working Party No. 3 on Co-operation and Enforcement

Data Screening Tools for Competition Investigations – Summaries of Contributions

28 November 2022

This document reproduces summaries of contributions submitted for Item 3 of the 136th OECD Working Party 3 meeting on 28 November 2022.

More documentation 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]

JT03508793

Table of contents

Summaries of contributions

This document contains summaries of the various written contributions received for the discussion on Data Screening Tools for Competition Investigations (136th OECD Working Party 3 meeting on 28 November 2022). When the authors did not submit their own summary, the OECD Competition Division Secretariat summarised the contribution. Summaries by the OECD Secretariat are indicated by an *.

Australia

The Australian Competition and Consumer Commission (ACCC) is actively exploring the use of data screening tools in competition investigations. This forms part of a suite of initiatives the ACCC is presently engaged in to drive proactive detection of competition issues.

Like many other competition regulators, the ACCC's initial focus is on identifying possible collusion in the context of public procurement. Public procurement has long been considered susceptible to cartel conduct due to the fact that procurement transactions are highly structured and transparent. The OECD estimates its member countries spend approximately 12% of their GDP via public procurement, and that the elimination of bid rigging could help reduce procurement prices by 20% or more.

The ACCC is in the process of building a "cartel screening tool" and obtaining procurement data on which to apply it. This is a collaborative project between investigators, economists, intelligence analysts and data scientists.

In this submission we provide some reflections about the two fundamental components that are needed when building a proactive detection capability. Firstly, a methodology, which encompasses a definition, the build, and collaboration. The second component is data, with a focus on access, quality, and advocacy.

BIAC

Business at OECD (BIAC) welcomes the opportunity to provide its views to the OECD Competition Committee WP3 for the roundtable on data screening tools in competition investigations. Indeed, the use of empirical methods to analyze economic data, such as price developments, bidding patterns and market share fluctuations to detect suspicious instances of cartel behavior or other anticompetitive conduct has over the past few years gained significant practical importance.

Despite its reservations that are further set out in these comments, BIAC is of the opinion that the prudent use of screening mechanisms may not only bring direct benefits in the form of more effective cartel enforcement but may also lead to indirect benefits, such as enhanced compliance. At the same time, however, the price screens which reflect parallel behavior should not automatically be equated with anticompetitive conduct, even when those tools reflect direct interaction with competitor pricing activity, as "price follower" strategies, decisions not to bid, and intentionally high bids, absent express agreement among the parties, all may constitute legitimate forms of long-run competitive behavior.

BIAC notes that the methodology underlying a variety of the structural screening methods is subject to criticism. The thrust of this critique is that the methodologies applied may not reliably indicate the presence of cartel or otherwise anticompetitive behavior and may give rise to both false positives and negatives. In this respect, BIAC is particularly concerned that pro-active detection techniques may give rise to false positives, which in turn may inflict significant cost on enforcers and businesses without benefit. At the extreme end of the spectrum, there is also a risk that the use of screens may result in disproportionate and excessively wide-ranging ex officio investigations.

Proactive enforcement tools, including a careful and informed use of economic methodologies such as tailored detection screens, may in some specific cases be efficient and may even be desirable, provided adequate procedural safeguards for the companies under scrutiny are in place. Accordingly, in BIAC's view, the proper question should not only be which techniques and approaches are best suited to detect anticompetitive conduct in specific settings, but also which procedural safeguards should be put in place to ensure that the voice of companies under investigation are taken into account in a timely and appropriate manner.

Brazil

The Brazilian antitrust screening instrument, Project Cérebro, was launched in 2014 to develop new investigative tools. Its implementation posed first challenges related to (i) hiring and training civil servants, (ii) database access, and (iii) creating an environment for innovation.

During its initial phase, the main guidelines for the project activities were focus (immediately testing the techniques in the available databases), transparency (always documenting the development and testing of the models), and partnerships (sharing techniques and models with bodies of criminal prosecution and external control).

In 2016, CADE began an investigation solely based on the results provided by Cérebro. After applying two screening models (Benford and Entropy), the authority identified 16 suspicious companies that attended 4,400 tenders. In 2020, after the Federal Police Department carried out Operation Meeting Point to serve search warrants in the head office of 14 companies, CADE initiated an administrative proceeding against 42 companies and 43 individuals involved in a large cartel in public and private procurement for outsourced services.

