UNIVERSITY OF BRISTOL

Peer reviewed version

License (if available):
Other

Link to published version (if available):
[10.1093/jeclap/lpz024](#)

[Link to publication record in Explore Bristol Research](#)
PDF-document

University of Bristol - Explore Bristol Research
General rights

'Screening for Cartels' in Public Procurement: Cheating at Solitaire to Sell Fool's Gold?

Dr Albert Sanchez-Graells[1]

Abstract

Despite growing global interest in the use of algorithmic behavioural screens, big data and machine learning to detect bid rigging in procurement markets, the UK's Competition and Markets Authority (CMA) was under no obligation to undertake a project in this area, much less to publish a bid-rigging algorithmic screening tool and make it generally available. Yet, in 2017 and under self-imposed pressure, the CMA released '*Screening for Cartels*' (SfC) as 'a tool to help procurers screen their tender data for signs of illegal bid-rigging activity' and has since been trying to raise its profile internationally. There is thus a possibility that the SfC tool is not only used by UK public buyers, but also disseminated and replicated in other jurisdictions seeking to implement 'tried and tested' solutions to screen for cartels. This paper argues that such a legal transplant would be undesirable.

In order to substantiate this main claim, and after critically assessing the tool, the paper tracks the origins of the indicators included in the SfC tool to show that its functionality is rather limited as compared with alternative models that were put to the CMA. The paper engages with the SfC tool's creation process to show how it is the result of poor policy-making based on the material dismissal of the recommendations of the consultants involved in its development, and that this has resulted in the mere illusion that big data and algorithmic screens are being used to detect bid rigging in the UK. The paper also shows that, as a result of the 'distributed model' used by the CMA, the algorithms underlying the SfC tool cannot improved through training, the publication of the SfC tool lowers the likelihood of some types of 'easy to spot cases' by signalling areas of 'cartel sophistication' that can bypass its tests and that, on the whole, the tool is simply not fit for purpose. This situation is detrimental to the public interest because reliance in a defective screening tool can create a false perception of competition for public contracts, and because it leads to immobilism that delays (or prevents) a much-needed engagement with the extant difficulties in developing a suitable algorithmic screen based on proper big data analytics. The paper concludes that competition or procurement authorities willing to adopt the SfC tool would be buying fool's gold and that the CMA was wrong to cheat at solitaire to expedite the deployment of a faulty tool.

Keywords

Competition, public procurement, bid rigging, artificial intelligence, algorithms, big data, data ownership, data governance, commercial sensitivity, commercial secrets, screening.

JEL keywords

C50, C80, D40, H57, K21, K23, K42

---

[1] Reader in Economic Law, University of Bristol Law School. Comments welcome: a.sanchez-graells@bristol.ac.uk. This paper was prepared for the V Annual Meeting of the Spanish Academic Competition Law Network, held at Carlos III University of Madrid (UC3M) Madrid, 30-31 May 2019.

1. Introduction

Given its ability to process very large volumes of information and to identify patterns that would generally escape human observation, artificial intelligence (AI) is expected to have very significant impacts in the field of law. In particular, there are high hopes for the application of AI to law enforcement. This is also the case in the area of competition law,[2] where algorithms can be used to develop cartel screening tools and to boost the functionality of the behavioural screens that have been developing in recent years.[3] AI-supported tools could combine empirical screens with learning algorithms in order to develop adaptable tools that would support enforcement agencies by allowing the detection of market anomalies on the basis of patterns and features of market interactions that are not necessarily taken into account under current investigative methods. AI would not replace human judgement, but it could expand the toolkit used to identify potential violations and launch *ex ante* investigations. This could be useful in markets prone to collusion (ie concentrated markets, based on repeated and highly transparent interactions), as these also tend to be the type of markets where data mining and supervised machine learning *could* lead to the proper operation of AI-supported tools. Moreover, these are markets where (timely) antitrust intervention *should* make a practical difference in terms of preserving or restoring a sufficient degree of competition.

Unsurprisingly, public procurement markets have been identified as a prime area for experimentation in the use of screens for anticompetitive behaviour. Procurement markets are particularly prone to collusion in the form of *bid rigging*.[4] It is also clear that the conditions of competition in procurement markets are largely determined by the design of the tender process and its technical and economical requirements,[5] which adds another dimension where the screens can facilitate *ex ante* avoidance of competition distortions, for example by triggering the redesign of tenders likely to distort or exclude competition.[6] Moreover, given

---

Note: all websites last accessed on 3 May 2019.

[2] It will be interesting to read the report resulting from the consultancy contract 'Artificial Intelligence Applied to Competition Enforcement', which the European Commission tendered in 2017 (COMP/2017/017) http://ec.europa.eu/competition/calls/exante_en.html.

[3] See eg M Huber and D Imhof, Machine Learning with Screens for Detecting Bid-Rigging Cartels (2018) University of Fribourg Faculty of Social and Economic Sciences Working Paper Series No 494, https://doc.rero.ch/record/308901/files/WP_SES_494.pdf. See also D Imhof, Empirical Methods for Detecting Bid-rigging Cartels (2018) University of Fribourg PhD Thesis, https://tel.archives-ouvertes.fr/tel-01963076. Regarding the broader theme of cartel screening tools, not necessarily based on artificial intelligence, see OECD, Policy Roundtable on ex officio cartel investigations and the use of screens to detect cartels (2013) http://www.oecd.org/daf/competition/exofficio-cartel-investigation-2013.pdf. For extended discussion and references to other relevant work, see J E Harrington Jr, 'Behavioral Screening and the Detection of Cartels' in C-D Ehlermann and I Atanasiu (eds), *European Competition Law Annual 2006: Enforcement of Prohibition of Cartels* (Hart, 2007) 51-68; C Lorenz, 'Screening markets for cartel detection: collusive markers in the CFD cartel-audit' (2008) 26(2) *European Journal of Law & Economics* 213-232; M Buşu and B Cîmpan, 'Screening methods for the detection of cartels – Literature review' (2014) 1(3) *Network Intelligence Studie*s 21-31; I Zlatcu and M-C Suciu, 'The Role of Economics in Cartel Detection. A Review of Cartel Screens' (2017) 6(3) *Journal of Economic Development, Environment and People* 16-26; C J Crede, 'A Structural Break Cartel Screen for Dating and Detecting Collusion' (2019) 54(3) *Review of Industrial Organization* 543-574.

