Olivier Armantier and Adam Copeland

Challenges in Identifying Interbank Loans

- Empirical analyses of the federal funds market often use the so-called "Furfine algorithm" to identify activity in the market at the most disaggregated level—individual loans between two specific banks.
- However, a formal test of the accuracy of the algorithm in identifying fed funds transactions shows that the algorithm may be ill-suited to this task.
- Given access to the identifiers used by two large banks to denote fed funds payments, the authors are able to compare a set of payments known to be fed funds transactions with the set of payments pegged as such by the algorithm.
- The authors find that for the 2007-11 period, an average of 81 percent of all pairs of payments identified by the algorithm are not, in fact, fed funds transactions conducted by the two banks, while an average of 23 percent of the banks' actual fed funds transactions are overlooked by the algorithm.

1. Introduction

The U.S. federal funds (fed funds) market is an interbank market for unsecured, mostly overnight loans of reserves held by banks at Federal Reserve Banks. It is an over-the-counter market where banks arrange trades either on their own on a bilateral basis or through brokers. Historically, the fed funds market has been a key financial market with major macroeconomic and monetary policy implications. In particular, the average fed funds market rate, known as the effective fed funds rate, has substantial influence on the terms at which commercial banks lend to businesses and individuals. Furthermore, the Federal Reserve implements monetary policy by creating conditions under which fed funds trade around a specific target or within a target range set by the Federal Open Market Committee (FOMC).[1]

The traditional source of data on the fed funds market is based on fed funds trades reported by the major fed funds brokers to the Federal Reserve Bank of New York (FRBNY). Using these data, various market-level interest rate statistics are calculated and published daily by the FRBNY. These

[1] Although other forms of short-term interbank lending may be informally referred to as "fed funds," we are solely concerned in this article with loans of reserves between eligible counterparties as officially defined as fed funds ▸

Olivier Armantier is an assistant vice president and Adam Copeland a research officer at the Federal Reserve Bank of New York.

Correspondence: adam.copeland@ny.frb.org

Note: For its analysis of interbank lending markets in the conduct of monetary policy, the Federal Reserve Bank of New York relies on different sources of data, not on an algorithm's output. Consequently, our results have no bearing on the Federal Reserve Bank of New York's operational understanding of interbank lending markets and its calculation of market-level measures, including the effective federal funds rate.

The authors thank Gara Afonso, Marco Cipriani, Todd Keister, Anna Kovner, Antoine Martin, Jamie McAndrews, David Skeie, and James Vickery for their help on this project, as well as Isaac Davis, Sha Lu, and Michael Walker for valuable research assistance. They also thank seminar participants at the Federal Reserve Bank of New York. The views expressed in this article are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

statistics—in particular, the effective fed funds rate—are used widely by policymakers, financial market participants, and researchers in academia.

An alternative source of data, used exclusively to conduct academic research, is inferred from algorithms based on the original work of Furfine (1999). Although there are now different versions of the original algorithm, they all seem to rely on the same principles. A number of recent empirical papers use a version of the Furfine algorithm's output to make important contributions. These papers assume, but do not formally test, the accuracy of the output of their algorithms. As we explain in more detail below, the main purpose of this article is to formally test these Furfine-based algorithms. Importantly, the results presented here do not extend to the traditional source of data collected by the FRBNY from the fed funds brokers. In particular, the results have no bearing on the ability of the Markets Group of the FRBNY to understand the fed funds market and to accurately calculate market-level measures, including the effective fed funds rate.

In this study, we focus on the revised Furfine algorithm used by the Research Group of the FRBNY. This algorithm exploits the fact that privately traded fed funds transactions are often settled over the Fedwire® Funds Service (Fedwire), the large-value real-time payments settlement system operated by the Federal Reserve.[2] As further explained in section 2, the algorithm searches all payments sent over Fedwire to identify the pairs of payments that look like fed funds loans. Specifically, the algorithm tries to identify first a "sent" payment from bank A to bank B on a given date for an amount that could reasonably constitute a loan principal, and then a "return" payment from bank B to bank A on the following day for an amount that could reasonably constitute the principal plus interest payment.

If the algorithm correctly identifies fed funds transactions with sufficient accuracy, then its output could be useful to academic economists in studying the fed funds market. Indeed, it would provide data at the lowest level of aggregation (that is, individual transactions between specific pairs of banks) that could help shed light on the underpinnings of the U.S. fed funds market. The algorithm's output is especially attractive when trying to explain the behavior of the market during the 2008-09 financial crisis, as well as the specific role played by individual banks. Indeed, there has been a surge in the number of papers that use the algorithm's output (we found eleven papers written in the past two years; all are listed in the References section of this article).[3]

An important question remains, however: To what extent does the algorithm identify individual fed funds transactions? Indeed, nothing guarantees that a pair of payments between two banks labeled by the algorithm as a fed funds transaction is indeed a fed funds loan between those two banks. In 2009, we started to test the algorithm's output. In this article, we report the outcome of a formal test assessing the ability of the algorithm to identify individual overnight fed funds transactions.

The basic methodology underlying the test, discussed more fully in section 3, may be summarized as follows. From the flow of payments a bank receives over Fedwire, its back office needs to be able to identify those corresponding to the fed funds transactions initiated by the front office. While back offices use a variety of strategies, at least two banks require their fed funds counterparties to incorporate a unique identifier into the message portion of the Fedwire payment. These two institutions, which are among the biggest banks and account for a large fraction of transactions in the fed funds market, gave us access to their unique identifier. As a result, we can flag every fed funds payment these two banks receive through Fedwire on a given day. To assess the quality of the algorithm, we can then compare the set of payments constructed with the unique identifiers to the set of transactions identified for these two banks by the algorithm. Our identification method rests on the hypothesis that the unique identifiers provided by the two banks are included in every fed funds transaction they settle over Fedwire. At the end of section 3.1, we present evidence supporting this hypothesis.

The outcome of the test is discouraging: In the first quarter of 2007, we estimate that 64 percent of all pairs of payments identified by the algorithm are not fed funds transactions conducted by the two banks (type I error), while 24 percent of the

Footnote 1 (continued)
by the Board of Governors of the Federal Reserve System in Regulation D (see http://www.federalreserve.gov/bankinforeg/reglisting.htm#D). See the FedPoint document at http://www. newyorkfed.org/aboutthefed/fedpoint/fed15.html for a concise definition of fed funds. Examples of papers considering similar definitions of fed funds are Hamilton (1996, 1997), Demiralp, Preslopsky, and Whitesell (2006), Afonso, Kovner, and Schoar (2011), and Afonso and Lagos (2012a, 2012b).

[2] Fed funds transactions can be settled over Fedwire, possibly settled over CHIPS (Clearing House Interbank Payments System, another high-value payments settlement system), or conducted on a bank's books. However, on the basis of conversations with industry participants, Bartolini, Hilton, and McAndrews (2008) report that fed funds loans settle almost exclusively over Fedwire as opposed to other payment services. Still, to the best of our knowledge, the exact extent to which Fed funds are primarily settled over Fedwire has not been established formally.

[3] The following papers use a version of the Furfine algorithm to varying degrees (although the main results may not depend on the algorithm's output): Ashcraft and Bleakley (2006), Ashcraft and Duffie (2007), Atalay and Bech (2010), Acharya and Skeie (2011), Ashcraft, McAndrews, and Skeie (2011), Bech et al. (2011), Afonso, Kovner, and Schoar (2011, 2013), Afonso and Lagos (2012a, 2012b), and Armantier et al. (2011). We are not implying that these authors did anything improper. Specific concerns about the algorithm only emerged recently. Furthermore, some of these papers explicitly discuss the potential problems with the algorithm.

fed funds transactions actually conducted by the two banks are not identified by the algorithm (type II error). This negative result seems to be robust with respect to the time period considered. If we go forward to the first quarter of 2011, the type I error is estimated to be 93 percent, while the type II error is estimated to be 17 percent. Although our results may not extend to every bank, we argue that they apply to the majority of the algorithm's output for at least two reasons. First, the two banks that provided their unique identifier are either senders or receivers for about three-tenths of all pairs of payments output by the algorithm over the 2007-11 period. Second, if we assume that the estimates of type I and type II errors generalize to other large banks with similar Fedwire activity, then our estimates apply to almost half of all pairs of transactions output by the algorithm. Consequently, we conclude that there is substantial doubt about the ability of the algorithm to produce transaction-level measures that characterize accurately and comprehensively the fed funds market.