Along its trajectory, Project Cérebro has enabled its team to find some points that can require particular attention in the coming years, which regard (i) development (to find an equilibrium between designing new filters and adapting those already tested and validated by experts), (ii) team (internal and external competition for professionals in the area), (iii) governance (mapping the real needs of the areas responsible for investigating cases), (iv) data (having prompt access to data and with the required levels of quality and granularity), and (v) boundaries (to assess the adoption of AI-based solutions to boost the efficiency of investigations or improve the detection of antitrust violations).

Canada

The Competition Bureau ("Bureau") uses various tools and approaches for cartel detection. This includes data screening algorithms, which can be used to find cartel-like conduct from bidding data. Data screening algorithms can be a proactive method of detecting and fighting cartels.

The main challenge for the Bureau in developing data screening algorithms is accessing bidding data. Technological advancements and improvements in data access and data portability will bring benefits to data screening algorithms.

Colombia

The Competition Authority's data analytics team has developed two digital screening tools: Sherlock and Inspector.

SIC has identified the need to harness the increasing data available and has developed tools that would aid investigators to assess relevant data systematically, thus increasing the rate of detection of anticompetitive disruptions to markets, whether they come from regulators or from market agents themselves.

The SIC has focused on enhancing the detection functions for industries and markets where data is not only publicly available but also offered in different forms. For example, public information available on the Colombian public procurement system or at the webpages of the regulators, depending on the case.

Sherlock is a project that seeks to support SIC's investigators when it comes to identifying signs or patterns that suggest potential anticompetitive behaviours with the data available from public procurement processes. First, Sherlock facilitates the access of investigators to public data available on the web, the software cleans, structures, and transforms the collected data and then loads it into a database within the SIC. Second, Sherlock involves automation search of the above mentioned signs and patterns within the bulk of public procurement data available online. The data collection is performed after the definition of variables relevant for investigators and identifiable via CCE's API called SOCRATA. Finally, at this moment the tool works generating alerts[1] and will have the capacity of providing informed red flags.

Inspector reviews the publication pages of the monitored entities in search of new regulatory projects, performs in the entire value chain, identifying possible new regulatory projects and issuing a consideration on the impact of such regulation on free competition in the market corresponding to the monitored entity. This analytic data tool streamline and facilitate the review process of regulatory projects issued by regulatory entities by the members of the Competition Advocacy Working Group of the Deputy Superintendence for Competition Protection.

- [1] When the awarded budget is equal to or less than 5% of the official budget
- When the same bidder is awarded three or more times in a selection process within the same State Entity (concurrence), identifying within the objects key words in accordance with the economic sectors
- When there is only one bidder identified in the process

Denmark

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.

Three complementary methods are used as screens to flag tenders and companies with potentially suspicious collusive bidding patterns: Statis-tical indicators derived from all the bids of a tender, machine learning screens which combine multiple indicators into a model, and company bidding pattern analysis, of both individual businesses and groups of companies.

The screens help differentiate between non-collusive bidding patterns and potentially unlawful, collusive bidding patterns. 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.

The development and use of Bid Viewer has multifaceted benefits be-yond identifying potentially suspicious behavior using the screens locally in Denmark. 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.

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.

Ecuador

This document presents the experience of the Superintendency for Market Power Control of Ecuador (SCPM) in the implementation of the association rule algorithm (a priori) for identifying leads of collusion in public procurement processes; in this sense, this paper details the methodology used and the justifications for choosing this methodology, the process that was carried out, the information sources used, the description of the algorithm functioning and the limitations of the implementation.

The experiences that are gathered in this document derive from the analysis carried out for the Market Study of the Public Procurement National System done by the SCPM during 2021 – 2022, which proposes the use of these quantitative methodologies due to the findings that neither the regulator nor the contracting entities within the scope, have tech tools that can issue alerts about different leads of collusive agreements that may arise on these processes; a situation that make the system vulnerable to possible anticompetitive practices.

In light of what has been mentioned, the a priori algorithm (tool proposed by the Superintendency) was used to identify the behavior patterns of bidders and winners of procurement processes that could provide leads regarding anticompetitive practices done by undertakings within the procurement process for market sharing. This algorithm of data mining allows the identification of sets of frequent elements within a database in order to generate association rules; in this particular case, the aim was to adapt the reasoning of the algorithm to the behavior of the bids in the process an of its winners, meaning, to adapt the condition of "if" and "then" to the data of the procurement processes.

