Unclassified **DAF/COMP(2013)20**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

23-Oct-2013
_____
English - Or. English

DIRECTORATE FOR FINANCIAL AND ENTERPRISE AFFAIRS
COMPETITION COMMITTEE

Cancels & replaces the same document of 23 October 2013

ROUNDTABLE ON EX OFFICIO CARTEL INVESTIGATIONS AND THE USE OF SCREENS TO DETECT CARTELS

-- Paper by Rosa Abrantes-Metz --

The note by the Secretariat is submitted FOR REFERENCE under item X of the forthcoming Competition Committee meeting to be held on 30-31 October 2013.

Please contact Mr Antonio Capobianco [Antonio.Capobianco@oecd.org] if you have any questions regarding this document.

JT03347081

**ROUNDTABLE ON *EX OFFICIO* CARTEL INVESTIGATIONS AND THE USE OF SCREENS
TO DETECT CARTELS**

Note by Rosa M. Abrantes-Metz[*]

1.   Motivation

1.   Despite the successes of cartel detection over the last twenty years, there are many who believe
that competition authorities have just started to scratch the surface. The focus of this note is to make the
case that proactive detection and deterrence policies need to be established, and that those policies should
be led by the use of empirical screens.

2.   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 identify whether collusion, manipulation or any other
type of cheating may exist in a particular market, who may be involved, and 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.

3.   Over the last few years, economic analysis in general, and empirical screens in particular, have
become increasingly important in cases of conspiracies and manipulations, a trend detailed for example in
Abrantes-Metz & Bajari (2009 and 2010), and Hüschelrath (2010). 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.[1]

4.   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 manipulation,
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.

---

[*]   Rosa M. Abrantes-Metz is a 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 note was 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 31st, 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.

[1]   Surveys of screening methodologies and their multiple applications can be found in Harrington (2006 and
2008) and Abrantes-Metz & Bajari (2009 and 2010). 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. Abrantes-Metz & Sokol (2013), and Abrantes-Metz, Bajari & Murphy
(2010) make the case for the use of screens in corporate antitrust compliance programs.

5.   As this note 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, such as on LIBOR which is perhaps the largest conspiracy and manipulation of its type ever uncovered.

6.   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 of technic are still fairly recent, particularly in antitrust. As a consequence, their worldwide adoption may take some time. Additionally, in some markets data may be difficult to gather which is necessary to employ screening. 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 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?"

7.   This note will directly address these and other concerns expressed on the use of screens. It is organized as follows. Section two reviews the basics of screening, both structural and empirical, and their significant successes. In section three I review leniency programs and their successes, and focus on the value added of screens with respect to leniency in section four. Section five illustrates the arguments made in the previous sections on the use of screens with the flagging of the LIBOR conspiracy and manipulation. A review and detailed discussion of the main criticisms on the use of screens is covered in section six, and section seven concludes.

2.   Screening Methodologies

2.1   Screening Basics

8.   There are essentially two different types of economic analyses pursued with the objective to flag the possibility of a conspiracy (Harrington (2008)). The first can be classified as a "structural approach" and 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.[2]

9.   The second type of economic analysis in cartel detection is empirical and uses what has 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 have been produced under an explicit agreement. An example of a market monitoring program combining both components is that put forward by the European Commission in 2007(European Commission (2007)).

---

[2]   A general list of these factors is further detailed in "Proof of Conspiracy under Antitrust Federal Law (2010)", in the chapter dedicated to the role of the economic expert. See Proof of Conspiracy under Antitrust Federal Law, American Bar Association Editions, April 2010, Chapter VIII. A non-exhaustive "check list" of characteristics that influence the susceptibility of a market to tacit or explicit collusion is: number of firms and market concentration; differences among competitors; product heterogeneity; demand volatility; barriers to entry; benefits of cheating; transparency and multi-market contact.

2.2 Structural Screening

10.   Examples of the literature contributing to the screening structural approach include research on factors related to cartel duration, such as Marquez (1992), Dick (1996), Suslow (2001, 2005), and Levenstein & Suslow (2006, 2010).

11.   Most recently, and following this literature, Abrantes-Metz, Connor and Metz (2013) estimate a cartel duration model and, among other results, find that cartels first detected by the U.S. Department of Justice or the European Commission tend to be longer-lived, likely because those detected by other jurisdictions are primarily follow-ups of related larger and older cartels first uncovered in the United States or Europe; bid-rigging cartels tend to be longer-lived than others, while cartels distributed across geographies tend to be shorter-lived; cartel durations are increasing in the size of a cartel's affected sales and sanctions; industries such as Petroleum & Coal, Finance & Insurance, and Food, Feed, Tobacco & Transportation have shorter-lived cartels, while industries such as Electronic Products have longer-lived cartels; the state of the economy can impact the duration of a cartel as well; and cartels where the leading firm has a market share of at least 40% have longer durations.

12.   Studies of this type focusing on the identification of factors correlated with cartel duration are important. They assist in screening for conspiracies through the identification of characteristics predicting the higher likelihood of a successful (measured by long duration) cartel, which in turn should be used by competition authorities worldwide to screen for anti-competitive behavior. The relevance of this type of studies to enhance antitrust enforcement has been recognized by many, including Grossman (2004), though comparisons across studies of this nature should be taken with care.

13.   An example of this type of approach was developed in Great Britain, at the Office of Fair Trading (OFT). The OFT empirically identified industry-level variables that predict cartel activity, using a data set of DOJ's price-fixing cases since 1994, and European Commission's price-fixing cases since 1990 (Grout and Sonderegger (2005)). The OFT tried to predict the incidence of cartels with industry level data, and concluded that industry turnover, cost measures, concentration measures, entry barriers, and employee costs, among others, help explain the probability of collusion in an industry.

14.   Another example of the use of a structural approach comes from The Netherlands competition authority, in which economic indicators are used to flag possible cartels, and have successfully done so on a regional shrimp cartel.

2.3 Empirical Screening

15.   Though important, a structural approach is typically insufficient particularly when quality data are available, and it may also lead to too many false positives. In cases where the existence of an alleged conspiracy is in dispute, a behavioral or outcomes oriented approach will be needed. This approach develops and implements (empirical) screens.

