Enhancing Financial Benchmarks
Comments on the OICU-IOSCO Consultation Report on Financial Benchmarks

10 February 2013

Rosa M. Abrantes-Metz and David S. Evans*

*Abrantes-Metz is Adjunct Associate Professor at the Stern School of Business, New York University and a Principal of Global Economics Group; Evans is Executive Director of the Jevons Institute for Competition Law and Economics and Visiting Professor at the University College London, Lecturer at the University of Chicago Law School, and Chairman, Global Economics Group. The views expressed are those of the authors and should not be attributed to affiliated institutions or their clients.
I. Summary

1. In this Comment we advance a number of guiding principles on matters of benchmark formation, administrative governance, screening, transparency and regulation, in response to the OICU-IOSCO Consultation Report on Financial Benchmarks dated January 2013 (“Report”). We are looking for practical solutions which discourage, to the extent possible, attempts at manipulation or fraud while providing the most useful information to the market. Our basic positions can be summarized as follows.

2. First, we believe that whenever possible benchmark indices should be based on actual transactions as opposed to quotes. This offers a number of advantages, not the least of which being its robustness (though not immunity) to fraud and manipulation. Where a transactions basis is not viable, the second-best alternative is to use committed quotes. The final option, uncommitted quotes of the type used for LIBOR, should not be adopted.

3. Second, we believe that some care should be taken in the manner of calculating the benchmark given the input data (either transactions or committed quotes). While averages are a common and certainly reasonable choice, there are alternatives – taking medians or modes for example – which may be more robust to manipulation.

4. Third, it must be understood that following the LIBOR scandal and associated investigations and liabilities, voluntary participation in a quote-based benchmark is unlikely to be attractive. Compulsory participation may be required, as tied to market participation.

5. Fourth, calls for “full transparency” – either of the calculating methodology or of the input data themselves – may be misplaced. Transparency facilitates cheating, and while some disclosures will certainly be necessary to inspire market confidence in a benchmark, full disclosures should be limited to the maximum extent possible.

6. Fifth, robust governance and supervision should be adopted to minimize the possibility of fraud and manipulation. The Administrator of the benchmark should be demonstrably disinterested in the benchmark itself. This may preclude public agencies from acting as Administrator, as well as trade associations, among others. Also, it must be recognized and accepted that submitters will almost always have material interests in the final benchmark result. Those interests must either be offset, as tends to happen naturally with transactions data, or proper incentives must be put in place to dissuade fraud, as can happen with an appropriate commitment mechanism for quotes.

7. Sixth, there is little to no role for regulatory direction of the content of the benchmark. The benchmark calculation methodology should be the proprietary interest of the private Administrator with an objective of reliability, continuity and accuracy. Regulation on how inputs must be formed and how the benchmark must be calculated risks introducing a wedge between the benchmark and the market it is meant to service. And it shouldn’t be strictly necessary if care is taken to base the index on proper inputs and if robust governance controls are adopted.
8. Finally, despite all the best efforts mentioned above, any benchmark should be routinely screened for attempts at manipulation. There will always be parties with very deep, material interests in the outcome of a benchmark index. No type of input data is immune from manipulation. No calculation methodology can be totally robust against fraud. No governance structure is perfect. Whatever structures are adopted, it is critical that ongoing reviews – so-called screens – take place to enhance the integrity of the benchmark.

9. We make a few brief remarks on our qualifications for presenting these recommendations and refer IOSCO to our attached curriculum vitae.

   a. Professor Rosa Abrantes-Metz is the co-author of a paper which identified, through econometric screening methods, possible problems with the LIBOR in 2008. Her paper addressed not only the possibility of manipulation but also collusion among the contributing banks.\(^1\) The U.K. House of Commons discussed Professor Abrantes-Metz’s various papers on LIBOR during its preliminary findings on July 3 2012, and in the subsequent testimonies of Mr. Bob Diamond and Mr. Paul Tucker.\(^2\) The U.K. House of Commons Treasury Committee has also cited her work in its findings in August and December of 2012, and so have other governmental investigators.\(^3\) Professor Abrantes-Metz specializes on conspiracies and manipulations and on the development of empirical screens to detect cheating and defend against such allegations. Professor Abrantes-Metz has a Ph.D. in Economics from the University of Chicago.

