

Unclassified

English - Or. English

24 May 2023

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

**Working Party No. 3 on Co-operation and Enforcement**

**Executive Summary of the Roundtable on Data Screening Tools for Competition Investigations**

**Annex to the Summary Record of the 136th meeting of Working Party 3**

28 November 2022

This Executive Summary by the OECD Secretariat contains the key findings from the Roundtable on Data Screening Tools for Competition Investigations, held by Working Party 3 on 28 November 2022.

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

Please contact Mr Antonio CAPOBIANCO if you have questions about this document.  
Email: [Antonio.CAPOBIANCO@oecd.org](mailto:Antonio.CAPOBIANCO@oecd.org)

**JT03519570**

## *Executive Summary of the Roundtable on Data Screening Tools for Competition Investigations*

By the Secretariat\*

On 28 November 2022, Working Party No. 3 held a [roundtable](#) to discuss data screening tools for competition investigations. Considering the [background note](#) prepared by the OECD Secretariat, the written contributions, as well as the discussion by delegates, the following key points emerged:

**1. It is difficult to assess the true extent of cartel screening given authorities do not always publicise in full their use of cartel screening. Nonetheless, it seems that there has been increasing interest in the use of cartel screening methods in recent years, by both competition authorities and even some private entities.**

Based on public statements from competition authorities it seems that there is an increasing interest in cartel screens. This is potentially due to increasing availability of digital procurement data, new methods in the academic literature, and a decline in the number of leniency applications which may indicate the need for more proactive cartel detection.

In 2013, the written contributions to the OECD roundtable on cartel screening suggested very few authorities applied cartel screens on a systematic basis, although many had considered them to a limited extent or on a more ad hoc basis. In 2016, 15 out of 27 competition authorities surveyed by the International Competition Network reported they were doing some cartel screening. In June 2022, the Danish Competition and Consumer Authority reported that they were co-developing the Bid Viewer cartel screening tool with several authorities and exchanging views on screens with over 15 authorities. In November 2022, there were 16 country contributions to this roundtable, most of which indicated the use, or planned use, of cartel screens. There were also several examples of the use of relatively sophisticated screens, as well as some examples of cases that successfully relied on output from cartel screens.

Many authorities, however, do not publicise their screening initiatives for fear that this would lead companies engaged in anti-competitive conduct to become more sophisticated and, ultimately, render the screen ineffective. Furthermore, authorities will typically only publicise the successful use of cartel screens, so less is known about the failed use of cartel screening tools. Thus, the true extent of screening is not known nor can easily be estimated. Nonetheless, there does seem to be an increasing interest in cartel screening.

Interest in cartel screens may have increased for three main reasons. First, cartel screening is more common due to improving access to data. The increase in data availability is due to proactive collection by competition authorities and the emergence of new technologies that make data easier to extract from hardcopy and digital sources. Second, cartel screen accuracy has improved due to new methods and technology, deriving from machine learning and data science. Third, the number of leniency applications, the single most important cartel detection method, has declined in recent years. This reduction in reactive cartel detection methods may be driving the need for authorities to use more proactive cartel detection tools.

---

\* This Executive Summary does not necessarily represent the consensus view of the Competition Committee. It does, however, encapsulate key points from the background note, the discussion, and the delegates' written submissions.

**2. Cartel screening methods in the academic literature have expanded considerably in recent years. Some of the most promising developments in the academic literature have been the combination of supervised machine-learning techniques with traditional behavioural screens, such as bid/price distribution methods. However, it may be difficult for authorities to collect the training data required to use supervised machine-learning methods. More commonly authorities rely on simpler screens or unsupervised machine learning methods that do not require training data.**

While there have been recent developments in cartel screening methods, the existing distinction between structural and behavioural screens remains. Structural screens seek to identify markets for which it is more likely that a cartel *will form*, while behavioural screening seeks to identify markets for which a cartel *has formed*.