The algorithm has an additional, perhaps insurmountable, problem: Even if it could correctly find every fed funds transaction, there is no guarantee that it correctly identifies the ultimate originator and beneficiary of a payment. Indeed, while Fedwire data list which bank is sending the payment over Fedwire, it is not at all clear whether that bank or one of that bank's correspondents is originating the payment. Similarly, the algorithm cannot guarantee the identity of the ultimate beneficiary of the payment. Although we are unaware of the exact extent of this problem, conversations with market participants suggest that having cash accounts at other (typically large) banks is not uncommon.[4] Not being able to identify with certainty the true counterparties of a Fedwire payment poses a fundamental challenge to constructing transaction-level or even bank-level estimates of fed funds activity.[5]

These negative results cast doubt on the robustness of empirical work that uses the output of Furfine-based algorithms at the transaction level. Our findings strongly suggest that, going forward, a better understanding of the federal funds market at a disaggregate level depends upon finding data, or improving (and validating) the Furfine algorithm, rather than using the current algorithm's output. Alternatively, researchers may want to forgo the lure of disaggregate measures of fed funds activity and use the transaction-based market-level statistics published by the FRBNY, which are based on data from fed funds brokers (for example, see Hamilton [1996, 1997]).

While our work focuses on the revised Furfine algorithm used by the Research Group of the FRBNY, slightly different versions of this algorithm are used by researchers outside the Bank. Demiralp, Preslopsky, and Whitesell (2006) and Bech, Klee, and Stebunovs (2012) use the same proprietary Fedwire data and a similar algorithm to create measures of overnight fed funds activity. We therefore expect their algorithm's output to suffer from the same problems we highlight in this study. Beyond fed funds, researchers have used algorithms based on Furfine (1999) to construct estimates of unsecured interbank lending. For instance, Kuo, Skeie, and Vickery (2012) have expanded the algorithm to identify loans with maturities longer than overnight. In addition, similar algorithms have been applied to Canadian and European data to identify overnight loans.[6] In particular, using data from TARGET2 (a large-value payments system for European banks), Arciero et al. (2013) conduct a test suggesting that their algorithm produces substantial type I errors but virtually no type II errors.

The test we conducted only demonstrates the inability of the algorithm to identify correctly individual overnight fed funds transactions conducted by two specific banks. Although we believe our results extend more generally, it is possible that the algorithm performs better for some specific types of banks. It is also possible that when the output of the algorithm is aggregated to the bank-to-bank level (that is, all transactions conducted between two banks), to the bank level (that is, all transactions conducted by a bank), or to the market level, it produces useful summary statistics to analyze the fed funds market. In the conclusion, we argue that our negative test results apply at the transaction, bank-to-bank, and bank levels, and we identify conditions under which the algorithm could be considered to produce accurate statistics at the market level. Finally, we discuss in the conclusion the possibility that, beyond fed funds, the algorithm output captures more general overnight interbank loans. Although we provide some evidence to support this hypothesis, we ultimately conclude that the algorithm cannot systematically recognize that a given pair of payments corresponds to an overnight interbank loan between two specific banks. In any case, the hypothesis that the algorithm's output captures overnight interbank loans would need to be formally tested in order to be validated. Until then, researchers and policymakers should be reluctant to use the algorithm's output as a proxy for interbank lending.

The remainder of the article is structured as follows. In section 2, we describe the algorithm and discuss its potential

[4] For example, foreign banks often have nostro accounts at domestic banks.

[5] Several of the papers mentioned in footnote 3 discuss this issue (such as Ashcraft and Bleakley [2006] and Afonso, Kovner, and Schoar [2011, 2013]). See also Furfine (1999).

[6] Hendry and Kamhi (2009), Allen et al. (2012), and Allen, Kastl, and Hortacsu (2012) make use of a similar algorithm applied to Canadian payments data. Millard and Polenghi (2004) and Acharya and Merrouche (2011) make use of a similar algorithm applied to U.K. payments data.

problems as a tool for identifying individual fed funds transactions. In section 3, we present the methodology underlying our test and report the outcome of the test. We conclude in section 4 with a discussion of our results' implications.

2. The Algorithm

2.1 Background

Fedwire is a real-time gross settlement system operated by the Federal Reserve. It enables depository institutions and other financial institutions to make large-value payments that are immediate and final.[7] To initiate a transfer through Fedwire, a participant must populate a number of fields in an electronic form specifying in particular the identity of the sending and receiving parties and the amount sent.

While data from Fedwire are not publicly available, some researchers within the Federal Reserve System have access to the transaction-level payments data. As part of this group of researchers, we can observe the universe of payments sent over Fedwire on any given day. However, we are only allowed to observe a subset of the message fields. Specifically, we observe the American Bankers Association (ABA) number of the sending and receiving banks, the amount sent, the time the payment was sent and received, a payment type code, and a payment business code. These last two fields give the bank sending the payment the opportunity to characterize the nature of the payment. Unfortunately, there are no industry-wide standards regarding the use of the payment type and business code fields. Consequently, the content of these two fields is not sufficient to determine unambiguously the nature of the payment sent.

To infer overnight fed funds transactions settled over Fedwire, Furfine (1999) proposed an algorithm that has been slightly adapted over the years by researchers at the FRBNY and the Federal Reserve Board. The current algorithm used by the FRBNY to produce some of its reports follows these general steps:

1. Transfers from or to a settlement institution (that is, CHIPS, CLS, or the Depository Trust Company) are dropped because loans to or from these institutions are not considered fed funds loans as defined by Regulation D.
2. On a given business day t, the algorithm considers every pair of banks $\{i,j\}$. Then, it constructs the set of possible send payments X_{ijt} consisting of all the transfers x_{ijt} from bank i to bank j on day t that are both greater than or equal to \$1 million and in increments of \$100,000. Each payment x_{ijt} in X_{ijt} is therefore considered to constitute the principal on a possible fed funds loan from bank i to bank j on day t.
3. For each payment x_{ijt} in the set X_{ijt} , the algorithm now constructs the set $Y(x_{ijt})$ of possible return payments the next business day (t+1). Specifically, every payment y_{jit+1} from bank j to bank i on day t+1 is evaluated to determine whether it could represent the principal x_{ijt} plus a plausible interest payment. To make this determination, the algorithm calculates the (annualized) interest rate implied by the pair of payments x_{ijt} and y_{jit+1}.[8] This implied interest rate is then compared with the range $[\underline{i}, \bar{i}]$, where $\underline{i}$ (respectively, $\bar{i}$) is the minimum (respectively, maximum) fed funds rate published by the FRBNY at date t minus (respectively, plus) 50 basis points.[9] If the implied interest rate is within the range $[\underline{i}, \bar{i}]$, then y_{jit+1} is included in the set $Y(x_{ijt})$ of possible return payments for x_{ijt}. Otherwise, y_{jit+1} is not considered a possible return payment for x_{ijt}.
4. Next, the algorithm determines the most likely return payment for each payment x_{ijt} in X_{ijt}. Three scenarios are possible. First, if there are no candidate return payments (that is, $Y(x_{ijt}) = \varnothing$), then x_{ijt} is not considered part of an overnight loan. Second, if there is a unique matching return payment (that is, $Y(x_{ijt})$ is a singleton), then x_{ijt} and the unique y_{jit+1} in $Y(x_{ijt})$ are linked and said to be an overnight loan. Third, if there are multiple candidate return payments (that is, $\dim[Y(x_{ijt})] > 1$), then the algorithm first computes the median interest rate implied by all the candidate payments in $Y(x_{ijt})$. The algorithm then chooses the return leg of the overnight loan with an implied interest rate that is closest to the median rate from above.[10] If linked to a send payment x_{ijt}, a return payment y_{jit+1} is then removed from consideration as a candidate match for all remaining send payments x'_{ijt} in X_{ijt}.[11]

[7] See Armantier, Arnold, and McAndrews (2008) for further details on Fedwire operations.

[8] This interest rate is equal to $((y_{jit+1}-x_{ijt})/x_{ijt})*(360/n)$, where n is the number of calendar days between business day t and t+1, while 360 is used to annualize an overnight loan, per convention in the fed funds market.

[9] Every day, the FRBNY conducts a survey of the four largest fed funds brokers. As mentioned in the introduction, the FRBNY uses this source of data to publish the mean, standard deviation, minimum, and maximum interest rates of brokered fed funds transactions for the prior day.

Currently, the minimum bound on an interest rate is the maximum of 0.9 basis point and the minimum fed funds rate reported by the FRBNY (using the data collected from fed funds brokers) minus 50 basis points. In the past, the minimum bound was the maximum of 1/32 and the minimum fed funds rate reported by the FRBNY minus 50 basis points. The absolute lower bound was pushed down from 1/32 to 0.009 percent because the extremely low nominal rates in recent times made interest rates below 1/32 plausible.

[10] In the case of ties, the algorithm chooses a return leg randomly among those with an implied interest rate closest to the median rate from above.

[11] The algorithm's output may differ depending upon the ordering of the x_{ijt} in the set X_{ijt}, because a matched return payment y_{jit+1} is removed from consideration, without replacement, as a candidate match for all remaining send payments x'_{ijt}. We have not yet studied how changes in the ordering of payments affect the algorithm's output.

