Federal Reserve Bank of New York
Staff Reports

A Sampling-Window Approach to Transactions-Based Libor Fixing

Darrell Duffie
David Skeie
James Vickery

Staff Report No. 596
February 2013

A Sampling-Window Approach to Transactions-Based Libor Fixing

Darrell Duffie, David Skeie, and James Vickery

Federal Reserve Bank of New York Staff Reports, no. 596
February 2013
JEL classification: G01, G10, G18, G28

Abstract

We examine the properties of a method for fixing Libor rates that is based on transactions data and multi-day sampling windows. The use of a sampling window may mitigate problems caused by thin transaction volumes in unsecured wholesale term funding markets. Using two partial data sets of loan transactions, we estimate how the use of different sampling windows could affect the statistical properties of Libor fixings at various maturities. Our methodology, which is based on a multiplicative estimate of sampling noise that avoids the need for interest rate data, uses only the timing and sizes of transactions. Limitations of this sampling-window approach are also discussed.

Key words: shadow banking, financial intermediation

Duffie: Stanford University (e-mail: duffie@stanford.edu). Skeie, Vickery: Federal Reserve Bank of New York (e-mail: david.skeie@ny.frb.org, james.vickery@ny.frb.org). The authors thank David Hou and Ali Palida for outstanding research assistance, as well as Spence Hilton, Antoine Martin, Jamie McAndrews, and Simon Potter for comments.

1. Introduction

This note considers an approach to constructing Libor fixings using transactions data and multi-day sampling windows.[1] For instance, one could fix the 3-month Libor rate on a given date as the average of the actual interest rates on all 3-month loans in the relevant historical sample whose transactions dates are within the trailing 10 business days. This "10-day sampling window" is merely for purposes of illustrating the concept. We will examine the influence of the sampling window on sampling noise and consider additional techniques for "fattening" the sample and weighting the data so as to reduce sampling noise and mitigate biases. We also consider the potential range of applications of this approach, and some of its disadvantages.

Libor provides an estimate of the interest rate at which major banks active in London may borrow from other banks on an unsecured basis. The British Bankers Association (BBA) currently reports Libor on a daily basis for 10 currencies and 15 maturities between overnight and one year.[2] These daily interest rate "fixings" are constructed based on bank submissions. Each of a panel of banks self-reports its own estimated hypothetical borrowing rates at each tenor. Notably, Libor is not currently computed directly from actual loan transaction rates. Published Libor rates are referenced in the settlement of many forms of financial contracts, including corporate bonds and loans, mortgages, as well as interest-rate futures, swaps and options.

Attention has recently focused on the potential to address shortcomings of the survey approach to Libor with a fixing method that is somehow based directly on actual loan transactions data. While advocating for the retention of a submission-based approach, the Wheatley Review of Libor (H.M. Treasury, 2012) recommends that Libor submissions should be "explicitly and transparently supported by transaction data." It also outlines guidelines for how this principle should be implemented in practice by Libor panel banks.[3] The judgment and expertise of submitting banks still plays a role under this approach.

An alternative would be to compute Libor directly as an average of individual transaction rates. One concern over such an approach, however, is the relative sparseness of daily interbank unsecured loan transactions at certain maturities,

[1] Libor stands for "London Interbank Offered Rate".

[2] The number of currencies and maturities is planned to be reduced in the future in line with the recommendations of the Wheatley Review of Libor (H.M. Treasury, 2012). See section 2.

[3] These guidelines (section 4.8 of the Wheatley Review) lay out a hierarchy of transaction types that banks should use when determining their submissions. Highest priority is given to transactions in the unsecured interbank deposit market, particularly those undertaken by the contributing bank. In the absence of relevant transaction data the guidelines suggest that expert judgment should be used to determine the bank's submission. They also state that "submissions may also include adjustments in consideration of other variables, to ensure the submission is representative of and consistent with the market for inter-bank deposits", such as placing less weight on non-representative transactions.

particularly during periods of financial stress. A fixing that is based on relatively few transactions could have excessive sampling noise and could also create a heightened incentive for some market participants to transact with the purpose of influencing the daily fixing. (In a stock-market context, Carhart, Kaniel, Musto, and Reed (2002) discuss evidence of transactions designed to "paint the tape.")

The Wheatley Report (H.M. Treasury, 2012) indicates that there are too few transactions to support Libor in many of the currency-maturity pairs for which Libor is currently reported.[4] We show, however, that at least for some of the more active U.S. dollar maturities, the use of a sampling-window approach would significantly reduce the noisiness of transactions-based average interest rates. This approach would also improve robustness to misreporting incentives. The approach could be exploited either as the basis for a new fixing rate for a subset of currencies and maturities, or as a source of additional information in judging the validity of other fixing methods.

We illustrate the transaction-window approach using two partial datasets measuring unsecured wholesale lending activity. The first is a historical dataset of brokered interbank loans. The second is a set of putative unsecured loans inferred from Fedwire payments using a statistical algorithm developed by staff of the Research Group of the Federal Reserve Bank of New York that extends the work of Furfine (1999). (See Kuo, Skeie, Vickery and Youle, 2013 for a detailed description of this algorithm.)

We note that while these datasets are useful for illustrating our approach, neither could be used in practice as the basis for constructing a transaction-based index of bank funding costs. In particular, we emphasize that the Kuo et al. statistical algorithm identifies term interbank loans with error. Historically, algorithms based on the work of Furfine have been used as a method of identifying overnight or term federal funds transactions. The Research Group of the Federal Reserve Bank of New York has recently concluded that the output of its algorithm based on the work of Furfine[5] may not be a reliable method of identifying federal funds transactions.[6] This paper therefore refers to the transactions that are identified using the Research Group's algorithm as overnight or term loans made or intermediated by banks. Use of the term "overnight or term loans made or intermediated by banks" in this paper to describe the output of the Research Group's algorithm is not intended to be and should not be understood to be a substitute for or to refer to federal funds transactions.

