CPI COMPETITION POLICY INTERNATIONAL

CPI Antitrust Chronicle
Oct 2014 (2)

Recent Successes of Screens for Conspiracies and Manipulations: Why Are There Still Skeptics?

Rosa M. Abrantes-Metz
Global Economics Group
NYU Stern School of Business

Recent Successes of Screens for Conspiracies and Manipulations: Why Are There Still Skeptics?

Rosa Abrantes-Metz[1]

I. INTRODUCTION

A veritable "who's who" of high profile financial benchmarks has been under investigation for years now, and likely for years to come. The first was USD LIBOR. In 2008 two *Wall Street Journal* articles reported possible manipulation intended to artificially depress the LIBOR rate based on an empirical screen. These reports were quickly followed by my own research presenting evidence of possible collusion among many of the participating banks well before the financial crisis. Investigations then extended to other "Ibors" including Euribor, Yen LIBOR, and TIBOR. To date banks have already been fined more than $6 billion, and more is expected.

After the "Ibors" came foreign exchange ("FX"), when in mid-2013 Bloomberg presented evidence of a possible manipulation based on screening of price movements. My own work on FX was contained in a December 2013 complaint filed in New York, which extended Bloomberg's analysis and showed further evidence of highly anomalous price spikes at key times of the day when certain benchmarks are set.

The London Gold Fixing was next. In December 2013 I wrote an Op-Ed arguing that the large price declines I observed around the time of the London pm and Silver fixings—when the "price of gold and silver" for the day are determined for the purposes of many derivative contracts—were consistent with collusion to manipulate these benchmarks. A *Bloomberg* article followed in February 2014 outlining additional results from my research on gold. Since then approximately 30 lawsuits have been filed in the United States alone. Last May the U.K.'s Financial Conduct Authority fined Barclays for gold manipulation.

In 2013 Bloomberg reported that the U.S. Commodity Futures Trading Commission ("CFTC") had found evidence of manipulation of ISDAfix, a key benchmark referenced in a number of swaps. My empirical collusion analysis on USD ISDAfix, contained in the complaint filed in September by the Alaska Electrical Pension Fund, shows that once again banks most likely colluded to move this benchmark to the benefit of their derivative positions. Four days after this complaint was filed, Bloomberg revealed that the CFTC is reported to have found evidence of criminal (collusive) behavior on USD ISDAfix submissions.

[1] Rosa M. Abrantes-Metz is a Managing Director in the Antitrust, Financial Regulation and Securities Practices at Global Economics Group, and an Adjunct Associate Professor of Economics at New York University's Stern School of Business; RAbrantes-Metz@GlobalEconomicsGroup.com. This article is a summary of the note Abrantes-Metz prepared for the Directorate for Financial and Enterprise Affairs, Competition Committee, Organization for Economic Cooperation and Development, for discussion at its meeting held on October 31, 2013, in Paris, France. The views expressed in this note belong solely to the author and should not be attributed to the organizations with whom she is affiliated or their clients.

There can be little doubt that major banks are colluding to manipulate many financial benchmarks. I have argued for years that these benchmarks are easy targets for abuse. It is often easy for a handful of banks to move these numbers; even a miniscule movement can give rise to millions in illegal (and risk-free) profits, and virtually nobody is watching. There should also be little doubt that monitoring the data through empirical screens is powerful and effective in identifying this behavior, and authorities around the world need to take the lead on regularly screening the markets to detect and deter manipulation and collusion.

Despite the successes of cartel detection over the last 20 years, even prior to the cartels uncovered since LIBOR, there are many as myself who believe that competition authorities have just started to scratch the surface, and that proactive detection and deterrence policies need to be established and lead by the use of empirical screens.

This article will directly discuss the value added of screens over leniency programs as well as the main concerns raised on the use of screens.

II. SCREENING BASICS

What is screening? The ability to flag unlawful behavior through economic and statistical analyses is commonly known as screening. A screen is a statistical test based on an econometric model and a theory of the alleged illegal behavior, designed to (i) identify whether collusion, manipulation, or any other type of cheating may exist in a particular market, (ii) who may be involved, and (iii) how long it may have lasted. Screens use commonly available data such as prices, bids, quotes, spreads, market shares, volumes, and other data to identify patterns that are anomalous or highly improbable.

There are essentially two different types of economic analyses pursued with an objective to flag the possibility of a conspiracy.[2] The first can be classified as a "structural approach;" it looks at the structure of the industry at hand "scoring" the likelihood of collusion based on factors such as homogenous product, too few competitors, stability of demand, and other commonly used collusive markers.[3]

The second type of economic analysis in cartel detection is empirical and uses what have become commonly known as "screens," or sometimes "empirical screens." These analyses use time-series, cross-sectional data, and/or panel data sets on variables that measure market outcomes—including prices, volumes, and market shares—to detect potential anticompetitive behavior. This is called a "behavioral" or "outcomes" approach, in which economists look at markets' and participants' behavior as translated into observable data and apply screens for conspiracies and manipulations to address whether the observed behavior is more or less likely to

[2] Joseph Harrington, *Detecting Cartels*, HANDBOOK OF ANTITRUST ECONOMICS (P. Buccirossi, ed. 2008).