[4] For extended discussion and further references, see A Sanchez-Graells, 'Prevention and Deterrence of Bid Rigging: A Look from the New EU Directive on Public Procurement' in G Racca & C Yukins (eds), *Integrity and Efficiency in Sustainable Public Contracts* (Bruylant, 2014) 171-198

[5] A Sanchez-Graells, *Public Procurement and the EU Competition Rules* (2nd ed, Hart, 2015) 60 ff.

[6] In a similar vein, see the arguments of R Porter in OECD, Summary of the workshop on cartel screening in the digital era (2018) https://one.oecd.org/document/DAF/COMP/M(2018)3/en/pdf

the high level of transparency that the EU (and other) public procurement rules impose,[7] it would not be unreasonable to expect sufficient data availability to enable AI-based competition screening, in particular in jurisdictions that have a fully digitalised electronic procurement system.[8] Not least because this data is recorded by the public sector itself, and increasingly published in public contract registries, which could provide the big data architecture required for the training and deployment of machine learning.

Seeking to harness the potential of behavioural screens to support its enforcement priorities in public procurement markets, in 2015, the UK's Competition and Markets Authority (CMA) launched a project that eventually resulted in the publication in 2017 of the '*Screening for Cartels*' tool (hereinafter, the 'SfC tool') 'to help procurers screen their tender data for signs of cartel behaviour'.[9] The tool is described as 'software [that] uses algorithms to spot unusual bidder behaviour and pricing patterns which may indicate that bid-rigging has taken place'.[10] The SfC tool is meant to operate as a 'self-help' tool for public buyers, who can then report instances of suspected collusion to the CMA for further investigation. The use of the SfC tool is supported by the UK anti-corruption strategy 2017-22, which encourages procurers to use it to help them in their fight against illegal cartel behaviour and to investigate and report such activity to the CMA.[11] The CMA is pursuing several advocacy initiatives to raise awareness of the existence of the tool and the risk of bid rigging among public buyers,[12] thus seeking to increase its uptake and the embedding of data analytics in standard procurement evaluation processes. The SfC tool is also being presented internationally as a major development (eg at an OECD workshop in January 2018[13]). There is thus a possibility that the SfC tool is not only used by UK public buyers, but also disseminated and replicated in other jurisdictions seeking to implement 'tried and tested' solutions to screen for cartels. This paper argues that such a legal transplant would be undesirable.

The paper starts by providing a detailed description and an initial assessment of the SfC tool (section 2). It then proceeds to track the origins of the indicators included in the SfC tool to show that its functionality is rather limited as compared with alternative models put to the CMA (section 3). The paper engages with the SfC tool's creation process to show how it is the result of poor policy-making based on the material dismissal of the recommendations of the consultants involved in its development, and that this has resulted in the mere illusion that big data and algorithmic screens are being used to detect bid rigging in the UK. The paper also shows that, as a result of the 'distributed model' used by the CMA, the algorithms underlying the SfC tool cannot be trained (that is, they are static), that the publication of the

---

[7] For background discussion and further references, see A Sanchez-Graells, 'Transparency and competition in public procurement: A comparative view on a difficult balance', in K-M Halonen, R Caranta and A Sanchez-Graells (eds), *Transparency in EU Procurements: Disclosure within public procurement and during contract execution*, vol 9 EPL Series (Edward Elgar, 2019) 33-56.

[8] The standard example is South Korea; see OECD (n 3) 141 ff. In the EU, it is generally accepted that the most advanced jurisdiction is Estonia.

[9] Competition and Markets Authority, Guidance: Screening for cartels: tool for procurers (2017) https://www.gov.uk/government/publications/screening-for-cartels-tool-for-procurers.

[10] Competition and Markets Authority, Press release: CMA launches digital tool to fight bid-rigging (2017), https://www.gov.uk/government/news/cma-launches-digital-tool-to-fight-bid-rigging.

[11] UK Department for International Development and Home Office, Policy paper: UK anti-corruption strategy 2017 to 2022 (2017) https://www.gov.uk/government/publications/uk-anti-corruption-strategy-2017-to-2022.

[12] Competition and Markets Authority, Collection: Public sector procurers: bid-rigging advice (2017) https://www.gov.uk/government/collections/bid-rigging-advice-and-tools-for-public-sector-procurers.

[13] OECD (n 6).

SfC tool lowers the likelihood of some types of 'easy to spot cases' by signalling areas of 'cartel sophistication' that can bypass its tests and that, on the whole, the tool is simply not fit for purpose (section 4). This situation is detrimental to the public interest because reliance in a defective screening tool can create a false perception of competition for public contracts, and because it leads to immobilism that delays (or prevents) a much-needed engagement with the extant difficulties in developing a suitable algorithmic screen based on proper big data analytics (section 5). The paper concludes that competition or procurement authorities willing to adopt the SfC tool would be buying fool's gold and that the CMA was wrong to cheat at solitaire to expedite the deployment of a faulty tool (section 6).

2. The *Screening for Cartels* tool: smoke and mirrors

The SfC tool is a software package—in other words, an app—that allows public buyers to input data about tender procedures to generate a report that comprises a number of red flags indicating potential indications of bid rigging based on algorithmic analysis. In particular, the report includes the output of 12 tests, each of which is assessed by a separate algorithm and displayed on the basis of a simplified pass/fail check, and which are then combined into a single weighted 'suspicion score'. The output is a rather easy to read and user-friendly report (see Figure 1). The CMA is careful to stress that '[t]he suspicion scores highlight which tenders are more likely than others to be suspect. So, a high score doesn't prove the existence of a cartel nor mean procurers should immediately contact the CMA, but could be cause to go back to the bid documents and look again and ask questions.'[14]

Figure 1. Sample report generated by the SfC tool

Tender	1) 3 bidders or less	2) only 1 bidder	3) V. high text similarity	4) winning outlier	5) made up prices	6) same author	7) low effort	8) similar prices	9) similar text and word count	10) typical cartel	11) fake prices and outlier	12) high prices and low effort	Suspicion score
Tender 1	Fail	Pass	Pass	Fail	Pass	Fail	Fail	Pass	Pass	Pass	Pass	Pass	280
Tender 2	Fail	Pass	Pass	Pass	Pass	Fail	Pass	Pass	Pass	Pass	Pass	Pass	220
Tender 3	Fail	Fail	Pass	Fail	Pass	Pass	Fail	Pass	Pass	Pass	Pass	Pass	110
Tender 4	Fail	Pass	Pass	Pass	Pass	Pass	Fail	Pass	Pass	Pass	Pass	Pass	60
Tender 5	Fail	Fail	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	50
Tender 6	Fail	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	20
Tender 7	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	0

Source: CMA (n 14).