It is worth mentioning that for the implementation of the algorithm, the SCPM used for its market study, already developed libraries, over which we carried out the necessary adjustments regarding public procurement information. Afterwards, we made processes for validating the data, which aimed to prove that the results of the algorithm are relevant and coherent with what is expected. In this phase of the process, two public procurement databases were used in order to carry out the respective controls.

Finally, some recommendations are exposed in order to obtain better results from the algorithm, as currently we have witnessed some limitations in its use and on its automatization; likewise, it is stated that the algorithm can generate interesting information for further analysis, determining that its use does not allow to obtain information about the presence of anticompetitive practices and its use is recommended for ex ante actions on identifying possible leads that could be investigated in due process.

France

The technological innovation brought about by digitisation can benefit competition authorities and strengthen and improve the tools they have to carry out their missions. Among these new tools, *data screening tools* can provide valuable assistance to the Investigation Services in the detection of anticompetitive practices: cartels, price fixing and geographic market sharing are just some of the practices for which these tools can be very effective. These tools can also be used to automate certain market surveillance tasks.

The introduction of data screening tools is generally based on two distinct but interconnected components: a data collection tool and a tool for viewing data and the relevant indicators associated with it.

Two main options exist for data collection.

APIs (*Application Programming Interfaces*) allow access to a wide range of data, both current and often historical, in a simple, secure and authenticated framework. They are ideal for long-term monitoring.

In France, a principle of open public data has been agreed and a large amount of data, as well as a set of public APIs, is available online.

However, not all data is necessarily available via an API and in this case the scraping method can help to retrieve data accessible online.

The introduction of a scraping method requires a certain technicality and preliminary identification work, with each solution being tailored according to the architecture of a page at a given time.

The visualisation module makes it possible to enhance the value of the data collected and avoid a "black box" effect during the monitoring period. The creation and implementation of appropriate indicators are crucial to the success of a screening tool and usually take place in coordination with the case handlers in charge of the file.

Italy

The Authority has experimented the use of ex-ante screening tools to identify suspect bid patterns, although such an activity is not undertaken in a systematic manner due to the difficulties in building a reliable dataset of public bids. Nevertheless, specific tenders are subject to ad-hoc screening, e.g. because of their high value. Moreover, bid screening is also undertaken after the opening of an investigation, for instance as a means to identify the scope of the anti-competitive agreement or to produce additional supporting evidence to prove the infringement. Some examples are described to illustrate how screening tools have been used for different purposes.

The contribution highlights some challenges in using such tools include: the reliability of the data (the need for completeness and quality of public tender datasets that are designed to detect corruption risks rather than collusion risks); the complexity of more advanced tools which are time and resource intensive; the risk of missing a broader collusive pattern in terms of products and geographic areas when focusing on individual tenders; issues related to the rights of defence and confidentiality; and the evidentiary standard as screening tests generally provide circumstantial evidence that rarely support alone the finding of infringement.

In the Authority's experience screening tests have also offered opportunities. First, they have proven very useful tools for refining enforcement priorities. Moreover, the Italian practice has shown that even simple screening tests can be very effective in detecting potential collusive behaviour and help strengthen the case by providing additional supporting evidence. Another opportunity is the use of the results of the screening process to inform and support future AGCM advocacy opinions on tender design. Finally, a more widespread use of screening tests may encourage a collusion risk-assessment of tenders by contracting stations.

Kazakhstan

One of the main tasks of the Agency for Protection and Development of Competition of the Republic of Kazakhstan (APDC, antimonopoly authority) is the prevention, detection and, investigation, suppression of violations of the legislation of the Republic of Kazakhstan in the field of competition protection. At the same time, the fundamental tool of the APDC in the implementation of the task is the collection of the necessary information and evidence base confirming or refuting the commission of a violation.

In practice, the analyses carried out according to the old scheme have a number of drawbacks associated with a lengthy procedure for information withdrawal (sending and processing numerous requests to market participants, authorized agencies, etc.), forming conclusions, and, as a result, making decisions after the fact, when the effect of such actions lose their relevance (individual cases reach 2-3 years). Today, the antimonopoly authority of Kazakhstan is focused on reengineering its activities with an emphasis on automating the processes of monitoring and analyzing commodity markets, including using Big Data.

