

Unclassified

English - Or. English

10 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 Ecuador

28 November 2022

This document reproduces a written contribution from Ecuador 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]

JT03507279

Ecuador

1. Background

1. The SCPM, following its attributions conferred by the Organic Law for the Regulation and Control of Market Power, did a market study on 2021 about the National Public Procurement System, where, within its quantitative analysis, some hypothesis and evaluation methodologies were issued about the characteristics of procurement processes that could allow to allege leads of collusion¹².
2. Even though the referred study established six (6) possible conditions that could be considered leads³⁴, this document (given the quantitative methodology applied to its analysis) we only reference the leads regarding “market sharing”, over which we applied, as methodology, an algorithm for the revision of rule associations that allow to identify cases with historic behavior patterns between bidders in public procurement processes.
3. In light of what has been mentioned, the following sections we will detail the processes developed by the SCPM for the implementation, validation, application and gathering of results of the referred methodology. Furthermore, we will detail the limitations of the analysis of the algorithm used and recommendations for its enhanced application.

2. Data sources

4. The quantitative analysis done in the market study to the National Public Procurement System, used as information source, the base of public contracts done by the different Ecuadorian public institutions, in the period 2015-2020. These contracts lie on the State’s Official Procurement System. It is important to note that, in virtue of the Ecuadorian normative, all public procurement processes must be registered in this system with the relevant information related to them, except for very specific cases established in the corresponding law.
5. Among the information contained in the mentioned database we can find, per process, the following variables: i) participating bidders, ii) adjudicated bidders, iii)

¹ On the market study, we used the definition of collusion in public procurement established by the OECD in the document “Guidelines for fighting Bid rigging in Public Procurement”, which defines it as such practices that restrict or limit competition between suppliers of the State, harming public procurement through higher prices or impeding the offer of products and services that might satisfy in a better way the needs of the contracting entities.

² Guidelines for fighting Bid rigging in Public Procurement: <https://www.oecd.org/daf/competition/42851044.pdf>

³ i) winner in first bid; ii) false bids; iii) price agreements; iv) societary links; v) market sharing; and, vi) subcontracting with collusive goals (maening, when a method of payment used are subcontracts among colluded parties).

⁴ Regarding the determinants of collusion in public procurement: the Chilean case. Juan Francisco Martínez S., David Escobar y Claudio Loyola.

economic offers presented, iv) referential amounts, v) adjudicated amounts⁵. The complete details of the variables that make up the database is attached on Annex 1.

3. Methodologies used

6. With the objective of identifying possible leads of a collusion under the modality of market sharing, we adapted the algorithm “*a priori*” for the study, a tool that was chosen because it allows the identification of behavior patterns in a data series. The objective of the algorithm was, mainly, to identify repetitive or alternating behaviors on bidders and winners data from the public procurement databases. There upon, we detail the methodology and the implementation of the algorithm to the data available for analysis in the study:

3.1. Association rules (*a priori* algorithm)

7. This data mining algorithm identifies sets of frequent elements on a database in order to generate rules of association. This algorithm has been widely used to identify frequent combinations on supermarket shopping. For these specific cases, the algorithm checks the purchases of several clients and identifies those products which have a larger probability of being purchased together, establishing the condition that “if” a client buys product “x”, there is a probability “then” that he will buy product “y”. For the case under analysis of the SCPM, we aimed to use the same reasoning, meaning conditions “if” and “then” to the behavior of the offers done by suppliers and winners of the procurement processes. Following this premise, the reasoning used in the algorithm was to identify the frequent combinations of the participants in the procurement processes and the winners.

8. The implementation of the algorithm *a priori* requires, as a condition, that the data to be used complies with the following conditions: i) that it represents the set of all products /events related with this problem; ii) the item is a product or an event⁶; and, iii) the data needs to be a set of transactions, where each transaction is a set of elements. In this line, within the exercise done in a preliminary manner a verification was done to check that the data available in the database of public procurement complies with the aforementioned conditions, out of which, we concluded (for this specific case) that such conditions were thoroughly complied; given that: i) the database used represents the totality of events of public procurement carried out; ii) it was feasible to treat each procurement process as an item, for which, it can be considered that each of them represents events, iii) public procurement processes are a transaction of purchase by the State and in each of these processes there is a set of participants and winners.

9. Hereinafter we detail the process that the algorithm follows and the results that it generates; it is important to notice that the market study was established as a tool to analyze the *a priori* algorithm in light that its process basically adapts to the defined goal, as it estimates the following aspects:

1. It identifies the frequent elements under predefined thresholds (support);

⁵ Oficio Nro. SERCOP-SDG-2021-0460-OF, 19 May, 2021

⁶ In this analysis, the ítem would be the participation of a supplier in a procurement process.