16.   Broadly speaking, empirical screens for collusion used in the literature employ two strategies. The first is to search for improbable events. This type of screen is similar to looking for a cheat in a casino. For example, the probability that a gambler at a Las Vegas casino will place a winning bet in a roulette is roughly .5 percent. During her shift, a roulette dealer may see a handful of players win 5, or even 7, times in a row. However, the probability of winning 20 times in a row is almost zero (though not impossible). If a pit boss sees this occur, he may not be able to prove that cheating has occurred, but he would be well advised to watch closely or risk losing a lot of money. One set of collusive screens generalizes this idea by looking for events that are improbable unless firms in the industry have coordinated their actions.

17.   The second type of screen uses the idea of a control group. A somewhat extreme example illustrates the idea. In the 1980s, organized crime in New York City operated a concrete club that rigged bids on contracts over $2 million. During the 1980s, the price of concrete was 70 percent higher in New York City than other U.S. cities. While it is true that the price of many goods and services is higher in New York City, relatively few of those prices are 70 percent higher than in other large cities. Prices that are anomalous compared to other markets suggest a competition problem. In this simple example, we are forming a control group for New York by using prices in other cities as a basis for comparison. But most collusion is not this blatant.

18.   Though screens can be very powerful, these are econometric tools, with all the usual caveats, and they may potentially be misused. In my view, we can think of two golden rules when designing and implementing screens that need to be kept in mind. These rules should be obvious when stated, but may sometimes be forgotten. First, "one size does not fit all," and second, "if you put garbage in, you get garbage out."

19.   The first screening rule states that a screen needs to be designed or at least adjusted to the situation at hand. Just because a given set of variables and model specification prove highly effective when estimating the demand for bread, it does not mean that those same variables or specifications work when estimating the demand for cars. In turn, the second screening rule states that, as is always the case in empirical work, a screen is only as good as the choices of what is put into it (Abrantes-Metz (2011) discusses in further detail issues related to the development of screens).

20.   In general, six requirements are key to developing and implementing a good screen:

   i.   An understanding of the market at hand, including the nature of competition and the potential incentives and opportunities to cheat—both internally and externally—to a firm;

   ii.   A view of the likely nature of cheating;

   iii. A view of how cheating will affect market outcomes, and in particular, the available data;

   iv.   A set of statistics based in an econometric model, capable of capturing both the implications of cheating as well as ordinary, natural relationships between key market variables;

   v.   The identification of an appropriate non-tainted benchmark against which the evidence of cheating can be compared; and

   vi.   Empirical or theoretical support for the screen.

21.   There are particular collusive markers which may be expected to occur during a price-fixing or bid-rigging conspiracy which include, but are not limited to:[3]

   i.   Prices that are high and have low variance; prices that are unresponsive to changing market conditions, and those representing a structural break which cannot be explained by regular market

---

[3]   The U.S. Department of Justice has identified a set of pricing patterns that are intended to help identify collusive behavior. *See* U.S. Dep't of Justice, Antitrust Division, *An Antitrust Primer: Price Fixing, Bid Rigging, and Market Allocation Schemes: What They Are and What to Look For*, at http://www.usdoj.gov/atr/public/guidelines/primer-ncu.htm ("DOJ Antitrust Primer").

conditions; or similarly, prices which do not seem to change in response to market structural breaks; prices which are not responsive to cost changes.

ii.   Bids across competitors which are highly correlated even after controlling for common legitimate factors among them such as costs and number of competitors; and bids which do not appropriately reflect costs.

iii.   Market shares that are stable over time among cartel members, also presenting negative serial correlation which reflect such stability; absence of rotation of customers among cartel members over time.

22.   Very recently in 2008, empirical screens flagged the possibility of a conspiracy and manipulation of the USD LIBOR ("LIBOR") by major worldwide banks, as in Wall Street Journal articles April and May 2008 on the topic, Abrantes-Metz, Kraten, Metz and Seow (2008, 2012), and followed by empirical evidence provided in Snider and Youle (2010), Abrantes-Metz, Judge and Villas-Boas (2011) and Abrantes-Metz & Metz (2012). The initial 2008 screens preceded worldwide investigations on the matter which became public in March 2011. Only late in the Spring of 2011, did one of the LIBOR panel contributing banks, UBS, file for leniency with the U.S. Department of Justice ("DOJ"). This is how screens should work and how they can supplement and enhance leniency applications. The example of LIBOR will be discussed in further detail in section 5.

23.   Screens main purpose should be for detection, but they can have multiple applications in litigation and also in the pre-litigation, as explained and exemplified in Abrantes-Metz and Bajari (2009 and 2010). The ongoing litigation on an alleged conspiracy and manipulation of LIBOR provides an example of the use of screens for detection by agencies worldwide, by plaintiffs, by defendants, and internally to companies in their compliance programs.

3.   Leniency Programs and their Successes

24.   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 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 prior. These were also accompanied by significant increases in penalties and more extensive use of jail time.

25.   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.

26.   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 (Riley (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 (Harrington (2010)).

27.   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.

28.   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 (Connor and Lande (2012)). 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 (Klawiter (2012)). These are factors that need to be taken seriously into account when designing and adjusting a leniency program.

4.   Limitations of Leniency and Value Added of Screens

29.   While leniency programs in the U.S. 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. Leniency programs do present shortcomings and opportunities to be enhanced and complemented by other detection tools such as screens.

30.   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 as 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 (Klawiter (2012)).

31.   This high level of concentration may in fact be driven by the DOJ's Leniency Plus program. Under this, 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.

32.   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.

33.   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.

34.   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 which 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.

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

36.   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 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?

37.   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.

38.   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 anti-cartel detection program needs to rely on additional proactive tools. Well-developed screens will pass a cost-benefit analysis and can be successfully implemented.

39.   A proactive policy by competition authorities could be composed of both structural and behavioral approaches, with the later centered on empirical screens (Friederiszick and Maier-Rigaud (2008)). It is recognized that structural screens may lead to too many false positives – because there usually are many omitted factors influencing cartel formation (Harrington (2010)), and as a consequence, only a small fraction of markets and agreements verifying the check list 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.

40.   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.

41.   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 which 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 anti-cartel programs worldwide.

