

# **Identifying Interbank Loans, Rates, and Claims Networks from Transactional Data**

---

**Carlos León, Jorge Cely and Carlos Cadena**

Carlos León, Jorge Cely and Carlos Cadena

### Identifying Interbank Loans, Rates, and Claims Networks from Transactional Data

**Abstract:** *Our objective is to identify interbank (i.e., non-collateralized) loans between financial institutions from Colombian large-value payment system data by implementing Furfine's method. After identifying interbank loans from transactional data, we obtain the interbank rates and claims without relying on financial institutions' reported data. Contrasting identified loans with those consolidated from financial institutions' reported data suggests the algorithm performs well, and it is robust to changes in its setup. The weighted average rate implicit in transactional data matches local interbank rate benchmarks strictly. From identified loans, we also build the interbank claims network. The three main outputs (i.e., the interbank loans, the rates, and the claims networks) are valuable for examining and monitoring the money market, for contrasting data reported by financial institutions, and as inputs in models of financial contagion and systemic risk.*

**Keywords:** *Furfine's method, interbank rate, IBR*

**JEL Classification:** E42, E44

### Identificando préstamos interbancarios, tasas y redes de acreencias a partir de datos transaccionales

**Resumen:** *Buscamos identificar los préstamos de fondos interbancarios (i.e., no colateralizados) a partir de información transaccional del sistema de pagos de alto valor por medio del método de Furfine. Con base en dichos préstamos, y sin recurrir a información reportada por las instituciones financieras, calculamos las tasas y los saldos interbancarios. El resultado del contraste de los préstamos identificados con aquellos reportados sugiere que el algoritmo se ajusta adecuadamente y que es robusto a cambios en su configuración. La tasa promedio ponderada implícita en la información transaccional concuerda con las tasas interbancarias de referencia del mercado local. Con los préstamos identificados también construimos la red de acreencias interbancarias. Los tres principales productos (i.e., los préstamos interbancarios, las tasas y la red de acreencias) son valiosos para examinar y monitorear el mercado monetario, para contrastar la información reportada por las instituciones financieras y como insumos en modelos de contagio financiero y riesgo sistémico.*

**Palabras clave:** *método de Furfine, tasa interbancaria, índice bancario de referencia*

**Clasificación JEL:** E42, E44

### Une méthode pour identifier les taux des prêts interbancaires et les montants nets à partir des données transactionnelles

**Résumé:** *Dans cet article nous avons cherché à identifier les prêts interbancaires (les non-collatéraux), à partir de l'information transactionnelle fournie par le système de paiements à haute valeur, à travers de la méthode de Furfine. Pour ce faire, nous prenons la base de données des prêts, sans faire recours à l'information rapportée par les institutions financières mêmes, afin de calculer les taux et les soldes interbancaires. Le résultat des prêts identifiés ont été ensuite comparés avec ceux qui sont rapporté par les institutions financières. Les résultats suggèrent que l'algorithme est bien ajusté et qu'il est robuste face aux changements dans sa configuration. En effet, le taux moyen pondéré implicite calculé à partir de l'information transactionnelle correspond aux taux de référence interbancaires du marché local. Les prêts identifiés permettent également construire un réseau des crédits interbancaires. Les trois produits financiers principaux (à savoir, les prêts interbancaires, les taux et le réseau des crédits interbancaires) sont précieux, d'une part, pour examiner et surveiller le marché monétaire, afin de vérifier les informations communiquées par les institutions financières et, d'autre part, ils constituent un input dans les estimations dérivées des modèles de contagion financière et du risque systémique.*

**Mots-clés:** *méthode de Furfine, taux interbancaria, indice de référence de la banque*

**Classification JEL:** E42, E44

## Identifying Interbank Loans, Rates, and Claims Networks from Transactional Data

Carlos León, Jorge Cely and Carlos Cadena \*

–Introduction. –I. The dataset. –II. The algorithm. –III. Main results. –Final remarks.  
–References.

doi: 10.17533/udea.le.n85a03

*Original manuscript received on 10 July 2015; final version accepted on 6 November 2015*

### Introduction

The interbank funds market plays a central role in monetary policy transmission, as it allows financial institutions to exchange central bank money in order to share liquidity risks (Fricke & Lux, 2014). For that reason, they are

---

\* *Carlos León*: Researcher, Financial Infrastructure Oversight Department, Banco de la República; Extramural Fellow CentER, Tilburg University. Address: Banco de la República, Carrera 7 #14-78, Bogotá, Colombia. Email: cleonrin@banrep.gov.co / carlosleonr@hotmail.com. [corresponding author]

*Jorge Cely*: Leader Professional, Operations and Market Development Department, Banco de la República. Email: jcelyfe@banrep.gov.co.