2. It generates association rules that satisfy the *confidence* restriction⁷.
10. In order to comply with this aspects, the algorithm basically follows these steps:
 - It generates all the trials of frequent elements. A set of elements is frequent when its occurrence exceeds the minimum support or a pre-established threshold; and,
 - Once all the sets of elements have been explored, only those sets of frequent elements are held and all the non-frequent elements are excluded.
11. It must be pointed out that, for the implementation of the algorithm in the analysis of the market study we used previously developed libraries⁸, over which we proceeded to do the adaptations to the structure and public procurement data; par example, a marker⁹ was established for the winners of the public procurement in order to differ them from the bidders. This with the objective that, as mentioned above, with this analysis we can observe leads of alternation between the participants and winners; in this sense, the algorithm must be able to identify the frequent elements of bidders that take part in processes and have as consequence or probability of occurrence a determined winner.

4. Limitations of the analysis

12. Regarding the limitations, it must be pointed that, even though the complete database of the public procurements represents an advantage for the general analysis of the market study, the application of the algorithm in a database of such magnitude, also derived in distortions and biases in the results, as the calculations of the algorithm were done in function of the repetitions of the observations of the database, which, by being complete (meaning, it registers all the purchases without accounting for which products or services were acquired in each of them), biased the results toward the combinations with products that are more frequently bought, leaving aside combinations that might result more representative (due to their high purchase amounts, for example) but happen in products of lesser purchase frequency, and therefore, were not reflected as frequent elements; therefore, par example and as detailed below, the results of the algorithm were biased toward office supplies due to the frequency of their purchase.
13. Despite this limitation happening on the analysis done in the market study, it might not be a serious problem on itself, because, as we will see afterwards, when applying the algorithm to one of the databases developed in the unit of Investigation and Control of Restrictive Agreements and Practices of the SCPM related to an investigation process on collusion that was sanctioned, the results of the algorithm were consistent with what was expected; it is important to mention that, in line to what was mentioned in the last paragraph, the calculations were applied to a smaller database of the investigated sector, for which the problem of observation repetition bias was not significative.
14. Currently, given the availability of data and the way of delivering them by the governing body on public procurement in Ecuador, this analysis cannot be done in an

⁷ Confidence refers to the conditional probability that a procurement process contains the combination of bidders or established winners.

⁸ Libraries is understood as public domain programming that was previously made for their use and adaptation.

⁹ This marker means that a letter was added to the ID number of the winning undertaking for them to be easily identifiable.

automatic and constant manner; on the contrary, it is necessary to feed and update the database on each occasion required to update the results.

5. Effective validations

15. During the implementation of the *a priori* algorithm, we carried out processes for control and validation in each different stage; these verifications were made, firstly with the complete public procurement database within the market study, and in a second instance (as mentioned before) we carried out trials to a database provided by the unit of Investigation and Control of Restrictive Agreements and Practices. Here we detail the actions taken in each of the moments:

1. **On the complete database of public procurement:** at this stage we observed the results generated by the algorithm, which were validated, compared and corroborated through manual review in the database; for such effect, we selected some of the results with higher scores following the parameters established in the algorithm. On the other hand, in function of the results of the algorithm, we separated from the complete database, those sets of data that were necessary to verify the results obtained.

This manual review allowed to observe that the algorithm is correctly specified, a situation that was met in satisfactory way during the process developed in the market study; nevertheless, in this stage we observed one of the limitations mentioned beforehand, regarding the fact that in the process for the selection of frequent elements important information about bids was lost specifically bids of greater scope (regarding the contracting amounts), these because the less frequent processes do not exceed the parameters established in both analysis and were therefore rejected.

2. **On the database used by the unit of Investigation and Control of Restrictive Agreements and Practices in a process of collusive agreements that were sanctioned:** the purpose of this validation phase was to identify if in a specific database (and therefore, of a smaller size compared to the database mentioned on the former stage) the results of the algorithm would be coherent with the expected results, meaning, that the variables listed as undertakings under investigation would be within the datasets with greater frequencies and probabilities of occurrence, as it will be presented below, this situation was met on the obtained results.

6. Results obtained

16. Among the results obtained, using both databases, the complete database of procurements and the database provided by the unit of investigation and Control of Restrictive Agreements and Practices; we can observe that the algorithm presents coherent and interesting results for further analysis. For Example, in the review of the exercise done on the complete database of public procurement, it is noticeable that the number of combinations that exceed the established threshold (both in support and confidence, meaning, the parameters that determine the threshold a set must surpass in order to be considered frequent) oscillate between 87 and 427 per year; out of this combinations, we can observe that they present, predominantly, a confidence of 10% and 40%, and therefore, such combination is present between 10% and 40% on which the undertaking associated with the condition “if” appears on the database. While these values might seem small given

the size of the database, this particularity might be corrected by doing the analysis over a more specific database.