42.   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, and LIBOR conspiracy and manipulation (leading to the launch of investigations of benchmarks worldwide), and most recently, the alleged conspiracy and manipulation in the foreign exchange markets – have two common components: they were all initially flagged by empirical screens, and none 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.

5. Lessons from LIBOR for Detection and Deterrence of Cartels

43. Worldwide investigations on the alleged conspiracy to manipulate the U.S. dollar London Interbank Offered Rate ("LIBOR") and other currency-denominated LIBOR, as well as similar benchmarks such as Euribor in Europe and TIBOR in Japan, have been making news. Most recently, investigations to Platts Oil Benchmarks, ISDa Fix benchmark for swaps and foreign exchange rates benchmarks have also been underway.

44. What is LIBOR? On a daily basis, the 16 (then) participating banks are surveyed by the British Bankers Association and submit sealed quotes which answer "[a]t what rate could you borrow funds, were you to do so by asking for and then accepting interbank offers in a reasonable market size just prior to 11:00 a.m. London time?" LIBOR is computed by averaging over the middle eight quotes, discarding the four highest and the four lowest.

45. From a structural point of view, it should have been clear to the relevant authorities and institutions that the LIBOR setting has many ingredients facilitating cheating, mainly:

  i. The number of surveyed banks is small for the dimension of this market, which facilitates coordination, even more when participants have multimarket contact, as favors can be traded across different markets;

  ii. Submissions are based on estimates rather than actual transactions, which are significantly easier to manipulate; in addition, they re not verifiable against any actual transactions, i.e., complete lack of an auditing process;

  iii. Quotes are submitted sealed but are publicly disclosed daily which facilitates potential coordination across banks, given that monitoring behavior and punishment of deviations from collusive agreements become easy to implement;

  iv. Banks' trading desks have financial interests benchmarked to LIBOR, and therefore have a direct interest in the value that LIBOR takes. Not only knowing which direction is LIBOR going to move on a particular day, ahead of the market, but also having the ability to move LIBOR to the desired level to benefit portfolio or trading positions, can provide the motive to manipulate and collude to benefit trading positions;

  v. Banks' quotes are important signals of their financial health, as none of them wants to admit it can't borrow cheaply; and

  vi. Banks oversaw themselves through their trade association, the British Banks Association (BBA).

46. Together, these provide the means, motive and opportunity to collude and manipulate LIBOR, the structural component of screening and described in the previous sections. In and of itself, this should have been an important red flag for possible illegal behavior, but it went by apparently unnoticed for decades, until empirical screening provided the additional evidence which apparently ended up getting the authorities' attention.

47. The empirical screening of LIBOR occurred in 2008, several decades after LIBOR was born. Arguably, LIBOR investigations were triggered by these findings. The WSJ first looked at LIBOR and

found suspicious patterns in April and May of 2008. Its screen showed that since January 2008, the banks' individual quotes were too low compared to what their credit default swap spreads indicated. Abrantes-Metz, Kraten, Metz and Seow followed with a working paper in August 2008 which noted that: LIBOR was completely constant day-in and day-out for many months prior to the financial crisis, while other comparable rates such as the Federal Funds Effective Rate were not; most banks' quotes were identical to each other, while other market indicators such as their pricing in the capital markets, indicated differences (even if slight) across banks; LIBOR did not respond to increasing risk at least since late Spring and early Summer of 2007. In fact, we saw signs of possible collusion dating back to prior to financial crisis.

48.   Both the WSJ articles and the Abrantes-Metz et al (2008) analyses used screens to flag such anomalies. Through settlement agreements with some of the banks, we know now that investigations on LIBOR were initiated in October of 2008, though they only became publicly known in March of 2011.

49.   Other research on LIBOR was then conducted by Snider & Youle (2009 and 2010), by Abrantes-Metz, Villas-Boas, & Judge (2011), and by Abrantes-Metz & Metz (2012). Snider and Youle present findings supporting their claim that banks' LIBOR quotes are difficult to rationalize by observable cost measures for the period under study, including a given bank's quotes in other currency panels. They also introduce a model in which banks' possession of LIBOR indexed contracts induces them to produce LIBOR quotes that are clustered around discontinuities. Using this model the authors show that there was a severe clustering in the USD LIBOR for the three month maturity throughout 2009. Snider and Youle further try to quantify gains from such behavior and present results showing large exposures to LIBOR by several banks through their interest rate derivative portfolios, allegedly enabling them to profit from the rapid decline of LIBOR starting late in the Summer of 2007. The authors argue that these exposures may have been the incentive behind a deliberate misreporting of LIBOR quotes by the banks.

50.   Abrantes-Metz, Villas-Boas & Judge (2011) show that Benford's Law, a mathematical law commonly used to detect fraud in other contexts, is violated for the USD LIBOR from early 2006 through the Summer of 2007 and with continuing anomalous patterns throughout 2009. In 2012, Abrantes-Metz & Metz followed with an analyses similar to bid-rigging to explain that, given that LIBOR quotes are submitted sealed, and that they are supposed to be idiosyncratic to each of the banks as forecasts of their own interbank borrowing costs, simultaneous moves by a large number of submitters from one day to the following to the exact same number are more consistent with explicit rather than with tacit collusion. This article is discussed in further detail in section six, as an example of our empirical screens can be used to attempted to distinguish explicit from tacit collusion.

51.   There seemed to have been a lack of understanding by the relevant authorities for the incentives and opportunities to cheat, and to the clearly anomalous data patterns. The internal auditing procedures at the banks applied to LIBOR submissions were either nonexistent or clearly inadequate. Had such internal mechanisms based on screens been in place, banks would have been able to detect and hopefully deter problem submissions, as argued in Abrantes-Metz & Sokol (2012). And had these same tools been applied by the relevant agencies, they would have flagged the anomalies perhaps years earlier, and before the successful screens applications by the WSJ and Abrantes-Metz et al (2008).

52.   In addition to detection, the critical role of screens in compliance is recognized and stressed in the settlements reached between the regulatory agencies and the three banks to date on this matter. In fact, as part of these settlements, authorities demanded banks to implement these monitoring and auditing tools internally.[4]

---

[4]   As an example, the Settlement Agreement between Barclays and the U.S. Department of Justice, the U.S. Commodities Futures Trading Commission and the U.K Financial Services Authority requires the

53.   It is reasonable to expect that these types of requirements for enhanced monitoring may become a reality in other areas as well, not just internally to companies, but externally also in the context of the ongoing reform of financial benchmarks worldwide lead by the U.S. CFTC and the U.K. Financial Conduct Authority ("FCA") through the International Organization of Securities Commissions ("IOSCO"). At least it is reasonable to expect that corporations will start to seriously consider the adoption of these methods, not only to protect themselves against cheating within the corporation, but also against possible cheating up- and down-stream.