   b. Professor David S. Evans has written widely on the financial services industry and on its regulation. He was an adviser to the U.S. House Financial Services Committee during 2009 and has testified before the U.S. Congress on financial services matters on several occasions. He has also written widely on competition policy and has testified before the European General Court and many U.S. Federal Courts. He is the Executive Director of the Jevons Institute for Competition Law and Economics at University College London where he also serves as a


II. Methodology

A. Inputs to the Benchmark: Transaction-based benchmarks are most preferred; committed quotes are second-best. Uncommitted quotes should not be used.

10. As discussed in the Report, there are essentially three different types of inputs that are collected which contribute to the various benchmarks: (i) data on actual transactions; (ii) committed quotes which compromise the submitter to the terms of possible transactions; and (iii) uncommitted quotes which do not compromise the submitter.

11. As a general principle transaction-based benchmarks are preferable for several reasons. First and foremost, they represent prices at which agents actually traded. Second, while not entirely immune, they are far less susceptible to collusion and manipulation than are quote-based benchmarks, since every transaction necessarily involves a willing buyer and seller with opposite price interests. Under normal circumstances, they would not both want an artificially inflated or deflated price.

12. But transaction-based benchmarks also present significant challenges. First, in particular markets transactions may not take place every day; indeed they may not take place for days in a row, even in normal economic times. In markets such as the interbank lending market there may be many days without any transactions during financially stressed times. Obviously the same benchmarks will not be defined when transactions do not occur, and that will tend to be precisely at such times when the market will most need reliable benchmarks. To maintain some continuity the benchmark may have to switch to other inputs when transactions are not occurring, but this may cause an undesirable and spurious volatility: the benchmark may change from one day to the next only because transactions either occur or don’t.

13. Second, even when there are transactions, their number may be very limited, either because the market is in general thinly traded or because the time period considered for the benchmark computation is short. If few transactions determine the benchmark, it may not be very representative of the overall market. It may also be more susceptible to manipulation, either through collusion among

---

4 As a statement of interest, Professor Abrantes-Metz has been retained by various plaintiffs that have filed or are considering filing lawsuits against the banks that participated in the setting of LIBOR and other comparable benchmarks; Professor Evans has worked for numerous clients in the financial services industry, including banks, but is not currently working for any party on issues related directly or indirectly to LIBOR or other comparable benchmarks. Abrantes-Metz is a Principal of Global Economics Group, LLC and Evans is Chairman of Global Economics Group.

5 Or they will have to be replaced by values generated through various techniques, for example, interpolation.

6 Some indexes are based on trades occurring in only the last few minutes of trading, and in some markets there may only be a few trades in such a short time period, or no trades at all.
several market players or even, possibly, through the efforts of just one.

14. If a transaction-based approach is not feasible for a particular market, it is worth considering whether that benchmark is truly essential, or whether a substitute benchmark may be developed if it is not already available. If there is no substitute available and if the benchmark is to be used for a significant notional amount, then the second-best option, a committed quote system, should be implemented.

15. A committed quote system is described in Abrantes-Metz & Evans ((2012a) and (2012b)) for LIBOR.7 Under this system the participants agree to transact at (or “near”) their quotes if they transact at all. It is important to note that the participants are not obliged to transact, they only agree they will not transact “far” from their quotes.8 Such a system is superior to the regulated option proposed by the Wheatley Review for LIBOR in which outsiders will dictate the rules by which each bank’s (uncommitted) LIBOR quote should be formed; this will likely produce the least risky quote rather than the most accurate quote. The committed quote system, on the other hand, will “increase [the] incentive to provide accurate inputs” as noted by the Report.9

16. The committed quote system will work only if participants know that their quotes will be verifiable against actual transactions, whenever these do occur. Hence it is critical that an institution such as a clearing house or an exchange exist in which all transactions are registered and compared against the submitted quotes. If this capacity to verify quotes is not available, and if screening methods to detect cheating are not employed, then any system not based on actual transactions will be severely undermined: without fear of detection the incentive to cheat will increase and the reliability of the index will be compromised.

