Is LIBOR Still Being Manipulated?: Identifying colluders with methods of detecting LIBOR tampering

Daniel S. Levy[1], Audrius Girnius, Karthik Padmanabhan, Timothy Tardiff, Tim Weithers, and Megan Yeretsian

Advanced Analytical Consulting Group, Inc.

Abstract

We analyze the one-month U.S. Dollar London Interbank Offer Rate (LIBOR) between January 1987 and February 2015 to determine whether there are signs of manipulation based on a previously published test that has appeared in a peer reviewed economics journal and been cited in the business press and legal filings of major financial disputes.[2] Our analysis starts with the period from February 1, 2014 to February 28, 2015, after the Intercontinental Exchange (ICE) took over the publication of LIBOR from the British Bankers' Association (BBA) and after significant time for reforms to be implemented related to the widely publicized LIBOR fixing messages between bankers, as detailed to the British government in the Wheatley Report.[3] We find that this previously published test still indicates the presence of LIBOR manipulation from February 2014 and into 2015. We then perform the previously published test for tracking the integrity of important market indicators, such as LIBOR, from 1987 through February 2015, and find that this test would nearly always trigger a finding of suspicious behavior, either indicating that LIBOR has been consistently manipulated since 1987 to 2015 or that the proposed test has no power to distinguish periods of LIBOR manipulation from periods of non-manipulation. We further discuss the nature of scientific evidence and how the use of non-scientific methods, if taken seriously, can lead to the misallocation of corporate, regulatory, enforcement, and potentially, in the case of such a widely use financial measure as LIBOR, global economic resources.

---

[1] Authors are affiliated with Advanced Analytical Consulting Group, Inc., 211 Congress Street, Boston Massachusetts, 02110. Corresponding author, Email at DanLevy@AACG.com. The authors thank Benjamin D. Williams and the participants of the U.S. Securities and Exchange Commission's Office of Litigation Economics Seminar for their helpful comments. The errors remain our own.

[2] The previously published test can be found in Rosa M. Abrantes-Metz, Sofia B. Villas-Boas & George Judge, "Tracking the Libor rate," *Applied Economics Letters*, 18 (2011):10, 893-899), also cited in legal cases as Rosa M. Abrantes-Metz and Sofia B. Villas-Boas, "Tracking the Libor rate," July 2010.

[3] Wheatley, Martin. "The Wheatley Review of LIBOR: Final Report." HM Treasury, UK, September 2012. http://cdn.hm-treasury.gov.uk/wheatley_review_libor_finalreport_280912.pdf.

December 27, 2016

I. Introduction

The London inter-bank offer rate (LIBOR) has been used around the world as a reference point and benchmark for the cost of funding faced by major banks.[4] It is also used for the settlement of the majority of interest rate derivative contracts and, therefore, has an impact on individuals and companies across every industry. LIBOR is relied upon due to its reputation as an effective measure of the cost of interbank borrowing. If LIBOR is not a valid reflection of the cost of borrowing, it sends the wrong, or a noisy, signal, or information, to the markets, which may cause market participants to make the wrong decisions. A reliable, publicly-available measure of the cost of borrowing at major banks is important for the efficient, and perhaps equitable, performance of markets and the economy.

Since 2011, various individuals and banks have been accused of manipulating LIBOR.[5] As the scale and scope of the LIBOR manipulation has been, and continues to be, revealed, the natural question arises of how market manipulations and other similar violations could be prevented. Various proposals have and will be made. First, additional regulations and proposed changes to the process of calculating LIBOR based on actual transacted inter-bank rates could limit the ability of banks to manipulate LIBOR.[6] Second, penalties and punishments, if of appropriate magnitude, may substantially inhibit undesirable reporting behavior.[7] Third, more precise

[4] LIBOR was published by the British Bankers Association until January 2014 when Intercontinental Exchange took oversight and publication. Much like the BBA, Intercontinental Exchange states the following: “ICE LIBOR is the primary benchmark for short term interest rates globally. It is written into standard derivative and loan documentation, such as the ISDA terms, and is used for an increasing range of retail products such as mortgages and student loans.
It is also used as a barometer to measure the health of the banking system and as a gauge of market expectation for future central bank interest rates. It is the basis for settlement of interest rate contracts on many of the world's major futures and options exchanges.” See “ICE LIBOR.” Accessed April 15, 2015.
https://www.theice.com/iba/libor#calculating.

[5] See for example, Charles Schwab Bank, N.A. v Bank of America Corporation *et al*, August 23, 2011; The City of Philadelphia v Bank of America *et al*, July 27, 2013. The Mayor and City Counsel of Baltimore *et al* v Credit Suisse Group AG *et al*, April 30, 2012.
[6] See for example Wheatley, *op. cit.*
[7] For an early influential work see Gary S. Becker, “Crime and Punishment: An Economic Approach” *Journal of Political Economy* Vol. 76, No. 2 (Mar. - Apr., 1968): 169-217.

detection of violations could help catch violations sooner and, in doing so, would reduce the incentive of potential perpetrators to attempt a rate manipulation.[8] Each of these three categories of monitoring and enforcing has the potential to improve the validity of LIBOR as a true measure of the interest rates that major banks face in the market. Of course, if the criterion used to identify potential manipulation is itself flawed, to the extent that it either identifies manipulation when it does not exist or misses manipulation when it does exist, then allocation, and disruptive economic, implications follow.[9] A test that has a high percentage of false positives could send regulators, enforcers and the court system down a path of useless action or, perhaps more damaging, provide false evidence to court proceedings, potentially leading to false findings of guilt. Alternatively, a test with a high percentage of false negatives would miss periods of LIBOR manipulation. If designed incorrectly, each of the three components of deterrence, including regulation, punishment, and detection, could have limited benefit, or even lead to further inefficiencies, misinformation and misallocation of resources.

This paper focuses on the third category listed above, detection. It analyzes, and provides evidence about a metric that has been published in peer reviewed economic literature, discussed in the popular press, cited in legal complaints of major cases, and proposed as a "real time and

---

[8] "The Scam Busters" in *The Economist*, "Free Exchange" column, December 15, 2012 noting "In a 2011 paper Rosa Abrantes-Metz of New York University's Stern School of Business and Sofia Villas-Boas and George Judge of the University of California, Berkeley, examined LIBOR data over rolling 6-month windows, and found that LIBOR was far likelier than another benchmark interest rate to depart from Benford patterns. A quick Benford test would have pointed to LIBOR anomalies long before one of the colluding banks chose to own up." http://www.economist.com/news/finance-and-economics/21568364-how-antitrust-economists-are-getting-better-spotting-cartels-scam-busters. Similarly, Rosa Abrantes-Metz, "How to Use Statistics to Seek Out Criminals" in *Bloomberg*, February 26, 2012, "The Libor studies were classic screens, in that they tested for divergence from normal statistical behavior or from markets thought to be functioning properly. To understand how screens work, consider one popular statistical tool: Benford's law. The law states that the digits in certain types of data from naturally occurring events follow a consistent pattern. The number 1 is by far the most frequent first digit, followed by 2, 3 and so on all the way to 9. The second significant digit is more evenly distributed, and so is the third digit. Such patterns have been observed in financial data such as stock prices, corporate revenue and interest rates. Libor submissions followed Benford's law closely for about 20 years, but began to diverge sharply in the mid-2000s." https://www.bloomberg.com/view/articles/2013-02-26/how-to-use-statistics-to-seek-out-criminals.

[9] A similar sentiment is expressed by Abrantes-Metz et al (2011) when noting the potential damage done by LIBORs that are not valid (Abrantes-Metz et al., *op. cit.*, p.893). Here we note that misallocation of resources can occur when a non-valid test is used to determine the validity of LIBOR rates.

not ex post" objective way to track the integrity of important economic indicators like LIBOR.[10] We first use this published method to examine LIBOR during the period starting February 1, 2014, after the date that the Intercontinental Exchange (ICE) took over the calculation of LIBOR from the British Bankers' Association (BBA) and after significant opportunity for reforms to be implemented, related to the widely publicized alleged LIBOR price-fixing as detailed to the British government in the Wheatley Report.[11] We find that this previously published test still indicates LIBOR manipulation from 2014 and into 2015. We examine this published method and extend the 6-month rolling test periods back to January 1987. We find that the previously published test produces a warning of LIBOR manipulation for 321 of the 333 6-month rolling periods tested. Contrary to Abrantes-Metz et al. (2011), we do not find that the proposed test supports the conclusion of no evidence of manipulation in 2005 until January 2006[12], the earliest periods tested in that work.