As mentioned above, the SfC tool is based on 12 indicators, which seek to test for suspicious signs in three key areas: (i) the number and pattern of tenders; (ii) pricing patterns; and (iii) technical analysis of metadata to track document origin and low endeavour submissions. Table 1 (below) details these indicators, their justification, the metrics they use and the default pass/fail thresholds, as well as the default weightings used by the SfC tool. It should be noted that the default thresholds and weightings can be altered in the app, so that each public buyer can adjust the tool to the circumstances of the relevant tender. That is why the descriptions include instructions on how to alter them.

---

[14] Competition and Markets Authority, Guidance: About the cartel screening tool (2017) https://www.gov.uk/government/publications/screening-for-cartels-tool-for-procurers/about-the-cartel-screening-tool.

Table 1. Content of the SfC tests

Indicator and Justification	Threshold basis	Default weighting
Num 1: 3 bidders or fewer Cartels usually operate in small groups, so a lower number of bidders might suggest a cartel	Currently set at 0.33 for three bidders or fewer, based on the formula 1/3 = 0.33. If you wanted to change it to X bidders or less you would need to follow the same formula: 1/X. Therefore for 5 bidders the threshold would be 1/5=0.2	A fairly low weighting to reflect that this feature is not particularly suspicious by itself
Num 2: Only 1 bidder Cartels can result in just 1 bid. This could be because the other competitors know that it is another bidder's turn to win	Currently set at 1. If you wanted to change it to only X bidders would need to follow this formula: 1/X. Therefore for 2 bidders the threshold would be 1/2=0.5	Since this score will compound with Test 1 a moderate weighting is given so the combined minimum score for a single bid is 50
Num 3: Very high text similarity in losing bids In a cartel, the losing bids could be written by the same person or copied and renamed because all these bidders have agreed that they will lose. This is one of two tests that analyse this issue	Tenders will fail this test if there is a high similarity between the text of bid submissions. Currently set at 10. The text similarity of the losing bids is measured using term frequency-inverse document frequency (tf-idf). This value is multiplied by 100 to bring the scale to a more readable range.	A very high weighting is suggested to ensure any tenders that fail this test are investigated further
Num 4: Winning bid is an outlier Where a cartel exists, the winning bidder can be significantly lower than the others	This test will fail if the winning bid is more than 1 standard deviation (SD) from the mean. This is the multiple of SDs the winning bid is below the mean bid, calculated by subtracting the winning bid from the mean and dividing this difference by the SD	A relatively low weighting is suggested because a winning outlier price is worthy of note but not necessarily suspicious
Num 5: Prices look made up This test uses Benford's Law – a frequency distribution of numbers in real-life data – to spot made up numbers	The distribution of numbers in each bid is compared to the Benford's Law distribution and the accuracy converted to a numerical value. The metric score is calculated from 10 divided by the mean accuracy value to give an indicator of pricing that appears to be artificially derived	A more significant weighting is given since this is a sign of artificially doctored numbers to rig the bid price
Num 6: Same authors Where two or more of the authors for the bids are the same, this would suggest that the same person wrote it for two different firms, suggesting some suspicious behaviour	This test will fail if two bid submissions have the same authors. The threshold for this cannot be changed. Returns 1 if there are multiple bids with the same author in the document metadata or 0 if not	Not very likely to occur, but very suspicious if it does, so a high weighting is given

Indicator	Threshold basis	Default weighting
Num 7: Low effort of losing bid This test looks at the level of effort put into the creation of a document. If the bidder has not put in a lot of effort into the bid, this could be because they already know that they are not going to win the bid	The effort is calculated by the number of edits of the document divided by the time editing it. A low ratio is considered to be a low effort bid and the percentage of low effort bids for a tender is calculated to give the metric score. This test will fail if more than half of the bidders have a low effort	A fairly significant weighting is given since this could mean the majority of losing bids didn't put effort in since they knew they would not win
Num 8: Similar prices across bids Where cartels operate, prices can be similar because bidders are trying not to stand out	Calculated as the mean bid price divided by the standard deviation of the prices. This test will fail if the standard deviation is less than 10% from the mean, so currently set at 15	A fairly significant weighting is given as this could show bidders are trying to manipulate the choice of the buyer
Num 9: Similar text and word count in losing submissions Where a cartel operates, the losing bids could be written by the same person or copied and renamed because they are all in agreement that they will lose	This test will fail if there is a high similarity between the text in the losing bids and if the word count is similar. The mean and standard deviation of the word counts of all the bids is measured and the standard deviation calculated as a percentage of the mean. If this value is low it shows a similar word count across the bids. The metric is calculated as 100 divided by the standard deviation percentage of the mean. Read in conjunction with test 3	A significant weighting is given since this is a sign of documents being shared between bidders
Num 10: Made-up prices and low number of bidders This combination test looks for both a low number of bidders and made up prices	See tests 1, 2 and 5 for details	The thresholds and weighting are fairly low as the test is not a sure sign of suspicious behaviour, but this test would contribute to a higher score which would amount to an overall suspicious tender
Num 11: Made-up prices and outlier This combination test looks for the winning bid being an outlier and made up prices	See tests 4 and 5 for details	As above
Num 12: Made-up prices and low effort This combination test looks for low effort losing bids and made up prices	See tests 5 and 7 for details	As above

Source: CMA, Screening for cartels tool user guide (2017).

As indicated in Table 1, there are 9 baseline indicators and 3 combined tests. The 9 baseline tests seek to measure: (1) limited competition as evidenced by a low (ie three or less) number of tenders, (2) absence of competition as evidenced by a single tender, (3) similarity in the text of the different offers, (4) big pricing differences between the winning and the rest of the tenders, (5) manipulation of prices, (6) identity of authors across tenders, (7) low effort in the preparation of losing tenders, (8) similarity of prices across tenders and (9) similarity of text and word count in losing tenders. The reasons for the use of each of these indicators are reflected in Table 1. However, it is not clear why there are multiple indicators for essentially the same circumstances. For example, indicators (1) and (2) both seek to identify reduced competition or its absence. The reasons for a penalisation of single bid procurements are not completely clear in terms of what it adds to the screening exercise. The relationship between other tests is also unclear. For example, both tests (3) and (9) are based on text similarity. It is unclear whether both tests are useful, but test (3) seems more general than (9) and the role of the wordcount in test (9) is unclear. Other aspects of the test are also unclear, such as the requirement for more than half of the bidders to fail the low effort test (7), which seems to assume that most of the bidders must be part of the bidding ring for it to be effective—without, however, justifying this assumption.