17. The third alternative is an uncommitted quote system, such as the current LIBOR approach. This is by far the least attractive alternative and is the most likely to be manipulated. We now know of course that LIBOR was, in fact, easily and repeatedly manipulated (or attempted to), so this is no longer a theoretical argument. Such an uncommitted quote system could only work with regular screening for possible cheating, for instance by comparing quotes against actual transactions every time transactions take place along with an independent and private benchmark Administrator, or instead with a heavy reliance on regulation to decide how quotes should be formed and supervisory administration as proposed for LIBOR in the Wheatley Review. But in the first instance, if regular

---


8 Criticisms of the committed quote system related to unwanted expansion of balance sheets by contributing institutions, or the fact that they may not all want to transact at all currencies and tenors all days, would only apply if submitters were always required to transact at the committed quote every time they submitted a quote. Instead, our proposal is to require submitters to transact at or around the committed quote, if and when they are to transact. Such a system would be immune to the two criticisms above.

9 Report, second paragraph, pp. 18.
transactions are available for screening and if independent private Administrators are also available, then why not move to a committed quotes benchmark (given that there aren’t enough transactions on every day to compute LIBOR simply based on actual prices)? In the second instance, regulatory interference can drive a wedge between the benchmark and the market it is supposed to represent, thus reducing its value. In general, an uncommitted quote system offers no advantages over, and several disadvantages relative to a committed quotes system, and of course, to a transaction based system. Uncommitted quotes should be discouraged.

18. Moreover, because of the likelihood that heavily regulated quote-based benchmarks will be uninformative, regulators should be careful not to adopt any policies that would deter market participants from relying on alternative benchmarks based on transactions or committed quotes, neither should they prevent competition among benchmark providers to develop more robust alternatives.

B. Calculation of the Benchmark: Some alternatives, such as medians or modes, may minimize the influence of outliers.

19. There is a wide range of methods to calculate benchmarks. Some use simple averages, perhaps trimming out the highest and the lowest values (“winsorizing”) to de-emphasize outliers. Averages are simple and typically smooth, and therefore provide greater stability to the index. On the other hand, any submitter, by itself, may influence or even manipulate the final result even when it is trimmed-out of the calculation.\textsuperscript{10}

20. Another common method is to take a transaction volume-weighted average over actual prices. If a market participant wants to manipulate the benchmark it must place a very large trade at an artificial price which, ordinarily, would be a risky and potentially costly way to influence a benchmark. Examples of these benchmarks are those for NYMEX commodities futures contracts. It is worth noting that these are based on transactions taking place only in the last few minutes of floor trading. Depending on the commodity, few such transactions may exist, raising the possibility of manipulation. Influential trades need not be absolutely large, but only relatively large next to the other trades in the short window of time considered for the benchmark inputs.

21. Other benchmarks are calculated as averages over a random selection of prices or quotes on a given day. This has the advantage of reducing the likelihood that cheating will succeed, but on the other hand potentially induces greater volatility in the index and may be less representative of the market as a whole.

22. There are alternatives to averaging. Other options include calculating the index as the median, mode or inside spread of all quotes. Everything else equal, these options are preferable as they not only reflect a more representative submission than an average or a randomly chosen number, but are also more robust to potential attempts to manipulate the benchmark.

\textsuperscript{10} See Abrantes-Metz and Evans (2012a), paragraph g, pp. 6-8.
C. Participation: With few benefits against real risks and costs, participation may be mandatory.

23. When the benchmark is not based on actual transactions but instead relies on quotes, it requires participation from enough market volunteers. That is particularly worrying after the launch of the LIBOR investigations and the current discussions on benchmark reform: why would submitters want to incur the associated costs (regulatory and administrative) and face the risk of future investigations for no apparent benefit whatsoever?