In the next section, II, we employ the test proposed in Abrantes-Metz et al. (2011) ("Abrantes-Metz"), as it purports, "to identify tampering and human influence on the market process".[13,14] The test is based on the comparison of the distribution of the second digit of LIBOR to that of the second digit of a distribution found in Benford (1938).[15] The shape of this distribution is discussed below and pictured in Figure 4. The basic idea for the test employed by Abrantes-Metz is that if LIBOR is tampered with, through mechanisms such as collusion, the digits of LIBOR will tend *not* to follow some "typical or natural" pattern. Abrantes-Metz states that "in many naturally occurring numerical data sets and in several financial data sets, the digits follow a

[10] *Ibid.*, p.893.
[11] Wheatley, *op. cit.*
[12] Abrantes-Metz et al., *op. cit.*, p.896.
[13] *Ibid.*, p.894.We note that all markets are the result of human influence, but that there can be cases were the form of human influence is a manipulation of that market which is illegal. It is these cases of inappropriate manipulation that Abrantes-Metz et al. are attempting to detect.
[14] Other tests, unrelated to Benford distributions, have been proposed to detect LIBOR manipulation. See for example Snider and Youle, SSRN-id2189015, December 2012. Rosa M. Abrantes-Metz, Michael Kraten, Albert D. Metz, Gim S. Seow, "Libor Manipulation?", *Journal of Law and Economics*, 36 (2012) 136–150, among others.
[15] For justification of comparing the second digit of LIBOR to the second digit of a Benford distribution see Abrantes-Metz et al., *op. cit.*, p. 894.

logarithmic weakly monotonic distribution."[16] This distribution of the first digit of a Benford distribution spans the range of digits from 1 to 9. In proposing to use this Benford distribution as a benchmark against which the observed LIBOR rate can be tested Abrantes-Metz note that LIBOR does *not* follow the Benford distribution. Therefore, as an alternative Abrantes-Metz tests the correspondence between the second digit of LIBOR (from 0 to 9) and the second digit of a Benford distribution (from 0 to 9). No further justification or citations about the validity of such a test were provided. Abrantes-Metz (2011) does not use the first two digits of the Benford distribution, but rather simply the second digit of the Benford distribution. Section III presents results of the same test for test periods running back to January 1987. The analysis in Section III (as well as Section II) is as much a test of the method proposed to detect LIBOR manipulations as is it a test for the manipulation itself. In Section IV, we discuss why our results differ from those found previously. In Section V, we discuss how a change in the duration of the test periods used by the previous test of LIBOR, based on the second digit of Benford, led those tests to identify a transition from a period of apparent non-manipulation to a period of apparent manipulation. Section VI evaluates Abrantes-Metz proposed method of detecting manipulation of LIBOR based on the tests performed there, Benford tests. In Section VII we apply the standard first digit Benford test to a set of 20 populations which Benford used in the development of the Benford distribution to shed light on whether Benford tests, in their most commonly used form, can generally be relied upon to detect "tampering and human influence."[17] In Section VIII we conclude and discuss the influence that non-valid, and unscientific, testing methods can have on efforts to detect tampering behavior, the impact on governmental regulators, regulatory actions, private legal actions, juries and the allocation of resources in the economy.

---

[16] *Ibid*., p.893.
[17] *Ibid*., p.894.

II. A Test of LIBOR Manipulation since the LIBOR Reforms

In February 2014, LIBOR calculations moved from the British Bankers' Association (BBA) to the Intercontinental Exchange (ICE). Further, the process of determining these rates has had significant time to implement changes to improve the validity of LIBOR. Have these changes allowed LIBOR to be reported without manipulation? To investigate this we employ a test that was used to detect manipulation in LIBOR during certain periods after the start of 2006. Following Abrantes-Metz, we construct 6-month rolling time periods and tested the distribution of the second digit of LIBOR against the distribution of the second digit of a Benford distribution. For February 1, 2014 through February 28, 2015, the daily posted 1-month USD LIBOR was grouped into 6-month rolling periods with start dates at the beginning of each month. For example, daily 1-month LIBORs for February 1, 2014 through July 31, 2014 are put into a 6-month period. The next 6-month rolling period started in March 1, 2014 and ran through end of August 2014. We continued the 6-month periods through September 1, 2014 – February 28, 2015. For the days in each of these 8 6-month rolling periods, we enumerate the number of times each value, 0 through 9, appears in the second digit of LIBOR. So, for example, if LIBOR was reported as 2.45 on a given day, we count that the value of 4 occurred as the second digit of LIBOR for that day. If the LIBOR was smaller than 1, for example 0.89 on a given day, we count that the value of 9 was the second digit of LIBOR of that day. The calculation is done this way to follow Benford (1938), which states that "If a decimal point of zero occurs before the first natural number it is ignored, for no attention is to be paid to magnitude other than that indicated by the first digit."[18] The number of times each digit, 0 through 9, is counted creates an empirical distribution. A chi-square[19] test is implemented to determine whether there is a statistically significant departure of the distribution of the second digit of LIBOR compared to the

---

[18] Benford, Frank. "The Law of Anomalous Numbers." Proceedings of the American Philosophical Society 78, no. 4 (1938): 551-72.

[19] Snedecor, George W. and Cochran, William G, *Statistical Methods*, 6th ed. (The Iowa State University Press, 1967), p.231. Mittelhammer, Ron C, Judge, George, G, and Miller, Douglas, J., *Econometric Foundations*, (Cambridge University Press, 2000), p.719; Abrantes-Metz et al., *op. cit.*, Tracking the Libor rate, fn. 3. We note that LIBOR rates exhibit significant serial correlation, adding additional considerations in the application of a chi-squared test.

distribution of the second digit of the Benford distribution.[20] Table 1 lists the chi-square values for each of the 8 6-month rolling time periods.

Table 1: Chi-square values for 6-month rolling test periods

Start Date	End Date	Chi-Squared Test Statistic
2/1/2014	7/31/2014	1068.59
3/1/2014	8/31/2014	1068.59
4/1/2014	9/30/2014	1077.90
5/1/2014	10/31/2014	1105.84
6/1/2014	11/30/2014	1184.81
7/1/2014	12/31/2014	929.42
8/1/2014	1/31/2015	635.20
9/1/2014	2/28/2015	421.51

Critical (99% or .01%) chi-square value for 10 categories = 21.67

We find that chi-squared statistics for each of the 8 6-month periods is greater than the critical value of 21.67. Therefore, the distribution of the second digit of LIBOR is statistically significantly different from that of the second digit of a Benford distribution in each of the eight periods tested. Further Abrantes-Metz mentions the magnitude of the chi-square statistics as a measure of the degree of divergence, with chi-squares for some periods over 800.[21] Some of the chi-squares in Table 1 are even larger, over 1000. Interpreting the finding in a manner consistent with Abrantes-Metz, the distribution of LIBOR, even after the opportunity for behavioral modification in light of regulatory actions, still does not follow what Abrantes-Metz had

[20] Figure 4 below lists the distribution of the second digit of a Benford Distribution. Also see Benford, *op. cit.*, p. 551-72.

[21] Abrantes-Metz et al., *op. cit.*, Tracking the Libor rate, p.897.

characterized as the "path that the LIBOR had followed for at least the prior 20 years" prior to the date of the LIBOR traders emails, and the divergence may be getting even worse.[22]

To further investigate why LIBOR is exhibiting this pattern that previous research has classified as an indication of tampering, we test, using data from 1987 through February 2015, whether the distribution of the second digit of LIBOR in 6-month rolling periods differed from that of the second digit of a Benford distribution. Previous research had found a statistically significant difference between the distribution of the second digit of LIBOR and that of a Benford distribution for certain periods between 2005 and 2008. For the overall time frame prior to that, from 1987 to 2005, Abrantes-Metz found *no* statistically significant evidence of tampering.

III. Tracking the Rate: Test of Departure from Benford Prior to LIBOR Reform

In this Section we perform the test employed by Abrantes-Metz to detect "tampering or human influence" in LIBOR, based on the LIBOR rates posted by Intercontinental Exchange (ICE) (before September 2012, published by Thompson Reuters on behalf of the British Bankers Association)[23] These are rates posted both prior to and after the reform of LIBOR. As above, the test is a chi-square test of the similarity of the second digits of LIBOR during 6-month rolling periods compared to the distribution of the second digits of a Benford distribution. The results of a chi-square test for the 6-month rolling periods with first months that start in January 1987 and the last period starting in September 2014 are presented in Figure 1. The blue data series represents the chi-square value for each 6-month period. The dashed, red line represents the critical value for a chi-square test based on 10 cells, corresponding to the ten possible values of

[22] *Ibid.*, p.897

[23] https://www.theice.com/iba/libor. We accessed LIBOR data through the Federal Reserve Economic Data site:

ICE Benchmark Administration Limited (IBA), 1-Month London Interbank Offered Rate (LIBOR), based on U.S. Dollar© [USD1MTD156N], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/USD1MTD156N, November 16, 2016.

ICE Benchmark Administration Limited (IBA), 3-Month London Interbank Offered Rate (LIBOR), based on U.S. Dollar© [USD3MTD156N], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/USD3MTD156N, November 16, 2016.

the second digits, 0-9, at a significance level of .01, which is 21.67.[24] The 6-month periods where the values of the chi-square test statistic (blue) are above the critical value of the chi-square (red) indicate a statistically significant difference between the distribution of the second digit of LIBOR and that of the second digit of a Benford distribution.

Figure 1: Test Statistic Measuring Equality of Benford and Actual Counts, 1-Month LIBOR