*Carlos Cadena*: Specialized Engineer, Financial Infrastructure Oversight Department, Banco de la República. Email: ccadensi@banrep.gov.co.

The opinions and statements in this article are the sole responsibility of the authors and represent neither those of Banco de la República and its Board of Directors, nor those of Tilburg University. Comments and suggestions from Hernando Vargas, Pamela Cardozo, Clara Lía Machado, Freddy Cepeda, Fabio Ortega, Constanza Martínez, Miguel Sarmiento, and Ricardo Mariño are appreciated. The authors also benefited from discussions with Ronald Heijmans and Richard Heuver from De Nederlandsche Bank, and with the technical staff of Banco de la República. Any remaining errors are the authors' own.

the focus of central banks' implementation of monetary policy and have a significant effect on the whole economy (Allen, Carletti, & Gale, 2009; p. 639). In addition, the absence of collaterals in the interbank funds market creates powerful incentives for participants to monitor each other, and thus it plays a key role as a source of market discipline (Rochet & Tirole, 1996; Furfine, 2001). Therefore, the interbank funds market is an important element for an efficiently functioning financial system (Heijmans, Heuver, & Walraven, 2010).

Reliable and comprehensive data from the interbank funds market is somewhat elusive, with restrictions on the consistency, granularity, and opportunity of the corresponding databases. This explains the numerous research articles that aim at identifying unsecured (i.e., non-collateralized) interbank funds loans from large-value payment systems' transactional data. The first of such articles is credited to Furfine (1999), who developed an algorithm for identifying interbank overnight loans for the US money market from Fedwire data. The procedure in Furfine (1999) is straightforward: matching payments in day  $t$  from Bank A to Bank B greater than one million dollars rounded to the nearest integer of \$100,000 (i.e., the loans), with payments from Bank B to Bank A in day  $t + 1$  (i.e., the overnight refund) such that the implicit interest rate between both payments is reasonable (i.e., it falls inside a  $\pm 50$  basis points corridor with respect to publicly available measures of the federal funds rate). This general approach is commonly referred as *Furfine's method* or *Furfine's algorithm*.

After Furfine (1999) several authors have attempted to apply similar algorithms. Demiralp, Preslopsky, and Whiteshell (2004) include smaller size loans (i.e., greater than \$50,000, with equal-sized increments) and include an interest rate 1/32-rounding rule for discarding interbank overnight transactions that are incompatible with market practices. Millard and Polenghi (2004) apply Furfine's algorithm to UK's large value payment system (CHAPS), including a threshold on one million pounds. Hendry and Kamhi (2007) apply Furfine's algorithm to data from the Canadian large value transfer system with a half a basis point rounding rule for filtering transactions. Some authors have considered interbank non-overnight transactions for the Dutch, Swiss, and Euro interbank markets (see Heijmans et al. 2010; Guggenheim, Kraenzlin & Schumacher, 2011; Arciero et al., 2013).

Regarding the validity of the Furfine's method, Armantier and Copeland (2012) have questioned the results obtained when implementing Furfine's method. On the other hand, Arciero et al. (2013) contrast results reported by Armantier and Copeland, and confirm the validity of Furfine's method conditional on a deep knowledge of the underlying data and the technical attributes of the system under analysis.

This paper implements the Furfine's method on an unusual dataset from the unique Colombian large-value payment system (*Cuentas de Depósito* – CUD). As explained in detail below, it is an unusual dataset because loans and refunds may be easily and accurately filtered out from large-value payment system's raw data. Our objectives are the following: First, to filter out interbank funds transactions in the Colombian market in order to identify loans of all maturities (up to 90 days) without relying on data reported by financial institutions;<sup>1</sup> second, to contrast the loans identified by our algorithm with those identified from reports by financial institutions, and to contrast our implicit interbank overnight interest rate ( $IIR_{ON}$ ) with the publicly available interbank overnight reference rate ( $IBR_{ON}$ ) and interbank overnight funds average rate ( $TIB$ ); and third, to construct the interbank claims networks, a key input for examining financial contagion under recent approaches to systemic risk and financial stability.

By accomplishing these three objectives, we expect to provide evidence on the usefulness of large-value payment systems' data for monitoring, overseeing, and analyzing the interbank funds market. Some uses are worth emphasizing. First, as this type of exercise provides the opportunity to evaluate the interbank funds market without relying on reports from financial institutions, lags and potential errors arising from consolidating, processing and transmitting reports by financial institutions and financial authorities may be conveniently avoided. As financial authorities have to rely on delayed and costly sources of data (i.e., reported data, on-site and off-site analyses), it is warranted to have potential alternatives (Kyriakopoulos et al., 2009).