24. If market participants cease providing quotes there are two potential problems. First, the benchmark may end up being based on a small number of submitters, making it less representative and more susceptible to collusion among submitters and manipulation. Second, it could result in the cessation of the benchmark altogether if all submitters drop out.

25. If it is determined that a benchmark based on actual transaction data is not viable, and it is necessary to use a committed or uncommitted quote systems, then it may be required for regulators to demand participants in a market to submit quotes and participate in the benchmarking process. This requirement could be limited to participants over a certain size, so long as there remain enough participants to limit sufficiently the likelihood of manipulation and collusion.

III. Transparency

26. There are two types of transparency in the context to benchmarks: transparency of the methodology and transparency of the inputs.

27. The optimal level of transparency will critically depend on the features of the benchmark itself, namely whether it is based on actual transactions or quotes, whether there is a large or small number of submitters, and whether an insufficiency of inputs is a common occurrence.

28. LIBOR provides the best example of what excessive transparency can do: the LIBOR formula is fully disclosed and so are its inputs in virtually real time. Such complete transparency makes collusion and manipulation both more attractive and easier to achieve.\textsuperscript{11} But if the market is to rely on a benchmark that is not fully transparent, a governance system must be in place which provides the market with the necessary confidence in the benchmark itself (discussed in Section IV).

A. Transparency of the Methodology: Complete transparency is neither required nor desired.

29. A fully transparent methodology would allow market participants to replicate the calculation of the benchmark given the input data. Many market participants (and academics) appreciate this property. Some may find it difficult to benchmark to an index the nature of which they do not fully understand.

\textsuperscript{11} The LIBOR setting has other problems as well. It is set by a fairly small number of banks which have direct interests in the benchmark, has a trade association as the benchmark Administrator which presents a clear conflict of interest, and suffered from a lack of monitoring and governance.
30. The downside of a fully transparent methodology is that attempts to collude or manipulate the benchmark are more likely to be successful not only because cheaters know the exact rule by which the benchmark is calculated, but also because they know there will be a limited role for discretion on the part of the benchmark Administrator to change the rule setting on any particular day.

31. As a general principle, the more susceptible the input data are to manipulation, the less transparent the methodology should be. If the input data are actual transactions, which are intrinsically difficult to manipulate, there is more room to disclose the benchmark methodology in full, because knowledge of the methodology is of smaller use to a would-be cheater (though manipulations do still occur). On the other hand, if the inputs are uncommitted quotes which can be altered almost at will, knowing the exact methodology of the benchmark would be tantamount to knowing exactly how to manipulate its level. In that case, keeping as much of the methodology in secret is almost the only defense against its being abused.

32. If, as we will argue below, benchmarks are to be run by private and independent Administrators, these institutions will own proprietary methodologies to compute these benchmarks. It therefore seems reasonable to insist that enough information be made publicly available to achieve the necessary market confidence in the benchmark, but without releasing the exact formula. This serves not only to protect the Administrators’ proprietary interests but also to complicate efforts at manipulation.

33. Markets like rules, they like to know what to expect under what conditions. Less than complete methodological transparency does not necessarily mean that markets will not have enough knowledge to trust the benchmark and to understand the basis of its computation. For example, it may be disclosed that the benchmark is based on an average of prices corresponding to transactions taking place on the last few minutes of trade, but not exactly disclose how such average is taken (i.e., whether it is weighted by volume or not, and whether all or only some of the prices are taken into account), or how many minutes are “the last few.”

34. When Administrator discretion is allowed (or required) to calculate the benchmark, it is critical that the conditions under which such discretion is applied be clearly disclosed to the public. Market participants should be able to understand what types of data the Administrator will look at when the usual inputs are missing, and how such alternative inputs are likely to be combined to generate the value for the benchmark for that day. That does not mean that every time that discretion is used by the Administrator that fact must be disclosed to the public. Rather, such discretion will need to be properly documented for the benefit of a regulatory or supervisory body.