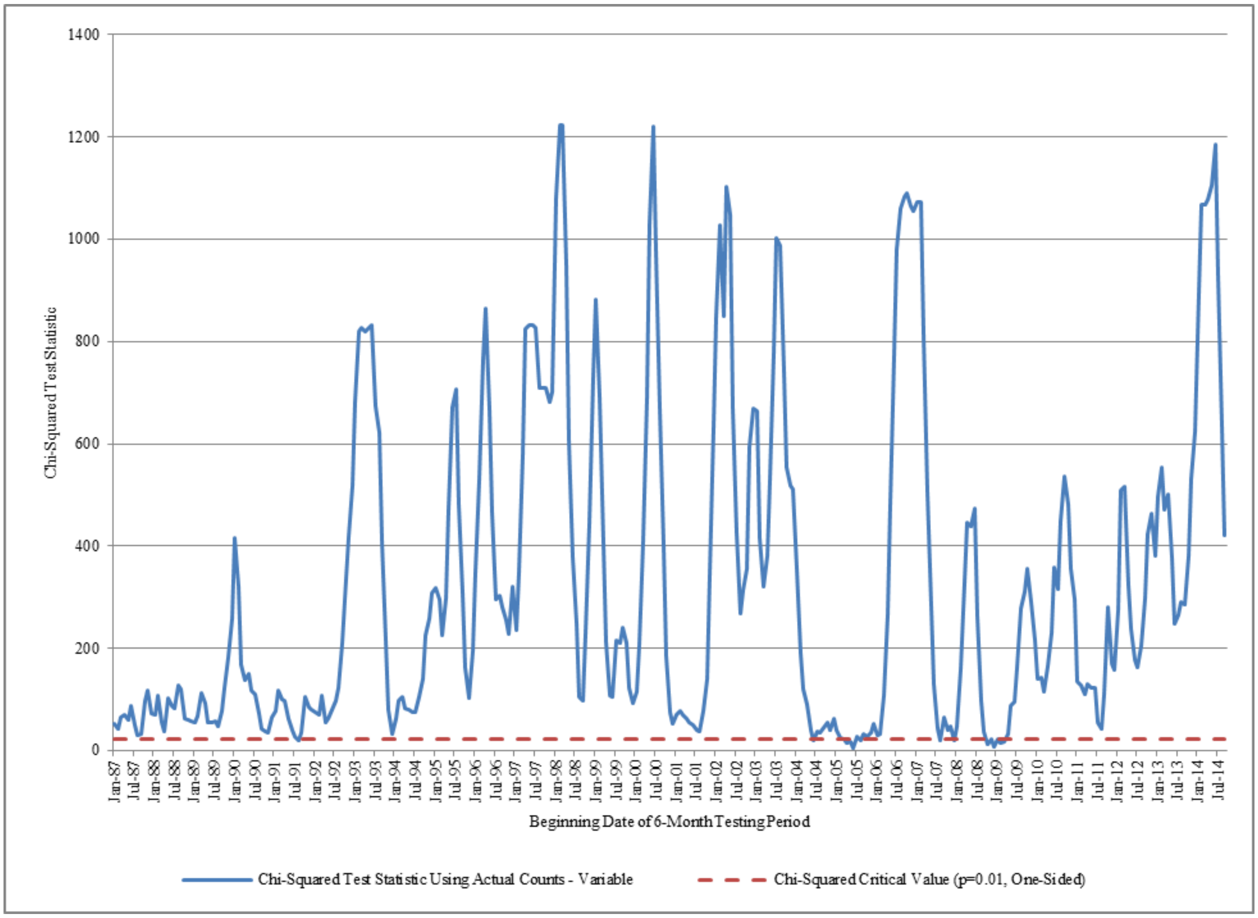


Figure 1 shows that based on a chi-square test on 6-month rolling periods the distribution of the second digit of LIBOR is statistically significantly different from the distribution of the second digit of a Benford distribution in the vast majority of periods. In fact, of the 333 6-month rolling periods between January 1987 and September 2014, 321 are statistically significant, meaning

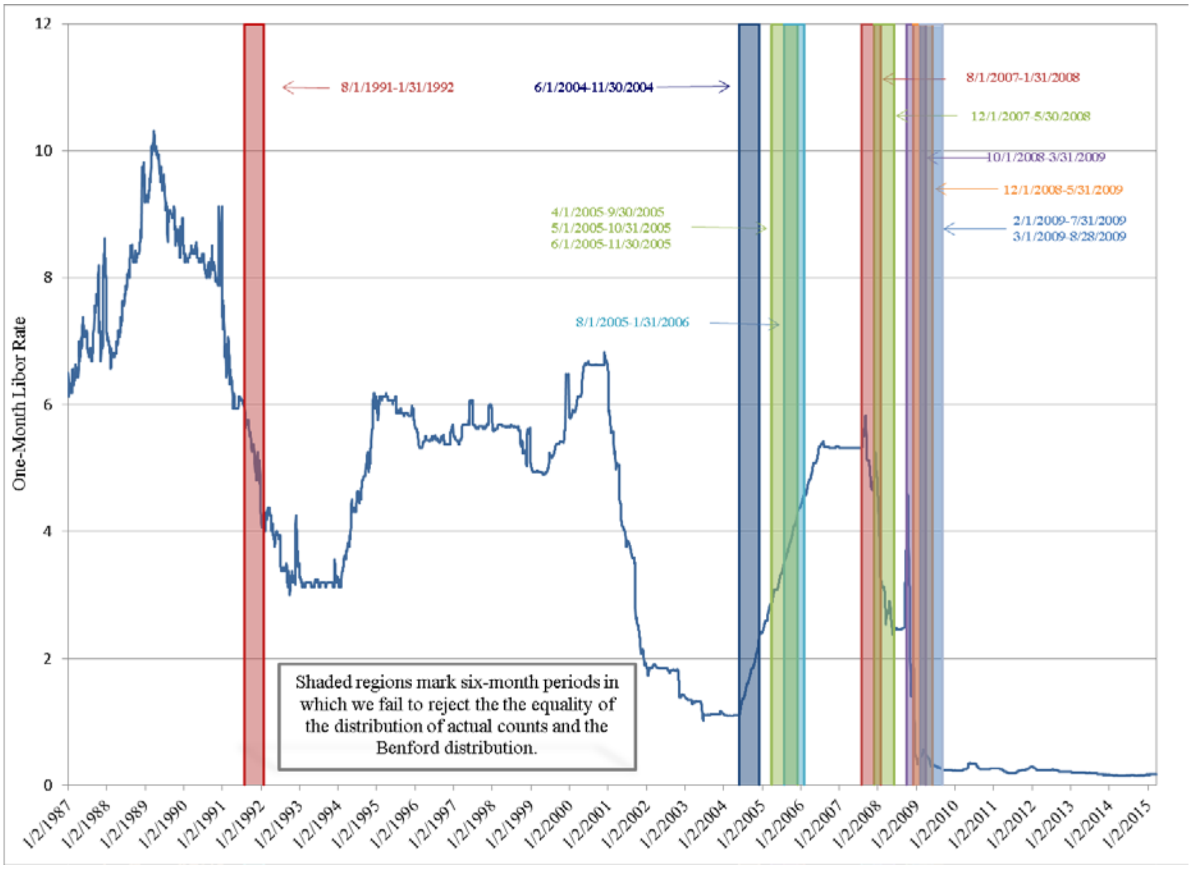
[24] Snedecor and Cochran, *op. cit.*, p.550-551.

that the second digit of LIBOR is statistically significantly different from that of a Benford distribution.[25] Further, to the extent that one uses the magnitude of the chi-square as a measure of the magnitude of the divergence, there have been high levels of divergence of the second digit of LIBOR from that of the Benford reference for decades, with the largest values in the late 1990s and early 2000. However, statistically significant difference span the entire period for which we analyzed data, since 1987.

Figure 2 highlights the 12 6-month periods where the second digit of LIBOR and the second digit of Benford do not statistically significantly differ. They are the 6-month periods with the following starting months: August 1991, June 2004, April 2005, May 2005, June 2005, August 2005, August 2007, December 2007, October 2008, December 2008, February 2009, and March 2009.

[25] We replicated this analysis for 3-month LIBOR rates and the results are broadly similar. For 3-month LIBOR rates there are 21, 3-month rolling periods in which we fail to reject the similarity of the Benford distribution and the empirical distribution of LIBOR second digits.

Figure 2: 6-Month Periods in which Actual Distribution of LIBOR Second-digit does not Differ from Benford Distribution



However, all of the other 6-month rolling periods exhibit a statistically significant difference between the distribution of the second digit of LIBOR and the second digit of Benford. Note that since these time periods used are 6-month rolling time periods, the periods overlap. For example, three of the 6-month time periods in which LIBOR does not exhibit a divergence from that of a Benford distribution, 4/1/2005-9/30/2005, 5/1/2005-10/31/2005, and 6/1/2005-11/30/2005, cover a combined 8 months.

If we were to believe a chi-square test on 6-month rolling periods of second digit of LIBOR against a second digit of a Benford distribution is a real time objective-predictor for "tampering and human influence,"[26] we would conclude that LIBOR had been manipulated nearly constantly

[26] Abrantes-Metz et al., *op. cit.*, Tracking the Libor rate, p.894.

from the start of 1987 through September 2014, calling for investigation. The test itself would potentially be used (misused) as evidence of a violation that some might bring before regulators, enforcers and the courts.

These results, exhibiting nearly constant statistically significant departure of the distribution of the second digit of LIBOR from that of the Benford reference, are strikingly different from those of Abrantes-Metz, which performed similar tests based on the same data through October 2008. That previous research found that the second digit of LIBOR initially, from 1987 to 2005, was not statistically significantly different from that of a Benford distribution, and that the statistically significant divergence developed after February 2006, "find[ing] that in two recent periods, Libor rates depart significantly from the expected Benford reference distribution."[27] Further, Abrantes-Metz concluded that this change in behavior after February 2006 from what "Libor had followed for at least the prior 20 years, raise[s] questions regarding the integrity and quality of its rate signals coming from individual banks and cry out for an answer."[28] In the next section, we answer that cry by investigating where our results differ from that previous analysis of manipulation of LIBOR and explain the reasons for the difference.

IV. Departure from Previous Findings

In this section, we describe where our results differ from the analysis by Abrantes-Metz of manipulation of LIBOR based on Benford distributions and investigate the reasons for the departure of our results from those of previous research. An understanding of these differences

[27] *Ibid*., p.893

[28] *Ibid*., p.897. For the period 1987-2005, Abrantes-Metz also finds that there is no statistical difference between the distribution of the second digit of the Federal Funds Rate and that of a Benford distribution. The value for the standard Pearson chi-squared statistic for that period 1987-2005 is 520.75, not 6.47, as listed in *Ibid*., p.895. As a result, we reject the equality of the empirical distribution of the Federal Funds Rate second digits and that of the Benford distribution, where Abrantes-Metz did not reject the hypothesis of equality.

illuminates not only why the results of the two studies differ, but also provides information about what a comparison of the second digit of LIBOR to the second digit of Benford measures.

Abrantes-Metz found that between 1987 and 2005 there was no statistically significant difference between the distribution of the second digit of LIBOR and that of a Benford distribution. That same research found that a difference arose sometime after February 2006 and continued for 18 6-month rolling periods.[29] Abrantes-Metz reports intermittent periods of statistically significant departure thereafter, through the end date of their last 6-month rolling period, in October 2008. Abrantes-Metz concludes that this pattern raises questions about the integrity of LIBOR.