54.   LIBOR is the perfect example of how screens can be effective, proactive, add value, lead to leniency applications and Leniency Plus applications in a variety of other markets, lead to uncovering of direct evidence, and have a material impact on cartel detection and deterrence.

   i.   Empirical screens first flagged the possibility of a manipulation and conspiracy of LIBOR and signaled the need for regulatory agencies to inquire;

   ii.   Had structural screens been applied early only, they would have raised various flags given the structure of the LIBOR setting and the incentives to cheat which would have prompted the use of empirical screens earlier on;

   iii.   Investigations were started shortly after the empirical evidence put forward through screening, in October of 2008;

   iv.   Given (iii) and the fact that the flagging was motivated by empirical screens, leniency applications were generated years into the investigation and a few short months after these became public (public knowledge of these investigations happened on March 15, 2011, through a disclosure in UBS's 10-k, to which UBS followed with a leniency application in May of the same year);

55.   In fact, it is likely that the LIBOR manipulation and conspiracy may never have been uncovered had screens not been used. It is particularly unlikely that it would have been identified through leniency – contributing banks would have had no incentive to report on any alleged cheating since profits derived from these may have been very substantial. They did not report for many years, while the alleged behavior seemed to have been ongoing, why would they do so at any other time?

   i.   LIBOR investigations have led to investigations into many other similar rates and benchmarks across the board, and to leniency and Leniency Plus applications as well;

   ii.   LIBOR investigations have led to a substantial effort worldwide though IOSCO to reform financial benchmarks, with the objective of enhancing their robustness and reliability, i.e., having in mind the structural screening approach, to change the structures and incentives, detection and deterrence of similar future behavior;

   iii.   Enhanced internal monitoring programs have become required by the regulatory agencies involved in the investigations, and these also require the use of screening tools;

56.   But how about improving anti-cartel policies? After all, and based on public information, illegal behavior may have been ongoing for a very long time and have been widespread, yet it was completely missed by all relevant agencies and regulatory bodies worldwide — shouldn't this imply that an

---

following procedures with respect to Barclay's LIBOR submissions (Barclay's Settlement Agreement, pages 37-38).

improvement is also needed at this level? If we were to take a similar position here as many competition authorities have with respect to other detection tools and how ineffective those may be in their views, what would the LIBOR scandal imply for the effectiveness of currently reactive anti-cartel policy tools?

57.   There have always been those who are naturally skeptical that simple empirical analyses can be brought to bear in complex markets. Hopefully, the LIBOR scandal will settle the question of whether screens should be more vigorously applied and move the discussion to how that needs to happen.

6.   Addressing Concerns Raised on Screens

58.   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.

59.   Some of the arguments also place screens to a standard of proof that no other screening methods in other areas of research neither any other tools 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 anti-cartel detection policy. The major concerns and counterarguments as presented below and should not dissuade agencies from actively adopting screens.

6.1   *"Screens Have High Error Rates, Erroneously Identifying Cartels Where None Exit"*

60.   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.

61.   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. And they certainly present limitations on addressing the *intention* of particular wrongdoing, though this may well depend on the specific case at hand.

62.   Examples already exist in which the power of screens is tested. One of such examples was performed by the Italian Competition Authority, as documented in Esposito and Ferrero (2006). In this paper, the authors tested the power of the variance screen for prices to detect previously known illegal conspiracies. In particular, they pose the question of whether a price variance screen (as initially proposed in Abrantes-Metz, Froeb, Geweke and Taylor (2006)) could have identified collusion in two well-known Italian cartel cases involving gasoline and diesel on the one hand and in baby milk on the other. They also ask whether such a screen could have correctly identified who was involved and during which time period. And the answer to both questions is "yes:" the screen would have correctly identified these two cartels before the Italian Competition Authority did. Another application in which the power of screens was demonstrated involved the German cement market, as discussed in Hüschelrath and Veith (2011). The authors show that buyers could have detected this cartel ahead of the launch of investigations, through the use of screens for structural breaks.

63.   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.

64.   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 do leniency applications (and other detection tools for that matter) really have. It seems to me that it is critical to have this information in order to disregard screens in favor of leniency or other tools.

65.   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?

6.2   *"Screens Cannot Distinguish Explicit from Tacit Collusion"*

66.   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?

67.   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 which may indicate whether such an outcome was more likely to have occurred through tacit or explicit collusion.

68.   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?

69.   Below I will discuss 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.

Example 1: USD LIBOR

70.   In order to clearly illustrate this important point, below I will present a detailed analysis on LIBOR, first put forward in Abrantes-Metz & Metz (2012), in which I am my co-author 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.

71.   In Abrantes-Metz & Metz (2012), we calculate the coefficient of variation for the determining set of LIBOR quotes each day;[5] which measures the dispersion or variability in the daily quotes of the participating banks. If the contributing banks were submitting unique quotes each day (which just happened to average to the same level day after day), the coefficient of variation would be large. But if the banks were all submitting essentially the same quote, it would be low, and in the extreme case where the middle eight quotes were identical, the coefficient of variation would be 0. Figure 1 from our paper presents this measure. It is clear that from early August 2006 through early August 2007, the middle eight quotes are essentially identical day in and day out. This, again, seems highly anomalous.