---

[4] For this reason, and because of their low usage, the Wheatley Review recommends discontinuing Libor for tenors of 4, 5, 7, 8, 10 and 11 months, and discontinuing Libor entirely for five currencies. Reporting of Libor is to continue for the US Dollar, Euro, Japanese Yen, UK Pound and Swiss Franc.

[5] It should be noted that for its calculation of the effective federal funds rate, the Federal Reserve Bank of New York relies on different sources of data, not on the algorithm output.

[6] The output of the algorithm may include transactions that are not fed funds trades and may discard transactions that are fed funds trades. Some evidence suggests that these types of errors in identifying fed funds trades by some banks may be large.

Given the limitations of existing datasets, a transaction-based index would require constructing a centralized and auditable repository of relevant interbank transactions. One possible methodology for reporting the necessary data is the Trade Reporting and Compliance Engine (TRACE), developed by FINRA for the reporting of individual trades in certain types of fixed-income securities.

In section 4 of this paper we also highlight a number of conceptual limitations of the sampling window approach, and consider potential solutions. One important issue is that a fixing based on a lagged moving window will reflect stale information during periods when market conditions change rapidly, for example after monetary policy announcements, or at the onset of a financial crisis. For applications that allow hindsight, such as ex-post corroboration of other methods for fixing Libor, a two-sided sampling window could be used, incorporating transaction data from both the days before and after the fixing date. This could mitigate the staleness. A two-sided sampling window is of course infeasible if the fixing needs to be publicly released in real time. A second potential concern is that the available sample of underlying wholesale loan transactions may be small even with a multi-day sampling window, particularly during periods of market stress. One way to mitigate this problem could be to consider a wider range of unsecured funding instruments when constructing the transaction-based index.

2. Wider Sampling Windows

Suppose there is a source of actual transactions data on large unsecured loans to banks in the desired borrower class. In case the volume of interbank loan transactions is viewed as insufficient, one may wish to consider a wider range of sources of unsecured "wholesale" funding to major banks, perhaps including certificates of deposit, commercial paper and so on.

Even for a global currency such as the U.S. dollar, there are extremely few large unsecured loan transactions at many of the maturities at which Libor is currently fixed. Even a sampling-window approach would not be reliable in such cases. Alternatives for these "sparsely populated" maturities include interpolation, improving the current survey-based approach, or a cessation of Libor fixings, as recommended by the Wheatley Review of Libor. Fortunately, the maturities at which there are few transactions suitable for determining a reference rate are also less important for applications. For example, there are relatively few derivatives, bonds, and other instruments that reference 9-month Libor. The most commonly referenced Libor rates in major currencies are those with maturities of one month, three months and, to a lesser extent, six months, as indicated by a survey appearing in the Wheatley Review of Libor. We focus on currencies and maturities for which the aggregate-sample transactions frequency is potentially sufficient to consider for a fixing, or for validation of a fixing.

Even for relatively active U.S. dollar loan maturities such as 1 month and 3 months, we will show that a substantial proportional reduction in the sampling noise associated with a transactions-based fixing can be achieved with the use of a rolling sampling window. This is not surprising, but the empirical magnitude of the effect is notable. Moreover, our methodology does not rely on access to the interest-rate data themselves, but rather on the times and sizes of transactions. Our approach is in the spirit of statistical filters that attempt to extract longer-frequency movements in time-series data (such as the Hodrick-Prescott filter or the Kalman filter).

This approach could be employed in at least two ways: 1) to provide a replacement to the current quote-based approach for determining the Libor fixing; 2) in corroboration of a quote-based or poll-based Libor fixing, for example as part of the process of strengthening oversight of Libor. In the first application, it would be necessary to use a one-sided lagging window, since the fixing would need to be announced in real time. For ex-post validation purposes, however, it would be possible to use a two-sided sampling window to construct the fixing, incorporating both past and future data.[7] Our numerical examples below focus on a one-sided window. From a statistical filtering point of view, a two-sided sampling window would lower average the degree of sampling error.

Our simple illustrative example is a fixing of the 3-month rate on a given date as the average of the rates on all 3-month loans in the relevant historical sample whose transactions date is within the trailing 10 business days. One may also wish to use a sampling window based on maturity. We elaborate and generalize as follows.

Suppose one wants to create an estimate R(t,m) of a "representative" m-month maturity loan rate on day t. Let S(t,m;w,d) be the subset of all loans in the entire relevant historical sample available on the fixing date t whose transaction date is within the trailing w days and whose maturity is within d days of m. One could fix R(t,m) as the volume-weighted average of the loan rates in this fixing sample S(t,m;w,d). For example, for a lag (w) of 10 days and a maturity window (d) of 5 days, the fixing sample for the three-month borrowing rate (that is, m = 3 months) on a given day (t), say March 15, 2013, would include all transactions in the relevant pool with loan origination dates between March 1, 2013 and March 15, 2013, inclusive (that is, lagging by no more than 10 business days), with loan maturities of three months plus or minus 5 business days. In choosing the lagging transaction-date window w and the centered maturity window d, one can trade off the benefit of increased sample size against the cost of biases associated with increasingly stale or off-maturity data. In the last section, we explore the benefits and costs of reducing the weights applied to the transactions according to the time lag, in order to mitigate staleness bias.

In practice, the relevant term loan maturities appear to be tightly concentrated around the standard maturities of 1 month, 3 month, and 6 months. It may be

[7] We thank Simon Potter for alerting us to this point.

argued that it is relatively pointless to use a non-trivial maturity window. On the other hand, fattening up the sample by including similar-maturity loan transactions would lower sampling noise somewhat and seems unlikely to create important biases. The use of a maturity window also lowers the potential incentive for loan market participants whose transactions are sampled to customize their maturity dates so as to avoid entering the fixing sample.