5. Finally, the algorithm determines whether the overnight loans identified should be considered fed funds or Eurodollars. If the send leg on the pair of transactions has been given a "CTR" business code, then the pair of transactions is deemed an overnight Eurodollars loan.[12] Otherwise, the pair of transactions is classified as an overnight fed funds loan.[13]

At the end of these steps, the algorithm's final output consists of a series of paired Fedwire payments labeled as fed funds loans. To get a sense of the amount of filtering done by the algorithm, the algorithm identified slightly more than 0.7 percent of the 493,000 Fedwire payments sent on an average day in the first quarter of 2011 as being a leg of a fed funds loan.

If the algorithm is perfectly accurate, then the pairs of payments identified should capture the entire population of individual overnight fed funds loans settled over Fedwire that day. The algorithm therefore produces data at the most granular level, that is, individual loans between two specific banks. From each pair of payments, several characteristics may be inferred, such as the loan's interest rate, duration, or time of repayments. While the algorithm's output has a variety of uses, FRBNY researchers have used it to calculate summary statistics that describe features of the fed funds markets (for example, average rates and volumes) at the bank-to-bank, bank, and market levels.

2.2 Potential Problems

The algorithm described above produces pairs of payments that are labeled overnight fed funds loans. Here, we describe the potential mistakes the algorithm may make that would generate false positives and false negatives.[14]

False positives are pairs of payments that are incorrectly categorized as fed funds activity between the two specific banks sending and receiving the payments over Fedwire. Beyond the obvious case of two completely random payments incorrectly paired by the algorithm, we can suggest four general reasons why the algorithm could generate false positives.

First, the pair of transactions could be a fed funds loan, but not between the two banks sending the payments over Fedwire. As noted in the introduction, the algorithm cannot distinguish between a bank sending or receiving a payment on its own behalf and a bank doing so on behalf of a correspondent. In such a case, the algorithm would have identified a legitimate fed funds loan, but attributed it to the incorrect bank(s). This type of misassignment of counterparties will not affect aggregate market-level analysis, but it may bias estimates of fed funds activity at the transaction, bank-to-bank, or bank level. While we know this type of correspondent banking activity does occur, we do not know how often it occurs and how large a share of total fed funds activity it represents.

Second, the pair of transactions could be an overnight unsecured loan different from a fed funds transaction as defined under Regulation D. Observe that these types of loans may not exclusively capture interbank lending. In particular, the algorithm could pick up loans conducted on behalf of wealth-management funds, hedge funds, or even firms outside the financial sector.

Third, the pair of payments could be related to a collateralized loan. For the vast majority of collateralized loans, the cash portion is not sent over Fedwire. There is potentially a concern, however, with tri-party repo transactions.[15] While the cash portion of these repo transactions typically moves around on the books of the clearing banks, there are cases when the cash portion of a tri-party repo transaction is sent and returned between the cash investor and the clearing bank over Fedwire. This payment activity could be picked up in the algorithm and incorrectly labeled as a fed funds transaction.

Fourth, the algorithm could identify a legitimate fed funds loan, but incorrectly link one of the two payments related to that transaction. Such an error may occur when the algorithm finds multiple candidates for one of the legs of the transaction. Instead of picking the payment corresponding to the actual fed funds transaction, the algorithm incorrectly selects an unrelated but similar payment. In most cases, this mismatch might not severely bias the most important characteristics of the fed funds transaction (that is, interest rate, amount), but it could affect other characteristics, such as the timing of transactions.

False negatives are actual overnight fed funds loans settled through Fedwire that are not identified by the algorithm. The constraints embedded in the algorithm could produce such errors in at least two ways: First, the algorithm requires the principal amount of fed funds loans to be greater than or equal to $1 million and in increments of $100,000. Actual fed funds activity in which the principal is less than $1 million or is not in an increment of $100,000 will be missed by the algorithm. Second, if there is considerable variability in the fed funds rates

[12] "CTR" stands for customer transfer, and is meant to designate that the beneficiary of the payment is not a bank.

[13] The motivation for using the CTR business code to differentiate fed funds loans from Eurodollars loans is based on internal work at the FRBNY. The classification, however, may include errors because the use of the CTR code by banks is neither mandatory nor an explicit industry standard.

[14] Some of the potential mistakes listed in this section have been previously discussed in, for example, Furfine (1999).

[15] See Copeland, Martin, and Walker (2010) for a description of the tri-party repo market.

across banks, the plus or minus 50 basis point range around the minimum and maximum fed funds rate published by the FRBNY might rule out actual fed funds activity.

In addition to the systematic problems that may arise with the algorithm, idiosyncratic difficulties exist. A bank, for example, may return the principal and interest associated with a fed funds loan in two separate payments. Likewise, fails can occur when a bank, perhaps because of operational difficulties, does not return the principal and interest the next day. According to a handful of industry participants, these events rarely occur. When they do occur, however, the algorithm will not identify the underlying fed funds activity. Finally, the objective of the algorithm is to identify fed funds activity settled through Fedwire. As a result, the algorithm cannot provide any information about fed funds loans settled outside Fedwire—for example, over other payment systems or on a bank's books.

3. Testing the Quality of the Algorithm

3.1 The Test's Methodology

From the perspective of a given bank, each of its fed funds transactions consists of two legs: a "send leg," in which the money flows from the bank to its counterparty, and a "receive leg," in which the money flows from the counterparty to the bank. When a bank sells fed funds, the send leg precedes the receive leg; when the bank purchases fed funds from a counterparty, the receive leg precedes the send leg. The perspective of the bank's counterparty is the mirror image—that is, the send leg for a bank that sells fed funds is the receive leg for the counterparty that purchases the fed funds.

Every day, banks may send and receive a large number of payments over Fedwire (more than 150,000 in some cases), a tiny portion of which correspond to fed funds transactions (typically less than 0.1 percent). Because banks must keep track in real time of every fed funds transaction they conduct, they have to be able to flag automatically a fed funds transaction from within the flow of Fedwire payments they receive. To do so, large banks typically require their fed funds counterparties to incorporate an identifier into the message portion of the Fedwire payment. Two of these banks voluntarily gave us access to their unique identifiers. Using these identifiers, we can locate the receive leg of every fed funds transaction the two banks have conducted by searching for the unique identifier within the message fields of all Fedwire payments they receive.[16] Unfortunately, we do not have access to the unique identifiers for the two banks' counterparties (except, of course, when these two banks interact with each other). Thus, we can identify only the receive legs but not the send legs of the fed funds transactions conducted by the two banks. Consequently, we do not know for sure the true interest rate associated with a receive leg of a fed funds transaction, because it takes both legs to infer unambiguously the interest rate of a fed funds loan. Although this limitation has no impact on our estimates of type I and type II errors, we will need to keep it in mind when studying the interest rates produced by the algorithm.