**Figure 1: There is Almost No Variation Across the Middle Eight
Quotes from August 2006 Through August 2007**
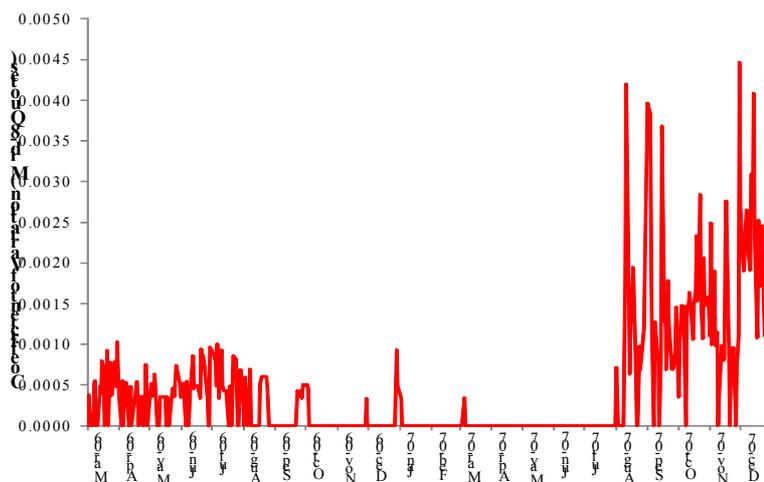
72.   How can we account for the stability of LIBOR and the convergence among the quotes? Let's consider some possible explanations. One possibility, of course, is that at least some of the banks were *explicitly colluding*. That is, they were in contact with each other and agreed in advance on a virtually identical quote to submit.

73.   Another possibility is that there was *tacit collusion*. Tacit collusion, or conscious parallelism, exists when there is strategic coordination between participants. Each participant attempts to optimize its strategy with consideration as to how the others may respond and adapts its actions based on learning from others' strategies. This can result in convergence of individual outcomes. Consider, for example, the phenomenon of equal gasoline prices across stations on each corner of an intersection. While that might be the result of explicit collusion, it might also be the result of tacit coordination.

74.   Yet a third possibility is that convergence across quotes represents a non-cooperative outcome. Perhaps the banks are each independently reacting to (or anticipating) a common market driver. Without any consideration of how the other banks may react, the banks may individually arrive at the same number. Perhaps, in other words, many of the participating banks truly had identical borrowing costs, or at least expected to have identical borrowing costs at particular maturities.

---

[5]   Recall that the coefficient of variation is the standard deviation divided by the mean. The mean, of course, is just the Libor level itself. We exclude from this computation the quotes we know did not enter the Libor computation on each given day.

75.   Each of these—explicit collusion, tacit collusion, or non-cooperative outcomes—may result in an empirically similar outcome in the LIBOR setting: convergence across quotes. The question we turn to now is whether they are empirically *identical*. Can screens draw any distinctions among these?

76.   Actually, these three different theories—explicit collusion, tacit collusion, and non-cooperative outcome—lead to subtle differences in empirical predictions in the LIBOR setting. Consider first the possibility that the convergence in quotes was simply a non-cooperative outcome; that the participating banks independently arrived at the same quote for some common legitimate fundamental reason. To accept this explanation we would have to understand why these banks, which differ in important ways, should nevertheless have identical borrowing costs. Not just "similar" borrowing costs, but *identical*. The member banks have varying sizes, varying asset portfolios of varying risk profiles, and varying liability structures. They participate to different degrees in different market segments. It seems highly unlikely that they should then have fundamentally identical borrowing costs, though we may expect them to be similar.

77.   We would make a second point. To accept the non-cooperative outcome explanation we need to understand why the quotes weren't always common. In other words, as shown in Figure 1 above, why did the coefficient of variation suddenly drop to 0 in the beginning of August of 2006? And why then did it abruptly increase after August of 2007? We know that the structural break on August 9, 2007 was due to particular events on that day which triggered the "official start" of the financial crisis.[6] On that day, LIBOR quotes drastically changed and became immediately different from each other. Which also begs the question of why the triggering of the crisis did not affect banks' quotes equally, if market conditions were the reason why they were identical previously in the first place. What fundamentally changed during the 12 months from August 2006 to August 2007 that could account for this independent convergence? We require an explanation as to why those costs would be identical for a 12-month window, but not before, and not after.

78.   So the data seem inconsistent with a non-cooperative outcome. That leaves us with two other possibilities. How might we empirically distinguish explicit from tacit collusion? If the banks slowly converged to the same common quote, and then repeated it over and over again, we would see results as in Figure 1 for the most part. But if the collusion were tacit, if the banks were "learning" from the strategic reaction of the other banks, then we would expect to see a transition period in Figure 1 in which the variation of intraday quotes would be decreasing towards zero. We don't. Instead, we see an abrupt transition to 0.

79.   Further complicating the tacit collusion argument is the fact that the quotes are sealed. Only after LIBOR is computed are the quotes made public. If the banks submitted largely the same quote day after day, and we saw that the other banks were learning and converging toward that common quote, we might more likely ascribe this to tacit, strategic collusion. But if we see that many banks submit a common (sealed) quote one day, and then submit a common but different (sealed) quote the next day, that is more difficult to understand as tacit behavior.

---

[6]   On August 9, 2007, there were major news on the realization of a liquidity and subprime crisis: (i) there was a "coordinated intervention" by the European Central Bank, the Federal Reserve Bank, and the Bank of Japan; (ii) AIG warned that defaults were spreading beyond the subprime sector; and (iii) BNP Paribas suspended three funds that held mortgage backed securities.

80.   Another justification we have heard from this behavior, and interestingly from regulators investigating LIBOR, is that "the quotes were the same because the banks were all benchmarking against one particular index." Really? Let's think this through. If they were benchmarking their quotes against another index, why didn't they do that before, why only now? And if they switch to such benchmarking on these dates, how did they suddenly come up with such an agreement from one day to the next to do so? Would that still require an explicit agreement among the banks to do so, given that the change in quotes patterns is so sudden?

81.   Table 1 below details the quotes of early August 2006. The quotes not highlighted represent those which were excluded from determining LIBOR on those days. Table 2 summarizes the quotes by listing the unique values that determined LIBOR on those days, and the number of banks that share that quote.