Second, transactional data grants financial authorities the ability to contrast surveys and reports by financial institutions. As the Libor panel scandal

---

<sup>1</sup> We do not include intraday interbank funds transactions.

demonstrated (see Guggenheim et al., 2010), the ability to contrast is rather limited when relying on reports and surveys only.<sup>2</sup> Also, as stressed by Kyr-  
iakopoulos et al. (2009), using financial transaction records is particularly  
useful to uncover financial misconduct and rogue trading.

Third, using transactional data allows monitoring the interbank overnight  
funds market in a continuous manner. Therefore, as suggested by Heijmans  
et al. (2010), it should be used as an early warning indicator tool, both at the  
macro level (the whole market) and at the individual (financial participants)  
level.

It is important to state that our implementation of Furfine's method is  
somewhat easier than in other markets. Interbank transactions settled in  
large-value payment systems are not labeled as such in most systems around  
the world (Heijmans et al., 2010), whereas the Colombian large-value pay-  
ment system obliges financial institutions to use specific codes when regis-  
tering interbank transactions. Before March 2013 there was a single code  
for interbank funds transactions, which did not allow distinguishing between  
loans and refunds. After March 2013, separate codes for intraday and non-  
intraday loans and refunds were enacted. Therefore, unlike most attempts to  
implement Furfine's method, our dataset is unusual as we already have iden-  
tified loans and refunds, and our task is to match converse transactions at  
a reasonable implicit rate in the future. As described below, such matching  
requires defining an area of interest rate plausibility (i.e., an interest rate cor-  
ridor) with respect to publicly available interbank reference rates, defining a  
procedure for solving multiple refund matches for a single loan, and defining  
transactions' maximum reliable time to maturity.

Three main contributions are worth stating. First, consistent with the lack  
of timely and granular information during the Great Financial Crisis, we pro-  
vide further evidence on the usefulness of large-value payment systems' data  
for monitoring, overseeing, and analyzing the interbank funds market. Sec-  
ond, we emphasize the importance of recent efforts by financial authorities to

---

<sup>2</sup> In March 2008, Libor panel banks were accused of avoiding negative signals about their  
refinancing conditions by misreporting interest rates.

attain detailed financial transactions datasets (e.g., Bank of England, 2015);<sup>3</sup> that is, turning our rather unusual dataset into a standard is now deemed advantageous for implementing monetary policy and financial stability. Third, by running the algorithm backwards (i.e., starting with the last date of the sample until reaching the first one) we attained a practical method for mitigating the over-identification of long-term interbank loans, an issue also encountered by Guggenheim et al. (2010) and Heijmans et al. (2010).<sup>4</sup> To the best of our knowledge, this backwards-run setup has not been attempted in related literature before, and may be worth exploring in further implementations of Furfine's algorithm.

## I. The Dataset

As usual in other implementations of Furfine's method, the data source is the local large-value payment system. This method relies on the premise that all interbank loans will eventually result in a converse payment between financial institutions in the large-value payment system. Such premise may be questionable whenever financial institutions tend to settle their interbank transactions outside the large-value payment system, say on the books of a common settlement bank, or by the physical delivery of cash or checks between them.<sup>5</sup>

---

<sup>3</sup> Bank of England's intention is to collect transaction-level data from sterling money markets, secured and unsecured, from all participating financial institutions. These efforts overlap with those by the Federal Reserve Bank of New York (i.e., Report of Selected Money Market Rates - FR2420) and the European Central Bank (i.e., Regulation No.1333/2014 concerning statistics on the money markets).

<sup>4</sup> Both research works deal with the over-identification of long-term loans by *a priori* privileging short-term maturities and/or privileging a set consisting of determined maturities (i.e., one-day, full week(s) and full month(s)).

<sup>5</sup> Incentives to settle interbank loans outside the large-value payment system may exist. For instance, the cost of transporting cash to/from remote locations may make lending and borrowing cash between financial institutions easier, cheaper, and more secure than transporting it from/to a major city. Also, due to the costs of sending payments in large-value payment systems, two non-banking financial institutions pertaining to the same conglomerate may find advantageous to ask their parent bank to settle a loan by modifying their books accordingly.

In Colombia, there is a single large-value payment system (CUD), which is owned and operated by the Central Bank (*Banco de la República*). This is the financial market infrastructure where all cash settlements (in local currency) take place, for all types of financial transactions—including interbank funds transactions. All types of financial institutions, banking and non-banking, local and foreign-owned, private and government-owned, are allowed to directly participate in the large-value payment system; that is, this is a non-tiered payment system with no settlement banks.