APDC, together with the Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan (BNS), has begun work on creating an analytical map of the state of competition. The analytical map of the state of competition is an information panel (dashboard).

Korea

With the transition to digitalization, there is an increasing interest among competition authorities in the use of datasets to detect cartel behaviors. Cartel conspiracies can take many forms, but bid-rigging showing apparent signs of collusion is the simplest type to be statistically assessed. In the context, the Korea Fair Trade Commission (KFTC) is operating a digital system for the collection and management of bid data, aimed at detecting bid-rigging in public procurement.

In January 2006, the KFTC started to operate an online system that collects and analyses public procurement data to screen for indicators of collusion and support the opening of investigations. The system is Bid-Rigging Indicator Analysis System (BRIAS). However, as the system uncovered only three cases by September 2015 raising an issue of the low utilization, an advancement project of BRIAS was launched in 2018.

The advancement project aimed to achieve stable connection of the system, with a new data sharing method that receives bid data through the Public Information Sharing Center (PISC) from other government agencies. In addition, the data pool has been expanded by combining contracts awarded through a two-phase Multiple Award System (MAS) and lowering the minimum amounts for contracts to be collected as BRIAS's raw data. Also, the reliability of screening outputs has been increased using a set of parameters tailored to each bid type. And the project enabled investigators to further group bids together, if having the same product or business entity, preceded by analysis on individual bids. It enhanced the data usability in investigations.

Currently, a total of 16 government agencies are transmitting bid data (such as tender offers) to BRIAS, and the compiled dataset is used for selecting target contracts to be checked and monitored every six months. Using the data, 26 bid-rigging cases were opened *ex-officio* from 2018 to 2021, of which nine were sanctioned with KRW 13.4 billon. In this way, the advancement project has delivered greater success to BRIAS.

Meanwhile, some challenges still remain for the better functioning of BRIAS. The system needs to be in line with the latest bid market trends and the minimum prices for the contracts to be collected should be phased out. And the BRIAS dataset needs to be expanded, combining local state-owned enterprises as data source. The set of parameters needs regular updates, and what procurement data the system's manager examines when selecting cases for *ex-officio* investigations should be clearly identified so that BRIAS can learn from the manager's insights.

Mexico

This written submission provides a general overview of Mexican Federal Economic Competition Commission (COFECE or Commission) screening activities conducted in two stages by the General Directorate of Market Intelligence (the Intelligence Unit). The first stage consists in the collection and processing of information, where sources should have a minimum degree of consistency, as well as be representative of the specific market. In the second stage, the Commission uses specific tools such as analytical applications, algorithms, and methods that "flag" or indicate potential irregular behaviours in the datasets.

COFECE's tools have been developed in-house due to the strategic nature of the information managed by the Intelligence Unit. This Unit is formed by staff from diverse backgrounds, including economics, statistics, mathematics, computer science, and data science. Since the creation of the Intelligence Unit, COFECE has have been able to develop several ex officio cases that are built with inputs including those that result from screening activities, mostly in public procurement. However, the Commission is also exploring applications to either detect or investigate other conducts.

COFECE collaborates with other competition agencies by exchanging experiences and best practices on screening activities, always abiding to common confidentiality provisions included in Memorandum of Understandings (MoU) signed by COFECE.

Romania

The scope of this contribution is to share a few lessons learned from Romanian Competition Council's journey towards using analytics, where after 5 years the competition authority is at the stage of refining the tools acquired and maybe where other competition authorities have already found success.

The contribution also provides some background information on RCC's earlier attempts at using digital screening, describes the components and functionalities of the Big Data platform recently implemented, and offers some considerations on the benefits and challenges of using such a platform by a competition authority.

Singapore

Singapore takes a serious view on bid-rigging, particularly, bid-rigging in government procurement. Government procurement is governed by the principles of transparency, open and fair competition and value-for-money. However, bid rigging in government procurement undermines these principles and results in harm to the wider society

To improve efficiency and reduce human error in an area of work that is typically manpower intensive, CCCS has developed a bid-rigging detection tool ("BRDT") to flag suspicious government tenders which may be subjected to bid rigging; and a document similarity tool ("DST") which can be applied to the suspicious tenders to identify similar tender documents between competitors which suggest bid rigging. The two tools are complementary and can be used in combination to identify suspicious government tenders and facilitate early intervention by CCCS. This submission provides further information on the tools and the learning points in developing such tools.