**Table 1: Individual Quotes for 1 Month USD LIBOR
from Early August 2006**

	2006				
	August 3	August 4	August 7	August 8	August 9
BTMU	5.410	5.430	5.370	5.370	5.330
Bank of America	5.400	5.420	5.380	5.370	5.325
Barclays	5.410	5.420	5.370	5.370	5.340
JPM Chase	5.410	5.420	5.380	5.370	5.330
Citi Bank	5.405	5.420	5.360	5.370	5.330
CSFB	5.405	5.420	5.360	5.370	5.330
Deutsche Bank	5.405	5.415	5.365	5.365	5.325
HBOS	5.410	5.420	5.350	5.370	5.330
HSBC	5.400	5.420	5.370	5.370	5.330
Lloyds	5.410	5.420	5.360	5.370	5.330
Norinchukin	5.410	5.420	5.370	5.370	5.340
Rabobank	5.405	5.415	5.370	5.370	5.330
Royal Bank of Canada	5.405	5.420	5.370	5.368	5.330
Royal Bank of Scotland	5.400	5.420	5.370	5.370	5.330
UBS AG	5.405	5.420	5.370	5.370	5.330
West LB	5.405	5.460	5.360	5.370	5.330

Table 2: The Distribution of "Determining" Quotes

		2006				
		August 3	August 4	August 7	August 8	August 9
Value 1		5.405	5.420	5.360	5.370	5.330
	Count	7	12	4	14	12
Value 2		5.410		5.365		
		6		1		
Value 3				5.370		
				8		

82.   For example, on August 3, seven banks submitted the quote 5.405 and six banks submitted the quote 5.410. But then on August 4, twelve banks submitted a quote of 5.420—different from any of the quotes submitted the previous day. On August 7 there were three different values that were part of the "middle eight" quotes on that day, 5.360 (submitted by four banks), 5.365 (submitted by one bank) and 5.370 (submitted by eight banks). And on August 8, fourteen of sixteen banks submitted a quote of 5.370.

83.   One might interpret that as tacit learning: eight banks submitted 5.370 on August 7, and then seven more did on August 8 (with one bank changing its quote away from 5.370). But let's consider the quotes of August 9. Twelve banks submitted a quote of 5.330. No bank had submitted that quote in the four days prior to August 9; in fact, no bank had submitted that quote since June 30, 2006 when two banks

16

had. Yet, somehow, twelve banks moved to it on that date. This seems inconsistent with the tacit collusion or learning theory, especially when we recall that the quotes are submitted sealed.

84. And just as was the case with the non-cooperative outcome theory, to accept the tacit collusion theory we would have to understand why there was strategic learning leading to convergence in the period August 2006 through August 2007, but not before, and not after. What was different about those twelve months?

85. So the data seem inconsistent with tacit collusion as well. That leaves us with explicit collusion as the most likely explanation. Of course, we have repeatedly explained in previous work that the patterns identified through screening are not proof of explicit collusion; but this explanation simply seems more likely than the alternatives. From August 9, 2006 onwards there are a few rare episodes in which a few banks submit slightly different quotes from the majority, but for the most part, and as evident in Figure 2, the group of banks effectively determining LIBOR submitted the same identical quote day in and day out for one year, until August 9, 2007.

Example 2: Undisclosed Major Financial Benchmark

86. Let's now look at a second example. Figure 2 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.

87. 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 2 days prior to that date.

88. Figure 2.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 3 setting dates. But on the 3 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 3 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 3 consecutive days, and not at least 15 days before or after those 3 consecutive days, never happens outside of those 3 setting days. These data immediately suggest a few questions. Since the differences in intraday variation among the two groups 15 days before and 15 days after the 3 dates are very similar, why are they so different during the 3 setting dates? And during the 3 setting dates, Group 2 continues 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?

89. Figure 2.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 3rd 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 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.

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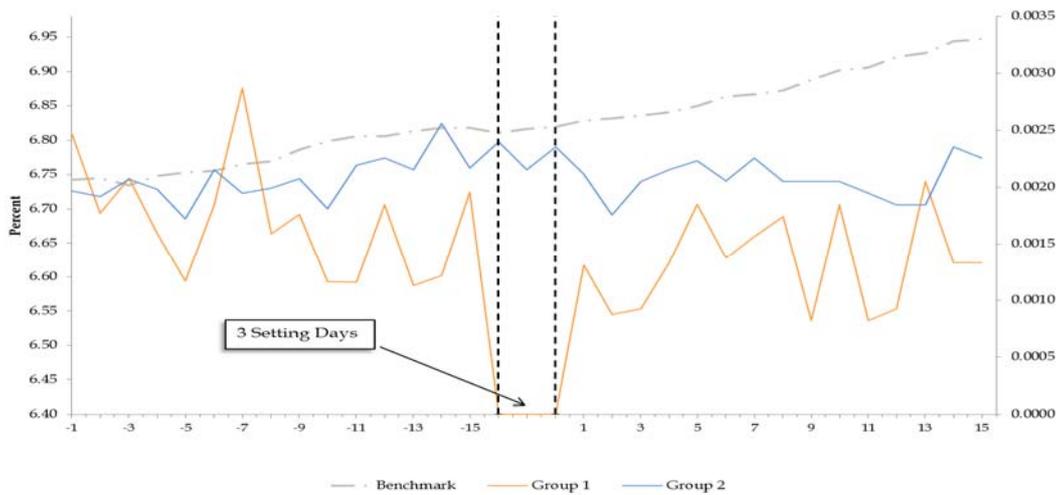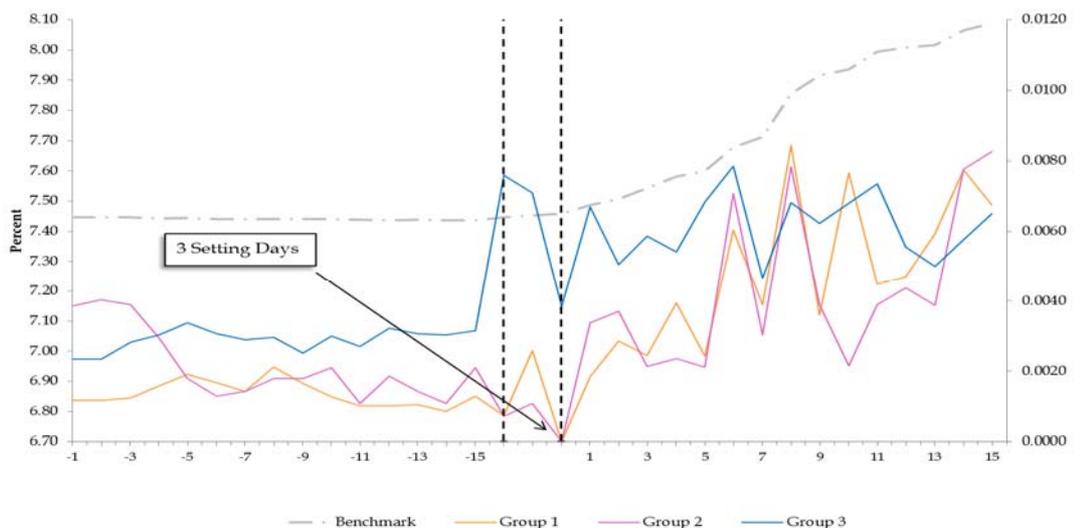Contracts Based on this Benchmark**