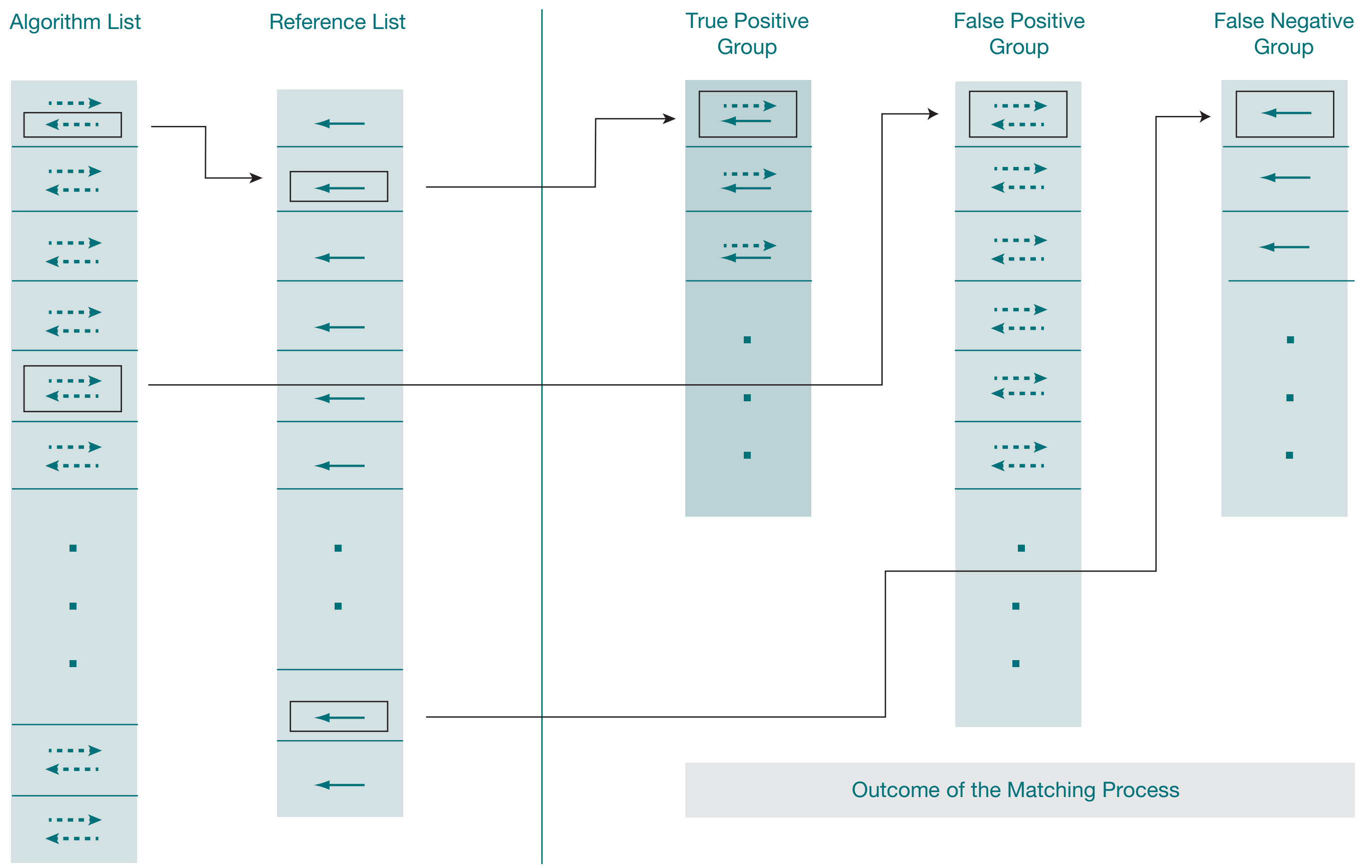
Our goal is to establish how well the algorithm identifies overnight fed funds transactions conducted by the two banks over Fedwire. To do so, we consider all possible pairs of payments $\{x_{ijt}, y_{jit+1}\}$ on consecutive business days between bank i and bank j, where bank i or j is one of the two banks for which we have a unique identifier. The null hypothesis is that $\{x_{ijt}, y_{jit+1}\}$ is not a fed funds loan, while the alternative hypothesis is that $\{x_{ijt}, y_{jit+1}\}$ is a fed funds loan. The algorithm can be seen as a test of the null hypothesis because it provides a method to decide which $\{x_{ijt}, y_{jit+1}\}$ should or should not be considered a fed funds loan. Because the presence of the unique identifier flags unambiguously which receive legs are, and which receive legs are not, part of a fed funds loan for our two banks, we can estimate when the algorithm incorrectly rejects the null hypothesis (type I error) and when the algorithm incorrectly accepts the null hypothesis (type II error). The method we use to construct these estimates consists of three steps (see the exhibit).

First, we run the algorithm for the two banks for every business day within a quarter. This gives us a list of paired payments, each consisting of a send leg and a receive leg. We call this the "algorithm list." Second, we construct another list of payments (the "reference list") by searching for the unique identifier over all the Fedwire payments the two banks received on every business day within the quarter. This reference list therefore consists of receive legs identifying all fed funds payments the banks received that quarter. Third, we compare the algorithm and reference lists, searching for matches. Specifically, we verify whether each of the receive legs in the reference list can be found in the algorithm list.

As illustrated in the exhibit, this matching process produces three different groups. The "true positive group" consists of every pair of payments in the algorithm list with a match in the reference list. The "false positive group" consists

[16] To be clear, the unique identifier is included in the receive leg of every fed-funds-related transaction conducted by the two banks, regardless of whether the two banks purchased or sold fed funds in that transaction.

The Test's Methodology



Notes: The "algorithm list" consists of all pairs of payments identified by the algorithm as federal funds loans. The "reference list" consists of all Fedwire payments with the unique identifier. The "true positive group" consists of every pair of payments in the algorithm list with a match in the reference list. The "false positive group" consists of every pair of payments in the algorithm list without a match in the reference list. The "false negative group" consists of every receive leg in the reference list without a match in the algorithm list. A dashed line indicates a send or a receive leg of a federal funds transaction identified by the algorithm. A solid line indicates a receive leg of a fed funds transaction with the unique identifier.

of every pair in the algorithm list without a match in the reference list. Finally, the "false negative group" consists of the receive legs in the reference list without a match in the algorithm list. The size of the false positive group relative to the size of the algorithm list gives us an estimate of the algorithm's type I error for the two banks. Similarly, the size of the false negative group relative to the size of the reference list gives us an estimate of the type II error for the two banks.[17]

[17] Technically, the type I error rate is the probability of the receiving leg not being part of a fed funds loan conditional on the algorithm labeling the receiving leg as part of a fed funds loan. The type II error rate is the probability of the algorithm not labeling the receiving leg as part of a fed funds loan conditional on the receiving leg being part of a fed funds loan.

This methodology, in fact, provides only a lower bound on the extent of type I errors for at least two reasons. First, we can test whether the algorithm correctly identifies the receive leg of a fed funds transaction but, because of the possibility of correspondent banking, we cannot confirm that the bank that sent the Fedwire payment is indeed the counterparty in the fed funds transaction. Second, a pair of payments is in the true positive group if it possesses the receive leg of an actual fed funds transaction. This does not imply, however, that the algorithm correctly identified the send leg of that fed funds transaction. As mentioned earlier, our methodology does not allow us to test this hypothesis. The consequences of such mismatches, however, should not be expected to be too severe. Although the mismatches may seriously affect some characteristics of the fed funds transactions (for example, the exact duration of the loan), in general they should not substantially affect the more important characteristics (that is, the amount loaned and the interest rate inferred). Indeed, by construction, the algorithm can only match an incorrect send leg to the receive leg of an actual fed funds transaction if the amount of this incorrect send leg is similar to the amount of a true send leg. As a result, we expect the interest rates inferred for the pairs of payments in the true positive group to be reasonably accurate.

In contrast, our methodology provides an upper bound on the extent of type II errors. Indeed, the two banks under consideration ask their counterparties to include the unique identifier for payments corresponding to any fed funds transactions, which include overnight as well as term fed funds transactions. As a result, some of the fed funds payments in the false negative group may not correspond to overnight loans, and our test's methodology may therefore exaggerate the extent of type II errors. Although we cannot quantify precisely the extent of this problem, conversations with fed funds traders at each of the two banks suggest that the number of term fed funds transactions they conduct is relatively small.

To conclude this section, we want to acknowledge that the validity of our test hinges on the fact that the unique identifiers provided by the two banks are included in every fed funds transaction they settle over Fedwire. Note that the validity of the unique identifiers has been confirmed at various points in time by different members of the two banks in question. Further, we were able to find independent evidence from a third, unrelated bank. Indeed, this third bank confirmed that a necessary condition to remain a fed funds counterparty to the two banks on which we base our test is that every fed funds payment sent over Fedwire must include the unique identifiers.[18]

3.2 Type I and Type II Errors

The results reported in Table 1 are discouraging. In the first quarter of 2007, the type I error produced by the algorithm is estimated to be 64 percent (18,633/29,077). While much lower, the estimated type II error, at 24 percent

Table 1

Estimates of Type I and Type II Errors for 2007:Q1

Algorithm List		Reference List
29,077		13,655
False Positive Group	True Positive Group	False Negative Group
18,633	10,444	3,211
Type I error: 64 percent		Type II error: 24 percent

Source: Authors' calculations, based on Fedwire data.

Note: The type I error is equal to the false positive group divided by the algorithm list; the type II error is equal to the false negative group divided by the reference list.

(3,211/13,655), is not inconsequential. To measure how well the algorithm performed through the recent financial crisis, we estimated the type I and type II errors for these two banks for the first quarters of each year between 2007 and 2011 (see Table 2).[19] The type I error is estimated to be higher as we go forward in time, reaching 93 percent in the first quarters of 2010 and 2011. Conversely, the type II error is estimated to be lower as we go forward in time, slightly declining to 17 percent in the first quarter of 2011.[20] On average, the type I error is estimated to be 81.4 percent from 2007 to 2011 and the average type II error is estimated to be 23.0 percent.

As noted earlier, type I errors may be the result of several factors (for example, the algorithm matches two completely unrelated payments or identifies a loan other than an overnight fed funds transaction). Although we are unable to trace back the source of these type I errors, we conjecture that correspondent banking, whereby the algorithm incorrectly assigns to our two banks fed funds transactions conducted on behalf of some of their clients, plays a major role.

In contrast, we can quantify some of the reasons behind type II errors. While we focus on the first quarter of 2007 for this analysis, similar results were found in the first quarter of 2011. First, the algorithm classifies some pairs of transactions as Eurodollars when they are in fact fed funds. Our results suggest that this occurs relatively frequently. In particular, out of the 3,211 fed funds transactions not recognized by the algorithm in the first quarter of 2007, 1,455, or 45 percent, had

[18] Ideally, we would have liked to double-check the validity of the hypothesis by comparing the transactions carrying unique identifiers with another source of data on fed funds transactions. However, we are not aware of such an alternative source. In particular, the data reported by depository institutions in the Federal Financial Institutions Examination Council's Consolidated Reports of Condition and Income and by bank holding companies in the Federal Reserve's FR Y-9C forms do not isolate fed funds transactions as defined by Regulation D. Instead, these filings report a broader measure of purchases and sales of unsecured funds among financial institutions.