35. Any changes to the methodology should also be disclosed to the public with an appropriate amount of lead time.

36. All of the above will only be reliable if the Administrator is independent and clear of conflicts with respect to the benchmark.
B. Transparency of the Inputs: Transaction inputs are generally already public; quote inputs should only be disclosed with a substantial delay, if at all.

37. As a general principle, the more a benchmark is based on quotes (committed or not), the less transparent its input data should be. When the benchmark is transaction-based, presumably such prices are collected on an exchange or clearing house, so these may properly be obtained by the public. But when based on quotes, full disclosure may facilitate stable collusion agreements (either tacit or explicit), provide a means for the participants to monitor each other, and punish deviations from a collusive agreement. LIBOR is a perfect example of this.

38. Furthermore, complete input transparency may create an extra incentive to manipulate quotes not out of a desire to change the final benchmark, but to signal something to the market. Again LIBOR may be an example of this: a reason as to why the banks may have distorted their quotes is that they were attempting to overstate their financial strength to the market.

39. Either individual quotes are to remain permanently sealed or at least be released only months after they are submitted, the latter recommended by Abrantes-Metz and Evans (2012a) and the Wheatley Review Second Report.

40. Again, the market will have confidence in the benchmark in the face of limited transparency only if the Administrator is independent and clear of conflicts.

IV. Governance & Regulation

A. Governance: Administrators should be free of conflicts of interest; submitters must be properly incentivized against fraud.

41. As discussed in the Report, a number of different entities may act as the Administrator of a benchmark. Those could be (i) trade associations such as the BBA for LIBOR and the EBF for Euribor; (ii) public entities; (iii) regulated exchanges or trading platforms; (iv) price reporting agencies, such as those responsible for calculating international commodity prices; (v) regulated firms such as banks or asset managers; and (vi) other commercial organizations such as data providers (Thompson Reuters, Markit, Bloomberg).

42. The involvement of trade associations in administering the benchmark should be discouraged. Not only do they present a clear conflict of interest, but they have also been found to provide the setting for potentially illegal exchanges of information among members.

43. Banks or asset managers may also have conflicts of interest with respect to benchmarks; their involvement should likewise be discouraged.

44. Instead, we should opt for private institutions which are clearly independent of the submitters and which have no direct financial interest at all in the benchmark, either in its level or volatility (or any of its other moments). Entities such as (iii),

---

12 Report, pp. 23.
13 Trade associations clearly have useful and legitimate purposes. But it is also true that they are often implicated in uncovered collusive agreements.
(iv) or (vi) identified in paragraph 41 above would appear to meet this standard. It is absolutely critical for benchmark reliability and market confidence that the Administrator be clearly and demonstrably independent. It is also advisable that it be a private company, since public interests may arise during periods of great financial distress for example and seek to influence critical benchmarks.

45. These requirements become increasingly important when there is less than full transparency, either of the benchmark methodology or its inputs. If an Administrator is permitted to exercise discretion in any form, market confidence can only be achieved if that Administrator is unambiguously disinterested in the result. This has not always been the case in practice.

46. It is almost inevitable that submitters will have a direct interest in the benchmark. They will very likely have transactions tied to the benchmark as they are presumably active market participants, and would potentially benefit from a higher or lower value for the benchmark on a given date. Furthermore, their interests may be difficult to identify and may change over time.

47. The benchmark system must acknowledge the perhaps substantial interests of the submitters. The LIBOR architecture for example failed to account for the possibility, let alone likelihood, that those providing quotes might have material interests in the final result. While submitters may always be presumed to have an interest, those interests can be managed or sometimes offset. Using actual transactions will tend to offset long interests against short interests. The interests do not disappear, but it is difficult for one to dominate at the expense of the other. A committed quote system can also discourage grossly fraudulent submissions, since such quotes may preclude the submitter from participating in the market.