[3] A general list of these factors is further detailed in *Proof of Conspiracy under Antitrust Federal Law*, AMERICAN BAR ASSOCIATION EDITIONS, Ch. VIII (April 2010, Chapter VIII). A non-exhaustive "check list" of characteristics that influence the susceptibility of a market to tacit or explicit collusion includes: number of firms and market concentration, differences among competitors, product heterogeneity, demand volatility, barriers to entry, benefits of cheating, transparency, and multi-market contact.

have been produced under an explicit agreement. An example of a market-monitoring program combining both components is that outlined by the European Commission in 2007.[4]

Over the last few years, economic analysis in general, and empirical screens in particular, have become increasingly important in cases of conspiracies and manipulations.[5] Competition authorities and other agencies worldwide have begun using screens to detect possible market conspiracies and manipulations, and defendants and plaintiffs have begun adopting them as well.[6]

Though screens have gained significant popularity and adoption over the last five to eight years, with a track record including successes such as the flagging of the LIBOR conspiracy and several other alleged benchmark manipulations, some competition authorities are still reluctant to adopt these empirical tools. Concerns often relate to "too many resources are required" or "lack of robustness" of current screens, or simply "screens don't work," among others discussed later in this note. In my view, these arguments against screens, though understandable, are misplaced.

As this article will make clear, the effectiveness of screens should, by now, largely be beyond dispute. Can screens flag illegal behavior or not? And have they already done so? Yes they can, and yes they have already flagged large-scale matters multiple times just in the last five years.

Why then are they not used more frequently by competition authorities worldwide, when agencies in many other areas use these tools for the detection of manipulation and other types of fraud? Well, one reason is that the use and successes of screens at the current level are still fairly recent, particularly in antitrust. As a consequence, their worldwide adoption may take some time. Additionally, in some markets data that is necessary to employ screening may be difficult to gather.

But, in my view, the most fundamental reason is a culture of cartel detection among some competition authorities which remains largely passive and reactive, rather than more aggressive

[4] European Commission, Implementing the New Methodology for Product Market and Sector Monitoring: Results of a First Sector Screening (2007). This is the accompanying document to "Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions."

[5] A trend detailed, for example, in Rosa Abrantes-Metz & Patrick Bajari, *Screens for Conspiracies and their Multiple Applications*, 24(1) ANTITRUST MAG. (Fall, 2009); Rosa Abrantes-Metz & Patrick Bajari, *Screens for Conspiracies and their Multiple Applications*, 6(2) COMPETITION POL'Y INT'L, 129-144 (2010); and Kai Hüschelrath, *Economist's Note: How are Cartels Detected? The Increasing Use of Proactive Methods to Establish Antitrust Infringements*, J. EUR. COMPETITION L. & PRACTICE, 1-7 (September, 2010).

[6] Surveys of screening methodologies and their multiple applications can be found in Harrington, *supra* note 2; Joe Harrington & Joe Chen, *Cartel Pricing Dynamics with Cost Variability and Endogenous Buyer Detection*, 24 INT'L J. INDUS. ORG. 1185-1212 (2006); and Abrantes-Metz & Bajari, *Id*. The use of these methods in antitrust litigation is detailed in the American Bar Association's *Proof of Conspiracy under Antitrust Federal Laws,* which specifically describes in Chapter VIII the role of the economic expert in proving a conspiracy and details the use of screens in this context. Rosa Abrantes-Metz & D. Daniel Sokol, *Antitrust Corporate Governance and Compliance*, HANDBOOK OF ANTITRUST ECONOMICS (forthcoming) and Rosa Abrantes-Metz, Patrick Bajari, & Joe Murphy, *Enhancing Compliance Programs through Antitrust Screening*, 4(5) *THE ANTITRUST COUNSELOR*, (September 2010) make the case for the use of screens in corporate antitrust compliance programs.

and proactive, sometimes taking the view that "if my leniency program is so successful, why should I try and do anything else to detect cartels?"

III. LENIENCY PROGRAMS AND THEIR SUCCESSES

Detecting cartel behavior is a difficult task for competition authorities. Traditional detection methods involved a proactive policy to evaluate pricing strategies, pricing stability, and opportunities for competitors to meet and reach agreements. Since the late 1990's, the success of the U.S. Department of Justice's ("DOJ's") corporate leniency program and the development of the Amnesty (or Leniency) Plus policy have for the most part replaced other methods of cartel detection existent in the United States. These were also accompanied by significant increases in penalties and more extensive use of jail time.

These programs have resulted in numerous criminal investigations in industries such as chemicals, computer components, and auto parts, among others. Billions of dollars in corporate fines have been generated since then, as well as incarceration of numerous corporate executives.

In Europe the experience is similar. After launching its leniency program in 1996, and a revised version in 2002, the European Commission received a large amount of leniency applications and provided partial or full leniency in 45 of 50 cartels from 1998 through 2007. South Africa receives on average three leniency applications a month (more than contemporaneous U.S. rate). In Spain, on February 28, 2008, the day the leniency program was launched, there were seven applications received.[7]

What contributes to increased detection due to leniency? There is an increased likelihood that the cartel will be uncovered by the authorities, there are clear and reasonable legal standards to prove the existence of a conspiracy, and penalties are large.

On the level of the penalties, some are of the view that these should be increased to at least five times the current levels in order to effectively deter cartel formation.[8] There are also concerns that the increased jail sentences imposed by the DOJ may make it less likely that non-U.S. citizens' defendants will continue to enter into plea agreements with the DOJ, preferring to take their chances with INTERPOL checks at airports and threats of extraction.[9] These are factors that need to be taken seriously into account when designing and adjusting a leniency program.