The academic literature has expanded considerably in recent years. It has mostly focused on behavioural screens. There are three broad approaches to behavioural screens using bid and price data: collusive markers, structural breaks and anomalies. A collusive marker is a pattern in the data that is more consistent with collusion than competition. A structural break is an abrupt change in the data generating process. An anomaly is a pattern in the data that is inexplicable or inconsistent with competition but may ultimately be found consistent with collusion.

The recent academic literature has adopted several cartel screening approaches. These include bid/price distribution-based methods, bid roundness, cointegration-based methods, conditional independence using bidding functions, difference-in-differences, historic bid fractions, network methods, probabilistic methods, regression discontinuity methods, structural break methods, suspicious clusters in geographic spatial data, and the combination of several screens.

Some of the most promising advances in cartel screening efficacy has resulted from the combination of machine learning with traditional cartel screens (such as price/bid distribution methods). Screens are based on assumptions of the way a market functions and the types of illegal conduct that may occur. This means that different screens are tailored to different potential types of violations and different markets, and that there is no single perfect screen able to identify all violations in all markets. However, supervised machine-learning techniques can be used to combine several cartel screens, determining the ideal weights for each of the screens, to optimise identification of illegal behaviour.

In practice, supervised machine-learning methods may be difficult to implement, as they require a relatively large amount of training data (which provides a mapping between cartel screen values and whether the bid was collusive or not), which may be difficult for competition authorities to gather (as an authority would have to collect data from previous known cartel cases). Thus, more commonly, authorities have relied on simpler screens or unsupervised machine learning methods that don't require training data.

**3. To develop and implement effective cartel screening tools, competition authorities require access to good quality data as well as staff with the necessary knowledge and skills. International cooperation can facilitate the exchange of knowledge and experience, and even trained machine-learning cartel screens or datasets.**

Despite progress in data collection, access to reliable data continues to be one of the main roadblocks to more frequent cartel screening. Even if data is available in digital form, there may be several databases, and/or data in pdf or other non-easily extractable format, so many authorities are trying to create editable databases. If domestic regulation allows, procurement and competition authorities can agree on the right format to keep procurement data (say in Excel templates with predetermined fields) to make it useful for cartel detection purposes, as well as making this available to competition authorities.

Economists already working at an authority may be able to implement cartel screening tools. However, several data collection methods and more complicated screens may require data science and/or computer science expertise. Several authorities have already started hiring “technologists”.

Authorities are facing similar issues regarding cartel screening. Thus, it makes sense to pool and share resources and knowledge, where possible. International fora like the OECD allow the exchange of knowledge and experience. Supervised machine learning models require training data, with larger training datasets improving the accuracy of the model. Thus, authorities may benefit from pooling data. Although, sharing training data across jurisdictions is still an active field of research. Datasets may be more difficult to share as they contain confidential information. But the use of synthetic data (which is artificial data generated from the original data based on the characteristics and structure of the original data), or anonymising data, may be useful as a means to overcome confidentiality issues. In any case, sharing trained algorithms or code should be easier as they do not reveal anything about confidential data. Tools such as BidViewer, are already being jointly developed and shared amongst authorities. Several other authorities offered to share their tools.

**4. Competition authorities usually use cartel screen results to open a case or launch a dawn raid. Cartel screens can also help to determine which cases to prioritise, for example, when there are several possible options and limited resources. The screen results are usually only indirect evidence and are normally supported by direct evidence obtained during the investigation. Although extremely rare, there have also been a few instances where cartel screens were used to close a case or support an infringement decision.**

The most common use of cartel screens is to start an investigation or issue a dawn raid order. Competition authorities can use screens to prioritise cases, identifying those that may have most chance of success.

There have been some limited examples where the court has relied on indirect evidence from screens as sufficient proof for an infringement. These instances are relatively rare though, as naturally the standard of proof required for an infringement is higher than that required to launch an investigation. There is a risk of false positives (as patterns consistent with collusion can have other innocent explanations). If an authority relies on a cartel screen as part of the evidence to close a case, they should disclose the model and underlying assumptions to the parties, to give them the opportunity to respond and challenge the conclusions.