Additionally, there are three combined tests: (10) Made-up prices and low number of bidders, which combines (1), (2) and (5); (11) Made-up prices and outlier, which combines (4) and (5); and (12) Made-up prices and low effort, which combines (5) and (7). The reasons for the inclusion of these combined tests based on the existence of seemingly made-up prices are unclear, as are the effects of their inclusion on the functional interaction of the different tests and the overall suspicion score. It seems, however, that this is redundant if none of tests (10), (11) or (12) can fail when (5) does not fail.

As mentioned above, the 12 indicators are then weighted to create an overall 'suspicion score'. However, there is no indication of the threshold at which 'suspicion scores' should be worrying and prompt further investigation. Figure 2 details the default weights recommended by the CMA, although each user of the tool is free to alter them to reflect their own views about the relevance of each of the tests in the specific context of a given tender.

Figure 2. Default weights used to calculate the SfC 'suspicion score'

Theme	Indicators chosen (suggested weighting)
Number and pattern of bidders	Low number of bidders (20) Single bid (30)
Suspicious pricing patterns	Winning price is outlier (20) Similar pricing across bids (20) Costs appear to be made up (40)
Low endeavour submissions	Same authors in two or more bids (200) Low endeavour losing bids (40) Similar text in losing bids (200)
Combination tests	Similar text and word count in losing bids (50) Low number of bidders and made up prices (20) Winning price is outlier and made up prices (10) Low endeavour losing bids and made up prices (10)

Source: CMA presentation, as summarised in OECD (n 6) 4-5.

It should be stressed that the thresholds for each pass/fail test (in Table 1) and the weightings for the calculation of the overall suspicion score (in Figure 2) have not been

determined through any algorithmic analysis, but are simply a judgement call of the CMA and its consultants. This means that these parameters of the 'suspicion score' do not reflect the outcome of the training of (meta)algorithms,[15] but are rather arbitrarily set—on the basis of some intuition or theory about the relevance of each of the tests for the identification of potential instances of bid rigging. However, this is not necessarily easy to perceive for users of the tool, who may assume that these weightings are the result of extensive testing and somehow reflect a ground truth about the relative importance of each of the tests (as further discussed below).[16] It should also be stressed that the CMA has not published any information concerning the precision, recall and F1 score of the different algorithms in the SfC tool—all of which are used to assess the extent to which the accuracy of an algorithm results in different levels of false positives and false negatives.[17] Thus, it is not possible to assess the extent to which the red flags are likely to provide an effective screening function in practice. However, the mere existence of the tool and the possibility of obtaining a 'suspicion score' (however potentially inaccurate) can create a false perception of competition for public contracts.[18]

In addition to the abovementioned doubts concerning the inclusion, weighting and interaction of the specific tests currently covered by the SfC tool, its capacity to reliably identify potential cases of bid rigging seems limited to the ability of the different tests to pick up on any of the following circumstances:

(i)   Limited competition (tests (1) and (2))
(ii)   Abnormal pricing (tests (4), (5), (8), (10), (11) and (12))
(iii)   Low effort and/or substantial identity across tenders (tests (3), (6), (7) and (9))

Each of these indicators seems problematic in its own terms, and all of them are likely to be limited by the fact that they only rely on intra-tender observability of the bid rigging conspiracy. In other words, no indicator builds on any information about market structure or dynamics, which is a major potential shortcoming of the screening tool (see below 3). Moreover, all of these tests are extremely simple—barring the technical complications underlying bid content comparisons—or vulnerable to specific market conditions.

Starting with the indicators aimed at assessing limited competition, it seems clear that three is an arbitrary threshold for the number of tenders that can be seen as generating sufficient competition. It is also rather obvious that no public buyer needs an algorithmic screen to identify instances where only one bid is received and thus prompt some in-depth assessment of the reasons for that limited participation. At the same time, such tests could be overemphasising the importance of the number of bids for the purposes of cartel screening,[19] as well as ignoring a broader trend towards lower number of tenders for public

---

[15] *Cfr* Huber and Imhof (n 3) 15 ff.
[16] Generally, on the importance of understanding how machine learning operates and the implications of the use of different techniques, see D Lehr and Pl Ohm, 'Playing with the Data: What Legal Scholars Should Learn About Machine Learning' (2017) 51(2) *University of California at Davis Law Review* 653-717.
[17] See Huber and Imhof (n 3) 2.
[18] Not least, because this is the message pushed by the CMA: 'If the test results are mainly green, and the suspicion scores are low (for example below 200), you should be confident that anti-competitive behaviour is unlikely to be happening between your prime suppliers'; CMA, Screening for cartels tool user guide (2017) 5.
[19] Huber and Imhof (n 3) 21 do not find the number of bidders statistically significant.

contracts for likely reasons other than collusion,[20] which could result in an excessive number of cases of suspected bid rigging (ie excessive false positives).

Regarding pricing tests, the metrics are also rather basic and sometimes based on opposing intuitions as to the workings of a cartel. For example, tests (4) and (8) are based on the contrary hypotheses that cartels will either create a big difference between winning and losing tenders (thus, making the winning bid an outlier, test (4)) or try to keep prices very close so that none stands out (thus, generating a clustering of prices, text (8)). The hypothesis underlying test (4) seems easy to dismiss.[21] Moreover, test (4) is vulnerable to instances of entry in a market, where the new entrant offering a higher discount over prior 'normally competitive' prices could unjustifiably be seen as the designated winner by colluding tenderers (again, leading to an excess of false positives). Tests (5), (10), (11) and (12) depend on a theory about the manipulation of prices that requires tenders to include a relatively large number of different line items, which makes these tests unfit for tenders based on an overall contract price or a limited breakdown of economic proposals. Moreover, recent scholarship has shown that, where intra-tender prices are used to screen for potential collusion, the relevant tests have to be based on rather complex assessments of price information, which can work well if the screen is plugged to a learning algorithm[22] (which is not the case of the SfC tool, see below section 4). Of all the price-based tests in the SfC tool, only test (8) seems capable of acting as a useful screen, although the threshold of 15 for the ratio between mean bid price and standard deviation is rather arbitrary.[23] Moreover, this test does not seem to receive a weighting such as 'to ensure any tenders that fail this test are investigated further' (*cf* test (4) in Table 1 and Figure 2) and could thus be offset by some of the other tests. Combined, these two aspects could neutralise the utility of this test.