[19] Because of technical limitations, the furthest back we can go to test the algorithm is 2007.

[20] We do not know why there are opposing trends in our estimates of the type I and type II errors. The total number of payments sent and received by these two banks over Fedwire is roughly flat over this time period. Further, the number of payments exceeding $1 million sent and received by these two banks over Fedwire is also roughly flat, except for a decline of 20 percent from the first quarter of 2008 to the first quarter of 2009.

Table 2
Estimates of Type I and Type II Errors over Time
Percent

	2007:Q1	2008:Q1	2009:Q1	2010:Q1	2011:Q1	Average
Type I	64	72	85	93	93	81.4
Type II	24	28	27	19	17	23.0

Source: Authors' calculations, based on Fedwire data.

been discarded by the algorithm as being Eurodollars.[21] Second, by construction, the algorithm ignores fed funds loans where the principal is less than $1 million. In the first quarter of 2007, there were 170 such small fed funds transactions, accounting for 5 percent of the 3,211 false negatives. These small overlooked fed funds transactions, however, account for only 0.07 percent of the false negatives in terms of dollar value. Third, the algorithm could have faced multiple-candidate receive legs and did not choose the correct receive leg with the identifier. This only happened in the case of 128 of the 3,211 false negatives (4 percent). Fourth and finally, even if a payment is above $1 million, the algorithm may not find a potential match because, for example, it is a term loan, or the negotiated interest rate is outside the range specified by the algorithm. For the first quarter of 2007, 1,458, or 45 percent, of the transactions fall into this category.

3.3 Is the Output of the Algorithm Biased?

Given the high rates of type I and type II errors, it would appear that the algorithm's transaction-level output is ill-suited to study the fed funds market, and more generally to conduct research. Nevertheless, it is possible that the algorithm's errors may be considered white noise, in which case the algorithm's output would be unbiased. Unfortunately, we find evidence that the algorithm does produce biased outputs along at least three dimensions: the set of counterparties, the distribution of amounts loaned, and the distribution of interest rates. Once again, we focus on the first quarter of 2007 for this analysis, but find that the algorithm produces similar biases in the first quarter of 2011.

We first examine the set of counterparties for both fed funds sold and fed funds purchased by the two banks in the first quarter of 2007.[22] For each of the two banks, we compare the top ten counterparties, as ranked by the number of transactions, for the reference and algorithm lists.[23] For both banks, only three of the top ten counterparties in the algorithm list also appear in the top ten counterparties in the reference list. When ranking counterparties by the total value of their transactions, for both banks we find that five of the top ten counterparties in the algorithm list also appear in the equivalent top ten counterparties in the reference list. This comparison illustrates the algorithm's poor performance in correctly identifying the most important counterparties of the two banks.

We now turn to quantities. In the reference list, we observe the amount of the receive legs of the fed funds loans conducted by the two banks. From the algorithm list, we construct a comparable set of amounts by extracting the receive leg from each pair of payments linked by the algorithm. As illustrated in Chart 1, the distributions of amounts differ across these two sets of payments. Specifically, the amounts in the reference list tend to be smaller than those in the algorithm list. In particular, the mean and median amounts in the reference list are $18.1 million and $72.5 million, as compared with $50 million and $143.8 million in the algorithm list. Using the Mann-Whitney U test, we can reject at the 1 percent significance level the null hypothesis that the distributions of amounts across both samples are equal (the Z-score is -54.8). We therefore find statistical evidence that the algorithm output is biased with respect to the amounts of fed funds loans. Similar biases are identified when we consider separately the amount of fed funds sold and the amount of fed funds purchased by the two banks (see Appendix Charts A1 and A2).

Finally, we consider interest rates. To compute the interest rate for a transaction in the reference list, we need to pair the receive leg with its send leg. As the latter is unobserved, the pairing can only be approximated. For the comparisons conducted below, we focus on the set of true positives in the first quarter of 2007, that is, the 10,444 send legs in the algorithm list that can be matched to a receive leg in the reference list. We can then compare the inferred interest rates from this set of transactions to the inferred interest rates in the algorithm list. In Charts 2 and 3, we plot the interest rate distributions

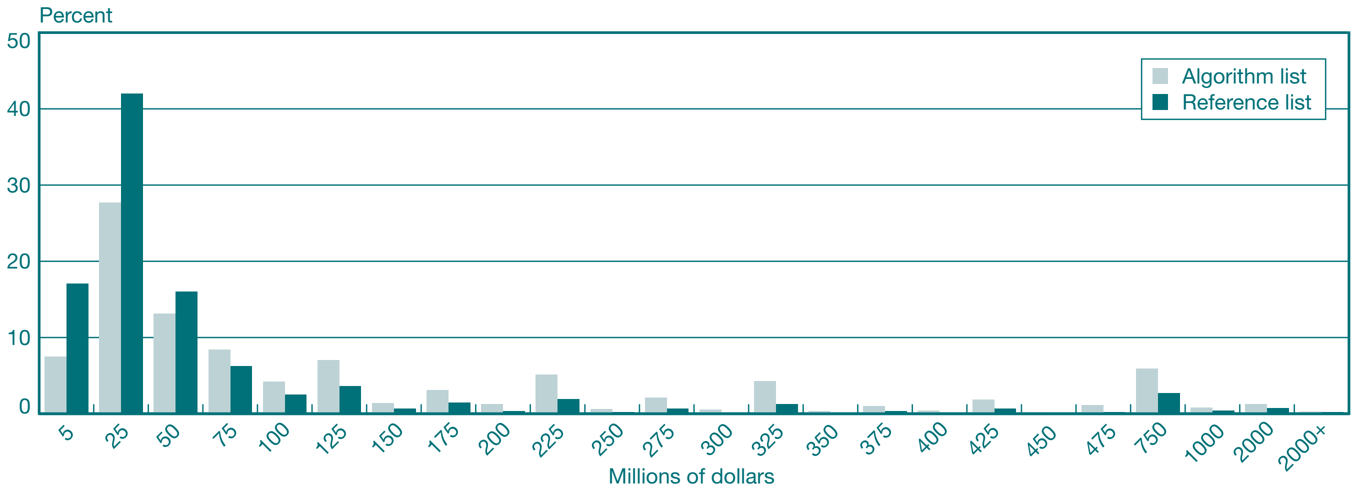
[21] In the first quarter of 2007, 32,647 pairs of payments were classified as Eurodollars instead of fed funds because the send leg had been given a "CTR" business code (see step 5 of the algorithm in section 2.1). We find that out of these 32,647 pairs of payments, only 1,455, or 4.5 percent, were in fact fed funds transactions. Our results therefore support the presumption that the "CTR" business code is an effective (albeit imperfect) way to distinguish Eurodollar from fed funds loans.

[22] Recall that neither the algorithm nor the unique identifiers for the two banks allow us to identify with certainty the fed funds counterparty of the banks. So instead of comparing counterparties, we may actually be comparing the correspondent banks of the true counterparties.

[23] According to the reference list, the top ten counterparties for each of the two banks account for, very roughly, two-tenths of the total number of fed funds transactions conducted by the two banks and one-half of their total value of fed funds activity.

Chart 1

Comparison of Transaction Amounts across the Algorithm and Reference Lists



Source: Authors' calculations, based on Fedwire data.

Notes: The comparison is conducted for 2007:Q1. For the algorithm list, amounts plotted are those in the receive leg of the paired payment transactions. The horizontal axis label is the amount bin's larger end point, except for "2000+," which denotes the bin with all payments greater than $2,000 million.

for the fed funds sold and purchased by the two banks. Our findings are similar to those for our analysis of amounts: the distributions of rates produced by the algorithm differ from the distributions of rates of the true positives. In particular, the median rates of fed funds sold and purchased are, respectively, 537 and 519 basis points for the true positives, while the median rates of fed funds sold and purchased are, respectively, 525 and 523 basis points for the algorithm list.[24] Using a Mann-Whitney U test, we can reject at the 1 percent significance level the null hypothesis that these distributions of rates are equal (the Z-score is -26.3 for fed funds sold and -33.3 for fed funds purchased). Because the algorithm is biased downward for fed funds sold and upward for fed funds purchased, these biases partially offset each other when the interest rates of fed funds sold and purchased by the two banks are combined. Nevertheless, even when fed funds sold and purchased are combined, there remain significant differences between the distribution of interest rates inferred from the algorithm and the distribution of interest rates from true positives (Appendix Chart A3). Hence, we find that the interest rates produced by the algorithm are statistically biased for fed funds sold and fed funds purchased—by 12 and 4 basis points, respectively. To gauge the economic magnitude of these biases, we note that over the same time period, the average spread between the overnight Libor rate (for U.S. dollars) and the one-month (six-month) Libor rate was 1.5 (5.7) basis points.