3. Empirical Methods and Results

In this section we present a proportional sampling-noise measure and our empirical evidence regarding how variation in the sampling window and other data filters affects the "thinness" of the data underlying a potential transaction-based Libor index.

3.1 Data sources

We do not have access to a comprehensive transaction-level database of unsecured wholesale loans. In the absence of such data, we illustrate our approach using two partial data sources:

1. A dataset of brokered interbank transactions from the period 2000-04.
2. Statistically inferred transactions based on interbank transfers of federal reserves passing over Fedwire Funds Service ("Fedwire"), a large-value payment system operated by the Federal Reserve, from the period 2007-12.

The first of these data sources was previously used by Bartolini, Hilton and McAndrews (2010) and obtained from BGC Brokers, one of the four largest U.S. interbank brokerage firms. These data represent one of the only direct transaction-level research datasets for US-dollar-denominated interbank loans available for research. However, this dataset has a number of limitations. First, the data are available only for a historical time period from January 1, 2000 until September 27, 2004. This sample pre-dates the 2007-08 financial crisis and the post-crisis period. Second, the data cover only brokered loans, which represent only a subset of the interbank market, and represent only trades negotiated through a single broker. The identities of trade counterparties are not provided. Finally, the data cover only interbank loans, and thus do not include other unsecured funding instruments (such as wholesale time deposits) that may be useful for constructing a transaction-based Libor fixing.

The second data source is a set of term loans made or intermediated by banks inferred from payments passing over Fedwire using a statistical algorithm developed in Kuo, Skeie, Vickery, and Youle (2012) (KSVY). The KSVY algorithm is a generalization of Furfine (1999), who applied the method to identify potential overnight loans, not term loans. The idea behind the KSVY algorithm is that most wholesale interbank loans are settled over a large-value payment system. In the case of US-dollar loans, this is likely to be either Fedwire or Clearing House Interbank Payments System ("CHIPS"). The KSVY algorithm searches for transaction pairs consisting of a "send" leg (from party A to party B) for a large round-lot amount, and a "return" leg (from B to A) on a subsequent date for a slightly larger amount, such that the implied annualized interest rate is a whole number of basis points and such that the transaction pair meets certain other characteristics. For the purposes of this paper, the algorithm is used to identify putative interbank transactions for which both the sending and return leg pass over Fedwire between January 1, 2007 and May 1, 2012.

The most important disadvantage of the KSVY inferences is that the set of identified transaction pairs are inferences, not direct observations of term loans. It is difficult to verify at this point how well or poorly these pairs correspond to actual unsecured transactions. KSVY do however present some tests suggesting that the results of the algorithm are informative. For example, KSVY show that prior to the onset of the financial crisis, the distribution of implied interest rates of these putative loans is clustered tightly around the Libor fixing rate, implying that the results are not statistical noise.

As we discussed in the introduction to this paper, it is important to emphasize that this method is subject to both Type-I and Type-II classification errors (failures to discover actual loans, and inferred loans that are not actual loans). One particular concern is that the proximate counterparties identified in the Fedwire data may be acting only as correspondents, rather than being the ultimate borrower and lender of funds. This is especially relevant if a user of the data wants to restrict their sample to a particular subset of borrowers. Notably, recent research by Armantier and Copeland (2012) concludes that the related *overnight* Furfine algorithm performs poorly in identifying overnight federal funds loans conducted by two large banks.[8] (Note: Federal funds loans are a subcategory of interbank loans which are not subject to U.S. reserve requirements.)

Given the issues described above, we emphasize that neither of the data sources we consider could reliably be used in practice as the basis for computing a transaction-based replacement for Libor. In practice, such a fixing would presumably require the creation of a record log of actual wholesale loans (whether restricted to interbank loans, or encompassing a wider set of unsecured instruments), which could be aggregated or audited by regulators or other outside parties.

In the meantime, however, in the absence of a suitable database of actual term interbank loans, an analysis of these two datasets provides at least a rough idea of the effect of the size of the sample window and other filters on the robustness of the sampling-window approach. Given the limitations of the data sources, we do not

[8] In part because of these concerns we do not make use of measured interest rates in this paper, for either data source. Instead, we restrict our use of these data sources to transaction times, maturities, and sizes.

present sampling-window estimates of the interbank rate itself, instead we focus on how a sampling window approach would affect the relative sampling noise associated with a transaction-based interbank index.

3.2 Results

Bearing in mind the important caveats described above, we use these two data sources to compute estimates of the relative sampling noise associated with an illustrative US-dollar index rate, for various data filters and maturities.

Figure 1 and Table 1 illustrate the effect of changing the sampling window for the implied *sample-volatility multiplier* V(t), a proportional sampling noise measure that is based on the number and relative sizes of loans in the fixing sample S(t,m;w,d). Specifically, V(t) is the square root of the sum of the squared dollar-size weights of the loans in S(t,m;w;d). For example, if the fixing sample S(t,m;w,d) includes two loans, of amounts $40 million and $60 million, then the relative size weights are 0.4 and 0.6. The sum of the squared weights is 0.16 + 0.36 = 0.52, so V(t) is 0.72.