Finally, all indicators of identity across tenders are potentially the most clearly suitable to capture instances of undue exchange of information between tenderers. However, test (6) on authorship of documents could be problematic where the contracting authority issues template documents that are later filled in by the tenderers, as they could all potentially have the contracting authority as the same author. The low effort test (7) could also be problematic where tenderers copy and paste information across documents or if they develop scripts to manipulate the data that the tests use. As mentioned above, the difference between tests (3) and (9) is not very clear. Moreover, these are the type of indicators of collusion that tenderers can control and, once the tests are disclosed, it seems likely that bid rigging conspirators will take measures to bypass the algorithm (see below, section 4).

All in all, it seems like the tests included in the SfC tool—with the only possible exception of test (8)—are not particularly well-suited to the effective and accurate screening of potential instances of bid rigging, except in the seemingly rare instances where the

---

[20] See eg regarding the situation in Finland, J Jääskeläinen and J Tukiainen, Anatomy of Public Procurement (2019) VATT Working papers Num 118, http://www.doria.fi/handle/10024/168335.

[21] Huber and Imhof (n 3) 8: 'cartel members do not make too low bids, since their aim is to raise the bid of the designated winner in order to extract a positive cartel rent'.

[22] Huber and Imhof (n 3) 2: 'two screens play a major role for detecting bid-rigging cartels, namely the ratio of the price difference between the second and (winning) first lowest bids to the average price difference among all losing bids and the coefficient of variation of bids in a tender. By far less important predictors are the number and skewness of bids'.

[23] Moreover, it does not make sense to set the threshold at 15 if the SfC intends for 'This test [to] fail if the standard deviation is less than 10% from the mean', as it should then be set at 10.

members of the bidding ring have made no effort to submit 'high quality' losing bids and to avoid identity and repetition in their submissions—and, more generally, that the tool seems to generate excessive false positives. This raises the question whether it would not be possible to have identified alternative indicators of potential collusion and, if so, how is it possible that they were not included in the SfC tool. The next two sections seek to answer these questions.

3. Ignoring the obvious: shortcomings of the *Screening for Cartels* tool due to lack of data

In trying to understand the reasons that led to the inclusion of the specific tests discussed above (section 2) in the SfC tool, it is useful to analyse the interim project report submitted by the CMA consultants Spend Network.[24] The report explains how, at the stage of proof of concept of a tool to screen cartels based on algorithms and after conducting a literature review, Spend Network developed and presented the CMA with a long list of potential tests or indicators. The proposed indicators included some of those finally adopted as part of the SfC tool—in particular, tests 1 to 8 except 4 (of which Spend Matters submitted an alternative version).[25] The initial list also included a few other potential tests, as described in Table 2.

Table 2. Additional tests developed by Spend Matters

Test	Description	Reason for acceptance or rejection
Market competitiveness*	A test to determine whether a tender or group of tenders are routinely operating at below the market average in terms of tender responses by comparing the tenders of one buyer to the wider market	Understanding the market is a vital to knowing whether a cartel could function in a given area. Comparing the tenders from one buyer to the wider market can reveal a low number of suppliers and high prices, key indicators of cartelisation
Over pricing*	A test to determine whether a tender is overpriced, using the average spend data for the chosen supplier or category	High prices are a significant indication of cartel activity
Document authorship metadata–PDF**	A test to determine if a document was authored by someone other than the supplier making the submission	Metadata rarely exists for PDF documents
Submission timing	A test to establish whether there is any unlikely pairing in the timing of bids that could be suspicious. If the timing of bid returns are clustered much more tightly than average it may indicate that suppliers are conversing before submitting their bids, or that bids were submitted by the same individual	It was feared that the inevitably higher volume of bid-submissions near deadlines could distort this analysis. This also relied on the data on submission timings being stored either by procurement portals or by the bid recipient

[24] Spend Network, Report on bid-rigging algorithms. The use of public sector procurement information to identify bid-rigging in public procurement markets (2016), included in the CMA, project update and response to Spend Network recommendations (2017), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/627670/Report_on_bid-rigging_algorithms.pdf.

[25] Indeed, instead of a test of outlier prices based on the difference between the winning bid and the mean bid, Spend Matters had proposed a test of 'Outlier pricing in a tender' aimed at identifying cases 'where a small number of bidders have "broken ranks" and submitted a normal market price against cartel members who have inflated their prices', so that the test would highlight 'how far prices have varied from the market'.

Test	Description	Reason for acceptance or rejection
Low endeavour submission (online questionnaire)	A test to count the number of log entries in a tender and to see if all the suppliers but one have made little or no effort when submitting their bids	This required bids to be submitted via online questions, a method not used by any of our partners
Suspicious pattern of awards: Repeated awards to incumbents	Geographic, sector and other pattern analysis to see if there is a pattern of incumbents retaining the business in a market	To make a valid analysis, this would have required data from more partners than we had and over a longer duration of time. It is very difficult to determine the geographic area over which a cartel operates we cannot define the set area of cartel activity. Data for the whole country is required before these geographic distinctions would be possible
Suspicious pattern of awards: Suppliers take turns	Geographic, sector and other pattern analysis to see if there is a cycle of a small group of suppliers winning the business in a market	To make a valid analysis, this would have required data from more partners than we had and over a longer duration of time. It is very difficult to determine the geographic area over which a cartel operates we cannot define the set area of cartel activity. Data for the whole country is required before these geographic distinctions would be possible
Suspicious pattern of awards: One supplier dominates	Geographic, sector and other pattern analysis to see if a single supplier is winning all of the business in a market	To make a valid analysis, this would have required data from more partners than we had and over a longer duration of time. It is very difficult to determine the geographic area over which a cartel operates we cannot define the set area of cartel activity. Data for the whole country is required before these geographic distinctions would be possible
Typo matching in bids	Analysis of typographic errors in the texts of bid submissions to see if the same mistake is repeated in bids from different suppliers to the same tender. This would indicate the same author for all submissions	A dictionary and any other relevant acronyms and names would need to be added to the text analysis tool. While this is a valid analysis, the text similarity algorithm was deemed preferable as it already identifies similar words and phrases in the documents but has a wider scope than just typographic errors
Submission source	Test to see if tenders from rival suppliers were submitted from the same IP	It would have been difficult to gather this data and the algorithm only identifies a very specific mistake by cartelists