These three comparisons provide statistical evidence of significant bias in the set of counterparties as well as the distributions of transaction amounts and interest rates inferred from the algorithm's output for our two banks. In other words, the algorithm's errors are not just white noise. Rather, the main characteristics of the pairs of payments produced by the algorithm seem to exhibit systematic biases. Further, the nature of these biases is such that they do not subside when the algorithm's output is aggregated to the bank-to-bank level, or at the bank level. Finally, the algorithm's errors and biases remain essentially unchanged when its implementation is slightly modified (for example, by relaxing the minimum $1 million loan amount or widening the range of possible interest rates).

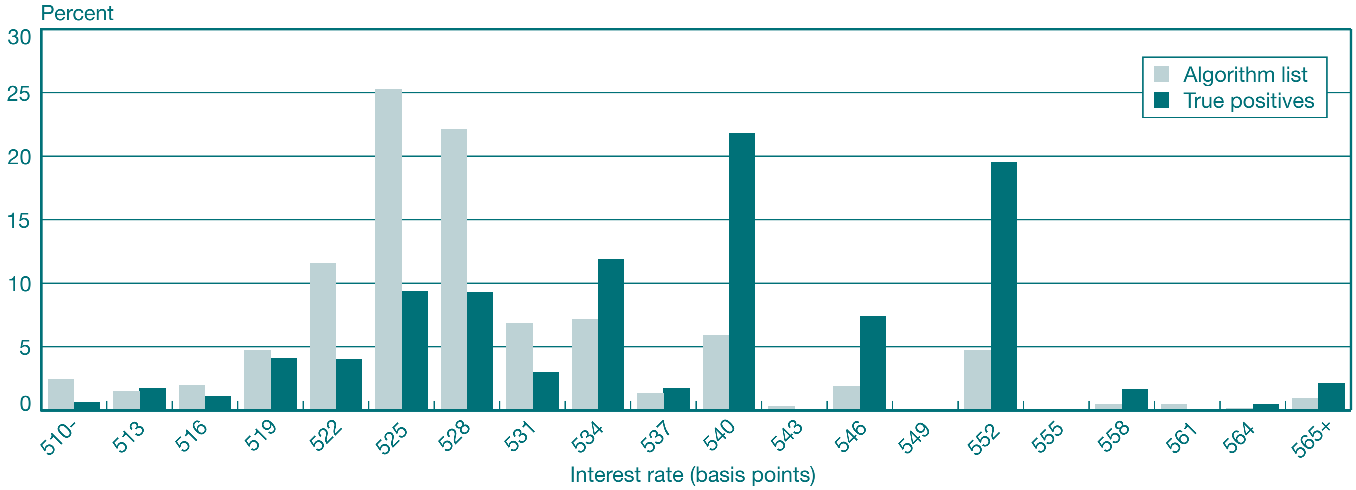
4. Discussion

Because the federal funds market has been one of the key financial markets in the United States, it has attracted considerable attention from researchers, especially after the 2008-09 financial crisis. Empirical analyses of this market have typically relied on transactions inferred by an algorithm

[24] In the first quarter of 2007, the target fed funds rate was 525 basis points.

Chart 2

Comparison of Interest Rates for Federal Funds Sold

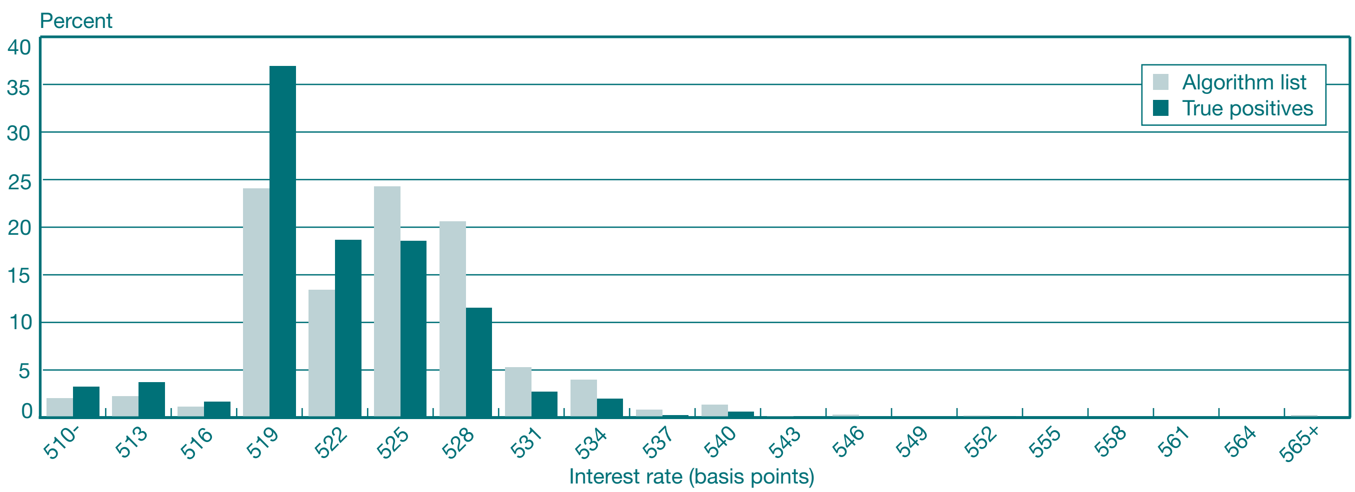


Source: Authors' calculations, based on Fedwire data.

Notes: The comparison is conducted for 2007:Q1. Note that the federal funds rate targeted by the Federal Open Market Committee in this quarter was 525 basis points. The horizontal axis label is the rate bin's larger end point, except for "565+," which denotes the bin with all interest rates greater than 565 basis points.

Chart 3

Comparison of Interest Rates for Federal Funds Purchased



Source: Authors' calculations, based on Fedwire data.

Notes: The comparison is conducted for 2007:Q1. Note that the federal funds rate targeted by the Federal Open Market Committee in this quarter was 525 basis points. The horizontal axis label is the rate bin's larger end point, except for "565+," which denotes the bin with all interest rates greater than 565 basis points.

comparable to the one used by the Federal Reserve Bank of New York. There is no guarantee, however, that this algorithm correctly identifies individual fed funds transactions.

In this article, we reported on a test aimed at assessing the transaction-level quality of the algorithm. For two large banks, among the more active in the fed funds market, we find the type I and type II errors to be large, averaging 81 percent and 23 percent, respectively, from 2007 to 2011. Further, we find evidence suggesting that these large errors cannot be considered white noise. Rather, they introduce significant biases in the computed rate and volume of fed funds activity, as well as in the set of counterparties. To be sure, we want to acknowledge that our study has possible limitations. In particular, our test applies only to fed funds as defined under Regulation D, and is based only on two banks. Despite these limitations, however, we argue below that our results have important implications.

4.1 How General Are Our Results?

The two institutions on which our test is based are large banks and so are not representative of all participants in the fed funds market. Hence, there is a possibility that the results of our test do not generalize to other fed funds participants. We provide two reasons, however, why we believe our results do, in fact, apply quite broadly. First, the two banks that provided their unique identifiers are either senders or receivers for a sizable share of all pairs of transactions that are output by the algorithm. Over the 2007-11 period, the two banks were involved, on average, with 29.4 percent of the algorithm's output. Our results, then, directly relate to a large fraction of the algorithm's output. Second, we believe it is reasonable to assume that our results are applicable to other large banks with similar Fedwire activity. We define large banks as those that receive or send over 800,000 payments a quarter (in a typical quarter, only nine or ten banks met this criterion). Assuming that our type I and type II errors generalize to these banks implies that, on average, 44.3 percent of the algorithm's output is affected (see Table 3).

The algorithm, however, may perform better for smaller banks. Indeed, these banks send fewer payments over Fedwire, and these payments may reflect fewer types of transactions. As a result, it might be easier for the algorithm to recognize fed funds transactions initiated by smaller banks. Although we cannot test it formally at this point, this hypothesis finds some support in the fact that there is separate preliminary evidence that the algorithm may perform well for some government-sponsored enterprises. If one can establish that

Table 3

Percent of Algorithm's Output to Which the Type I and Type II Error Estimates Apply

	2007:Q1	2008:Q1	2009:Q1	2010:Q1	2011:Q1	Average
Two banks	29.0	25.0	28.0	31.4	33.6	29.4
Large banks	39.5	37.4	40.7	49.4	54.7	44.3

Source: Authors' calculations, based on Fedwire data.