IV. LIMITATIONS OF LENIENCY AND VALUE ADDED OF SCREENS

While leniency programs in the United States and in Europe have been very successful for competition authorities, many cartels may still remain undetected—how many, we do not really know, we can only guess. But the very fact that leniency applications continue to be filed at high rates is suggestive evidence that collusion still occurs. Therefore, leniency programs do present shortcomings and opportunities to be enhanced and complemented by other detection tools such as screens.

---

[7] Joseph Harrington, *Leniency Programs: Past Experiences and Future Challenges*, Instituto Milenio SCI (2010).

[8] John Connor & Robert Lande, *Cartels as Rational Business Strategy: Crime Pays*, 34 CARDOZO L. REV. 437-490 (2012).

[9] Donald Klawiter, *Conspiracy Screens: Practical Defense Perspectives*, 3(1) CPI ANTITRUST CHRON. (March, 2012).

A clear observation from the track record of leniency programs is that investigations initiated by leniency seem to be concentrated in a handful of industries. For example, in the late 1990's, there was a concentration of investigations in the food and feed industry, followed by the vitamins industry, and then by the chemical industry. The degree of concentration has only increased throughout the 2000's, focusing heavily on the electronics/computer parts industry, the air cargo/passenger industry, the automotive parts industry, and, most recently, the financial services industry.[10]

This high level of concentration may in fact be driven by the DOJ's Leniency Plus program. Under this program, if a company being investigated for one cartel voluntarily provides evidence of a different cartel, it will not only receive full leniency for the second cartel but also an additional discount against any fines owing to the first violation. This creates an incentive for companies to be the first to report other cartels: Companies know that the likelihood these will be uncovered down the road is very high given the ongoing investigations in the first market, so better to be the first to report in the second and third markets.

The DOJ's Leniency Plus program has enhanced cartel detection and has also helped focus scarce investigatory resources more efficiently. But, on the other hand, by focusing such resources on these industries, it is arguably leaving cartels in other industries undetected.

From a deterrence and detection perspective, resources can and should be devoted to carefully developed and implemented economic analyses, namely screens, so that markets not delineated by Leniency Plus can also be analyzed for possible collusion. This would, at relatively low cost, enhance both leniency and Leniency Plus programs.

Moreover, it is possible that leniency programs suffer from an additional selection bias, beyond the one described above due to Leniency Plus. Leniency is likely to become a more attractive option precisely when the cartel is close to being discovered, or close to breaking up, possibly because it has become fairly ineffective. In short, leniency is less likely to be successful in identifying very profitable and effective (meaning, socially costly) cartels. Screens, on the other hand, are more likely to detect cartels that have the most market impact, i.e., the most effective and profitable cartels. Hence, screens may enhance cartel detection and leniency in general through the detection of potentially more harmful cartels in other industries. It is not just a matter of how many cartels are detected by screens, but which.

Additionally, screens will enhance deterrence. If likely cheaters know that their market data are being monitored, they will know there is an enhanced probability they will get caught. This may ultimately discourage the cartel from being formed in the first place.

Can cartel members also become more sophisticated and learn how to disguise their behavior more effectively, thus beating the screen? Yes, they can, but if a cartel is to be successful from its members' perspective and increase profits, it *must* have a market impact, and that impact must (at least in principle) be detectable. It is true that such market impact may be more difficult to detect and may require enhanced screening tools and continuous improvement of such tools. But the need for enhanced screening should not in any way deter authorities from

[10] *Id.*

using these tools. Investigatory tools are continuously improved in all other areas to detect infractions and crimes, why should it be any different with cartels in antitrust?

My view is that it should not be different, but the fact is that it has been different. One justification is that antitrust authorities worldwide have, for the most part to date, taken a reactive policy position to cartel detection when almost exclusively relying on leniency programs. True they were proactive in designing such programs, but the nature of these programs is, in and of itself, passive. Authorities wait for a cheater to voluntarily step forward and apply for leniency. Notwithstanding the importance of these programs, they are passive and should not be the almost exclusive tool to detect cartels. The detection of illegal behavior is usually more proactive in other areas of our society, and so should it be in antitrust.

Other concerns with leniency programs relate to their lack of flexibility. These programs are clearly delineated with little discretion to give full or partial leniency to an application. Cartels are viewed as very serious infractions, among the worst forms of competitive misconduct, but authorities around the world are willing to let many wrongdoers be forgiven. Though some degree of leniency may well be needed, a modern anticartel detection program needs to rely on additional proactive tools. Well-developed screens will pass a cost-benefit analysis and can be successfully implemented.

A proactive policy by competition authorities could be composed of both structural and behavioral approaches, with the latter centered on empirical screens.[11] It is recognized that structural screens may lead to too many false positives—because there usually are many omitted factors influencing cartel formation[12] and, as a consequence, only a small fraction of markets and agreements verifying the checklist of factors likely to correlate with collusion may be engaging in a cartel. Despite this, structural screens may still add value, particularly when paired with empirical screens.

The purpose of screening is not to deliver the final evidence based on which colluders will be convicted, but instead to identify markets where empirical red flags are raised and which are worth further investigations. In doing so effectively, screens will induce cartel members to come forward and file for leniency, and they will also assist in deterring cartel formation.

Recognizing the limitations of leniency programs and the advantages of screens, several antitrust authorities have started to search for alternative approaches to detecting conspiracies. Screening methods and leniency programs exhibit strong complementarities with respect to cartel deterrence and detection that can produce synergetic outcomes. The use of these complementary tools is, in my view, the natural next step to identifying markets where collusion may have existed, or is currently underway, and in strengthening anticartel programs worldwide.