Test	Description	Reason for acceptance or rejection
Login source	Where submissions are made via webforms, it may be possible for the same supplier to be submitting data on behalf of other suppliers. Monitoring IP addresses will test for this	It would have been difficult to gather this data and the algorithm only identifies a very specific mistake by cartelists
Submission authorship tests	Use of pre-existing algorithms to test language patterns that indicate texts are written by the same author	The text similarity algorithm already identifies similar words and phrases in the documents
Losing bid quality analysis	A test that analyses patterns in the submissions of a single supplier to determine whether or not they are deliberately submitting poor quality bids at different times	This depends on data for the same supplier across multiple tenders and Spend Network could not guarantee access to this sort of data. Bid quality is already being tested using metadata
Tender competitiveness	A test that predicts the number of responses based on the nature of the tender specification	Tender specifications can vary for a wide variety of reasons depending on what is being procured. Accounting for this sort of variation makes it difficult to set benchmarks for quality, therefore making it difficult to run this test
Losing price analysis	A test that monitors pricing data to see if suppliers are bidding to lose by comparing the pricing variations between bids	Depends on data for the same supplier across multiple tenders. Pricing data is already being analysed by various other algorithms
Pricing immunity	Prices in certain area remaining stable or increase, while prices elsewhere are falling. Testing this requires pattern analysis to check for a distortion of prices in a specific market	This would require data over a longer period of time than we had available. Pricing and markets were covered by other tests

Source: Spend Network (n 24). *Note that these two tests were initially adopted by the CMA, but later rejected due to lack of data. **Note that the test included in the final SfC tool is based on a proposal to take into account authorship of Word documents, which can thus reduce its scope.

The reasons behind the rejection of these additional tests mainly point towards two main issues. First, the fact that electronic procurement is still not a reality despite the rules in Directive 2014/24/EU[26] and its UK transposition through the *Public Contracts Regulations 2015*,[27] which prevents analysis based on web forms and IP addresses. Second, and more relevant, the inexistence of a reliable (centralised) repository of procurement data from which information can be extracted by enforcement agencies to develop and test the algorithms. The entire project of development of the SfC tool depended on the disclosure of procurement information by a small number of public sector entities.[28] The Spend Network report is very clear as to both the problems that this created in terms of development and testing of the algorithms, and the need to develop an alternative system for the future. It

---

[26] Directive 2014/24/EU of the European Parliament and of the Council of 26 February 2014 on public procurement and repealing Directive 2004/18/EC [2014] OJ L 94/65.
[27] SI 2015 No. 102.
[28] Spend Network (n 24) 13-14.

seems remarkable that all potential tests aimed at screening for the most common bid rigging strategies (such as cover bidding, bid rotation or bid suppression[29]) and to develop market benchmarks to assess abnormality of tendering prices were rejected due to a lack of data. The impact of these design strategies on the overall effectiveness of the SfC tool is hard to establish—as there is a possibility that none of the algorithms designed to analyse that information would have produced useful predictions or classifications—but that seems unlikely. What is clearer is that having developed those screens would have given the CMA (or even the public buyers themselves) a clear method to follow up with a more refined analysis of the suspected cases of bid rigging, which requires some sort of market-based benchmark in order to reach a better understanding of the situation and in order to make a decision whether to open an investigation for breach of competition law.[30]

Even within the reduced scope of the SfC tool as published, the report stresses the negative impact of the lack of data in the design of the algorithms related to the selected tests (above section 2). Some passages are worth highlighting, where Spend Network cautions that

> In general, more data was needed to further test and calibrate the algorithms. The example of the market competitiveness algorithm [which had been selected for inclusion in the tool but could not be developed] demonstrates how important this is to create a viable tool. A picture of a typical market and its competitiveness could not be created with the current data.[31]

This led Spend Networks to make very clear recommendations. In particular, Recommendation 1 indicated that the CMA should 'continue testing and calibrating the tool using more data'. In particular the recommendation stressed that

> To reduce false positives and calibrate the tool in different markets, a greater volume of data is required. The pricing in some areas will be markedly different from pricing in others without this being suspicious, so more understanding of how to set thresholds in different regions is needed. Similarly, some markets will vary in how they score on certain metrics without this being suspicious. For example, aerospace will have a smaller number of suppliers than web design. The tool is variable to take account of these differences, but we will need data on each market to calibrate these metrics.
>
> Data needed to calibrate the tool:
>
> - More councils from a wider range of regions
> - More years' worth of data
> - More markets, expanding our research away from construction …
> - More construction data to calibrate our current tests
> - Cartel data:
>   - From existing CMA cartel cases or cases investigated by global competition authorities
>   - Test that the tool can identify cartel activity and calibrate it if not.

[29] See OECD, 2012 Recommendation on Fighting Bid Rigging in Public Procurement (2012) https://www.oecd.org/daf/competition/oecdrecommendationonfightingbidrigginginpublicprocurement.htm.
[30] Along the same lines, Huber and Imhof (n 3) 3, who propose the application of the cover-bidding screen developed by D Imhof, Y Karagök and S Rutz, 'Screening for Bid-rigging – Does it Work?' (2017) CRESE Working Papers 2017-09, https://crese.univ-fcomte.fr/uploads/wp/WP-2017-09.pdf.
[31] Spend Network (n 24) 18.

Moreover, Recommendation 3 urged the CMA to 'create a platform to collect and process the data'.[32] It is relevant to contextualise these recommendations, which were made by Spend Network after having tried to work with data from five councils, which provided data on 55 tenders, with a total of 237 documents[33]—although some of it ended up being unusable. It is also relevant to stress that one of the main conclusions of the interim report was that 'the tool needs to be calibrated with a wider range of data, including cartel data, <u>before it can be a useful means of identifying bid-rigging</u>',[34] which is a rather telling conclusion by the consultants hired to develop the tool.

The following section reports how these recommendations were largely dismissed (or only aesthetically followed) by the CMA and assesses some of the implications of the deployment of underdeveloped algorithms through the publication of the SfC tool.

4. Succumbing to the hype: rushed deployment of the *Screening for Cartels* tool

In view of the recommendations and conclusions of the Spend Network report discussed in the previous section, one would be forgiven for having expected a serious reconsideration of the project by the CMA. In particular because there was no external obligation (eg under international or EU law, or even under domestic UK law) for the CMA to push ahead with the project[35] and, much less, to publish a bid-rigging algorithmic screening tool and make it generally available. However, contrary to any such expectation, in July of 2017, the CMA published the tool. Given such swift evolution from the situation of underdevelopment of the tool reported by Spend Network only in March of 2016,[36] it is worth considering the extent to which the CMA followed the recommendations. It will become obvious that, by and large, the CMA succumbed to the hype around the use screens to enforce competition law and materially dismissed the recommendations without providing adequate justifications.