There are several methodological differences between our analysis and that of Abrantes-Metz, but there are two major differences that lead us to a different conclusion about the departure of the distribution of LIBOR's second digit from that of Benford's. First, our analysis uses a standard Pearson's chi-square to test the equality of the Benford distribution of second digits and the actual distribution of second LIBOR digits.[30] The formula for this test can be found in many statistical texts, and, as applied to the test of LIBOR, is as follows:[31]

Eq 1:

$$\chi^2 = \sum_{i=0}^{9} \frac{(n_i - np_{iB})^2}{np_{iB}},$$

n_i = Observed count of second digits equal to *i*.
n = Number of days for which a LIBOR quote is available in the testing period.
p_{iB} = Percent of the population of second digit=*i* under the Benford distribution.

This is the same calculation listed in footnote 3 of Abrantes-Metz as the test statistic used throughout their paper, with slightly different notation. However, it is not the formula *implemented* by Abrantes-Metz in its analysis. By replicating the results found in Abrantes-Metz

---

[29] *Ibid*., p.895.
[30] Johnson, Richard A. and Gouri K. Bhattacharyya, *Statistics: Principles and Methods*, 5th ed. (John Wiley & Sons, 2006), 508.
[31] Snedecor and Cochran, *op. cit.*, p.231. Mittelhammer et al., *op. cit.*, p.719.

we have determined that the formula below, Eq. 2, nearly precisely replicates the test statistic presented in Abrantes-Metz for every month of their reported chi-square test.[32, 33, 34]

Eq 2: $$\chi^2 = \sum_{i=0}^{9} \frac{((100*n_i/n)-100*p_{iB})^2}{100*p_{iB}} = 100 \sum_{i=0}^{9} \frac{((n_i/n)-p_{iB})^2}{p_{iB}},$$

n_i = Observed count of second-digits equal to *i.*
n = Number of days for which a LIBOR quote is available in the testing period.
p_{iB} = Percent of second-digit=*i* in population under the Benford distribution.

Noting that $n_i = n * (n_i/n)$, Eq. 1 is converted into Eq. 2 by setting all *n* = 100 except for the *n* in (n_i/n), which is left as the actual number of days in the test period. The resulting Eq. 2 is equivalent to using the formula of a chi-square but replacing the frequencies required in a chi-square test with percentages, that is percents multiplied by 100. The resulting formula is not a chi-square test and none of the critical values of the chi-square test apply to it. Abrantes-Metz says that a goal of the paper was to construct a "scale-invariant" measure; Abrantes-Metz

[32]Abrantes-Metz did not reject tests of equality between the Benford distribution and the daily 1-month LIBOR rates over the period 1987 to 2005, but "[S]tarting in February 2006, and continuing for 18 six-month periods in Tables 2 and 3, the theoretical and empirical frequencies diverge and the chi-square distance measures escalate to χ^2values over 800 and thus indicating significant statistical difference and major departures from the Benford's SD distribution." Abrantes-Metz et al., *op. cit.,* Tracking the Libor rate, p. 896-897. We found the same results using the percentage based formula.
Our calculation of the "percentage chi-square" differed slightly from those found in Abrantes-Metz. The minor differences between Abrantes-Metz's results and ours may be due to how Abrantes-Metz constructed the data. For example, Abrantes-Metz appears to have included weekend LIBORs based on the LIBOR that existed on the previous active LIBOR market day. We included LIBOR rates only for days when LIBOR was computed. There may be some other differences in the calculation as well, but which result in the same conclusions about statistical significance. The percentage based chi-square statistics calculated using Eq 2 of this paper for March 2007-Aug 2007 equals 642.10, for April 2007-Sep 2007 equals 405.62, for May 2007-Oct 2007 equals 219.99, and for Jun 2007-Nov 2007 equals 100.12. In Abrantes-Metz et al., *op. cit.,* Tracking the Libor rate, p. 898, Table 2 lists 639.60, 410.42, 222.86, and 98.68 as the chi-square statistics for these time periods respectively.
[33] During a phone call with George Judge in February, 2013, Prof. Judge said that he had no knowledge of how the chi-square statistic reported in Abrantes-Metz was calculated because he did not participate in the implementation of the statistical tests in the paper, but that if the reported test statistic was calculated based on percents rather than on raw counts, n, it was performed incorrectly.
[34] On December 15, 2016 separate email requests were made to Dr. Abrantes-Metz and Dr. Villas-Boas (Corresponding Author) for the programs used to perform the calculations in Abrantes-Metz et al., *op. cit.,* Tracking the Libor rate. Both Authors responded that the programs were not available.

constructed one of its test periods to have 228 months while most of the others were 6 months. Cleansing the analysis from any impact of differing numbers of days could have been achieved by breaking the 228-month period into rolling 6-month periods as Abrantes-Metz did in the later periods. This allows for comparability of scale of observations and has the added benefit of permitting the use of valid statistical methods. This latter approach of cleansing the analysis of any impact of differing numbers of observations, by using test periods of equal size combined with a standard chi-square calculation is used in this paper, such as in Sections II and III. Abrantes-Metz states that the tests implemented in the paper were "the traditional chi-square goodness-of-fit tests" which was explicitly listed in the footnote as the standard Pearson chi-squared.[35] Abrantes-Metz did not describe how some scale-invariant form of chi-squared was achieved by substituting percentages in place of frequencies or what the statistical properties of this test statistic are.[36] However, since raw counts are used in a valid Pearson's chi-square test, the values listed as χ^2 throughout Abrantes-Metz are not distributed chi-square and critical

---

[35] Abrantes-Metz et al., *op. cit*., Tracking the Libor rate, p.895.

[36] The substitution of frequencies for percentages in a chi-square equation not only invalidates the use of the standard chi-square critical values such as in Snedecor and Cochran, *op. cit*., p. 550-551, it does not create a valid scale-invariant statistic to test which test periods are more divergent from a reference distribution such as Benford. This percentage based measure would produce a scale invariant statistic when the test periods have equal numbers of observations. However, even then the statistic would not be a chi-square. To demonstrate this, consider a simulation performed 100,000 times where in each simulation a pair of random samples is taken from a second digit of Benford distribution. The first sample has 4,800 observations, approximately what Abrantes-Metz had in the first period where they claim no statistical significant difference from Benford. The second sample has 130 observations, similar to the number of observations in the 6-month test periods. For each of the 100,000 simulations the "scale invariant chi-square" is calculated individually for the sample of 4,800 observations and for the sample of 130 observations, based on the distribution of the second digit of the Benford Distribution. Because these pairs are drawn from the same second digit of Benford distribution, if this percentage test were scale-invariant the distribution of the test statistics ("percentage statistic") for these pairs of 100,000 draws, the 4,800 sample size and the 130 sample size, should have the same distribution. But they do not. The percentage statistic for the 100,000 simulations of 4,800 samples has a median statistic of 0.174. The 1 percent cut off is only 0.449. The percentage test statistic for the 100,000 simulations of 130 samples has a median statistic of 6.428. The 1 percent cut-off for the 100,000 simulations for 130 observations is 16.757. In both set of simulations an appropriately calculated chi-square statistic had a 1 percent cut-off of approximately 21.67, as a chi-square table would list for 9 degrees of freedom. Clearly when using this altered version of the chi-square statistic the sample size has an impact on the resulting statistic; it is not scale-invariant. The samples with 130 observations have a statistic that is about 36.9 =4,800/130 times larger than those with 4,800 observations. Comparing other types of distribution to that of some Benford derived distribution reveal other biases.

values of a chi-square distribution do not apply, unless by chance some of the periods tested happened to have 100 LIBOR reporting days.[37] We believe none in Abrantes-Metz did.

For the individual 6-month rolling periods that Abrantes-Metz tests, there are about 130 days of LIBOR data, rather than 100 days, in each period. Replacing the actual number of days, approximately 130, with the value 100 will reduce the calculated chi-square by a factor of .769=100/130. However, for the period 1987-2005, which Abrantes-Metz tests as one single period, there are approximately 4,800 days. The calculated value of the standard Pearson chi-squared statistic for that period, 1987-2005, is 639.90, not 13.89, as listed in Abrantes-Metz.[38] The great disparity in the actual chi-square results from the incorrect implementation of the chi-square test found in Abrantes-Metz, reducing the actual chi-squared by about 48 times, or a factor of approximately 100/4800 =.021.

Based on the standard Pearson's chi-square statistic of 639.90, we reject the equality of the Benford second digit distribution to the distribution of LIBOR second digits for the single 19 year test period from 1987-2005 used by Abrantes-Metz. That is, the finding that there was no statistically significant divergence between the second digit of LIBOR and that of a Benford distribution is the result of errant statistical methods. Based on Benford tests implemented by Abrantes-Metz, but using the correct form of a chi-square test, even when employing the single 19-year test period used by Abrantes-Metz, there is evidence of LIBOR manipulation, based on the proposed Benford test, for the aggregate period 1987-2005. This period between 1987 and 2005 is the period for which Abrantes-Metz concluded there was no manipulation, based on that paper's invalid test statistic based on percentages.

The second major difference between the analysis found here and that of Abrantes-Metz is that our analysis, above, is based on consistent test time periods, a chi-square test on 6-month rolling periods. In contrast, the Abrantes-Metz analysis starts with a single 19-year period (228-months)

[37] Of course in the case where all of the observation periods have 100 days there would be no interest in a scale-invariant chi-square test because the scales would be equal. So while Abrantes-Metz seems to have had the goal of achieving a scale-invariant statistic by altering the Pearson chi-square test statistic, that goal was not achieved.

[38] Abrantes-Metz, *op. cit.*, Tracking the Libor rate 895.