I will end this section with one important point. Some of the largest conspiracies, manipulations, and frauds uncovered to date—Madoff's Ponzi Scheme, the NASDAQ alleged conspiracy, stock options backdating and springloading, LIBOR conspiracy and manipulation (leading to the launch of investigations of benchmarks worldwide), and, most recently, the

[11] Hans Friederiszick & Frank Maier-Rigaud, *Triggering Inspections Ex-Officio: Moving Beyond a Passive EU Cartel Policy*, 4(1) J. COMPETITION L. & ECON. 89-113 (2008).

[12] *Supra* note 8.

alleged conspiracy and manipulation in the foreign exchange markets—have two common components: they were all initially flagged by empirical screens, and none were flagged by competition authorities or other relevant agencies. Instead, these screens were developed by reporters, academics, consultants, and market experts. Competition authorities should adopt similar techniques to deter and detect cartels; after all, they are already being used to detect manipulations and fraud in other contexts by other agencies.

V. ADDRESSING CONCERNS RAISED ON SCREENS

This section focuses on the main concerns that I commonly hear on why some agencies may not use empirical screens. It addresses each of these concerns separately and explains why, though some of these are understandable, they may still reflect a misconception of either screening techniques or/and of how screens have been recommended to be used.

Some of the arguments also require for screens a standard of proof that no other screening methods in other areas of research, or other tools used to detect cartels, are required to satisfy, representing in my view unfair arguments to dismiss screens. Or simply, some of the arguments may just as well reflect a culture of reactive anticartel detection policy. The major concerns and counterarguments as presented below should not dissuade agencies from actively adopting screens.

A. "Screens Have High Error Rates, Erroneously Identifying Cartels Where None Exist"

Even a screen based on a solid theory of cheating and properly designed and implemented can still produce erroneous conclusions, just as is the case with any other statistical test: It may indicate that cheating may have existed where one did not (type I error), or it may fail to flag cheating which did exist (type II error). Again, just as we would not argue that statistical tests are useless because they have margins of error, we should not do the same to screens. The hope is that types I and II errors will not occur with high likelihood, though there is certainly a trade-off between the two.

Screens are very useful for flagging or identifying unusual patterns in outcomes but they should not be expected to provide the final proof that any wrongdoing did or did not take place.

Examples already exist in which the power of screens has been tested. One such example was performed by the Italian Competition Authority.[13] Here, the authors tested the power of a variance screen for prices to detect previously known illegal conspiracies. In particular, they posed the question of whether a price variance screen[14] could have identified collusion in two well-known Italian cartel cases involving gasoline and diesel, on the one hand, and baby milk on the other. They also asked whether such a screen could have correctly identified who was involved and during which time period. And the answer to both questions was "yes:" the screen would have correctly identified these two cartels before the Italian Competition Authority did.

[13] As documented in F. Esposito & Massimo Ferrero, *Variance Screens for Detecting Collusion: An Application to Two Cartel Cases in Italy*, Working Paper (2006).

[14] As initially proposed in Rosa Abrantes-Metz, Luke Froeb, John Geweke, & Christopher Taylor, *A Variance Screen for Collusion*, 24 INT'L J. INDUS. ORG. 467-486 (2006).

Another application in which the power of screens was demonstrated involved the German cement market.[15] The authors show that buyers could have detected this cartel ahead of the launch of investigations through the use of screens for structural breaks.

It is true that more of these studies should be undertaken, but we need to be conscious that the types of screens here described were developed and/or implemented over the last five to ten years, while more attention to these methods has only been more significant over the last five years. Not a whole lot more of testing could have occurred in this short time period.

In addition, how can we really compare screening errors to leniency errors, when we also do not know the latter? Do we know how many of the leniency applications received by the agencies are actually pursued and proven to be successful? In order to dismiss screens "because they have too many false positives," we would need to know how many such errors leniency applications (and other detection tools for that matter) really produce. It seems to me that it is critical to have this information in order to disregard screens in favor of leniency or other tools due to a concern about errors.

Yes, cartel screening will have some errors, but doesn't medical screening have as well? Screening is the beginning, not the end of an inquiry. We do not say that medical screens lack value, even if they contain a certain rate of errors—why should we hold cartel screening to an inexplicably high standard not satisfied by any other screening procedures?

B. "Screens Cannot Distinguish Explicit from Tacit Collusion"

Related to the point above is the concern that screens may have difficulty in distinguishing legal tacit collusion from illegal explicit collusion. This is a fair point, but it is important to stress that it is not universally true; it depends on the situation at hand. But even in those cases where it is true, why should that disqualify screens? Why hold antitrust screens to a standard that no other screens—such as medical—are required to attain? When a medical screen based on an ultrasound detects a node on a thyroid, that raises a flag and the doctor requires further investigation, namely, a biopsy. The biopsy will ultimately determine (with some error rate) whether the node is malignant or benign. Does anyone want to argue that because the initial screen cannot distinguish malignant from benign, it is therefore useless? Then why make the analogous argument about an antitrust screen?

Back to the industrial organization argument. Almost exclusively, this criticism of screens refers to the well-known result that the same market equilibrium of reduced output and higher price can be attained either through a tacit understanding between competitors or through illegal collusion. Though this is obviously true, as an argument against screens it forgets that a good screen will examine not only *what* equilibrium was attained but also *how* it was attained. It is the dynamics towards the new equilibrium of lower output and higher price that may indicate whether such an outcome was more likely to have occurred through tacit or explicit collusion.