First, let's consider the issue of the insufficiency of data. As mentioned above, the tool had been initially developed by Spend Matters on the basis of circa 200 usable documents relating to around 50 tenders. At the time of first publication of the SfC tool, the CMA indicated that it had 'used the data from over 100 tenders involving nearly 500 bids to refine the tests'.[37] While this increase may seem very relevant because it doubles the number of tenders and more than doubles the number of documents, in the context of developing algorithms and trying to deploy artificial intelligence, this increase in data is likely to be insufficient. There are no hard and fast rules about the volume of data needed to train an algorithm but, where classification problems are concerned and, in particular, anomaly detection problems are tackled, it can be argued that data in the tenths of thousands would be necessary for the negative class (ie tenders not affected by bid rigging), while a small

---

[32] In particular, Spend Network advocated for the CMA to support the roll out of the Open Contracting Data Standard. Its adaptation to the EU procurement rules is available at http://standard.open-contracting.org/profiles/eu/master/en/.

[33] Spend Network (n 24) 14.

[34] Ibid 28 (emphasis added).

[35] Except, perhaps, some peer pressure derived from developments in other OECD countries. See OECD, Fighting bid rigging in public procurement. Report on implementing the OECD Recommendation (2016) 13-14, https://www.oecd.org/daf/competition/Fighting-bid-rigging-in-public-procurement-2016-implementation-report.pdf.

[36] Despite the publication in July, it seems that the revision of the SfC tool was finalised by March of 2017, as reflected in the Competition and Markets Authority, Annual Report and Accounts 2016/17 (2017) 59, https://www.gov.uk/government/publications/cma-annual-report-and-accounts-2016-to-2017.

[37] CMA (n 14).

number of cases would suffice for the positive class (ie confirmed cases of bid rigging). Otherwise, the algorithms can suffer a range of problems that reduce their ability to both adequately fit the data and make accurate classifications. Given the lack of information about the SfC tool's evaluation metrics (precision, recall, F1 score, etc), it is not possible to assess the extent to which the revised tool has improved over the March 2016 version. However, a reasonable position to take would be to doubt that a tool developed on the basis of such a small sample can work very effectively and, in particular, that it can generalize well.[38] This could be the reason behind the 'flexibility' of the tool concerning, in particular, its threshold values and the weighting of the different tests—as, in all likelihood, the public buyer's guesses are as good as the CMA's at this stage.

Second, the CMA renounced the possibility of pushing for the development of an adequate data architecture for public procurement in the UK. In particular, the CMA's reaction to Spend Network's third recommendation was that

> Reflecting the situation in the UK, the tool has developed in a way that does not require the CMA to have an ongoing co-ordinating role. However, should others want to share experience and/or create such a central platform we would encourage that.[39]

This should, again, be put in context. Initially, the SfC tool was developed to be applied by the CMA to screen for potential bid rigging cases across the UK's procurement markets (ie a centralised model). However, as a result of legal issues surrounding ownership and access to procurement data, the CMA decided to change tack and finally published the SfC tool as part of a revised 'distributed model'. The CMA explained this change in the following terms:

> The CMA has made the tool freely available to procuring authorities as a downloadable app to use within their own systems. This model (as opposed to a CMA-hosted tool) reflects the UK's distributed procurement arrangements. Putting the tool directly in the hands of the procurers allows the tool to take on a life of its own; authorities and e-procurement platforms will be able to absorb this sort of data analysis process into their internal systems rather than having to use a separate, centralised system. This, in turn, increases the scope for future tailoring and development of the tool. This distribution model should help to build awareness of cartel risks across procurement officials but still result in intelligence sharing and enforcement leads for the CMA. Crucially, this also means that procurers do not need to share their data for analysis. The reluctance of authorities to share data was a key obstacle to the early stage development of the tool.[40]

The implications of this change of model are several and run rather deep. There are at least three that require some additional discussion. First, this 'distributed model' that the CMA presents as flexible, enabling 'the tool to take on a life of its own' and as 'increas[ing] the scope for future tailoring and development of the tool' actually creates rigidity. As mentioned above, the main parameters of the algorithms in the SfC tool (that is, the thresholds for the pass/fail tests and the weights of the different tests' contribution to the overall 'suspicion score') have been determined in a rather arbitrary manner by the CMA and its consultants—or, at best, on the basis of the limited 'training' of the algorithms based on the information related to the abovementioned 100 tenders. From the moment of publication

---

[38] As a benchmark, a recent study based on artificial intelligence to develop cartel screening tools used data for close to 500 tenders, which included 300 cases of confirmed cartels and almost 200 of tenders post-cartel detection.
[39] CMA (n 24) 2.
[40] Ibid 1.

of the SfC tool, there is no centralisation of information and, consequently, any new data fed to the algorithm by a given contracting authority will remain unknown to the parallel versions of the algorithm used by different contracting authorities. That means that the algorithms are fundamentally static because it will be unlikely that any given authority will carry out this analysis over a sufficient number of tenders so as to improve the parameters significantly and, even in that case, it will only use data referring to its own tenders, which will potentially introduce additional biases (not in the machine learning meaning of the term) in the way 'its' version of the algorithm works. Moreover, the CMA will have difficulty comparing scores coming from different implementations of the same tool, as it will have been altered by different data and as each contracting authority is given discretion to alter all parameters relevant to the screen.

Second, it seems rather naïve for the CMA to indicate that

This distribution model carries the risk that, by making the workings of the tool visible, the greater understanding of cartel detection techniques might be abused. But we think the potential benefits of the tool being taken on, development and promoted outweigh this risk.[41]

This does not make much sense in relation to those aspects of the SfC tool that tenderers can control without deviating from their bid rigging conspiracy and, in particular, all those indicators concerned with low effort and/or substantial identity across tenders (above, section 2). Those indicators respond to cases that could be considered 'easy to spot' in the sense that identical cases have been identified in the past without the aid of artificial intelligence because the identity of the content of the bids is relatively easy to perceive. The only 'newer' methods for detection concern the analysis of metadata and specific techniques to assess identity of content. In that regard, the publication of the tool largely neutralises its effectiveness by indicating the relatively small increments in the sophistication of the implementation of a cartel that can bypass the tests in the SfC tool. Whether this is very relevant in practice or not remains unknown,[42] but even a small reduction in the chances to spot sloppy cartels certainly seems to run contrary to the CMA's hopes for a positive trade-off derived from the publication of the SfC tool, not least due to limited take up of the tool by public buyers.[43]

Overall, the SfC tool can hardly be considered fit for purpose and whether it will lead to any increase in the number of uncovered bid rigging conspiracies in the UK remains to be seen. What is most problematic in the short run is that the publication of the SfC tool seems to have led the CMA to consider that its job is done, as it announced already in 2017 that it

---

[41] Ibid 2.