If one were to assume that, conditional on "fundamental" loan-market information, the individual loan rates in a given day's fixing sample are uncorrelated and have the same standard deviation D(t), then the fixing R(t,m) has a conditional standard deviation of D(t)V(t). Under these conditions, in the above example of a fixing sample with two loans of amounts $40 million and $60 million, the sample volatility multiplier of 0.72 means that the associated size-weighted average interest rate has a standard deviation that is 72% of that for a fixing rate based on a single loan transaction. These statistical assumptions do not apply in practice and we do not rely on them, but the sample-volatility multiplier V(t) nevertheless gives us a good idea of the relative effect of the length of the sampling window on the robustness of the sample. A relatively high sampling volatility multiplier V(t) means that there are relatively few loans dominating the sample, and therefore little opportunity for "diversification" of the sampling noise. At its maximum, for the case of a single sampled loan, V(t) = 1. As the number of loans becomes large and the fraction of any one loan size relative to the total quantity of loans becomes small, V(t) approaches zero, by the law of large numbers. We emphasize that V(t) says nothing about the levels or volatilities of interest rates in the inferred-loan sample. Rather, V(t) is determined entirely by the number and relative sizes of the loans in the fixing sample for date t.

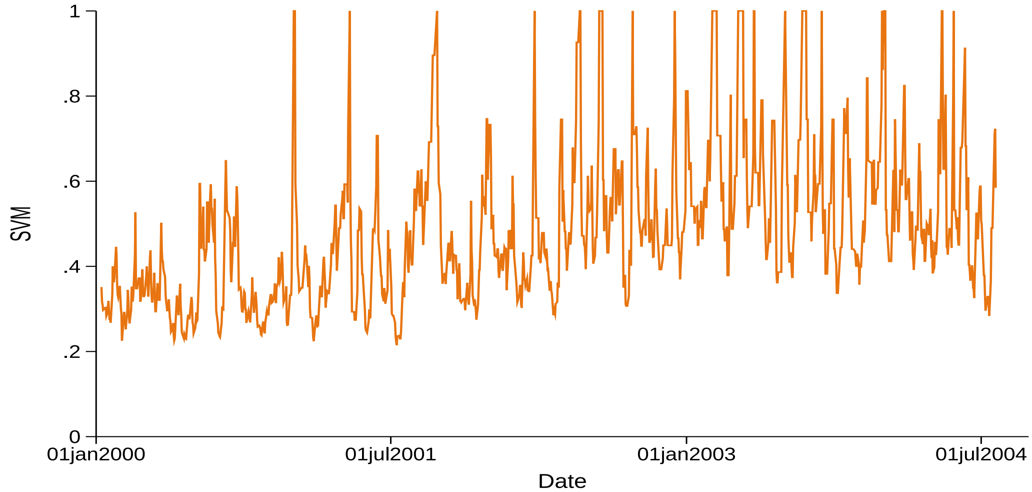
With interest-rate data from actual transactions, one could also directly study the sample standard deviations of the rates in the fixing samples, and the effect of the sampling window on biases and relative noise. Given the potential for misclassification using the KSVY algorithm, we avoid using the inferred loan interest rates.

For the 3-month maturity, Figure 1 below plots the time series of V(t) based on a 10-day sampling window from the two transaction-level data sources.[9]

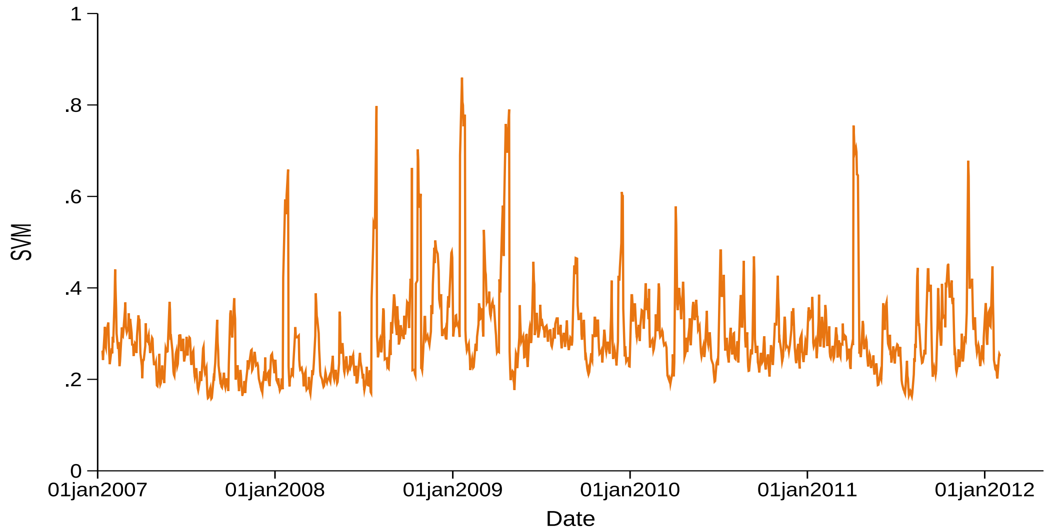
Figure 1: Time-series plot of V(t)

The daily sample volatility multiplier V(t) for 3-month maturity loans. The sample is based on a minimum transaction size of $25m and a 10-day sampling window. The sample period is 2000-2004 for the brokered data, and 2007-2012 for the Fedwire inferences.

A. Brokered interbank data



B. Fedwire inferences



[9] The brokered data sample used to construct Figure 1 as well as subsequent figures and tables includes both Eurodollar and term Federal funds inferred transactions (as discussed in Bartolini et al., 2010, the dataset includes a flag which indicates the transaction type; we retain both categories). Similarly, for the Fedwire inferences, we present results based on the entire dataset of interbank loan inferences, rather than attempting to restrict the sample to a particular loan type.

Figure 1 shows substantial variation over time in the daily sample-volatility multiplier V(t), for both data sources. The sample-volatility multiplier measured from the brokered data is consistently higher than that for Fedwire-inferred data. This is not surprising, given that the brokered data capture only a small segment of the market (those brokered interbank loans intermediated by a single broker). The difference in V(t) between the two data sources could also partially reflect false "matches" in the Fedwire inferences, differences in the sample period, and other factors.