**Figure 2.B: Intraday Coefficient of Variation of Bank Quotes
for Undisclosed Benchmark around Setting Dates for Derivatives
Contracts Based on this Benchmark**

Example 3: Canadian Bid-Rigging

91.   In September of 2012, a Canadian reporter flagged the possibility of a bid-rigging, market allocation and price fixing conspiracy in road construction in Montreal (Montreal Gazete (2012)). She noticed that winners frequently subcontracted losers, and that the same companies tended to win in the same regions repeatedly, among other patterns. Granted this could have been reached through tacit agreements, at least in principle, but the frequency of unexpected patterns was very high. But 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?

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

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

93.   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 detailed earlier in this note, they may well target cartels with differing market effects, and represent a proactive versus a reactive anti-cartel policies.

94.   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 were 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.

95.   These analyses do not have to involve a massive amount of resources. For example, I find it unlikely that the analysis conducted by the 2008 WSJ articles on LIBOR took more than a few days to put together, or used more than a handful of data series. My 2008 paper on LIBOR also only took less than one week for data to be compiled and results calculated. The Canadian reported 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 could reporters and academics run them to flag such large cases?

96.   These are clear examples of screens which did not take long and in many cases were performed by reporters, not expert economists. If they could do this, why can't competition authorities do the same?

97.   Of course screens can on occasion 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 (Ragazzo (2012)). Screens were used to select those cases in which market evidence was the most significant. Out of hundreds of possibilities, 10 were selected through screening, and among those 10, direct evidence of collusion was found for 6 cases. A success rate of 60% 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 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.

98.   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 (Mena-Labarthe (2012)). The effort did take significant resources, but it was worth it in the view of the Commission, and furthermore, it was recognized in court has constituting valid evidence for the collusive claim put forward by the Commission.

99.   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. None of the examples provided above were technically challenging, or else how could reporters have performed most of these analyses?

100.   Another example of a simple first stage analysis is contained in my recent paper on the aluminum market dislocation, from September 2013 (Abrantes-Metz (2013a)). This paper discusses the aluminum market, the structure,the incentives and opportunity to manipulate and collude. For the most part it simply plots the data on prices, quantities and other measures – it cannot get any easier than that – and evaluates them through an understanding of the market itself. Despite its simplicity, I believe the paper makes a reasonably compelling argument for possible collusion.

101.   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.

6.4   *"Screens Lack Robustness, Why Should I Use Them?"*

102.   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. 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.

103.   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.

104.   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 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.

105.   Screens are no different in this regard. It is not a fair criticism to blame the screen when it delivers different results if applied to the levels of variables rather than to their growth rates, when using different benchmarks or different time periods, or 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 them 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.

106.   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.

107.   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 (Abrantes-Metz, Froeb, Geweke and Taylor (20006); Athey, Bagwell and Sanchirico (2004); Harrington and Chen (2006)). 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.

6.5   *"Why Use Screens if Cartel Members Will Learn to Beat Them"*

108.   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 US 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?

109.   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.

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

110.   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 U.S., and data on bids for many procurements are publicly and easily available across the country. It bares 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.

111.   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 is currently under investigation by the U.S. 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.

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

112.   I am hoping that, if the reader is following this note 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.

6.8 *"Screens are Very Popular in Academia, But They Do not Work in the Real World"*

113.   How can this argument against screens be credibly made given all of the significant examples of their successes? LIBOR is the latest, but not the last, I am sure. In just the last two weeks Bloomberg has flagged an alleged conspiracy and manipulation in the FX markets (Vaughan and Finch (2013)). Other very real world examples include the Canadian matter flagged by a reporter and the Mexican and the Brazilian successes, all discussed above.

114.   Most recently, the SEC has initiated investigations of hedge funds generated through their screening programs (Lewis (2012)). 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.

115.   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 just state it and not make specious arguments against screens.

6.9 *"Screens are Used in All Other Areas, but Antitrust Screening is Harder"*

116.   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.

117.   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 section 6.6, there are still many industries ready to be screened.

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

118.   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.

119.   Did leniency programs "work" really well 40 years ago? It does not seem so, but most likely they work well now. Those 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.

120.   Why would antitrust agencies differ from other agencies which 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, 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.

7.   Final Remarks

121.   This note'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 the LIBOR screening success to illustrate to the reader why screens must be used. The LIBOR scandal provides valuable lessons on cartel detection and deterrence that must be understood by the relevant antitrust agencies and internalized.

122.   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.

123.   It is my hope that this note will contribute to a stimulating discussion at the upcoming OECD meeting and to the further adoption of screening methodologies in antitrust around the world, following the very successful examples from the Mexican and the Brazilian Competition Authorities.

REFERENCES

ABA, Proof of Conspiracy under Antitrust Federal Laws, 2010.

Abrantes-Metz, Rosa, 2013a. Aluminum Market Dislocation: Incentives, Structure and Reform. Working Paper.

Abrantes-Metz, 2013b. Proactive versus Reactive Anti-Cartel Policy: The Role of Empirical Screens. Working Paper.

Abrantes-Metz, Rosa, Connor, John and Metz, Albert, 2013. The Determinants of Cartel Duration. Working Paper.

Abrantes-Metz, Rosa and Sokol, D. Daniel, 2013. Antitrust Corporate Governance and Compliance. Handbook of Antitrust Economics, forthcoming.

Abrantes-Metz, Rosa and Sokol, D. Daniel, 2012. Lessons from Libor for Detection and Deterrence of Cartel Wrongdoing." Harvard Business Law Review Online 3, pp 10-16.

Abrantes-Metz, Rosa and Metz, Albert, 2012. How Far Can Screens Go in Detecting Explicit Collusion? New Evidence From the Libor Setting, Competition Policy International Antitrust Chronicle, March.

