Organisation for Economic Co-operation and Development

DAF/COMP/LACCF(2024)8

Unclassified English - Or. English

5 August 2024

DIRECTORATE FOR FINANCIAL AND ENTERPRISE AFFAIRS
COMPETITION COMMITTEE

Latin American and Caribbean Competition Forum

LATIN AMERICAN AND CARIBBEAN COMPETITION FORUM - Session III: Ex Officio Investigations

- Contribution from Spain -

9-10 October 2024

The attached document from Spain is circulated to the Latin American and Caribbean Competition Forum FOR DISCUSSION under Session III at its forthcoming meeting to be held on 9-10 October 2024 in Santo Domingo, Dominican Republic.

Marcelo Guimarães, Competition Expert - Marcelo.Guimaraes@oecd.org

JT03548191

Session III: Ex Officio Investigations

– Contribution from Spain –

Introduction[1]

1. One of the main priorities of the Comisión Nacional de los Mercados y la Competencia (Spanish National Markets and Competition Commission, CNMC) is to adapt its competition policy to the current trend of digitalisation and use of artificial intelligence (AI). As early as 2018, it launched a series of specific initiatives in this area, which have since undergone significant developments and, in some aspects, have seen notable progress.

2. Since then, the CNMC has had a unit specifically in charge of this, the Economic Intelligence Unit (UIE in Spanish). Among other activities, the Unit uses advanced data analysis tools and systems, including those based on AI, in order to detect not only new forms of anti-competitive behaviour (such as algorithmic), but also traditional ones, which are becoming more subtle and refined every day. The Unit thus underpins the CNMC's capabilities to improve ex officio investigation, and makes other investigations even more robust.

3. The creation and maintenance of a growing database covering all aspects of public procurement – recently including detailed factual data on companies and individuals linked to them, thanks to the agreement signed between the CNMC and the Association of Property and Business Registrars of Spain – has also been an important area of activity for the UIE. This provides its data analysis tools with a more solid basis for producing extremely relevant results.

4. The CNMC is thus honouring its commitment to applying new techniques to ensure that an adequate level of competition is safeguarded in all markets, with a focus on public procurement, due to its significant impact on the availability of resources aimed at increasing the welfare of citizens. The scope of this contribution will be to briefly describe some of these systems and tools, and their development status.

1. Detecting anti-competitive behaviour: The data

5. Algorithms and other IT-based tools can be used for anti-competitive purposes by means of software tools of varying degrees of sophistication (including those known as pricing algorithms). But anti-competitive behaviour can also manifest itself in other ways, such as self-preferencing or discrimination (e.g. in rankings or comparison algorithms). These risks appear most frequently in digital business models, where algorithmic tools are often used, but they can also occur in other sectors of economic activity.

6. As mentioned in the introduction, in order to address these challenges, the CNMC introduced digitalisation as a priority in its Strategic Plan,[2] giving rise to different initiatives

[1] This contribution deals with advanced tools applied to ex officio investigations, which is the central theme of Session III of the Latin American and Caribbean Competition Forum, to be held in October 2024. This contribution has been prepared by CNMC staff and should not be considered an official position, except for references to official documents approved by the CNMC.

[2] https://www.cnmc.es/sobre-la-cnmc/plan-estrategico

in its annual Action Plans.[3] One such initiative was to create an Economic Intelligence Unit (UIE) to take advantage of new digital techniques to help detect different types of anti-competitive behaviour, including when such behaviour is facilitated by algorithms or other advanced tools.

7. The UIE has a multidisciplinary character that is reflected in its staff: alongside competition law specialists, staff include data scientists and analysts, economists, engineers and other staff with expertise in information and telecommunications technologies (ICT). This gives it a better overview of the renewed competitive threats it must face, as well as providing it with broad capabilities to achieve its goals.

8. Starting as a very small initial team, the UIE has experienced remarkable growth in recent years, at a rate of 40%. This clearly demonstrates the CNMC's growing interest in matters that are the focus of the UIE's work, and a firm commitment to provide it with the necessary resources.