Table 1 shows the median across the period of the sample volatility multiplier V(t), for various maturities and sampling windows lags, normalized by the median of V(t) for 3-month maturity loans and a sample window lag of 10 days. We varied the sampling window from two days to 20 days, and considered maturities of 1, 3, and 6 months. (The normalizing cell associated with a 10-day sampling window and 3-month maturity thus always shows a value of 1.) The table also reports summary statistics from the two data sources.

Table 1: Relative values of V(t) for different maturities and sampling windows

Median values of the sample volatility multiplier V(t), for various combinations of lag window and maturity, normalized by the median value of V(t) for a lag window of 10 days and a maturity of 3 months. The sample period is 2000-2004 for the brokered data, and 2007-2012 for the Fedwire inferences.

Brokered interbank loans

Lag window (days)	Maturity: 1 month	3 months	6 months
2	1.04	1.57	2.22
5	0.81	1.31	1.66
10	0.61	1.00	1.36
15	0.51	0.85	1.17
20	0.46	0.77	1.05

Fedwire inferences

Lag window (days)	Maturity: 1 month	3 months	6 months
2	1.16	1.68	2.50
5	0.88	1.33	2.13
10	0.67	1.00	1.63
15	0.56	0.84	1.37
20	0.50	0.76	1.23

Table 1 shows that in both data sources, the sampling noise as measured by V(t) is significantly greater at longer maturities and for shorter sampling windows. For both data sources, V(t) is two to three times larger for six-month loans than for one-month loans. This is natural in part from the fact that longer-term loans roll over less often than shorter-term loans. (That is, the ratio of the *flow* of loans to the *stock* of loans is lower in steady state for longer-maturity loans.) In any case, our preliminary results suggest caution over whether it would be possible to construct a robust Libor fixing from underlying loan transactions for longer-term loans such as six months.

Table 2 presents summary statistics of the data used to construct the sampling-window Libor index. For both data sources, the average across the sample period of the number of inferred 3-month loan transactions within a 10-day sampling window is low, 8 and 25 transactions respectively for the brokered data and Fedwire inferences. Again, care should be taken in interpreting these statistics given that neither data source is comprehensive.

Table 2: Summary statistics (10 day window, 3 month maturity)

Summary statistics for the estimated sample volatility multiplier V(t), as well as the number of transactions within the 10 day sampling window, and the average transaction size. Sample period is 2000-04 for the brokered data, and 2007-12 for the Fedwire inferences. p10, p25 etc. refers to percentiles of the relevant distribution.

Brokered data

	Mean	p10	p25	p50	p75	p90	StDev
SVM	0.48	0.29	0.35	0.45	0.56	0.71	0.17
# of Transactions in Window	8.13	2	4	7	11	16	5.65
Transaction Size ($mm)	78.69	25	40	50	100	150	59.47

Fedwire inferences

	Mean	p10	p25	p50	p75	p90	StDev
SVM	0.30	0.21	0.24	0.28	0.34	0.41	0.10
# of Transactions in Window	25.45	13	18	24	31	41	10.59
Transaction Size ($mm)	110.81	25	38	54	110	246	213.74

3.3 Alternative specifications

We have experimented with various other data filters. In the appendix, we present two variations. The first considers a minimum transaction size of $100 million, rather than $25 million. Applying this higher size cutoff inevitably reduces the number of eligible transactions at any point in time, and thus raises V(t). One bears in mind, however, that the "root-mean-squared" definition of V(t) implies that a loan of size \$100 million has a relative impact on V(t) that is 16 times that of a \$25 million loan, when both sizes are present in a fixing sample.

Secondly, we have experimented with an approach in which more weight is given to transactions closer to date *t*. See section 4 below for a discussion.

In unreported calculations, we also experimented with expanding the width of the maturity window (by five days in each direction). We found that this has only a small effect on the number of eligible transactions.

4. Some Disadvantages of This Approach, and Their Mitigation

In this section we discuss some important potential disadvantages of a fixing that is based on a sampling-window approach: (i) the effect of using lagged data on the timeliness of the resulting Libor fixing, (ii) the risk of a lack of underlying transactions data, even within a sampling window, and (iii) possible calendar-date effects. We also consider some mitigants of these problems.

A first disadvantage of the sampling-window approach is that the fixing announced on a given day would be based in part on lagged data that may no longer be representative of market conditions. That is, the fixing rate could be somewhat stale during periods of rapid changes in market conditions, for example around the times of significant central-bank monetary policy announcements, or at the onset of a financial crisis or other period in which bank funding costs are shifting rapidly, such as August 9, 2007 and the period following it. The information that market participants and regulators learn from the resulting "Libor" report could therefore be stale.

There is no single "true" interbank borrowing rate, and no sampling method is perfect. One may wish to compare the bias and sampling noise of the sampling-window transactions-based approach that we have described with those of other feasible methods, including the current method for fixing Libor.

For applications involving bond or swap contracts, the staleness introduced by a sampling window measured in days is relatively unimportant. After all, an investor holding a position in swaps or floating-rate notes is concerned with the level of 3-month loan rates that is generally likely to prevail several years into the future, and is probably not so interested in variation in 3-month loan rates within a small time window that begins in several years.

Apart from its role in financial contracting, Libor is also useful for assessing current market conditions. However, even during the recent financial crisis, Kuo, Skeie and Vickery (2012) show that movements in Libor overall commove quite closely with a number of other publicly available indices (such as secondary-market CD rates and Eurodollar yields reported in the Federal Reserve H.15 report). These alternative indices, which would be more sensitive to short-term market shocks, would remain available to policymakers and market participants.

We also note that in terms of revealing information to market participants, a sampling-window fixing approach allows the recovery of most of the "fresh" market information that is present in the underlying data. Given that the difference between the fixing rate on day t and that on the previous day t-1 is caused by dropping observations from date t-w (for a lag window of w) and adding observations from the latest date t, observers can approximately invert the moving-average procedure so as to estimate the implied average rate of transactions that occurred on the latest available date. Of course, it would also be possible to simply release the average transaction rate for each day, as discussed further below.