Notes: "Two banks" are the two institutions on which our tests are based. "Large banks" are those that sent and received more than 800,000 payments in the relevant quarter. The same nine banks met this criterion every quarter in the table, including the two banks at the center of our analysis. A tenth bank met this criterion in the first quarters of 2007, 2008, and 2011, although the identity of this tenth bank is not the same across the three quarters.

the algorithm is only inaccurate for a few large banks, then a possible remedy could be to exclude these banks from any empirical analysis. Still, we see at least three problems with this approach. First, ignoring at least a third of all transactions output by the algorithm would prevent any comprehensive analysis of the fed funds market. Second, one would have to show that excluding banks in a nonrandom way does not introduce biases in the algorithm output. Third, this approach would not only exclude the fed funds transactions conducted by these large banks, but also those involving their smaller clients as part of correspondent banking. As a result, excluding a few large banks may not permit an accurate analysis of the fed funds transactions conducted by smaller banks.

4.2 Does Aggregating the Algorithm's Output Make It More Precise?

Our test suggests that the algorithm is unlikely to identify individual fed funds transactions correctly. However, if aggregated to the bank-to-bank level, the bank level, or the market level, could the algorithm's output be useful to study the fed funds market? In part because the algorithm cannot identify the ultimate originator or beneficiary of a fed funds transaction, we do not think that the algorithm can provide, in general, meaningful measures at the bank-to-bank or the bank level. In particular, the algorithm will attribute 1) more transactions to large banks that serve as intermediaries, and 2) fewer transactions to small banks using correspondent banks. Because small and large banks may transact fed funds at different rates, the average rate identified by the algorithm for those banks may be biased.

Our analysis provides little evidence that the algorithm may or may not provide accurate market-level measures of fed funds activities. Nevertheless, we note that correspondent banking may possibly be the major source of type I errors in our test. In other words, the algorithm may correctly identify fed funds transactions but attribute them to the wrong originator or beneficiary. If this is the case, then the algorithm would produce unbiased market-level data on the distribution of rates and volumes of fed funds. The algorithm's output would then be a useful complement to the data obtained through brokers by the FRBNY, because it would cover fed funds transactions arranged both through brokers and privately between banks. To confirm this hypothesis, however, further work is necessary to test whether the algorithm's type I errors are almost exclusively produced by correspondent banking.

4.3 Does the Algorithm's Output Capture More General Interbank Overnight Loans?

While the available evidence points to the algorithm's output being imprecise measures of fed funds activity at the transaction and bank levels, the algorithm may still be of value if it captures a broader type of overnight funding. This would follow if most of the false positives identified in our test were indeed loans, but simply not fed funds loans (for example, if they were loans to financial institutions other than banks). This hypothesis finds support in the fact that 89 percent of the transactions paired by the algorithm in first quarter of 2007 are found to have inferred interest rates that, once rounded, can be considered to be in whole basis points or 32nds of an interest rate.[25] Discussions with market participants suggest that overnight unsecured loans are typically traded in these discrete amounts, suggesting that the pairing of transactions by the algorithm is not random.[26]

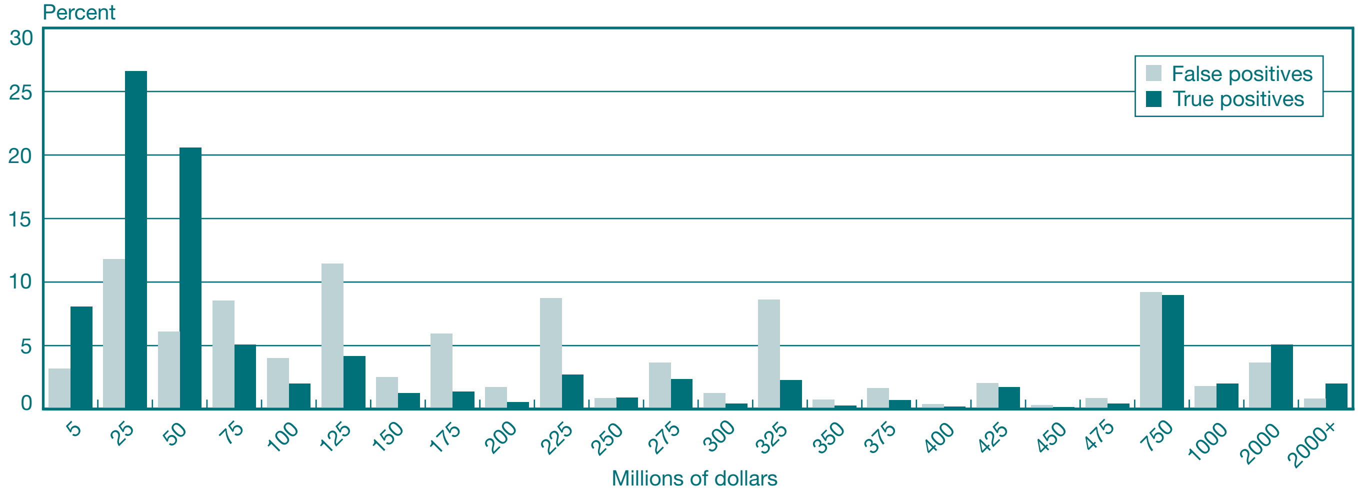
We note, however, that even if the algorithm correctly identifies loans, it may not accurately identify interbank loans. This would be the case in particular if loans are placed on behalf of bank clients that are outside the banking system or even the financial sector. Furthermore, even in the case of an interbank loan, the algorithm cannot guarantee the identity of the originator and the beneficiary because of the possibility of correspondent banking. More generally, the hypothesis that the algorithm's output captures overnight interbank loans would need to be formally tested in order to be validated. Until then, we believe that the algorithm's output should not be used as a proxy for interbank lending.

In conclusion, our results raise serious concerns about the appropriateness of using the algorithm's output to study the fed funds market. As a consequence, it raises questions about the validity of empirical results previously obtained using the algorithm's output. Finally, our analysis underscores the need to validate formally, prior to any analysis, that the indirect inferences produced by an algorithm are accurate.

[25] The dollar amount a bank can send to another bank over Fedwire is constrained to be rounded to the nearest cent. Because of rounding, the interest rate agreed upon by the banks when agreeing to a trade may differ from the interest rate we compute from the payment flows. Hence, when checking whether an implied interest rate is in whole basis points, we account for rounding. We do this by computing the implied interest rate when the principal and interest payment amount is increased by one cent and then when the amount is decreased by one cent. If these two inferred interest rates straddle an interest rate in whole basis points or 32nds of an interest rate, then we say that the algorithm's implied interest rate is consistent with a loan with an interest rate in whole basis points or 32nds of an interest rate.

[26] Substantiating these claims by market participants, we found that the interest rates of brokered fed funds trades between February 11, 2002, and September 24, 2004, provided by BGC Brokers, were all in whole basis points or 32nds of an interest rate. See Bartolini, Hilton, and McAndrews (2008) for details on these data.

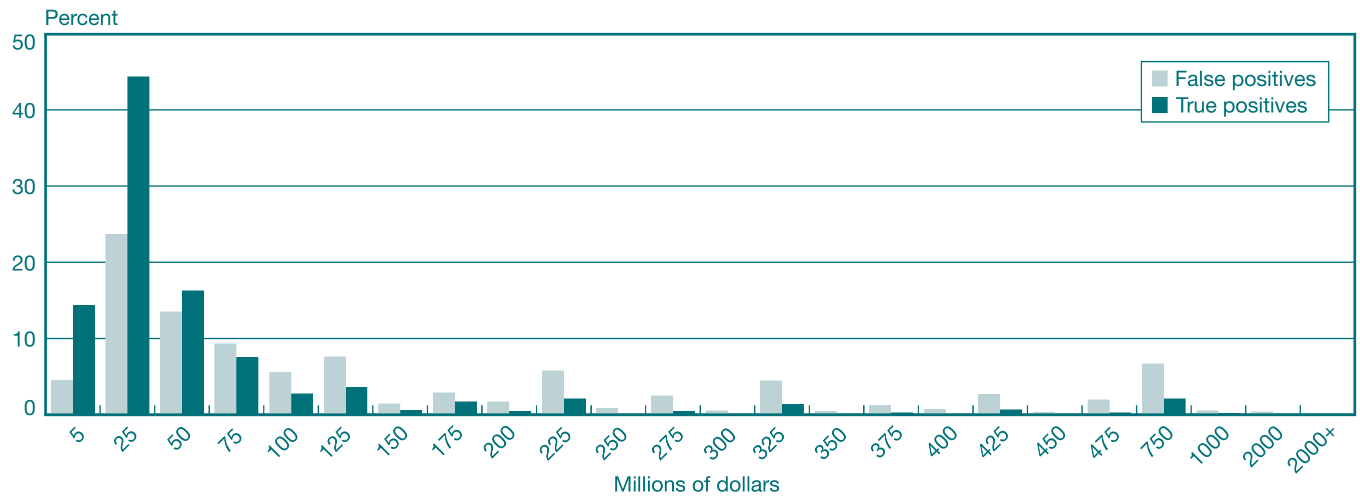
Chart A1
Comparison of Transaction Amounts of Federal Funds Sold



Source: Authors' calculations, based on Fedwire data.