[15] As discussed in Kal Hüschelrath & Tobias Veith, Cartel Detection in Procurement Markets, Center for European Economics Research Discussion, Discussion Paper No. 11-066 (2011).

In cartel screening, even if some of the flagged cases are later found to reflect tacit rather than explicit collusion, and only a few are actually illegal, what is the problem with that? Why should false positives invalidate antitrust screening but not medical screening?

Below I will outline three examples of screens to explain how an empirical look at the data can assist placing higher likelihood that the underlying behavior was due to explicit rather than to tacit collusion.

1. Example 1: USD LIBOR

Abrantes-Metz & Metz [16] addresses this issue directly and attempt to sort out whether explicit rather than tacit collusion was more likely on LIBOR setting. While I have always argued that a purely empirical analysis of market outcomes can never be the final proof of illegal behavior, under particular circumstances screens can indeed provide additional evidence to assess the more likely form of collusion. In this article we argue that the dynamics by which daily individual contributor quotes for USD LIBOR move from being different from each other to completely identical to each other to the last decimal point—and across most of the members in the space of a couple of days—is more indicative of explicit rather than tacit collusion. Additional details are contained in the article.

2. Example 2: Undisclosed Major Financial Benchmark

Let's now look at a second example. Figure 1 below refers to the intraday variation of daily submissions of different groups of banks, for a known (and here undisclosed) benchmark currently under investigation worldwide for possible collusion and manipulation. The numbers charted have been altered and do not reflect the actual numbers in the data, but the patterns of the statistics have been preserved.

In this setting, there are again multiple banks submitting quotes for the computation of this financial benchmark, in a manner comparable to LIBOR. There are particular dates, which I call "setting dates," which are relevant for many of the contributing banks. These are dates in which resets and maturities occur for derivatives contracts based on this financial benchmark, and there is therefore potentially an incentive to manipulate the financial benchmark upwards or downwards depending on the banks' positions and clients. These dates are set regularly throughout each calendar year. The third of the set of three consecutive dates is the most important, but some of the contracts may in fact target up to two days prior to that date.

Figure 1.A shows the clear pattern in the submitted quotes. There is a set of banks, called "Group 1," whose quotes are very different from each other immediately before and immediately after the three setting dates. But on the three setting dates, these quotes are completely equal to each other! Notice how the quotes for "Group 2" representing all other banks do not show the same patterns. Only Group 1 changes its quotes exactly on the three relevant dates, not one day before, and not one day after. Furthermore, it is observed that such convergence in quotes among Group 1 for only three consecutive days, and not at least 15 days before or after those three consecutive days, never happens outside of those three setting days.

[16] Rosa Abrantes-Metz & Albert Metz, *How Far Can Screens Go in Detecting Explicit Collusion? New Evidence From the Libor Setting*, 3(1) CPI ANTITRUST CHRON. (March, 2012).

These data immediately suggest a few questions. Since the differences in intraday variation between the two groups 15 days before and 15 days after the three dates are very similar, why are they so different during the three setting dates? And during the three setting dates, why does Group 2 continue with the same level of dispersion of quotes while Group 1 has all equal quotes? How likely is it that this would have happened without explicit communication among Group 1 banks?

Figure 1.B below shows another of such dates, in which Group 1 and Group 2 represent two sets of banks with completely equal quotes within each of the respective groups, only on the third and most important of the setting dates, and neither before nor after. Furthermore, that does not happen for all other banks represented in Group 3. And, in addition, Groups 1 and 2 may be attempting to influence the financial benchmark in differing directions, potentially one group trying to move it upwards, while the other is trying to move it downwards. Again, this is highly unlikely to occur without explicit communication, when it only happens on these key dates. And these patterns recur throughout these relevant dates over several years, until it is disclosed that authorities are investigating this benchmark and starting to find evidence of collusion and attempted manipulation.

These results point to means, motive, and opportunity to collude and manipulate, with the empirical evidence to support. Certainly the patterns here observed are much more likely to be the product of explicit rather than of tacit collusion.

Figure 1.A: Intraday Coefficient of Variation of Bank Quotes for Undisclosed Benchmark around Setting Dates for Derivatives Contracts Based on this Benchmark