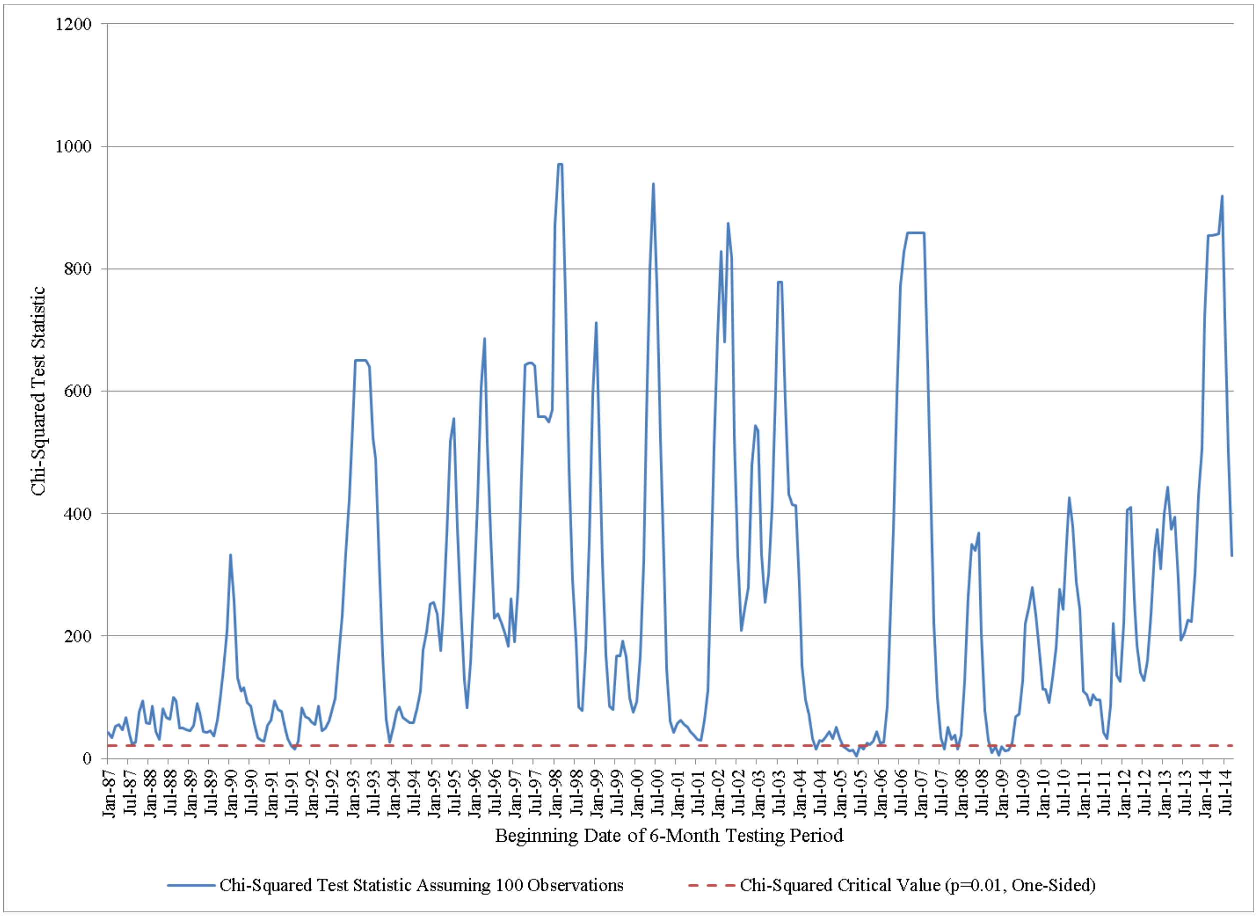
ending in 2005 and then changes, in the time period that starts in August 2005 and thereafter, to 6-month rolling time periods. Based on the single 228-month testing period (and the invalid construction of a chi-squared test), Abrantes-Metz finds no statistically significant difference between the distribution of the second digit of LIBOR and that of the Benford reference. Beginning in August 2005, Abrantes-Metz switches the testing period to 6-month rolling periods and finds, starting in February 2006, 18 consecutive periods of statistically significant differences between the distribution of the second digit of LIBOR and that of a Benford distribution.

This proximity in Abrantes-Metz between the date at which the authors switch the length of the testing period from 228 months to 6-month rolling periods and the start of statistically significant differences between the distribution of the second digit of LIBOR and that of the Benford reference is much more than a mere coincidence.

Our analysis described above, in Figures 1 and 2, shows that if a consistent 6-month rolling test period is used, the distribution of the second digit of LIBOR is almost always found to be statistically different from that of the Benford reference, starting as far back as 1987. This is true regardless of whether we use the correct Pearson chi-square test or the test statistic based on percentages that Abrantes-Metz appears to use.[39] Figure 3 replicates Figure 1, but in order to test whether the shift in the testing periods from 6-month to 228-month, alone, has created this finding of a change in the LIBOR behavior over time, we hold other aspects of the Abrantes-Metz paper constant, namely we uses percentages instead of actual frequencies in what is otherwise a chi-square formula. Again, the blue series plots the test statistic for rolling 6-month periods. The red lists the critical value for a standard chi-square test based on 10 cells, corresponding to the ten possible values of the second digits, 0-9, at a significance level of .01.

[39] We note that we have not found this second type of chi-square test based on percentages in the statistical literature.

Figure 3: Test Statistic Measuring Equality of Benford and Empirical Distribution (Using *n*= 100 or Percentages), 1-Month LIBOR



Even using the percentage test statistic created by Abrantes-Metz, the consistent use of 6-month rolling periods across the entire time period finds that there are only 18 6-month rolling periods out of 333, with starting dates between January 1987 and September 2014 in which test statistics are below 21.67, the critical value of a chi-square for 10 categories. If we were to use the percentage-based test statistic, as does Abrantes-Metz, it would reject the null hypothesis that the empirical distribution of LIBOR second digits follows that of the Benford expected distribution in 315 of 333 test periods. The starting dates of the 18 6-month periods where the second digit of LIBOR is not statistically significantly different from the second digit of a Benford distribution are July 1991, August 1991, June 2004, February 2005, March 2005, April 2005, May 2005, June 2005, July 2005, August 2005, August 2007, December 2007, October 2008, November 2008, December 2008, January 2009, February 2009, and March 2009.

This means that neither the behavior of the banks nor tampering nor “human influence” in the setting of LIBOR, as asserted by Abrantes-Metz, led LIBOR to transition from an apparent consistency[40] with Benford between 1987 through 2005 to a period of apparent intermittent statistically significance divergence from that of a Benford distribution sometime after 2005. There was no such transition from non-statistically significant difference to statistically significant difference. Instead, the change that Abrantes-Metz identified was simply a product of change in the method of testing, in the form of the change in the length of the test periods, before and after January 2006 combined with a non-valid test statistic.

We have already shown that a consistent use of 6-month testing periods combined with the implementation of the standard Pearson chi-square statistic produces nearly uniform findings of statistically significant differences between the distribution of the second digit of LIBOR and that of the Benford reference. In addition, we have shown that the change of length of the testing period of a 228-month period prior to 2006 to 6-month periods in 2006 and after, creates the finding of a change in the behavior of LIBOR in comparison to Benford, even when maintaining the use of the non-standard test statistic, based on percentages, as used in Abrantes-Metz.

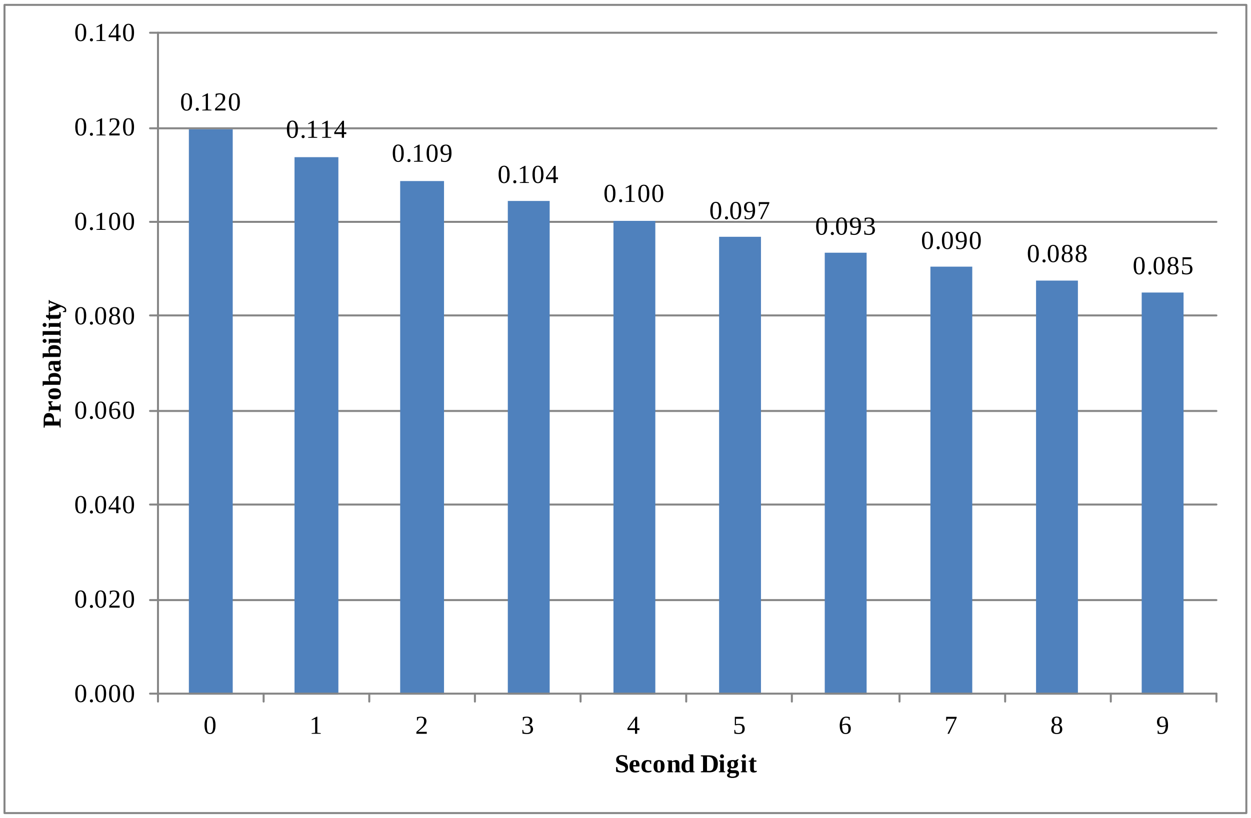
We next turn to why the shorter time periods creates the divergence between the distribution of values found in the second digit of Benford and that of the second digit of the distribution of LIBOR.