17. In the Table N°1 we present information provided by the algorithm, regarding the analysis of the total of public procurements; this information points out that: i) support: the proportion of times that the undertaking (or combination) appear on column “if” within the database; ii) confidence: the proportion of times on which the combination on column “if” and “then” appear in a set within the database; and, iii) the lift value, that is interpreted in a way on which if it surpass 1 there is a probability that conditions “if” and “then” are being met. For example, we can observe that undertaking “0102087798001 – MUÑOZ BRAVO FREDDY ARTURO” appears on 3,7% of procurement processes and that, on 23% of the times that this undertaking appears on the database he wins a bid when taking part with operator “1792125375001 – COMSUPPLIES S.A.”.

Table 1. List of combinations obtained through the association rule¹⁰

YEAR	IF	THEN	support	confidence	lift ¹¹
2017	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO, 1792125375001 - COMSUPPLIES S.A	3,7%	23,0%	6,2879
	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO, 1792440874001 - WANDA TECNOLOGIA CIA. LTDA.	3,3%	21,0%	6,2879
	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO, 1709766453001 - CAIZA CERON MARCO ANTONIO	2,7%	17,1%	6,2879
	1792125375001 - COMSUPPLIES S.A	1792125375001 - COMSUPPLIES S.A, 1792440874001 - WANDA TECNOLOGIA CIA. LTDA.	2,5%	69,5%	3,4875
	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO, 1792125375001 - COMSUPPLIES S.A, 1792440874001 - WANDA TECNOLOGIA CIA. LTDA.	2,5%	15,6%	6,2879
	1792125375001 - COMSUPPLIES S.A	1709766453001 - CAIZA CERON MARCO ANTONIO, 1792125375001 - COMSUPPLIES S.A	2,1%	57,6%	3,5409
2018	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO	1792125375001 - COMSUPPLIES S.A	2,6%	66,7%	3,6503
	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO, 1792125375001 - COMSUPPLIES S.A	2,6%	23,6%	9,0517
	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO	1792775671001 - TONERPLUSS S.A.	2,0%	52,1%	4,7672
	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO	0102087798001 - MUÑOZ BRAVO FREDDY ARTURO, 1792775671001 - TONERPLUSS S.A.	2,0%	18,5%	9,0517
2019	0190123626001 - ASEGURADORA DEL SUR C. A.	0990064474001 - SEGUROS SUCRE S.A.	3,2%	87,0%	8,4790
	0190123626001 - ASEGURADORA DEL SUR C. A.	0990064474001 - SEGUROS SUCRE S.A., 0190123626001 - ASEGURADORA DEL SUR C. A.	3,2%	39,8%	12,3448
2020	0190123626001 - ASEGURADORA DEL SUR C. A.	0990064474001 - SEGUROS SUCRE S.A.	4,6%	95,7%	7,6386
	0190123626001 - ASEGURADORA DEL SUR C. A.	0990064474001 - SEGUROS SUCRE S.A., 0190123626001 - ASEGURADORA DEL SUR C. A.	4,6%	45,2%	9,8921
	0990064474001 - SEGUROS SUCRE S.A.	1790551350001 - SEGUROS ALIANZA S A	2,3%	18,1%	7,7814
	0990064474001 - SEGUROS SUCRE S.A.	1790551350001 - SEGUROS ALIANZA S A,	2,3%	18,1%	7,7814
	0990064474001 - SEGUROS SUCRE S.A.	1790551350001 - SEGUROS ALIANZA S A			

Note: Winners are identified with orange highlight in the table

Source: Database provided by SERCOP 2021.

Elaboration: National Direction of Market Studies

18. On the other hand, regarding the analysis done over the database of the Unit of Investigation and Control of Restrictive Agreements and Practices of the SCPM, we could observe that in the first six (6) combinations with larger support value, the undertakings

¹⁰ In the table, we interpret the columns “if” and “then”, as the condition that when determined bidders appear in the database, the bidders enlisted on the “then” column determined bidders appear as well.

that were under investigation¹² will show up. In these results we can observe that undertaking “1701965798001 ARIAS NANCY YOLANDA DE JESUS” appears on 7,9% of the procurement processes within the database and that in 24,4% of the times that this undertaking appears on the database, it wins the bid when the undertaking “1712329554001 NUÑEZ MORILLO JORGE EDUARDO” also bids. It is important to notice that both undertakings were under investigation by the SCPM.