To the best of our knowledge, there is no information about the significance of interbank loans and refunds settled outside the local large-value payment system. We presume that the non-tiered access scheme of the Colombian large-value payment system, and the corresponding absence of settlement banks, should make the settlement of transactions in the books of a common financial institution unnecessary—therefore rare. However, as there are no limitations to these kinds of external settlements, or to the settlement of transactions by the physical delivery of cash or checks, we cannot rule out that some interbank loans and refunds are settled outside the large-value payment system. All in all, we expect external settlements of interbank loans and their refunds to be unimportant.

The local large-value payment system obliges financial institutions to classify their transactions based on a set of codes determined by the Central Bank. Regarding the interbank funds market transactions, after March 2013 financial institutions use distinct codes for registering interbank loans and refunds when executing the transfer of funds. Nonetheless, there is no tracking or earmarking between each loan and its corresponding refund.

Financial institutions do not register all interbank funds transactions in the large-value payment system. Interbank funds transactions corresponding to the interbank reference rate formation program (*Índice Bancario de Referencia-IBR*) are registered in the large-value payment system by a second financial market infrastructure, *Depósito Central de Valores* (DCV), which is the securities settlement system and the central securities depository for sovereign securities.



Hence, unlike most implementations of Furfine's method, the dataset is already filtered out. All non-interbank funds transactions may be easily discarded, and the remaining data are already classified as a loan or as a refund. The dataset consists of all interbank funds transactions between 1 April, 2013 and 30 December, 2014 (i.e., 428 business days); so intraday interbank transactions are not considered. This dataset comprises 27,200 interbank funds transactions: 13,504 correspond to interbank transactions registered by financial institutions directly, whereas 13,696 correspond to interbank transactions registered by DCV on behalf of the interbank reference rate formation program. The value of the interbank transactions corresponding to the reference rate formation program (i.e., registered in CUD by DCV) represents about 21.6% during the period under analysis.

Some features shape the Colombian interbank market. As with the large-value payment system, direct access is open to all types of financial institutions (i.e., banking and non-banking, local and foreign-owned, private and government-owned); that is, it is a bilateral, non-brokered, interbank market. Nevertheless, only a fraction of banking institutions lends and borrows in the interbank market on a regular basis. Most financial institutions, typically small in size, lend and borrow against collateral only. Accordingly, the contribution of unsecured borrowing to total borrowing between financial institutions is rather low, about 9.68% (see Banco de la República, 2015).<sup>6</sup> Hence, as reported by Martínez and León (2015), the Colombian interbank market appears to be a case of liquidity cross-underinsurance (see Castiglionesi & Wagner, 2013), in which strong negative externalities arising from counterparties' failure creates incentives not to provide liquidity in the absence of collateral or any other guarantee scheme.

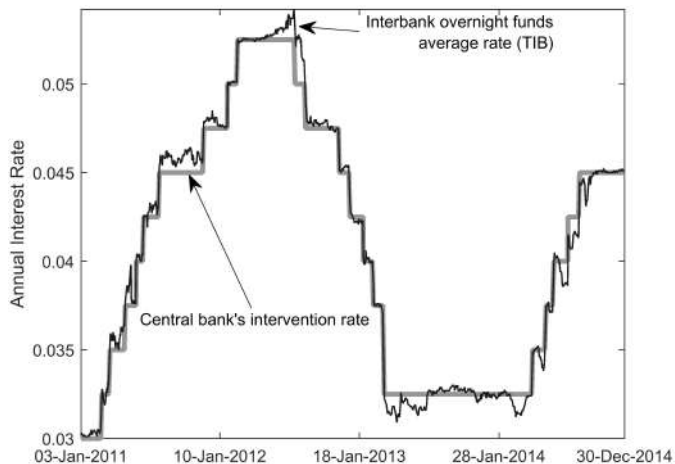
Both the interbank overnight funds average rate (*TIB*) and the target rate for implementation of Central Bank's monetary policy display a strong linear dependence, resulting in a 0.99 correlation for the 2011-2014 period. Also, the average absolute difference between these two rates is about 5 basis points, whereas the maximum absolute difference is 30 basis points. Therefore, as

---

<sup>6</sup> Based on 2014 figures, collateralized lending between financial institutions (i.e., sell/buy backs and repos) account for about 90.32% of money market transactions. Interbank (i.e., non-collateralized) lending accounts for the residual (9.68%).

exhibited in Figure 1, the interbank overnight funds rate follows Central Bank’s rates closely –as intended by monetary policy.

**Figure 1.** *The interbank overnight funds average rate (TIB) and Central Bank’s intervention rate*



Source: authors’ calculations.

## II. The algorithm

Unlike Furfine (1999) and most of the related literature on Furfine’s method, due to the unusual features of our dataset, our algorithm does not filter interbank transactions out from (raw) large-value payment data. Such features of the dataset avoid us making some assumptions for identifying loans and refunds. For instance, following the algorithm’s setup by Arciero et al. (2013), we do not need to define minimum loan values and increments in loan values to filter out potential loans. As making such assumptions is critical for minimizing false negatives and false positives (Arciero et al., 2013), working with filtered data should mitigate some serious sources of error.