[42] For example, it could be argued that this type of cases is also quite likely to fail to result in a finding of bid rigging, unless they are coupled by additional robust evidence of cartellisation. At least, this is the approach recently followed by the French Autorite de la Concurrence in its decision n° 19-D-06 of 19 April 2019, http://www.autoritedelaconcurrence.fr/pdf/avis/19D06.pdf.

[43] The CMA (n 13) 5 reported in January 2018 that it had received requests to access the SfC tool by 29 contracting authorities. An update in October 2018 increased the number to 71 public buyers. Given that there are over 12,000 local authorities in the UK, as well as a large number of other regional and sectoral buyers, this can hardly be considered extensive use of the tool. It is also remarkable that, by October 2018, access had been sought by almost as many private organisations, international organisations, charities and university researchers. See C Robins and T Nwaogu, 'Cartels - Lifting the lid on anti-competitive conduct', presentation at the East of England Local Government Association's Procurement and Commissioning Masterclass 2018, https://www.eelga.gov.uk/documents/conferences/2018/procurement%20masterclass%2019.10.18/clive_robins_ppt.pdf.

was 'not currently planning to do any further work on analysing data tested by the tool, or in further developing the tool itself',[44] without any subsequent development to the contrary.[45] This comes to generate the mere illusion that big data and algorithmic screens are being used to detect bid rigging in the UK. However, the limited functionality and utility of the tools (above Section 2), the lost opportunity of developing additional screens that allowed for complementary market-based analysis aimed at checking suspected collusion against tests based on the most common bid rigging strategies (section 3) and the abovementioned impossibility for the current 'distributed model' to result in a significant improvement of the SfC tool (barring any new initiative to pool data) leads to immobilism that delays (or prevents) a much-needed engagement with the extant difficulties in developing a suitable algorithmic screen based on proper big data analytics. This undesirable situation is assessed in some more detail in the following section.

5. Undesirability of rushed policy-making and broader governance concerns

The discussion in previous sections has shown how, despite its appearance, the SfC tool is rather limited. Despite the fact that the SfC tool is presented as being underpinned by algorithms, the better view is that most of the tests that it carries out are simple statistical or arithmetic checks, that the analysis of the metadata that it carries out could largely be performed manually and that the 'suspicion score' that it generates has an unproven—and in my view, likely very low—ability to adequately and accurately identify instances of potential collusion in public procurement. On the whole, then, the SfC tool is a relatively convenient automatized tool for the conduct of those relatively simple checks, but it is presented as a much superior and flexible software solution. This may be misleading, in particular to public buyers with limited IT training.

However, the CMA rushed to its packaging and publication and all too easily abandoned the prospect of a centralised screening function to avoid having to deal with complex data acquisition difficulties. This seems like rushed policy-making, which is somehow reflective of a lack of the capacity to accept that not all projects are successful and the unwillingness to write off sunk costs or, alternatively, to patiently continue investing in the development of complex solutions. This raises a number of questions, such as whether enforcement authorities, and regulators more generally, have a duty to be particularly attentive to avoid being blinded by new shiny things, such as the hype surrounding algorithms and machine learning.[46]

It is also clear that the absence of proper big data on procurement contracts prevented the development of more sophisticated algorithms and that the underlying data governance issues were simply brushed under the carpet by the CMA. This raises the question

---

[44] CMA (n 24) 2.

[45] See Competition and Markets Authority, Annual Plan 2018/19 (2018) 16, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/704594/Annual_Plan-201819.pdf. The SfC tool is not mentioned in the CMA's Annual Plan 2019/20, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/778629/AnnualPlan-201920-FINAL-TRACKED.pdf.

[46] Similar considerations apply to the even more hyped field of blockchain. For extended discussion from a governance perspective, see M Finck, *Blockchain Regulation and Governance in Europe* (CUP, 2018). For a broader ethical framework surrounding the deployment of artificial intelligence solutions, see European Commission High-Level Expert Group on AI, Ethics guidelines for trustworthy AI (2019) https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.

whether enforcement agencies have a duty to raise these issues to competent government departments—such as the Government Digital Services, or possibly the National Audit Office, in the UK—and to also push for a review of their data access powers. It seems odd that public sector entities can refuse to share data with the competition enforcement agency of the same jurisdiction, in particular where anonymisation and random reordering of the data is possible. More generally, the discussion in previous sections shows the need to engage with a global assessment of public data ownership and the need to invest in the development of proper data infrastructure to facilitate the activities of enforcement agencies, not only in competition and other market-oriented regulatory fields. Unless such data infrastructure is developed and its reliability is ensured in a sustainable manner, any efforts to deploy data analytics—and in particular complex machine learning algorithms based on proper training on big data—will be wasted. The SfC tool case study should be used to push for changes in this front, but this seems like an extremely unlikely prospect.

6. Conclusion: cheating at solitaire to sell fool's gold

This paper has critically assessed the *Screening for Cartels* tool published by the UK's Competition and Markets Authority in 2017. The paper has shown how, despite its appearance, the tool is rather limited and mostly dependent on arbitrary thresholds and weightings that significantly erode the trust that one can put in the accuracy of its tests and their ability to properly identify potentially problematic instances of bid rigging in procurement. The paper has also shown how the potential for the development of a superior tool was curtailed by the unavailability of proper big data to train sophisticated algorithms able to deploy better tests on the likelihood of collusion in public tenders.

On the whole, the paper has shown how, by abandoning the prospects of a centralised, powerful screening tool and by publishing instead a distributed, weak tool, the Competition and Markets Authority cheated at solitaire. There was no need for a rushed change of strategy or for the publication of the results of the project, which could have rather been shelved or, better, further refined. This experience should also have prompted broader reconsideration of the inexistence of the required underpinning big data architecture. However, the tool was published and is now being promoted both nationally and internationally. This paper closes with a simple warning. Caveat emptor. The *Screening for Cartels* tool may well be fool's gold.