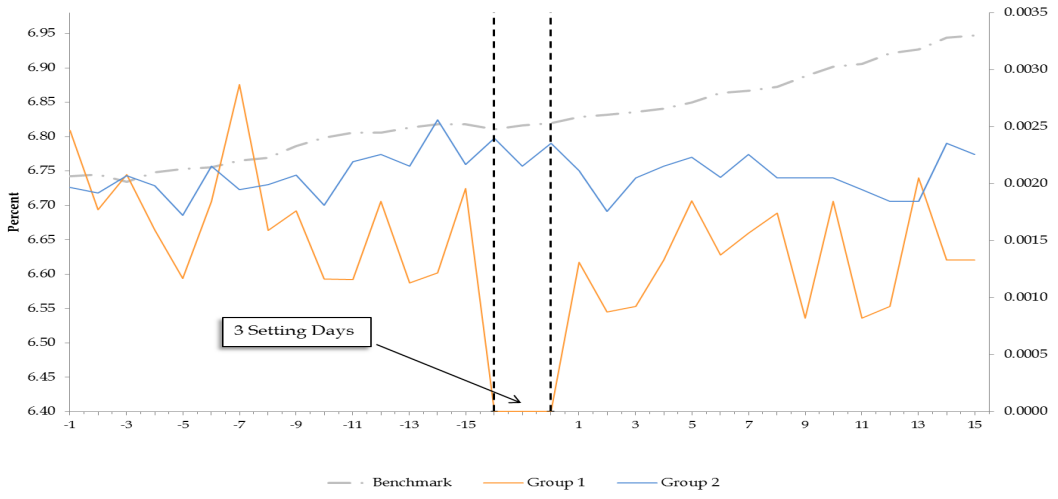
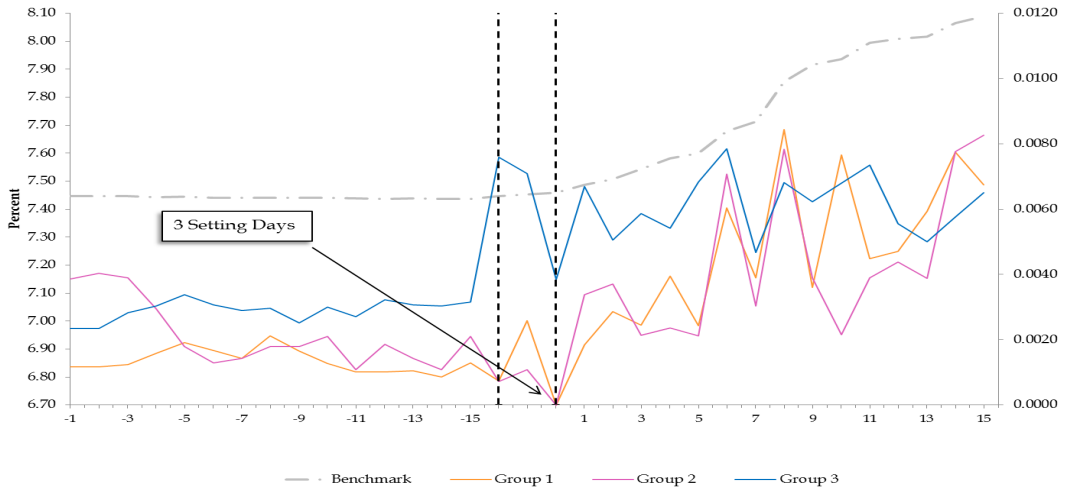


Figure 1.B: Intraday Coefficient of Variation of Bank Quotes for Undisclosed Benchmark around Setting Dates for Derivatives Contracts Based on this Benchmark



3. Example 3: Canadian Construction Bid-Rigging

In September 2012, a Canadian reporter flagged the possibility of a bid-rigging, market allocation, and price-fixing conspiracy in road construction in Montreal.[17] Among other patterns she noticed that winners frequently subcontracted losers, and that the same companies tended to win in the same regions repeatedly. Granted this could have been reached through tacit agreements, at least in principle, but the frequency of unexpected patterns was very high. She also ran a simple and very intuitive screen and found important empirical evidence in favor of explicit collusion: that presumably independent competitors submitted sealed bids to an agency and provided the same contact number in their bids! Unexpected certainly, and almost impossible. This clearly points to explicit coordination. But if the reporter was easily able to find this information through publicly available data, why didn't the agency flag this abnormality when the bids were submitted?

Needless to say that these findings generated an investigation and several resignations have already occurred.

C. "Screens are Very Resource Intensive, They Do Not Pass a Cost-Benefit Analysis"

It is true that screens will consume resources, but so will any other type of work. It is also true that screens will require more resources than sitting and waiting for a leniency application to be filed, but again, and as explained in detail earlier in this article, they may well target cartels with differing market effects, and represent proactive versus reactive anticartel policies.

[17] *Montreal construction bids are paved with questions*, MONTREAL GAZETTE (Sept. 15, 2012), *available at* http://www.montrealgazette.com/news/Montreal+construction+bids+paved+with+questions/7248408/story.html.

It has never been my recommendation to implement screens in every market and at every moment in time. That would likely be neither productive nor efficient. The development and implementation of screens has to be smart, focused, and strategic. First of all, data need to be available, otherwise empirical screens are not feasible. Second, screens should be applied to markets or industries where the likelihood of collusion is higher (to either already exist or to emerge), either due to the features of the industry itself or due to recent history or some other type of prior, including a lead from someone or a complaint. And finally, resources need to be put in place to appropriately develop a screen that fits to the situation at hand.

These analyses do not have to involve a massive amount of resources. For example, my 2008 paper on LIBOR took less than one week for data to be compiled and results calculated. The Canadian reporter who flagged the cartel in road construction in Canada only studied the data by herself for a few weeks and put forward an analysis that seemed convincing to authorities, so much so that an investigation was launched (and evidence of wrongdoing seems to have been obtained shortly thereafter). If screens are so resource intensive, how can reporters and academics run them to flag such large cases?

These are clear examples of screens that did not take long to analyze and, in many cases, were performed by reporters, not expert economists. If they are able to undertake this, why aren't competition authorities as well?

Of course screens can also be resource intensive. That was the experience of the Brazilian competition authority, CADE, when it used screens to select from hundreds of complaints of alleged localized gasoline cartels.[18] Screens were used to select those cases in which market evidence was the most significant. Out of hundreds of possibilities, ten were selected through screening, and among those ten, direct evidence of collusion was found for six cases. A success rate of 60 percent would normally be considered quite good.