Yet, some steps in the setup of the algorithm remain: First, defining the implicit interest rates’ plausibility (i.e., the interest rate corridor); second, excluding plausible rates that do not conform to market practices; third, defining

a solution procedure in case of multiple refund matches for a single loan; and fourth, defining the transactions' maximum reliable time to maturity.

Regarding the first remaining step, as not all financial institutions can trade liquidity at the same rates, an area of plausible interest rates has to be defined (Heijmans et al., 2010). Related literature favors a symmetrical or quasi-symmetrical corridor around a representative publicly available interbank rate. For instance, for the US Furfine (1999) uses a corridor of 50 basis points below (above) the minimum (maximum) of each day's federal funds' 11:00 rate, closing rate, and value-weighted funds rate. Demiralp et al. (2004) uses a wider  $\pm 100$  basis points corridor with a minimum rate of  $1/32$ . Heijmans et al. (2010) use a  $\pm 100$  basis points corridor with respect to Eonia (European Overnight Index Average) and Euribor throughout the European Central Bank's liquidity injection and interest rate decrease (i.e., September 2008-September 2009), and a  $\pm 50$  basis point corridor for the rest of their sample. Arciero et al. (2013) use three corridors ( $\pm 25$ ,  $\pm 50$  and  $\pm 200$  basis points) with respect to Eonia.

The area of plausibility should be chosen in such a way that it minimizes the probability of Type 1 and Type 2 errors (Heijmans et al., 2010). Type 1 error results when a transaction is mistakenly identified as an interbank transaction (i.e., a false positive), whereas Type 2 error results when an interbank funds transaction is erroneously omitted (i.e., a false negative). In our case the occurrence of Type 1 error is low due to the features of the dataset (i.e., interbank funds transactions only). As for Type 2 error, omitting actual interbank funds transactions may result from the choice of the area of plausibility: The narrower the corridor, the more likely it is to misclassify actual interbank funds transactions as implausible.

As in our case the occurrence of Type 1 error is expected to be low due to the features of the dataset, choosing a significantly wide area of plausibility may definitely minimize Type 2 error. However, a wide corridor may cause a Type 3 error, in which the algorithm yields a "wrong match" from multiple matches (see Arciero et al., 2013). Therefore, a too lax plausibility area is to be avoided as it may result in an unwarranted recurrence of multiple potential refund matches for a single loan.

Akin to Arciero et al. (2013), we use several corridors in order to test whether the results, along with the occurrence of Type 2 and 3 errors, are robust to the change of plausibility area. We use three corridors of  $\pm 50$ ,  $\pm 100$  and  $\pm 200$  basis points with respect to the maximum and minimum publicly available interbank reference rates (*IBR*) for the overnight, one-month and three-month time to maturities. A minimum plausible interest rate is set at 1.00 %.

As for solving for multiple potential matches, we use a recursive procedure. In case two or more potential refunds are available as plausible matches for a single loan, the algorithm estimates the interbank funds market term structure for all maturities from overnight to 90 days based on the three available interbank reference rates (i.e., overnight *IBR*, one-month *IBR*, and three-month *IBR*). The estimation of the term structure is done by cubic spline interpolation. The resulting term structure is used as a benchmark for deciding which of the competing plausible rates is the closest in absolute terms (i.e., lowest absolute spread) to the interpolated market price for liquidity. Afterwards, if two or more potential refunds persist, the one with the lowest time to maturity is chosen. This is consistent with interbank funds market maturities in the Colombian case, in which most transactions take place in the very short term, usually overnight. Finally, if the competing plausible refunds share the same absolute difference with respect to the interbank funds term structure, and have the same time to maturity, we use a *first in first out* (FIFO) rule that privileges the first occurring refund in the day.

Finally, we determine the maximum reliable time to maturity for the loans. Data from the Financial Superintendency disclose that the maximum time to maturity reported by financial institutions is 35 days during the period under analysis, in May 29, 2013. Therefore, one of our scenarios consists of limiting the maturity of potential refunds to 35 days. However, as one the advantages of using transactions data is to contrast other sources of information, and because longer maturities occurring in the future should not be disregarded, we also use a 90-day maximum reliable time to maturity scenario. Comparing both maturity scenarios will be useful for testing whether the algorithm is robust to different specifications.

Note that in designing the setup of the algorithm we do not consider excluding plausible rates that do not conform to market practices. This is a step that uses market practices or anecdotal evidence (e.g., rounding rules, minimum loan values, typical increments in value) to discard implausibly complicated loans or interest rates that do not follow standard interbank funds transactions (see Arciero et al., 2004; Demiralp et al., 2004). As before, because in our case the dataset is limited to interbank funds transactions, such exclusion procedure is unwarranted.