Table 2. List of combinations obtained through association rules

IF	THEN	support	confidence	lift
1701965798001 ARIAS NANCY YOLANDA DE JESUS	1701965798001 ARIAS NANCY YOLANDA DE JESUS, 1712329554001 NUÑEZ MORILLO JORGE EDUARDO	0,079137	0,244444	3,088889
1712462611001 RIVERA ARIAS ALBA IRENE	1708763006001, 1712462611001 RIVERA ARIAS ALBA IRENE	0,035971	0,16129	3,202765
1712462611001 RIVERA ARIAS ALBA IRENE	1709403172001, 1712462611001 RIVERA ARIAS ALBA IRENE	0,028777	0,129032	3,587097
1701965798001 ARIAS NANCY YOLANDA DE JESUS	1703345825001 MORILLO CRUZ MARIA GLADYS DE LOURDES, 1701965798001 ARIAS NANCY YOLANDA DE JESUS	0,021583	0,115385	4,009615
1701965798001 ARIAS NANCY YOLANDA DE JESUS	1701965798001 ARIAS NANCY YOLANDA DE JESUS, 1709403172001	0,021583	0,115385	4,009615
1703345825001 MORILLO CRUZ MARIA GLADYS DE LOURDES	1712462611001 RIVERA ARIAS ALBA IRENE, 1703345825001 MORILLO CRUZ MARIA GLADYS DE LOURDES	0,021583	0,166667	7,722222

Source: Database provided by the Unit of Investigation and Control of Restrictive Agreements and Practices
Elaboration: National Direction of Market Studies

7. Conclusions

19. The implementation of algorithms on the available information on the market, included on the information about public procurement, allows the automatization of some search criteria and the identification of leads in behaviors of participants that might be relevant, such as collusion.

20. The application of algorithms, such as the one referred in the present document, allows the identification of patterns and behaviors in procurement processes, the same that will allow to gather leads about contracts and undertakings that might require a closer look.

21. Quantitative analysis tools mentioned in this document do not provide a definitive criteria about agreements between competitors, but might be helpful for, alongside other investigation elements, analyzing some behaviors that might be taking place in different markets.

22. The implementation of tools such as the one mentioned in this document are becoming ever more relevant, given that in the Ecuadorian case it is evident that the regulatory body as the public contracting entities within the scope of their competencies, do not have the tech tools necessary to alert about the different leads of collusive agreements that might arise in such processes, having as a result, a system vulnerable to possible anticompetitive practices.

¹² Undertakings under investigation are detailed with the detail of their business name

8. Recommendations for the application and enhancement of the algorithm

23. For a better application, and therefore, better results from the algorithm, its application and trials must be done over specific datasets (meaning, delimited markets) with the aim of avoiding that, due to differences in the purchase frequency of specific products and services, valuable information for the analysis of results is lost.

24. Even though the results of the algorithm might generate interesting results, their interpretation and use must be complementary to a keen knowledge of the sector, this due to interpretations of the obtained results that do not get close to become leads of collusion and might be better explained by the conditions and the structure of the specific market.

Annex A.

N°	Variable	Description	Detail
1	Unique Identifier for the process	Unique Identifier for the process	Digits
3	RUC of the entity	Identifier of the contracting entity	13 digits
4	Business name of the entity	Business name of the entity	Name
5	Publication date	Date of publication of the procurement process	dd/mm/yy
6	Type of process	Type of process	Name
7	Central classification of products	Central classification of products within the process	Digits
10	Number of bidders	Number of bidders within the process	Digits
11	Budget	Amount budgeted	Digits
12	Adjudicated amount	Adjudicated amount	Digits
13	Supplier RUC	Supplier RUC	13 digits
14	Supplier Name	Supplier name	Name
15	Adjudication date	Adjudication date	dd/mm/yy
16	First bid	Bid adjudicated on first attempt	Dummy: 1: suppliers that bid and won on first attempt; 0: contrary.
17	Number of bids	Number of bids done in a process	Digits
18	Adjudicated suppliers	Number of adjudicated bidders in each process	Digits
19	Excess offers	Process with more than 4 bids	Dummy: 1: number of bids made equal or higher than 4; 0: contrary.
20	Consortium	Association of enterprises or entities with the common interests to enter jointly in a project or business.	Dummy: 1: suppliers that make up a consortium; 0: contrary.
23	Province of the contracting entity	Name of the 24 provinces	Name
24	Electronic reverse auction	Type of procedure for procurement	Dummy: 1: procedure is reverse auction 0: contrary