It did take a significant amount of resources to get there, but the other three options were either (i) use even more resources to investigate all of the hundreds of complaints; (ii) randomly select complaints to investigate, which would have been an inferior course of action; or (iii) investigate no cases and just wait for all of these to file leniency application. CADE took the right decision, appropriately applied screens, and consequently had a high success rate. CADE is of the opinion that these screens were cost-effective, as they helped focus resources on those cases with the highest likelihood of a market effect. Other agencies should follow the example.

Another successful example of screening was that of the Mexican Competition Authority which used bid-rigging and price-fixing screens to provide empirical evidence for an alleged cartel in pharmaceuticals in Mexico.[19] The effort did take significant resources, but it was worth it in the view of the Commission and, furthermore, the results were recognized in court as constituting valid evidence for the collusive claim put forward by the Commission.

Though screens can at times become technical and numerically challenging, that does not always have to be the case, especially at the beginning of the process. Most of the examples

[18] Carlos Ragazzo, *Screens in the Gas Retail Market,* 3(1) CPI ANTITRUST CHRON. (March, 2012).

[19] Carlos Mena-Labarthe, *Mexican Experience in Screens for Bid-Rigging*, 3(1) CPI ANTITRUST CHRON. (March, 2012).

provided in this article were not technically challenging, or how could these have successfully been used by reporters for a few days of work?

The purpose of screening is not to deliver the final evidence based on which colluders will be convicted, but instead to identify markets where empirical red flags are raised and which are worth further investigations. They do not have to be overly technical, complicated, or demanding. They do have to be smart, focused, strategic, and appropriately designed for the case at hand.

D. "Screens Lack Robustness, Why Should I Use Them?"

I have always been and will continue to be a believer in the power and the role of empirical screens in conspiracy and manipulation cases, and use them in most litigation cases I am involved in, either for defendants or for plaintiffs. But it is important to remember that these are econometric tools, with all the usual caveats, and they may potentially be misused. Screens, just like any empirical technique, can be effective only when properly applied; otherwise they risk producing nonsense.

The two golden rules for screens are (i) "one size does not fit all," and (ii) "if you put garbage in, you get garbage out." Take the first rule. A screen needs to be designed or at least adjusted to the situation at hand. Just because a given set of variables and model specifications prove highly effective when estimating the demand for bread does not mean that those same variables or specifications work when estimating the demand for cars. We typically do not say "this exercise of using the demand for bread to estimate the demand for cars is useless and imprecise" and therefore abandon econometrics altogether for estimating demand equations.

The observation that a model of bread demand does not make a good model of car demand does not represent an argument against econometrics in general. The basic idea of setting up an equation that explains quantity demanded as a function of price and other relevant demand-side variables, and estimating it using appropriate econometric techniques, remains valid. Instead, the lesson we should take is that we need to think clearly about what we want to estimate, the characteristics of the market at hand, and the appropriate set of variables and demand formulation to use so that the technique can be appropriately tailored to the case at hand.

Screens are no different in this regard. It is not a fair criticism to blame the screen when it delivers different results (i) if applied to the levels of variables rather than to their growth rates, (ii) when using different benchmarks or different time periods, or (iii) when either controlling or not for changes in other factors. Such choices represent fundamentals of the empirical specification, they are not "variations on a theme," and just as no "regression model" would be robust across all of these conditions neither should we expect "screens" to be. These are key decisions to be made when applying an existing screen, for whatever purpose it will be used.

This also leads us to the second golden rule: As is always the case in empirical work, a screen is only as good as the choices of what is put into it. Expertise is needed when developing and applying a screen. It is critical that the appropriate choices are taken based on sound justifications when designing and implementing an empirical approach to a conspiracy and manipulation case, or any other case for that matter. Screens are powerful, but they are not so powerful that they work "everywhere" across "any" data set.

A proper screen should have a theory of collusion underpinning it. For example, there is a significant amount of theoretical and empirical evidence that collusion is likely to induce decreased price volatility, under particular circumstances.[20] But that does not mean that all types of collusion are expected to have that effect on prices. Certainly, it is reasonable to expect that when cartelists are fixing prices, they will, to the extent they are successful, likely induce lower price volatility than would otherwise obtain, due to the nature of their agreement. But it does not directly follow that all types of conspiracies will induce price stability. Does that mean that a variance screen to detect collusion lacks usefulness and power? No, not at all. But it does mean that we need to know how and when to use it, and to appropriately take into account relevant market conditions.

E. "Why Use Screens if Cartel Members Will Learn to Beat Them"

The argument has been put forward that since it is possible for conspirators and manipulators to learn how to disguise market outcomes to avoid screens, it is futile to begin using screens in the first place. In any other areas of the law, we do not take this position—just because criminals may sometimes outsmart the detection tool does not render enforcement moot nor prevent agencies from improving their detection tools. Furthermore, was it not true that the U.S. leniency program did not work that well 30 or 40 years ago? Did the DOJ therefore abandon it? On the contrary, it worked to improve the program so that it could become more effective. Why should standards be any different for screening?

Competition authorities should use screens and keep on improving them, but in doing so, they should maintain a degree of non-transparency with respect to the screens used. Notice though that a well-designed screen will go to the core of the conspiracy so that it focuses on the key feature that would be altered by the collusion (if successful). Some screens are more robust in this regard than others. But even if well-designed and implemented, a screen still has a margin of error and may produce an erroneous result because, among other reasons, conspirators have learned how to beat it. If so, it is preferable to continually improve the screen and enhance its detection power rather than to abandon it altogether. We should not let the perfect be the enemy of the good.