### III. Main results

The main results of the implementation of Furfine's method for the Colombian interbank funds market are presented in three subsections. The first subsection compares the results obtained with different scenarios of plausible interest rates (i.e., corridors) and different transactions' maximum reliable time to maturities. After selecting the most convenient scenario for analytical purposes, the second subsection contrasts the resulting identified loans with those identified from financial institutions' reported data. The third subsection contrasts our implicit interbank overnight interest rate ( $IIR_{ON}$ ) with the publicly available interbank overnight reference rate (*Índice Bancario de Referencia Overnight* –  $IBR_{ON}$ ) and interbank overnight funds average rate (*Tasa Interbancaria* –  $TIB$ ). The last subsection presents a sample of the interbank claims networks available as a byproduct of the algorithm.

#### A. Examining the scenarios

Six different scenarios are designed to examine the algorithm and its robustness to its setup. Different choices of plausible interest rates (i.e., the corridor) and transactions' maximum reliable time to maturities are considered for the following scenarios:

- $S_{(50|90)}$ :  $\pm 50$  basis points corridor, 90-day maximum maturity
- $S_{(100|90)}$ :  $\pm 100$  basis points corridor, 90-day maximum maturity
- $S_{(200|90)}$ :  $\pm 200$  basis points corridor, 90-day maximum maturity

- $S_{(50|35)}$ :  $\pm 50$  basis points corridor, 35-day maximum maturity
- $S_{(100|35)}$ :  $\pm 100$  basis points corridor, 35-day maximum maturity
- $S_{(200|35)}$ :  $\pm 200$  basis points corridor, 35-day maximum maturity

The algorithm was run forwards, starting with transactions occurring in April 1, 2013, day by day, through December 30, 2014. Although our dataset covers the period April 1, 2013-December 30, 2014, we discard some observations at the end of the sample when comparing among scenarios. As the first three scenarios consider a 90-day maximum time to maturity, comparisons will correspond to April 1, 2013-September 30, 2014 period (i.e., the last 3 months of observations are discarded). All rates are annualized rates with an actual/365 day-count convention.<sup>7</sup> A lower bound for the corridor is set to 1.00% in all scenarios.

Table 1 compares how the scenarios matched actual transactions (by number of transactions and their value), and presents how they differ in terms of implicit interest rates and maturities.

Table 1 shows that the number and value of transactions matched from the large-value payment system's data are rather representative, slightly above 83% and 89%, respectively. Interestingly, changes in the interest corridor and the maximum reliable maturity do not affect the number and value of matched transactions: increasing the corridor by 150 basis points and the maximum reliable maturity by 55 days results in a (negligible) 0.35% and 0.60% increase in the number and value of matched transactions. This points out that some traits of the data may explain the fraction of transactions that could not be captured by the algorithm. These traits may include the separate refund of principal and interests; the aggregation of two to more refunds into a single settlement; and the settlement of the loan (or the refund) outside the large-value payment system (e.g., by the physical delivery of cash or checks).

---

<sup>7</sup> The local day-count convention for the interbank market is actual/360 annualized rates. We use the actual/365 annualized rate convention because it is the one used by the Central Bank for calculating the *TIB*, and the one required by the Financial Superintendency for financial institutions reports.

Table 1 also shows that the main overall outcomes of the algorithm are rather robust to its setup. There are trivial differences in the number and value of transactions matched from actual data, in the average implicit interest rates, and in the average time to maturities. However, there are non-negligible differences in the range (i.e., from minimum to maximum) of implicit interest rates and time to maturities. Moreover, all scenarios corresponding to the 90-day maximum reliable maturity tend to over-identify transactions, especially when the corridor widens. For instance, although the maximum observed time to maturity reported by the Financial Superintendency for the period under analysis is 35 days, the first three scenarios identified interbank funds transactions at maturities of 49, 85 and 90 days, in which the wider the corridor the longer the maturity. This concurs with Arciero et al.'s (2013) statement about how the amount of noise (i.e., falsely identified loans) tends to increase with the selected maximum reliable maturity. Again, as emphasized by Arciero et al., a deep knowledge of the underlying data and technical details of the system is essential to avoid spurious results.