V. Why Switching Duration of Test Period Switches Statistical Significance

The distribution of second digits of a Benford distribution, presented in Figure 4, is relatively flat over the integers from 0 to 9.[41]

[40] Consistent in the sense of not being statistically significantly different.
[41] Benford, *op. cit.*, p.551-72.

Figure 4: Benford Distribution of Second Digits

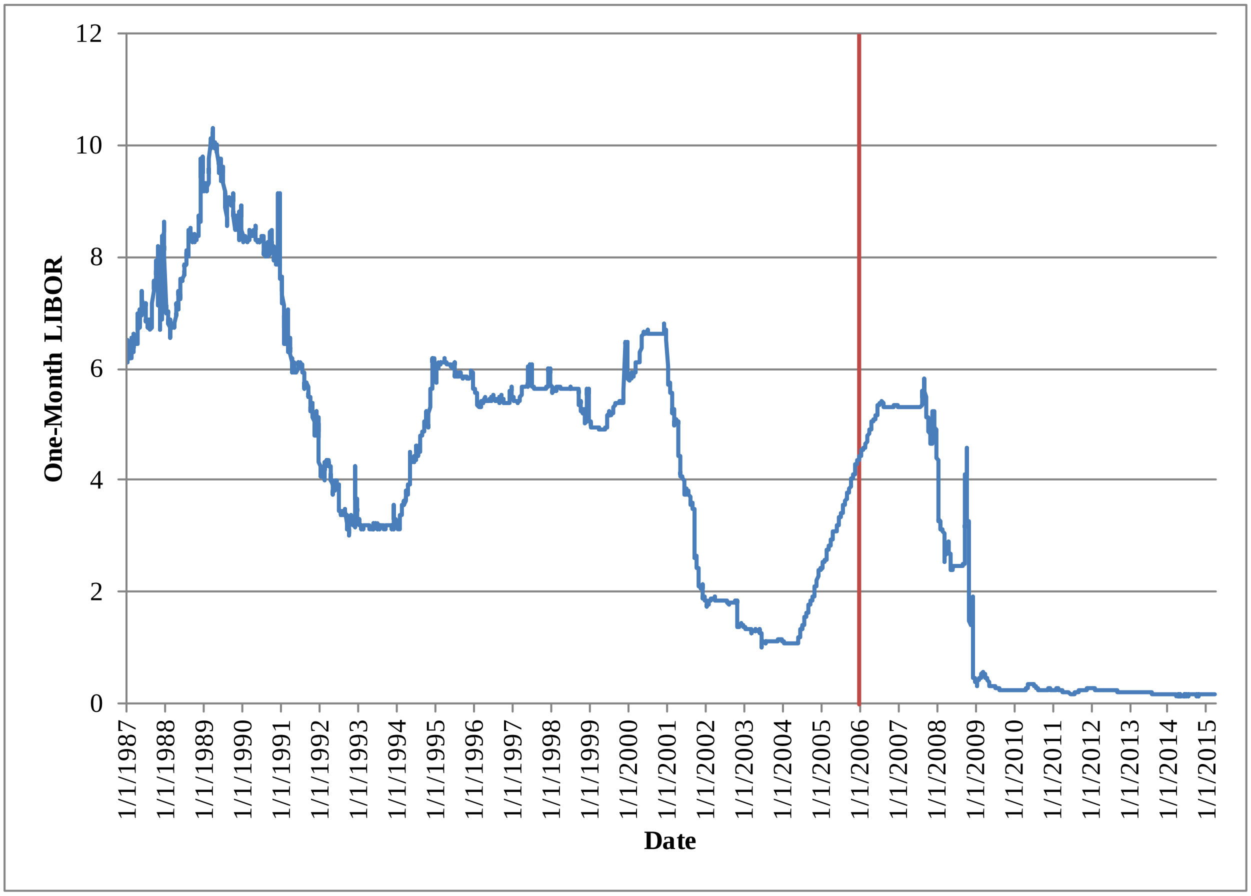


LIBOR in its untampered state should reflect the outcome of a market. It is not a randomly generated number. In periods of stable interest rates and risks, it is likely to change slowly and perhaps less frequently, containing serial correlation.

To the extent that LIBOR is relatively stable during a given test window, its first digit will be less likely to span the range of digits from 1 to 9. This point is so obvious that in Abrantes-Metz it was the justification Abrantes-Metz offered for not using the first digit of Benford and instead using the second digit of Benford. But the obviousness that LIBOR is not distributed like a Benford distribution in the first digit does not provide any support that the second digit of LIBOR should be distributed like the second digit of Benford. If during a given test period, LIBOR is relatively stable, the second digit of LIBOR will be less likely to span as many digits from 0 to 9 in as even a manner as does the second digit of a Benford distribution, shown in Figure 4. Equivalently, the shorter the testing period employed, the greater the likelihood that LIBOR will not span the range of values, and if it does, will not span them like the second digit of the relatively flat Benford distribution. Abrantes-Metz provides no research or evidence suggesting that LIBOR in its un-tampered state is generated in some fashion that would cause the second digit of multiple days of LIBOR to be distributed like the second digit of a Benford distribution.

It is, however, possible to understand why employing long testing periods such 228-month periods would allow the distribution of the second digit of LIBOR to more closely conform to the second digit of Benford than would shorter testing periods, such as 6-month intervals. Figure 5 shows the 1-month LIBOR, reported each day from January 1987 to February 2015.

Figure 5: 1-Month LIBOR Fixing Rate

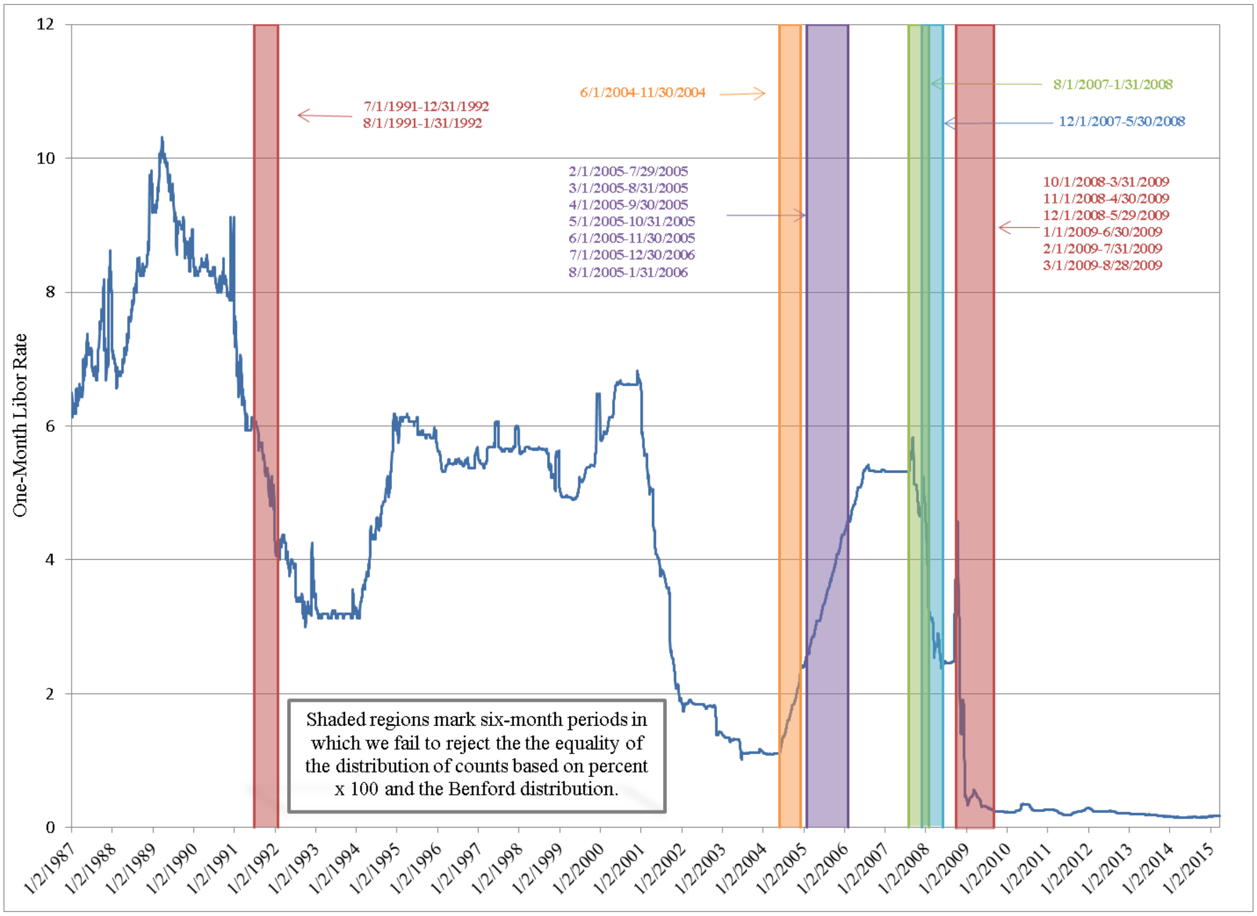


Over the period from 1987 through the end of 2005, LIBOR ranges from a high of over 10% to a low of just over 1%. There are periods of more rapid change and other periods of relative stability. As mentioned above, taken as a whole over the entire period from the start of 1987 through the end of 2005, Abrantes-Metz finds that the second digit of LIBOR conforms relatively well to the distribution of the second digit of a Benford distribution based on the invalid test statistic use therein.[42] But based on test periods of 6-months, we found all but 18 6-month periods show a statistically significant deviation from Benford based on the Abrantes-Metz test statistic.[43] For a number of these 6-month periods the second digit of LIBOR does not move very much, simply because LIBOR is not changing much during these periods. The shaded regions in Figure 6 mark the 18 6-month periods in which the distribution of the second digit of LIBOR was found not to be statistically significantly different from that of Benford, using a test statistic based on percentages. As Figure 6 shows, these tended to be in periods of more rapid change in LIBOR. So even during the period that Abrantes-Metz characterized as consistent with Benford, the only 6-month test periods where the distribution of the second digit LIBOR was consistent with that of a Benford distribution are during periods of rapid change in LIBOR even when we use the same test statistic created by Abrantes-Metz.