B. Regulation: Regulations should be restricted to compliance, not content.

48. In principle, regulation may apply at two levels: regulation of the submissions, or regulation of the Administrative functions. Regulatory influence – as distinct from supervision – of submissions is not appropriate when the benchmark is based on actual transactions. Actual prices should be what participants in the market agree to; there should be no role for regulatory direction of the content of these submissions.14

49. Some would argue that when the benchmark is based on quotes, either committed or uncommitted, regulation of these submissions may be warranted. This is the recommendation of the Wheatley Review for LIBOR, for example. In our view, input submissions should not be regulated, in the sense of regulatory direction of content. Instead, incentives need to be structured to enhance the likelihood of truthful submissions. Examples of such incentive structures are committed quotes coupled with verification and screening.

50. It is important to stress that when we say “regulation does not apply,” we do not mean that submitters or Administrators should be literally unsupervised. Both submitters and Administrators may be required to develop and comply with certain rules and procedures. Compliance would need to be verified, and that verification may be conducted by an independent auditor or, perhaps, a regulatory

---

14 There may be a role for regulatory supervision, that the appropriate prices are collected as intended.
body. Instead, we mean that regulators should not have a say in the content or methods of the Administrator or submitters. They should not, for example, determine the methodology for constructing the benchmark.

51. As long as the Administrator is independent, private and competent it should have the latitude to set the benchmark as it sees fit. It should have a detailed code of conduct and auditable procedures. It should meet regularly with regulators, auditors or supervisors to discuss any issues related to the benchmark, including the existence of unusual patterns in the benchmark and in the submissions. But the intellectual content of the benchmark should be its own, without interference or direction from outside parties.

V. Benchmark Screening

52. A critical component to ensuring the truthfulness and reliability of any benchmark is the ability to detect not only errors, but also any attempt at manipulation. This is true regardless of what mitigating steps may be taken. This is even more important where submissions are based on quotes, either committed or uncommitted, rather than on transaction data (though instances do occur of actual price manipulation with the ultimate objective of manipulating the benchmark).

53. The Report recognizes that Administrators need to be able to check the validity of the submissions by accessing relevant data. It also recognizes that an adequate Audit Trail needs to be kept, and admits that at least some of the benchmarks reviewed in preparation of the Report monitor submissions for non-deliberate errors but do not indicate whether any efforts are undertaken to detect other types of irregularities, including attempts to manipulate the index.  

54. It is necessary to delineate and implement clear and consistent policies for screening submissions against cheating. Otherwise the fear of getting caught is significantly reduced, and the reliability and accuracy of the benchmark subsequently impaired. The benchmark will then either go unused, or, if used, may lead to potentially serious market distortions.

55. Third parties should monitor the behavior of the benchmark on a regular basis through the use of empirical screens in order to detect attempted (or successful) manipulation.

56. LIBOR provides the perfect example of what is likely to happen in the absence of regular screening. Outside researchers and reporters identified patterns in LIBOR as unusual and unexpected at least as early as 2008. The illegal behavior

---

15 Report, pp. 16-17.
16 In some cases Administrators already use third parties to collect and/or calculate the benchmark. The screening and monitoring services should be provided by still other parties with no association or conflict with the benchmark.
currently being uncovered in LIBOR could have been detected earlier by the responsible authorities had empirical screens been routinely employed. It seems reasonable for the market to expect that, given the hundreds of trillions of dollars of transactions and contracts based on LIBOR, someone somewhere would have been responsible for monitoring the submissions. Yet this does not seem to have been the case.

57. Had screening been regularly employed to monitor LIBOR, it may also have deterred illegal behavior from occurring in the first place. Even if it didn’t deter every bank, it may have deterred some, meaning that the degree of final manipulation likely would have been less.

58. Screens have proven capable of detecting illegal behavior, and they can also deter such behavior from getting started. Though screens may not be able to detect all illegal behavior, they are more likely to detect manipulation the more impactful and consequential it is. This is how screens worked for LIBOR, and that is how they can work for other benchmarks.