Abrantes-Metz, Rosa, Kraten, Michael, Metz, Albert and Seow, Gim, 2012. LIBOR Manipulation? Journal of Banking and Finance, 36, pp. 136-150, first draft dated August 4, 2008 and available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1201389.

Abrantes-Metz, Rosa, 2011. Design and Implementation of Screens and Their Use by Defendants. Competition Policy International Antitrust Chronicle, September.

Abrantes-Metz, Rosa, Villas-Boas, Sofia and Judge, George, 2011. Tracking the Libor Rate. Applied Economic Letters, January, pp. 1-7.

Abrantes-Metz, Rosa and Bajari, Patrick, 2010. Screens for Conspiracies and their Multiple Applications. Competition Policy International, 6(2), pp. 129-144.

Abrantes-Metz, Rosa, 2010. The Power of Screens to Trigger Investigations, Securities Litigation Report, 10(10), November.

Abrantes-Metz, Rosa and Bajari, Patrick, 2009. Screens for Conspiracies and their Multiple Applications. Antitrust Magazine, 24 (1), Fall.

Abrantes-Metz, Rosa, Froeb, Luke, Geweke, John and Taylor, Christopher, 2006. A Variance Screen for Collusion. International Journal of Industrial Organization, 24, pp. 467-486.

Athey, Susan, Bagwell, Kyle and Sanchirico, Chris, 2004. Collusion and Price Rigidity. Review of Economic Studies, 71(2), pp. 317-349.

Chile's Fiscalia Nacional Economica, 2012. Programas de Cumplimiento de la Normativa de Libre Competencia, Material de Promocion No 3, June, available at www.fne.gob.cl.

Christie, William and Shultz, Paul, 1994. Why do NASDAQ market makers avoid Odd-Eighth Quotes? Journal of Finance, 49(50), pp. 1813-1840.

Connor, John and Lande, Robert, 2012. Cartels as Rational Business Strategy: Crime Pays. Cardozo Law Review, 34, pp. 437-490.

Dick, Andrew, 1996. When are cartel stable contracts? Journal of Law and Economics, 39(1), pp. 241-283.

European Commission, 2007. Implementing the New Methodology for Product Market and Sector Monitoring: Results of a First Sector Screening." 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."

Friederiszick, Hans and Maier-Rigaud, Frank, 2008. Triggering Inspections Ex-Officio: Moving Beyond a Passive EU Cartel Policy, Journal of Competition Law and Economics, 4(1), pp. 89-113.

Grout, Paul and Sonderegger, Silvia, 2005. Predicting Cartels, Office of Fair Trading, Economic Discussion Paper, March.

Grossman, Peter, 2004. How Cartels Endure and How They Fail. Studies in Industrial Economics, Editions.

Harrington, Joseph, 2010. Leniency Programs: Past Experiences and Future Challenges. Instituto Milenio SCI.

Harrington, Joseph, 2008. Detecting Cartels. Handbook of Antitrust Economics, P. Buccirossi Editor.

Harrington, Joseph, 2007. Behavioral Screening and the Detection of Cartels, European Competition Law Annual 2006: Enforcement of Prohibition of Cartels, Claus-Dieter Ehlermann and Isabela Atanasiu Editors.

Harrington, Joe and Chen, Joe, 2006. Cartel Pricing Dynamics with Cost Variability and Endogenous Buyer Detection. International Journal of Industrial Organization, 24, pp. 1185-1212.

Hüschelrath, Kai, 2010. Economist's Note: How are Cartels Detected? The Increasing Use of Proactive Methods to Establish Antitrust Infringements. Journal of European Competition Law and Practice, September, pp. 1-7.

Laitenberger, Ulrich and Hüschelrath, Kai, 2011. The Adoption of Screening Tools by Competition Authorities, Competition Policy International Antitrust Chronicle, September.

Levenstein, Margaret and Suslow, Valerie, 2006. What Determines Cartel Success? Journal of Economic Literature, 44(1), pp. 43-95.

Levenstein, Margaret and Suslow, Valerie, 2004. Studies of Cartel Stability: A Comparison of Methodological Approaches. In Peter Grossman, Editor, How Cartels Endure and How they Fail: Studies of Industrial Collusion, United Kingdom, Edward Elgar, pp. 9-52.

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

Klawiter, Donald, 2012. Conspiracy Screens: Practical Defense Perspectives. Competition Policy International, March.

Klawiter, Donald, 2011. Enhancing International Cartel Enforcement: Some Modest Suggestions. Competition Policy International, September.

Marquez, 1992. Life Expectancy of International Cartels: An Empirical Analysis. Board of Governors of the Federal Reserve System, Discussion Paper.

Mena-Labarthe, Carlos, 2012. Mexican Experience in Screens for Bid-Rigging. Competition Policy International Antitrust Chronicle, March.

Mollenkamp, Carrick and Norman, Lawrence, 2008. British bankers group steps up review of widely used Libor, Wall Street Journal, April 17, C7.

Mollenkamp, Carrick and Whitehouse, Mark, 2008. Study casts doubt on key rate. Wall Street Journal, May 29, A1.

Ragazzo, Carlos, 2012. Screens in the Gas Retail Market: The Brazilian Experience, Competition Policy International Antitrust Chronicle, March.

Snider, Connan and Youle, Thomas, 2010. Does the Libor Reflect Banks' Borrowing Costs? Working Paper.

Snider, Connan and Youle, Thomas, 2009. Diagnosing the Libor: Strategic Manipulation Member Portfolio Positions. Working Paper.

Suslow, Valerie, 2005. Cartel Duration: Empirical Evidence from Inter-war International Cartels. Industrial and Corporate Change, 14(5), pp. 705-744.

The Montreal Gazette, 2012. Montreal construction bids are paved with questions," published September 15, available at http://www.montrealgazette.com/news/Montreal+construction+bids+paved+with+questions/7248408/story.html.

Vaugham, Liam and Finch, Gavin, 2013. Currency Spikes at 4. P.M. London Provide Rigging Clues, Bloomberg, published August 27, 2013, available at http://www.bloomberg.com/news/2013-08-27/currency-spikes-at-4-p-m-in-london-provide-rigging-clues.html.