[42] Note that if the correct formula for the Pearson chi-square test statistic is calculated for the period from 1987-2005, it is 639.90 for the Benford second digit distribution and 727.07 for the uniform distribution. As a result, we reject the equality of the empirical distribution of LIBOR second digits to both the Benford and uniform distributions.

[43] As noted above, when using the Pearson chi-square test based on actual counts, rather than assuming 100 observations, there are 12 6-month periods for which we fail to reject the Benford distribution.

Figure 6: 6-Month Periods in which Actual Distribution of LIBOR Second-digit does not Differ from Benford Distribution (Chi –Squared Statistic assumes n=100 or Using Percentages)



These tests against Benford's second digits could very well simply be identifying some of the periods of comparatively greater movement in LIBOR as not statistically significantly different from Benford, which may be unrelated to any issues of tampering or inappropriate human intervention in the setting of LIBOR. If we cut the test periods further, from 6-month periods to 3-month rolling periods, there are only 4 periods in which we fail to reject the equality of the Benford and the empirical distribution of second digits of LIBOR.

VI. What Can We Conclude about LIBOR Manipulation Using Benford?

Summarizing, the set of tests performed so far provides a clear picture of whether the tests based on the second digit of Benford provide any indication of LIBOR manipulation. First, using a consistent testing method based on the rolling 6-month periods with start dates beginning in January 1987 through September 2014, we reject the null hypothesis that the LIBOR second digits follow the second digit of a Benford distribution in over 95% of the periods.[44] This is true whether we use the standard Pearson chi-square test (using frequencies) or the Abrantes-Metz test statistic (using percentages).[45] It is also true if we limit the data to the time period used in Abrantes-Metz from January 1987 to October 2008.

Second, even when using multiyear test periods, the similarity between the distribution of the second digit of LIBOR and that of a Benford distribution is rejected when using the standard Pearson chi-square test. This includes the longer multiyear test periods from 1987-2005, from 2006-2008, from 2006-2011, or from 2006-February 2015 for which the chi-squares all indicate statistically significant differences between the distribution of the second digit of LIBOR and the second digit of a Benford distribution.[46]

Third, if we use the invalid test statistic of Abrantes-Metz, but use consistent periods -- either shorter periods like 3-months, 6-months or longer periods like 5-years, we reject that the second digit of LIBOR is distributed like the second digit of a Benford distribution both before and after 2006. For example, if we test consistent 5-year periods from 2001-2005 and 2006-2010, based on the invalid test statistic, we reject that the LIBOR second digit was distributed like that of a

[44] It is important to note that the statistical tests performed here are on overlapping periods. Any finding of statistical significance in a single period must be considered in the context of the hundreds of statistical tests being performed.
[45] As noted above we have been unable to locate a reference to this chi-square test based on percentages in the statistical literature.
[46] From Jan1987-Dec2005 chi-square equals 639.90, from Jan 2006-Dec2008 chi-square equals 741.67, from 2006-2011 chi-square equals 616.68, or from 2006-February 2015 chi-square equals 622.71. Each is statistically significant.

Benford distribution in both periods.[47] Of course the test statistic used in Abrantes-Metz (2011) has unknown properties, so rejection criteria and outcomes have an unknown meaning.

Fourth, Abrantes-Metz constructed its evidence that the distribution of the second digit of LIBOR was inconsistent with the distribution of the second digit of Benford only after 2006 by combining *both* the non-standard test statistic with a change in the length of the test period in the middle of the analysis, at approximately the point where Abrantes-Metz found the switch from consistency with Benford to inconsistency with Benford. This combination of test statistics and change in length of test periods produces the finding that the second digit of LIBOR was consistent with the second digit of Benford from 1987 through 2005 and then this consistency vanished around September 2005. Uniformly using 3-month, 6-month or long time periods, (e.g. Jan 1987-Dec2005, Jan2006-Oct2008, or Jan2006-Feb2015)[48] combined with standard chi-square statistics, produces a rejection of the second digit of LIBOR being consistent with the second digit of Benford over the vast majority of the time period tested. If the invalid test statistic employed by Abrantes-Metz is used on the very long 19 year period of 1987-2005, it gives the false impression that LIBOR is consistent with the second digit of Benford, while the period after 2005, such as Jan2006-Dec2010 or Jan2006– Feb2015 is not[49]

In sum, the Abrantes-Metz findings are extremely sensitive to the length of periods chosen and are produced by using an invalid test statistic that substitutes percentages in place of frequencies in what is otherwise a chi-square formula. Calculated correctly, using a standard Pearson chi-square test and periods with a consistent period length, the second digit Benford test as suggested by Abrantes-Metz for a real time test of LIBOR would reject equality between the second digit of Benford and the second digit of LIBOR 96.4% of the 6-month rolling periods with start dates from January 1987 through September 2014. Further a standard chi-squared statistic will, for the

[47] The percentage based test statistic for Jan 2001-Dec 2005 equals 37.92, for Jan 2006-Dec 2010 equals 58.23 and for Jan 2006 –February 2015 equals 26.90. Each is statistically significant.

[48] For 3-month rolling periods, we reject the 2nd digit Benford distribution in 329 of the 336 periods tested. For 6-month rolling periods, we reject the 2nd digit Benford distribution in 321 of the 333 periods tested. For the period 1987-2005 the test statistic is 639.9, for Jan 2006-Oct 2008 it is 768.94, and for Jan 2006-Feb 2015 it is 622.71.

[49] If a test statistic based on a percentage is used, the period of 1987-2005 (19 years) gives the false impression that LIBOR is consistent with the second digit of Benford, while the period 2006-2010 (5 years) does not. Test statistic based on percentages for 1987-2005 equals 13.32 and for 2006-2010 equals 58.23.

longer periods we tested, reject the similarity between LIBOR and the Benford distribution. If inconsistency between the second digit of LIBOR and the second digit of the Benford distribution is a warning of tampering, as suggested by Abrantes-Metz, then the warning bell has been ringing almost non-stop for more than 25 years. And this same alarm continued to sound into 2015, after extensive investigations and resulting reforms.[50] But, as we have shown, the alarm is likely not sounding for the behavior of LIBOR or those who set it, but rather for the methods used to test LIBOR.

VII. Much Ado About Benford

So far we have found that the distribution of the second digit of LIBOR does not follow the distribution of a second digit of Benford for most of the past 25 years. Is there any reason to suspect that this inconsistency with the second digit of Benford's distribution means that there has been long-term LIBOR tampering going on? There is no empirical evidence found in Abrantes-Metz to suggest that the distribution of the second digit of an *un*-tampered LIBOR should follow that of Benford. Further, Abrantes-Metz does not provide citations to previous literature that the distribution of the second digit of Benford should be a useful benchmark for the second digit of an un-tampered LIBOR. In addition, Benford's distribution is not a mathematical law like π, measuring the relationship between the circumference and the diameter of every circle, or a law of nature, like gravity, which once proven can be relied on uniformly. It may be an empirical finding that may hold true in some cases, for some periods of time. It may or may not apply to LIBOR.[51] Benford's distribution reflects an empirical tendency that Benford

[50] The 8 6-month rolling test periods from the initiation of ICE LIBOR in February 2014 through September 2014 all exhibit a statistically significant difference in the second digit of LIBOR as compared to that of a second digit of Benford distribution. In addition, when tested as a single period from February 2014 through February 2015 the second digit of ICE LIBOR is statistically significantly different from that of a second digit of Benford distribution (chi-square is 1486.377 with 272 days and 10 categories).

[51] For a test of various theoretical distributions against the Benford Distribution see Leemis, Lawrence M,, Bruce W. Schmeiser, Diane L. Evans, "Survival Distributions Satisfying Benford's Law," *The American Statistician,* 54, no. 3 (August 2000). This paper used a chi-squared divided by the number of observations as a measure of the

observed.[52] He did not claim that it applied universally, perhaps not even generally. Benford did not suggest, at least not in his frequently cited publication, that deviations from the empirical tendency that he observed provided an indication of tampering. Others have suggested Benford's distribution as a reference measure for tampering based on studies that show a persistent regularity between Benford's distribution and the distribution of some measure of interest.[53] Of course, one does not need Benford's distribution to test for a pattern in one period and then check for departures from that pattern across time or geographies or other dimensions; one can do that by checking for statistically significant change over time or geographies with time series data or data from across geographies or other dimensions. Such a test would have to be done, even if applying a Benford distribution to some data during some "clean" period to show that the data followed the proposed benchmark distribution, Benford or otherwise. Based on enough previous empirical evidence of consistency, and further theoretical development, those sorts of tests may be useful. However, to determine the right amount of weight to put on a test of number patterns against Benford, without such testing of a clean base period, it may be instructive to consider how Benford developed his distribution.