One could reduce the bias associated with staleness by weighting the data within the fixing sample based on the time lag, using weights that decay with the lag, say exponentially. In order to illustrate the impact on sampling noise of de-weighting stale data, we explored the effect of an exponential decay in transaction weights that gives observations with a 10-day lag only 50% of the weight applied to observations on the current day. (This corresponds to a weight factor of 0.933 raised to the power of the number of days lagging.)

This degree of de-weighting of stale transactions causes a relatively small degradation in sampling noise.[10] For example, for 3-month inferred transactions obtained from Fedwire data for 2007-2012, we saw in Table 2 that the mean sample volatility multiplier is 0.30. With a weight decay factor of 0.933 per day of lag (50% de-weighting of 10-day old observations), the same data are associated with a mean sample volatility multiplier of 0.31, about 3% higher. The estimated effects on sampling noise of de-weighting stale data are similarly muted in all of the cases that we have examined, as demonstrated in additional charts and tables found in the appendices. It is to be cautioned that these results are preliminary and only for illustrative purposes.

In addition to publishing the sampling-window-based fixing rate, one could also publish some properties of the underlying data, such as the daily average rate, the daily number of transactions, or the sample-volatility measure. While financial contracts would presumably be tied to the fixing rate, other published information based on the sample could provide additional useful information and could

[10] In order to gain some intuition for the limited impact of decaying weights on the sample volatility multiplier, consider a relatively adverse case in which the transactions are concentrated at the first and last date of a 10-day sample window. Two equally sized transactions at each end of the 10-day sampling window, without decay, would have a sample volatility multiplier of $V(t)= (0.5^2 +0.5^2)^{0.5} =0.707$. With weights decaying proportionately by a factor of 0.933 per day, or 50% over 10 days, we would have $V(t) =[(0.5/k)^2 + (0.5 \times 0.5/k)^2]^{0.5}$, where $k=0.5+0.25=0.75$, implying $V(t)= 0.74$. So, indeed, even in this relatively extreme situation, the elevation of the sample volatility multiplier V(t) due to decay is only about 5%.

potentially be used in contracting, for example in order to allow financial contracts to be tied to market liquidity or to the quality of the fixing sample.

A second disadvantage of a sampling-window approach is that it is not guaranteed to produce reliable results under all market conditions. If there are too few transactions at a given maturity to provide even a reasonable estimate of major-bank borrowing rates, market participants will nevertheless require a reference rate on which to base the settlement of derivatives and floating-rate loan contracts. For the U.S. dollar market, our results based on a limited data set suggest some hope for the feasibility of transaction-based fixing, using sampling windows, for 1-month and 3-month maturities.

In any case, one may wish to introduce robustness safeguards in the definition of the fixing sample S(t,m;w,d), such as expanding the fixing sample whenever there is insufficient data for a reliable fixing. For instance, one could take the sample window to be a fixed number of days or the minimum number of days necessary to include a given volume of transactions, whichever is greater.[11]

As an alternative to fixing Libor based on unsecured borrowing rates, it has been suggested that Libor might be replaced with a benchmark rate based on secured lending transactions. Prominent among the suggested secured interest rates is "GCF repo," whose market is described by Fleming and Garbade (2003).[12] This approach would introduce several potential complications, however. First, for GCF repo, there remain robustness concerns over whether there is a sufficient volume of GCF repo transactions at the relevant maturities. Second, GCF repo rates are only indirectly connected to banks' unsecured cost of funds, which reduces the usefulness of GCF repo as the basis for an index rate for financial contracting. For commercial banks and bank holding companies, unsecured borrowing is generally a much larger source of overall funding than secured borrowing. Unsecured borrowing is also traditionally the primary source of funding on the margin. (For securities dealers, secured borrowing is a larger source of funding and a more typical marginal source of funding, relative to banks.) Further, Libor-based swaps are heavily used for risk-management and price discovery for the unsecured debt of non-financial corporations. Basing Libor on a secured borrowing rate would reduce its usefulness here as well. Third, using a secured financing rate such as GCF repo raises the

[11] A related concern is that a Libor fixing based on a sampling window approach could become distorted around key calendar dates, such as the end of a quarter or calendar year. Counterparties may for example lengthen or shorten the maturity of otherwise standard contracts to influence whether they cover particular financial statement dates, for window-dressing purposes or for other reasons. This could affect the set of contracts whose maturities lie within a given range (d) around a standard maturity such as one month or three months. In our examples, we set this date range to be constant, but it may be necessary to adjust d in such situations.

[12] The DTCC publishes an average overnight GCF repo rate for three types of collateral: Treasuries, agency MBS, and agency debt. Trading in futures linked to these indices began in July 2012. See https://globalderivatives.nyx.com/nyse-liffe-us/dtcc-gcf-repo-index-futures/settlement-procedures

question of how to treat legacy Libor-based financial contracts, of which there are enormous quantities. A counterparty receiving Libor on a legacy contract would not willingly receive instead the GCF repo rate, which is typically much lower. Replacing “legacy Libor” with an approximation of unsecured rates that are estimated from secured financing rates would likely lead to a substantial amount of contractual dispute.

This also raises the possibility of two parallel markets, at least during a transition period, with “legacy” and “new” benchmarks based on unsecured and secured (repo) rates, respectively. The associated transition would be awkward and lengthy, and involve splitting liquidity across the two markets with an attendant loss in market efficiency. In any case, a sampling-window approach could also be used for term repo rates, provided there are sufficient data.

The Wheatley Report (H.M. Treasury, 2012) reviews other alternative approaches and benchmarks, such as the overnight index swap rate (OIS), and provides a description of their advantages and disadvantages.