9. Among the duties assigned to the UIE, it is worth highlighting its work in promoting ex officio investigation, as well as its support in investigating other cases: enhancing and strengthening evidence of suspected wrongdoing, adding information to existing complaints or information from informants, planning home inspections through open source information (open source intelligence [OSINT] activity), etc.

10. It is essential to collect the right data in order to undertake these functions, and in particular to be able to use specific techniques and analysis, such as supervised machine learning, graph analysis and other advanced techniques related to data management and processing.

11. For this reason, from 2015, i.e. before the UIE was created, the CNMC worked on building a public procurement database, developing specific algorithms for data extraction and data cleansing, and categorising data by quality levels with respect to all participating bids and bidders.

12. In the area of public procurement, the CNMC database is the result of combining different sources. Firstly, the Public Sector Procurement Platform is a centralised and fully updated electronic platform that publishes all calls for tenders and their results. It includes data from regional and local contracting bodies, by interconnecting the platforms of the various administrations and public entities to establish a single platform that centralises the publication of public sector procurement. This repository is subsequently completed with data from contracting bodies that are not integrated within the Public Sector Procurement Platform, centralised procurement contracts and non-standardised documents associated with tenders, such as award notices or possible amendments to the tender.

13. In addition, the CNMC also downloads information on framework agreements, *encargos a medios propios* (orders to own means, from one government agency to another), centralised contracts, and other centralised procedures, as well as data on small contracts.

14. At the date of writing this contribution, **the CNMC's public procurement database contains more than 3.5 million contracts and over 4 million lots**.

15. Therefore, this database could be considered one of the most complete in Spain, with structured data on all levels and types of public administration tenders, including the documents related to these contracts (unstructured data). It should be noted that this database includes not only the data relating to the winning bids, as is more usual, but also

[3] https://www.cnmc.es/sobre-la-cnmc/plan-de-actuacion

the data relating to the rest of the bids submitted to the tenders, thanks to a private channel set up for this competition authority.

16. It is also worth highlighting that this database is confidential, as it contains information that is not available to the general public. Its use is therefore limited to the investigative activity of the CNMC's Competition Directorate.

17. Furthermore, in 2023, another relevant source of data became available, thanks to the agreement signed between the CNMC and the Association of Property and Business Registrars of Spain. Under this agreement, the CNMC has access to all the official economic-financial information of any companies that participate in public tenders, such as the company's address and its business premises, the main shareholders and/or persons in charge of the company, successor companies if the company is dissolved, etc.

18. Using the combined data from all these sources, the CNMC has created a comprehensive digital database of public procurement in Spain, which has recently made it possible to develop a set of powerful tools that use these consolidated data, free of errors and duplications. These tools are currently helping the UIE to detect possible cases of bid rigging.

2. First generation detection tools

19. Turning now to the panoply of tools available, the first group can be described as first generation, since they make use of conventional techniques to a greater or lesser degree, albeit with a high level of sophistication.

20. These are mainly statistical and econometric techniques implemented through business intelligence tools for data visualisation and analysis.

21. Indeed, a wide range of tools are available for processing these data: from statistical techniques (such as sunburst charts, or radial diagrams that, when sorted by characteristics, make it possible to visualise information more rapidly and to detect potential patterns shared between companies), to more sophisticated developments involving AI, which we will mention in the following paragraph about second generation tools.

3. Second generation detection tools

22. In addition to these representations and analyses adapted to efficiently handle large amounts of data, providing our analysts with the ability to discover trends that could otherwise go unnoticed, the UIE is also working on methods to incorporate new sciences in the field of AI, such as machine learning, natural language processing techniques or rudimentary generative AI processing techniques.

23. The first development in this regard, the ***gUIE.me* system**, is a powerful search tool specialising in tenders, supported by a web interface, developed entirely by UIE staff.