Notes: The comparison is conducted for 2007:Q1. The principal amount of the federal funds sale is graphed. The horizontal axis label is the amount bin's larger end point, except for "2000+," which denotes the bin with all payments greater than $2,000 million. A Mann-Whitney U test rejects the null hypothesis that the distribution of amounts across false positives and true positives is equal at the 1 percent significance level (the Z-score is -15.0).

Chart A2
Comparison of Transaction Amounts of Federal Funds Purchased

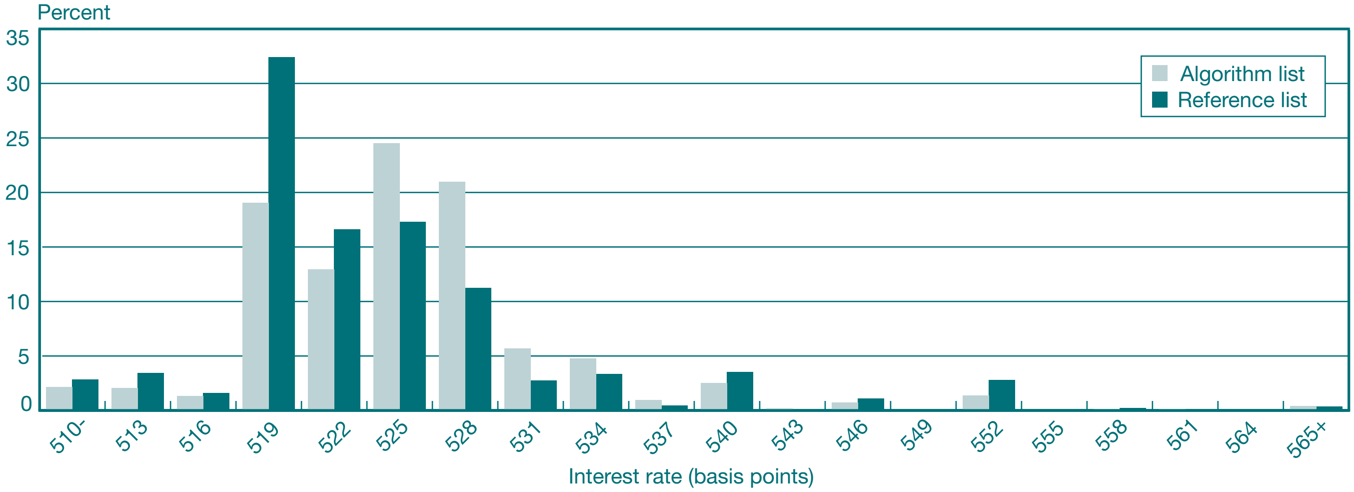


Source: Authors' calculations, based on Fedwire data.

Notes: The comparison is conducted for 2007:Q1. The principal amount of the federal funds purchased is graphed. The horizontal axis label is the amount bin's larger end point, except for "2000+," which denotes the bin with all payments greater than $2,000 million. A Mann-Whitney U test rejects the null hypothesis that the distribution of amounts across false positives and true positives is equal at the 1 percent significance level (the Z-score is -53.0).

Chart A3

Comparison of Interest Rates across the Algorithm and Reference Lists When Federal Funds Sold and Federal Funds Purchased Are Combined



Source: Authors' calculations, based on Fedwire data.

Notes: The comparison is conducted for 2007:Q1. For the reference list, interest rates were inferred for only those transactions in the set of true positives. Note that the federal funds rate targeted by the Federal Open Market Committee in this quarter was 525 basis points. The horizontal axis label is the rate bin's larger end point, except for "565+," which denotes the bin with all interest rates greater than 565 basis points.

References

Acharya, V., and O. Merrouche. 2011. "Precautionary Hoarding of Liquidity and Inter-Bank Markets: Evidence from the Sub-Prime Crisis." Review of Finance 17, no. 1: 107-60.

Acharya, V., and D. Skeie. 2011. "A Model of Liquidity Hoarding and Term Premia in Inter-Bank Markets." Journal of Monetary Economics 58, no. 5: 436-47.

Afonso, G., A. Kovner, and A. Schoar. 2011. "Stressed, not Frozen: The Federal Funds Market in the Financial Crisis." Journal of Finance 66, no. 4: 1109-39.

———. 2013. "The Importance of Trading Relationships in the Interbank Lending Market." Federal Reserve Bank of New York Staff Reports, no. 620, May.

Afonso, G., and R. Lagos. 2012a. "Trade Dynamics in the Market for Federal Funds." Federal Reserve Bank of New York Staff Reports, no. 549, February.

———. 2012b. "An Empirical Study of Trade Dynamics in the Fed Funds Market." Federal Reserve Bank of New York Staff Reports, no. 550, February.

Allen, J., J. Chapman, F. Echenique, and M. Shum. 2012. "Efficiency and Bargaining Power in the Interbank Loan Market." Bank of Canada Working Paper no. 2012-29.

Allen, J., J. Kastl, and A. Hortacsu. 2012. "Crisis Management: Analyzing Default Risk and Liquidity Demand during Financial Stress." Unpublished paper.

Arciero, L., R. Heijmans, R. Heuver, M. Massarenti, C. Picillo, and F. Vacirca. 2013. "How to Measure the Unsecured Money Market? The Eurosystem's Implementation and Validation Using TARGET2 Data." De Nederlandsche Bank Working Paper no. 369.

Armantier, O., E. Ghysels, A. Sarkar, and J. Shrader. 2011. "Stigma in Financial Markets: Evidence from Liquidity Auctions and Discount Window Borrowing during the Crisis." Federal Reserve Bank of New York Staff Reports, no. 483, January.

Armantier, O., J. McAndrews, and J. Shrader. 2008. "Changes in the Timing Distribution of Fedwire Funds Transfers." Federal Reserve Bank of New York Economic Policy Review 14, no. 2, September.

Ashcraft, A., and H. Bleakley. 2006. "On the Market Discipline of Informationally Opaque Firms: Evidence from Bank Borrowers in the Federal Funds Market." Federal Reserve Bank of New York Staff Reports, no. 257, August.

Ashcraft, A., and D. Duffie. 2007. "Systemic Illiquidity in the Federal Funds Market." American Economic Review: Papers and Proceedings 97, no. 2: 221-5.

Ashcraft, A., J. McAndrews, and D. Skeie. 2011. "Precautionary Reserves and the Interbank Market." Journal of Money, Credit, and Banking 43, no. 7: 311-48.

Atalay, E., and M. Bech. 2010. "The Topology of the Federal Funds Market." Physica A: Statistical Mechanics and Its Applications 389, no. 22.

Bartolini, L., S. Hilton, and J. McAndrews. 2008. "Settlement Delays in the Money Market." Federal Reserve Bank of New York Staff Reports, no. 319, March.

Bech, M. L., C. T. Bergstrom, R. J. Garratt, and M. Rosvall. 2011."Mapping Change in the Federal Funds Market." Federal Reserve Bank of New York Staff Reports, no. 507, August.

Bech, M., E. Klee, and V. Stebunovs. 2012. "Arbitrage, Liquidity, and Exit: The Repo and Federal Funds Markets before, during, and Emerging from the Financial Crisis." Board of Governors of the Federal Reserve System Finance and Economics Discussion Series, no. 2012-21.

Copeland, A., A. Martin, and M. Walker. 2010. "The Tri-Party Repo Market before the 2010 Reforms." Federal Reserve Bank of New York Staff Reports, no. 477, November.

Demiralp, S., B. Preslopsky, and W. Whitesell. 2006. "Overnight Interbank Loan Markets." Journal of Economics and Business 58, no. 1: 67-83.

Furfine. C. 1999. "The Microstructure of the Federal Funds Market." Financial Markets, Institutions, and Instruments 8, no. 5: 24-44.

Hamilton, J. 1996. "The Daily Market for Federal Funds." Journal of Political Economy 104, no. 1: 26-56.

References (Continued)

———. 1997. "Measuring the Liquidity Effect." American Economic Review 87, no. 1: 80-97.

Hendry, S., and N. Kamhi. 2009. "Uncollateralized Overnight Loans Settled in LVTS." Applied Finance Economics 19, no. 1: 869-80.

Kuo, D., D. Skeie, and J. Vickery. 2012. "A Comparison of Libor to Other Measures of Bank Borrowing Costs." Unpublished paper, Federal Reserve Bank of New York.

Millard, S., and M. Polenghi. 2004. "The Relationship between the Overnight Interbank Unsecured Loan Market and the CHAPS Sterling System." Bank of England Quarterly Bulletin, Spring: 42-7.