References

Armantier, Olivier and Adam Copeland (2012). "Assessing the Quality of "Furfine-based" Algorithms," Federal Reserve Bank of New York Staff Reports, No. 575 October.

Bartolini, Leonardo, Spence Hilton and James J. McAndrews (2010). "Settlement Delays in the Money Market", *Journal of Banking and Finance* 34, 934-945.

Carhart, Mark, Ron Kaniel, David Musto, and Adam Reed (2002) "Leaning for the Tape: Evidence of Gaming Behavior in Equity Mutual Funds," *Journal of Finance* 57, 661-693.

Fleming, Michael and Kenneth Garbade (2003). "The Repurchase Agreement Refined: GCF Repo," Federal Reserve Bank of New York *Current Issues in Economics and Finance*, 9 (June).

Furfine, Craig (1999) "The Microstructure of the Federal Funds Market*," Financial Markets, Institutions & Instruments* 8, 24-44.

Kuo, Dennis, David Skeie, James Vickery (2012), "A Comparison of Libor to Other Measures of Bank Borrowing Costs," Working Paper, Federal Reserve Bank of New York.

Kuo, Dennis, David Skeie, James Vickery, and Thomas Youle (2012), "Identifying Term Interbank Loans from Fedwire Payments Data," Working Paper, Federal Reserve Bank of New York.

H.M. Treasury (2012). "The Wheatley Review of Libor: Final Report," H.M. Treasury, London, September, 2012.

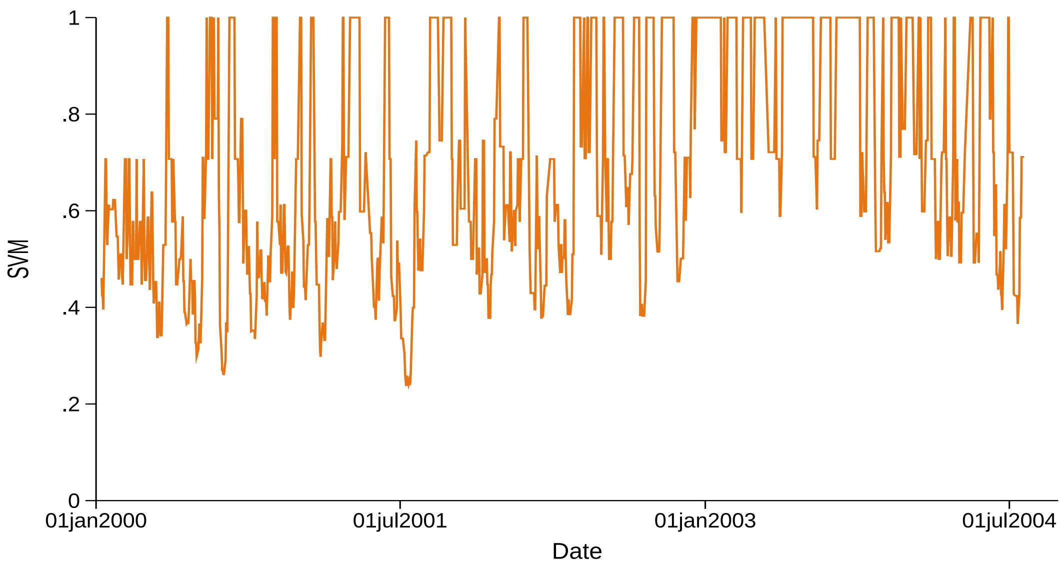
Appendix: Other Data Filters

Appendix A1. Minimum transaction size of $100m (rather than $25m)

The statistics shown here are computed for the same data as those underlying Figure 1 and Table 1, with the exception that the transactions sizes have a minimum of $100m, rather than a minimum of $25m.

Figure A1. Time-series plot of V(t)

i. Brokered data



ii. Fedwire inferences

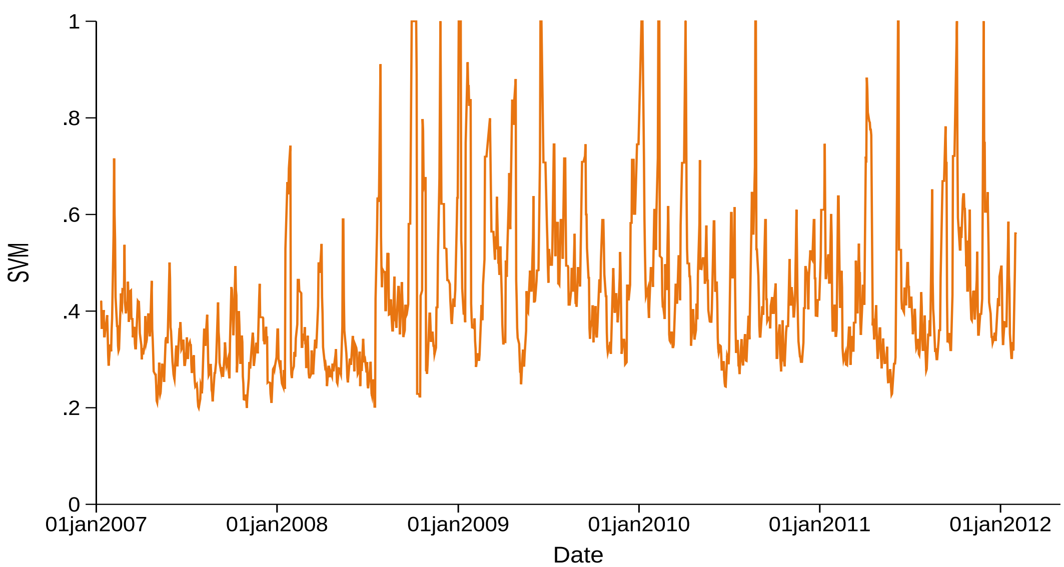


Table A1. Statistics for V(t)