**Table 1.** *Comparison of selected scenarios*

|                                      |                      | $S_{(50 90)}$ | $S_{(100 90)}$ | $S_{(200 90)}$ | $S_{(50 35)}$ | $S_{(100 35)}$ | $S_{(200 35)}$ |
|--------------------------------------|----------------------|---------------|----------------|----------------|---------------|----------------|----------------|
| Matched <sup>a</sup>                 | Transactions         | 83.45%        | 83.49%         | 83.80%         | 83.45%        | 83.49%         | 83.78%         |
|                                      | Value                | 89.51%        | 89.54%         | 90.11%         | 89.51%        | 89.54%         | 90.10%         |
| Implicit interest rate               | Average <sup>b</sup> | 3.41%         | 3.41%          | 3.40%          | 3.41%         | 3.41%          | 3.40%          |
|                                      | Minimum              | 2.78%         | 2.28%          | 1.20%          | 2.78%         | 2.78%          | 1.20%          |
|                                      | Maximum              | 4.72%         | 4.72%          | 6.51%          | 4.72%         | 4.72%          | 4.72%          |
| Time to maturity (days) <sup>c</sup> | Average <sup>b</sup> | 2.63          | 2.64           | 2.68           | 2.62          | 2.63           | 2.64           |
|                                      | Minimum              | 1             | 1              | 1              | 1             | 1              | 1              |
|                                      | Maximum              | 49            | 85             | 90             | 35            | 35             | 35             |

*Note:* <sup>a</sup> Matched transactions and matched value correspond to the number of actual transactions and value of transactions (in percentage) that are identified by the algorithm, respectively. <sup>b</sup> Simple average. <sup>c</sup> Calendar days.

*Source:* authors' calculations.

To examine the robustness of the algorithm to its setup, we attempted a straightforward test. As mentioned, the algorithm was run forwards, meaning that we start matching interbank loans occurring in April 1 with all refunds happening after April 1, then matching loans occurring in April 2, until –recursively– reaching the last date of our sample (i.e., September 30, 2014). If the maximum reliable time to maturity is 90 days, this type of procedure allows the algorithm to try to match April 1’s loans with all refunds occurring between April 2 and July 1, even though most of these refunds correspond to loans originated well after April 1. That is, by running the algorithm forwards the bulk of potential refunds corresponds to loans that occurred after April 2, and thus we should expect to be over-identifying long-term interbank loans (see Guggenheim et al., 2011; Heijmans et al., 2010).

Put another way, when run forwards the algorithm works with the sum of the distribution of refunds for each day within the maximum reliable time to maturity, with the resulting distribution of refunds departing from the observed distribution of loans time to maturity. Hence, the algorithm is forced to find matches from a set of refunds that is mostly constituted by refunds corresponding to loans occurring afterwards. On the other hand, when run backwards, starting with the last date of the sample until reaching the first one, the algorithm will try to find matches from a set of refunds that has already been used to match other loans, thus its distribution should be closer to loans observed time to maturity.<sup>8</sup>

In short, we start matching loans contracted in September 30, 2014 with the refunds registered after that date; afterwards, we match the loans contracted in September 29, 2014 with the refunds registered and available (i.e., not already matched before) after that date; and so on until reaching the first

---

<sup>8</sup> For instance, in a hypothetical case in which loans are limited to 1- and 2-day maturities, and the maximum reliable time to maturity is 10 days, potential refunds from the forwards-run algorithm for any given loan are the sum of all 1- and 2-day refunds occurring in the next 10 days. This results in a distribution of potential refunds maturities that does not reflect the actual 1- and 2-day maturities of loans, and that may cause the over-identification of long-term maturities (i.e., beyond 1- and 2-day maturities). If run backwards, the distribution of potential refunds maturities will tend to display the actual 1- and 2-day maturities of loans because the other (i.e., longer) maturities should have been used by the algorithm before.



day of the sample (April 1, 2013). By running the algorithm backwards, short-term matches will be more likely because spurious long-term potential refunds will tend to be scarce as they were already matched by short-term plausible transactions. Hence, running the algorithm backwards should mitigate the over-identification of long-maturity transactions and of false positives (i.e., Type 1 error), whilst the results should converge to those attained by defining an adequate maximum reliable time to maturity.

The backwards scenarios will be identified as follows:

- $Z_{(50|90)}$ : backwards,  $\pm 50$  basis points corridor, 90-day maximum maturity
- $Z_{(100|90)}$ : backwards,  $\pm 100$  basis points corridor, 90-day maximum maturity
- $Z_{(200|90)}$ : backwards,  $\pm 200$  basis points corridor, 90-day maximum maturity
- $Z_{(50|35)}$ : backwards,  $\pm 50$  basis points corridor, 35-day maximum maturity
- $Z_{(100|35)}$ : backwards,  $\pm 100$  basis points corridor, 35-day maximum maturity
- $Z_{(200|35)}$ : backwards,  $\pm 200$  basis points corridor, 35-day maximum maturity