24. It is noteworthy that gUIE.me performs entity extraction tasks (keywords of interest, or search elements) from the complete texts of tenders (unstructured data). This undoubtedly means it has greater potential than that of a simple search within a relational database, however powerful that may be.

25. Graph-oriented databases, based on the calculation of graph metrics, have also been used as a detection tool. These databases store nodes and relationships, instead of tables and documents. The related metrics represent the degree of correlation (centrality and

community detection, in terms of graph theory) between activity sectors, companies and company leaders themselves. These parameters make it possible to characterise groups of people or companies of interest (clusters) that can be the subject of more in-depth research.

26. In addition, more recently, access to data on unsuccessful bids (losing bids) has paved the way for the use of techniques that provide a decisive boost to pure ex officio investigation.

27. More specifically, this has led to the development and implementation of the BRAVA system (Bid Rigging Algorithm for Vigilance in Antitrust).

28. **BRAVA** is a classification tool **based on supervised machine learning** that uses different AI models to classify the different bids submitted to a tender as likely to be collusive or competitive. The success rates obtained when these models were tested are largely satisfactory and at least 90%. The system employs supervised learning based on input variables (from the public procurement and commercial registry databases mentioned above) and on more than 20 coded statistical screening variables (screens), some based on the most recognised international academic works, and others developed from scratch by the UIE based on its experience in corruption and collusion.

29. It should be noted that some of the indicators derived from graph analysis have also been used as an additional source of screens, to be used as inputs in the machine learning process.

30. After this process, the data from bids in which collusion has been found to be present, labelled accordingly, together with data from other bids that are considered competitive, presented as a contrast, are further divided into a training set and a test set, in order to advance the supervised machine learning process through different iterations. Thus, the level of accuracy achieved by the classifier can be considered as optimal, which makes it possible to apply it to the complete set of input data in order to classify them as collusive or competitive bids.

31. In practice, this serves a twofold function in determining collusion:

- It allows the CNMC to focus on specific tenders, such as those included in a communication from a contracting authority or a whistle-blower (including those of an anonymous nature).
- It makes it possible to trigger purely ex officio analyses on certain sectors, especially in markets suspected of harbouring stable cartels (whose members are less prone to defection and therefore less susceptible to leniency programmes).

32. Although, due to its nature, BRAVA will always be subject to continuous and intensive development, the tool is already in the full production stage for its first-time implementation and the results are very promising.

33. In any case, some of the main avenues for developing the system in the immediate future include processing the complete texts of technical bids submitted for tender by both the successful bidders and the losers. This may result in detecting similar patterns between them, or even not making an award if it seems that bids have been agreed with competitors. These collusive patterns, obtained using AI-based natural language processing techniques, will enhance the classifier's powers of detection.

34. Other avenues to develop in the medium and short term will include unsupervised learning (stepping up the use of variants of the algorithm *a priori* to detect the habitual simultaneous presence of alleged competitors, an algorithm that is already being used in the current system), market segmentation (necessary to better target ex officio actions) and

improving the detection of algorithmic collusion (also based on AI techniques implemented by cartel members to enter into their anti-competitive agreements more covertly).

Main conclusions

35. Antitrust enforcement is always complex, given the usually secretive nature of anti-competitive agreements, but in digital markets it is particularly challenging where algorithms and AI are concerned. Competition policy tools are flexible enough to adapt to digitalisation-driven disruption, but some challenges remain.

36. In order to respond to these and other challenges, the CNMC has also given strategic priority to monitoring anti-competitive conduct driven by algorithms and digital business models, by adapting and refining its enforcement tools. There is a real need for advanced computational tools based on similar techniques to ensure a more level playing field between companies and competition authorities. Therefore, the CNMC is currently directing certain efforts towards this goal.

37. Both its powerful first generation tools, supported by extensive and comprehensive public procurement and company databases, and its new AI-based BRAVA system, recently created by the CNMC's aforementioned Economic Intelligence Unit, seem to represent a determined effort in the fight against bid rigging, one of the most relevant and widespread collusive activities.