In his original paper, Benford lists 20 groups for which he counted frequencies. These include the following, as Benford enumerated them: newspaper items; pressure loss, air flow; h.p. lost in airflow; street addresses from the *American Men of Science* directory; American League, 1936; black body radiation; x-ray voltage; items in *Readers Digest* (except dates and page numbers); area of rivers; death rates; cost data, concrete; $n^1 ... n^8$,n!; design data generators; population, USA; drainage rate of rivers; n^{-1}, $n^{1/2}$…; molecular weights; specific heats; physical constants, and atomic weights.[54] From this group of 20 series, Benford observed a pattern and then developed formulas that approximated this pattern. If Benford's distribution is useful as a

---

comparative fit of various theoretical distributions against Benford. It does not use this statistic for a hypothesis test. Further it is a comparison of various theoretical distributions and hence there was no number of observed items. In this case, there is no distinction between the observed and the expected. In this case, for comparing fit across various distributions, dropping the *n* has no impact on the relative order of which distribution fits best.

[52] Benford, *op. cit.*, p.551-72.

[53] See, for example. Cindy Durtschi, William Hillison, and Carl Pacini, "The Effective use of Benford's Law to Assist in Detecting Fraud in Accounting Data," *Journal of Forensic Accounting*, 17-34.; Wendy K. Tam Cho. and Brian J. Gaines. "Breaking the (Benford) Law Statistical Fraud Detection in Campaign Finance," *The American Statistician*, 61(3) (2007): 218-223.

[54] Benford, *op. cit*., p.553, Table I.

detection tool for manipulation, it seems logical that these 20 series that Benford used to identify his Benford's law should follow Benford's law. Since Benford listed his counts for each of these categories in his publication, we can test whether the 20 component series that went into the construction of Benford's distribution are themselves consistent with Benford's distribution. To do this we use a Pearson's chi-square test.[55] The results are presented in Table 2.

Table 2: Test of Equality of Theoretical Benford First Digit Distribution and Examples from Benford (1938)

	Chi-Squared Test Statistic		P-Value
Rivers, Area	4.96		0.76
Population, USA	118.63	***	0.00
Physical Constants	24.44	***	0.00
Newspaper Items	0.16		1.00
Specific Heat	111.21	***	0.00
Pressure Lost, Air Flow	1.27		1.00
H.P. Lost in Air Flow	3.46		0.90
Molecular Weights	125.76	***	0.00
Drainage Rate of Rivers	11.14		0.19
Atomic Weight	17.25	**	0.03
$n^{-1}, n^{1/2}\ldots$	440.76	***	0.00
Design Data, Generators	19.21	**	0.01
Reader's Digest	3.23		0.92
Cost Data, Concrete	15.60	**	0.05
X-Ray Volts	5.43		0.71
American League, 1936	14.60	*	0.07
Black Body Radiation	9.52		0.30
Addresses	1.30		1.00
$n^1 \ldots n^8, n!$	24.99	***	0.00
Death Rates	7.55		0.48

Note: *** denotes items for which we reject the null hypothesis that the empirical distribution is consistent with Benford at the 1% level. ** denotes rejection of the null hypothesis at the 5% level. * denoted rejection of the null hypothesis at the 10% level.

[55] This is the standard chi-Square test and is the same as listed in note 3 of Abrantes-Metz et al., *op. cit.* Tracking the Libor rate,, but which was *not* actually empirically implemented in that paper.

The Table 2 shows that of the 20 series that Benford used to construct Benford's law, 6 are statistically significantly different from Benford's law at the 1% level. At the 5% level, 9 of Benford's series are statistically significantly different from the Benford distribution. At the 10% level, 10 of Benford's 20 series presented in his paper are statistically significantly different from the Benford distribution. As a detection device Benford's law would identify that between 6 and 10 of the series, depending on the significance level used, out of 20 component series in Benford (1938) that went into the identification of Benford's distribution, violated what is now called Benford's law. Clearly Benford's law is not a universal law. Further, it does not even appear to be a broadly applicable empirical tendency. This does not mean that Benford's distribution might not fit the empirical patterns of some data series. But it does demonstrate that without significant empirical support or other theoretical justification, departures from Benford's law may well be the norm for many distributions rather than the exception.

VIII. Benford, Evidence, and Why Scientific Methods Matter

Scientific methods to test behavior are critical to enforcement actions and the legal system. Scientific, empirical tests of collusion have been used in antitrust litigations for decades.[56] Techniques that achieve this scientific status can drive significant enforcement efforts and are used to determine guilt and innocence. In the instance of the LIBOR manipulations, scientific tests used to detect manipulation or tampering can guide regulators, enforcement agencies and private parties to unleash vast resources in the pursuit of potential perpetrators. By identifying, or increasing the probability of identifying perpetrators, these scientific tests can save enforcement

[56] William S. Comanor and Mark A. Schankerman note that techniques to screen for collusive behavior have been employed by federal agencies since at least 1936. Comanor, William S., Mark A. Schankerman, "Identical Bids and Cartel Behavior," *The Bell Journal of Economics*, 7, No. 1, (Spring, 1976) 281-286. Comanor and Schankerman writing in 1976, point out that the government policy since 1936, of relying on identical bidding as a common, extant objective, predictive way to identify collusive behavior to proactively identify collusion "was surely misplaced" in highly concentrated industries. This points to the fact that regulators have known of the need for an objective, predictive way to identify collusive behavior for nearly a century, perhaps longer. It also demonstrates that fact that the extent to which objective predictive tests of collusion do not exist, results from the difficulty in identifying accurate, objective predictive tests, when such screens are scrutinized seriously.

and legal resources. Further, by increasing the probability of finding perpetrators, the incentives for malfeasance are reduced, which in turn improves the accuracy of LIBOR as a true measure of the cost of borrowing of major banks.

However, unscientific tests cloaked in the trappings of science, including those that clear the hurdles of publication in peer reviewed journal, waste the resources of regulators, enforcement agencies and the courts by providing incorrect information and signaling infractions where there are none. These unscientific tests allow perpetrators to go unnoticed and snare the innocent in traps meant for others. In this case, unscientific methods used to detect manipulation, not only increase the chance to damage the innocent and give leave to the guilty, they also damage international financial markets and worldwide economic resource allocation by diverting the attention and clouding the vision of regulators, enforcers and participants in the legal system who need to separate the guilty from the innocent.

In this case, there is no evidence that the second digit of Benford's distribution provides any meaningful information about manipulation of LIBOR. While the impact of LIBOR manipulation continues to be assessed, the test proposed by Abrantes-Metz, Villa-Boaz, and Judge based on the second digit of Benford's distribution cannot distinguish periods of manipulation from periods of non-manipulation and provides no information about manipulation, tampering or collusion, on the part of the panel banks that set LIBOR.

Works Cited

Abrantes-Metz, Rosa M., Sofia B. Villas-Boas, and George Judge. "Tracking the Libor rate." *Applied Economics Letters* 18, no. 10 (2011): 893–899.

Abrantes-Metz, Rosa M., "How to Use Statistics to Seek Out Criminals" in *Bloomberg*, February 26, 2012. https://www.bloomberg.com/view/articles/2013-02-26/how-to-use-statistics-to-seek-out-criminals.

Becker, Gary S. “Crime and Punishment: An Economic Approach” *Journal of Political Economy* 76, no. 2 (March-April 1968): 169–217.

Benford, Frank. “The Law of Anomalous Numbers.” *Proceedings of the American Philosophical Society* 78, no. 4 (March 1938): 551–572.

Board of Governors of the Federal Reserve System (US), Effective Federal Funds Rate [DFF], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/DFF, November 16, 2016.

Comanor, William S., Mark A. Schankerman, “Identical Bids in Cartel Behavior,” *The Bell Journal of Economics*, 7, No. 1, (Spring, 1976) 281-286.

Durtschi, Cindy, William Hillison, and Pacini Carl. “The Effective Use of Benford’s Law to Assist in Detecting Fraud in Accounting Data.” *Journal of Forensic Accounting* 5 (2004): 17–34.

“Free Exchange: The Scam Busters.” *The Economist*, December 15, 2012. http://www.economist.com/news/finance-and-economics/21568364-how-antitrust-economists-are-getting-better-spotting-cartels-scam-busters.

ICE Benchmark Administration Limited (IBA), 1-Month London Interbank Offered Rate (LIBOR), based on U.S. Dollar© [USD1MTD156N], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/USD1MTD156N, November 16, 2016.

ICE Benchmark Administration Limited (IBA), 3-Month London Interbank Offered Rate (LIBOR), based on U.S. Dollar© [USD3MTD156N], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/USD3MTD156N, November 16, 2016

Johnson, Richard A., and Gouri K. Bhattacharyya. *Statistics: Principles and Methods*. 5th ed. John Wiley & Sons, 2006.

Leemis, Lawrence M,, Bruce W. Schmeiser, Diane L. Evans, “Survival Distributions Satisfying Benford’s Law,” *The American Statistician*, 54, no. 3 (August 2000).

Tam Cho, Wendy K., and Brian J. Gaines. “Breaking the (Benford) Law: Statistical Fraud Detection in Campaign Finance.” *The American Statistician* 61, no. 3 (August 2007).

Wheatley, Martin. “The Wheatley Review of LIBOR: Final Report.” HM Treasury, UK, September 2012. http://cdn.hm-treasury.gov.uk/wheatley_review_libor_finalreport_280912.pdf.

Cases Cited

Charles Schwab Bank, N.A. v Bank of America Corporation *et al*, MDL 2262, August 23, 2011

The City of Philadelphia v Bank of America *et al*, Case 2:13-cv-04352-LFR, July 27, 2013.

The Mayor and City Counsel of Baltimore *et al* v Credit Suisse Group AG *et al*,
MDL No. 2262. April 30, 2012