The BRDT improves the coverage and efficiency of the detection process by automating the bid rigging detection process through a set of indicators based on common signs of bid rigging and their corresponding expected signs in bidding data. The BRDT has allowed CCCS to expand our coverage and increase sensitivity to suspicious tenders, while saving manhours in a recent investigation.

The DST uses text analytics to compare similarity of documents. This allows CCCS to more efficiently detect similarities in tender proposals of competitors which may be indicative of bid-rigging and replaces the manual scanning process by CCCS officers which is time consuming and potentially prone to human error.

In developing the BRDT and DST, it is important to consider these tools as investments that require sacrifices in the short term in order to reap the benefits in the longer term. As the development of sophisticated tools does not fall traditionally within the expertise of a competition authority, it is useful to consider tapping into expertise outside of CCCS. The deployment of the tools is also an important aspect to consider, including relevant staff training and outreach to encourage adoption in other government agencies. This has been a learning journey for CCCS for several years, and we will continue to test, apply and improve the BRDT and DST to enhance our enforcement in bid rigging in Singapore, as well as sharing our tools with other competition authorities internationally.

Spain

The increasing availability of data and development of new tools are driving a change in Competition Authorities willing to take advantage of this new reality. This becomes especially important concerning public procurement and bid rigging detection.

Even if Spain is in a relatively good position in terms of availability of tender data, there remain issues which needed to be addressed, such as the need to aggregate sources, data cleaning, or the unstructured format of documents, which make prior treatment necessary for their correct use.

Since 2015, CNMC has been working on a public procurement database, trying to reduce these hurdles, and categorizing the data by quality levels.

Once the data are clean, multiple tools can be applied on them, both in a case-by-case scenario, as well as in a general and aggregated manner relying on more sophisticated techniques.

In Spain, access to data concerning losing bids opens the way for machine learning techniques that decisively enhance ex officio detection.

*Switzerland**

Considering the harm of bid rigging and the number of cases opened, COMCO wanted to develop a pro-active method to uncover and prosecute cartels. It initiated a long-term project in 2008 to fight bid-rigging cartels, which included developing a detection method or a so-called "screening" method.

To be applied quickly, a detection method should require few data. Therefore, COMCO decided to focus only on publicly available data.

COMCO developed its detection method in a heuristic way by observing the past behavior of cartels. The screening method developed by COMCO is reliable, easy to understand even for non-economists, parsimonious in its data requirement and can potentially be applied at a large scale.

One of the key success factors in developing the screening method was to observe past bid-rigging behaviors and their effect on the distribution of the bids to form an understanding on collusive markers. Those collusive markers could be compared to benchmark values derived from previous investigations. Nevertheless, the threshold values should be treated as an imperfect approximation and applied with a careful "human" judgement.

COMCO applied the developed screening method on bidding data of the construction sector in the Canton of St. Gallen. Based on the indications delivered by the application of the screening method, COMCO opened an investigation in 2013 and sanctioned the involved firms in 2016. All firms identified by the screening method have been found guilty of bid-rigging conspiracy.

COMCO decided to inform the public on its use of a screening tool and, to a larger extent, on how the tool functions. In that decision, COMCO weighted the positive aspect of increased deterrence against the adverse effect of an intelligent cartel trying to evade the detection by the screening method. In the opinion of COMCO, this risk weights less than the advantage of increased deterrence fostered by publicity.

In the future, COMCO intends to collaborate with other competition agencies in order to build more performant screening tools based on algorithms.

United States

DOJ formed its Procurement Collusion Strike Force (PCSF) in 2019 to coordinate a national response to antitrust crimes relating to public procurement at all levels of government. The PCSF is an interagency partnership consisting of federal prosecutors and national law enforcement partners tasked with detecting and preventing fraud, waste, and abuse.

Because of the decentralized nature of government procurement in the US, at the federal, state, and local level, data tied to procurement is fragmented and owned by a variety of government agencies, often using different procurement and data collection processes. As a result, the goal of the PCSF is not to build a universal data analytics program, but instead to build analytics tools that increase detection of collusion across all levels of government.

The PSCF advises government agencies on how to use procurement data in building their own tools. The PCSF also trains data scientists, analysts, auditors, and investigators on suspicious patterns and red flags that indicate collusion. Additionally, the PCSF advocates for collection and retention of pre-award data.