Median across the sample period of the sample volatility multiplier V(t) for the maturity and sampling window length shown, normalized by the median of V(t) for a sampling window of 10 days and maturity of 3 months.

i. Brokered data

		Maturity		
		1 month	3 months	6 months
Lag window (days)	2	1.00	1.41	1.41
	5	0.79	1.41	1.41
	10	0.54	1.00	1.41
	15	0.44	0.78	1.41
	20	0.38	0.70	1.41

ii. Fedwire inferences

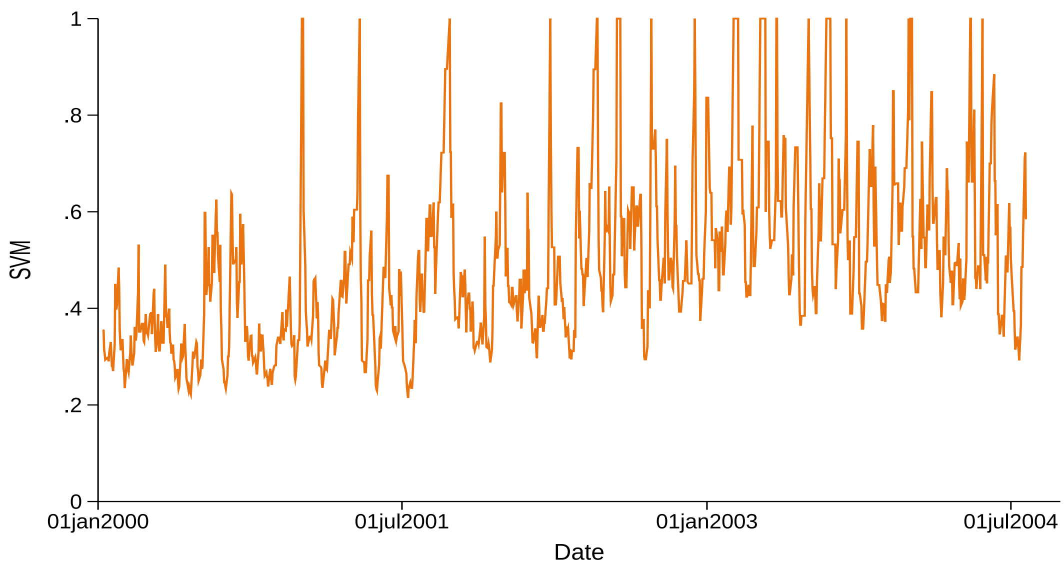
		Maturity		
		1 month	3 months	6 months
Lag window (days)	2	1.22	1.73	2.42
	5	0.92	1.42	2.42
	10	0.66	1.00	1.71
	15	0.54	0.83	1.54
	20	0.48	0.74	1.43

Appendix A2. Using exponential decay

The statistics shown in Figure A2 and Table A2 are calculated using the same samples as those of Figure 1 and Table 1, except that we incorporate exponential decay over the sampling window.

Figure A2. Time-series plot of V(t)

(i) Brokered data



(ii) Fedwire inferences

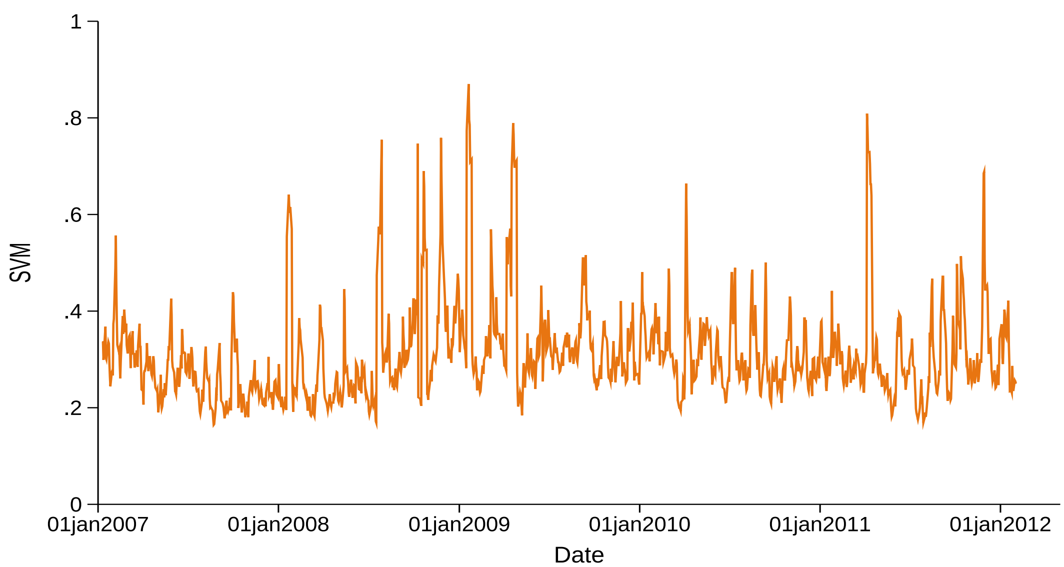


Table A2. Statistics for V(t)

Median for the period of the sample volatility multiplier V(t) for the indicated maturity and sampling window length shown, normalized by the median for a sampling window of 10 days and maturity of 3 months.

(i) Brokered data

		Maturity		
		1 month	3 months	6 months
Lag window (days)	2	1.03	1.55	2.19
	5	0.80	1.28	1.63
	10	0.61	1.00	1.35
	15	0.52	0.86	1.18
	20	0.47	0.78	1.09

(ii) Fedwire inferences

		Maturity		
		1 month	3 months	6 months
Lag window (days)	2	1.15	1.67	2.48
	5	0.88	1.31	2.10
	10	0.67	1.00	1.63
	15	0.57	0.86	1.38
	20	0.53	0.79	1.27