F. "Screens Are Very Limited Due to Data Restrictions, and We Cannot Subpoena Companies Just So We Can Screen Them"

It is true that the type of screens here discussed, i.e., empirical screens, can only be applied when relevant data are available. Though this restriction excludes some industries, there are many industries for which enough data for a screen are available. In addition, proxies for costs can often be obtained from the Bureau of Labor and Statistics in the United States, and data on bids for many procurements are publicly and easily available across the country. It bears repeating that reporters and academics have used publicly available data in the Canadian matter and LIBOR, and in several other cases. There are many opportunities to pursue which do not require agencies to subpoena companies for data.

[20] Abrantes-Metz, et al., *supra* note 15; Susan Athey, Kyle Bagwell, & Chris Sanchirico, *Collusion and Price Rigidity*, 71(2) REV. ECON. STUDIES, 317-349 (1998); Harrington & Chen, *supra* note 6.

But for industries for which data are not available, competition authorities need to take an active role in requiring basic data (such as prices and quantities) to be collected and made publicly available. There is still a significant number of industries which are completely obscure on basic market statistics—see, for example, the fracking industry (pressure pumping) which provides services to all oil and natural gas companies worldwide and seems to be under investigation by the DOJ. I am not aware of any publicly available data on prices, quantities, or market shares for this industry. All that exists are a few companies putting out surveys to the operators in the industry to get a sense of what these statistics may be, but such data are not approved by the operators themselves. The lack of information in an industry like this is conducive to anticompetitive practices, as it cannot be monitored through screening or any other way. Responsible agencies need to take a stand and start requiring such data to be reported, collected, and available publicly, even if not freely.

G. "My Leniency Program Works so Well, Why Should I Bother Engaging in Screening?"

I am hoping that, if the reader is following this article closely, this question would already have been clearly answered in section 4. Going further, agencies such as the U.S. Securities and Exchange Commission run a whistleblower program and screening programs in parallel. These approaches are complements, not substitutes. I stress again that LIBOR was first flagged by screening; only years later were leniency applications filed. Leniency Plus applications have likely begun to be filed, and many additional conspiracies and manipulations are currently being investigated around the world, all as a consequence of the initial flagging of LIBOR through screening.

H. "Screens are Very Popular in Academia, But They Do Not Work in the Real World"

How can this argument against screens be credibly made given all of the significant examples of their successes, namely on financial benchmarks?

Most recently, the SEC has initiated investigations of hedge funds generated through their screening programs.[21] Other examples include flagging the Madoff Ponzi scheme years ahead of investigations, as well as cases involving stock option backdating. Are these (and others) not "real world" enough? True that reporters, academics, and market experts have been the main players proving the successes of screens. But that doesn't mean screens are hopelessly academic; rather it may indicate the passivity of some agencies.

I would maintain that none of the arguments commonly advanced against screens really hold up. Why then are screens not more universally adopted? The reason must be beyond any of the ones that I have been discussing here. It could simply be "this is our culture, and we will just not screen." If that is the real reason, then better to put it forward instead of pointing to specious arguments against screens.

[21] Craig Lewis, Risk Modeling at the SEC: The Accounting Quality Model. Securities and Exchange Commission, Speech by the Chief Economist, (December 2012), *available at* Securities and Exchange Commission.

I. "Screens are Used in All Other Areas, but Antitrust Screening is Harder"

Antitrust screening is not more difficult than financial market screening. In fact, it is a whole lot easier. Financial prices such as stock and commodity prices can move apparently randomly, but the price of cars or bread is not comparably volatile. When flagging an abnormal price movement in those markets it is likely easier to dismiss other possible explanations; that may not be the case when studying the price of oil.

The markets of interest to competition authorities often have less data available than those of interest to financial regulators. Still, that situation can start to change and, as discussed in above, there are still many industries ready to be screened.

J. "Screens Just Don't Work. We Used them 40 Years Ago and They Did Not Work"

There is hardly anything in our lives that can be compared, in terms of performance and effectiveness, to what an earlier generation of the same products was 40 years ago. Our medical screens are much better now than they were 40 years ago, and we are thankful for those developments. We are thankful, in other words, that earlier generations did not dismiss medical screens out of hand simply because they weren't perfect.

Did leniency programs "work" really well 40 years ago? It does not seem so, but most likely they work well now. Those older programs were developed and improved—they were made better. I am fairly confident that, compared with the state of the art 40 years ago: (i) the screening technology is significantly better, (ii) data are more readily available, and (iii) computing power and data mining algorithms are incomparably more advanced.

Why would antitrust agencies differ from other agencies that currently use screening? The U.S. Internal Revenue Office, the U.S. Department of Transportation, the U.S. Commodities Futures Trading Commission, the U.S. Securities and Exchange Commission, and the U.S. Federal Trade Commission, among others, all use screens of one type or another. All agencies are resource constrained. They must believe these are not wasteful initiatives, and they have already produced meaningful investigations.

VI. FINAL REMARKS

This article's objective is to make the case that competition authorities worldwide need to be proactive when detecting and deterring cartels, with a focus on screening methodologies. It addresses in detail what I believe to be the key arguments put forward against screening in antitrust and uses recent screening success to illustrate to the reader why screens must be used.

We should focus our discussions on how to start implementing these methods in the best and most efficient way, given resource constraints. The evidence is clear that appropriately developed and implemented screens do work.