The backwards-run scenarios in Table 2 exhibit some interesting features. First, as before, the number and value of transactions matched from the large-value payment system's data are rather representative, slightly above 83% and 89%, respectively, and they are robust to changes in the setup of the algorithm. Second, the main outcomes of the scenarios are robust in cross-section, with trivial differences in the average implicit interest rates, average time to maturity, minimum time to maturity, and number and value of captured transactions. Third, concurrent with forwards-run scenarios in Table 1, the widest corridor ( $\pm 200$  basis points) results in non-negligible differences in the maximum and minimum implicit interest rates, and the maximum maturities. Fourth, as expected, the over-identification of long-maturities is mitigated in the backwards-run scenarios:  $Z_{(50|90)}$  and  $Z_{(100|90)}$

result in a maximum time to maturity equal to that reported by the Financial Superintendency for the sample under analysis (i.e., 35 days in 29 May, 2013), whereas analogous forwards-run  $S_{(50|90)}$  and  $S_{(100|90)}$  resulted in 49 and 85 days maturities. Avoiding the over-identification of long-term transactions by perfectly matching the maximum time to maturity in the sample is a straightforward empirical test of the benefits of running the algorithm backwards.<sup>9</sup>

**Table 2.** Comparison of backwards-run selected scenarios<sup>d</sup>

|                                      |                      | $Z_{(50 90)}$ | $Z_{(100 90)}$ | $Z_{(200 90)}$ | $Z_{(50 35)}$ | $Z_{(100 35)}$ | $Z_{(200 35)}$ |
|--------------------------------------|----------------------|---------------|----------------|----------------|---------------|----------------|----------------|
| Matched <sup>a</sup>                 | Transactions         | 83.45%        | 83.49%         | 83.79%         | 83.45%        | 83.49%         | 83.78%         |
|                                      | Value                | 89.51%        | 89.54%         | 90.11%         | 89.51%        | 89.54%         | 90.10%         |
| Implicit interest rate               | Average <sup>b</sup> | 3.41%         | 3.41%          | 3.40%          | 3.41%         | 3.41%          | 3.40%          |
|                                      | Minimum              | 2.78%         | 2.28%          | 1.42%          | 2.78%         | 2.28%          | 1.42%          |
|                                      | Maximum              | 4.72%         | 4.72%          | 6.37%          | 4.72%         | 4.72%          | 6.11%          |
| Time to maturity (days) <sup>c</sup> | Average <sup>b</sup> | 2.62          | 2.63           | 2.64           | 2.62          | 2.63           | 2.64           |
|                                      | Minimum              | 1             | 1              | 1              | 1             | 1              | 1              |
|                                      | Maximum              | 35            | 35             | 80             | 35            | 35             | 35             |

*Note:* <sup>a</sup> Matched transactions and matched value correspond to the number of actual transactions and value of transactions (in percentage) that are identified by the algorithm, respectively. <sup>b</sup> Simple average. <sup>c</sup> Calendar days. <sup>d</sup> Instead of running the algorithm from the first to the last date, we run the algorithm backwards to mitigate the over-identification of long maturity transactions and examine the robustness of results.

*Source:* authors' calculations.

Table 3 exhibits how the algorithm for the selected scenario ( $Z_{(100|90)}$ ) solved for matches in transactional data. Single matches (i.e., only one refund is available as a plausible match) are the most common solution (94.45%). Multiple matches are uncommon (5.55% by number of transactions, 13.09% by transactions' total value), and most of them are solved by the *first in first out* rule (4.33% and 10.84%, respectively) –after the lowest time to maturity

<sup>9</sup> However,  $Z_{(200|90)}$ , which is the scenario with the widest corridor ( $\pm 200$  basis points) under a lax time to maturity limit (90 days), still results in maturities well beyond reported ones (i.e., 85). This is expected because a wide corridor will tend to yield multiple potential refund matches and to over-identify transactions.

and the closest to the term structure rules have been exhausted. That is, for the selected dataset, the algorithm encounters multiple matches occasionally, and thus results are rather robust to the selection of solutions for multiple matches.

**Table 3.** *Solving for single and multiple matches*

| Type of match <sup>a</sup>     | Transactions <sup>b</sup> | Value <sup>c</sup> |
|--------------------------------|---------------------------|--------------------|
| Single                         | 94.45%                    | 86.91%             |
| Multiple                       | 5.55%                     | 13.09%             |
| <i>Lowest absolute spread</i>  | <i>1.23%</i>              | <i>2.25%</i>       |
| <i>Lowest time to maturity</i> | <i>0.00%</i>              | <i>0.00%</i>       |
| <i>First in-first out</i>      | <i>4.33%</i>              | <i>10.84%</i>      |

*Note:* <sup>a</sup>Based on the recursive procedure for solving multiple potential matches in Section II. <sup>b</sup>Calculated as the proportion of total loans in the sample. <sup>c</sup>Calculated as the proportion of total loans value in the sample. Single matches are the most common solution.

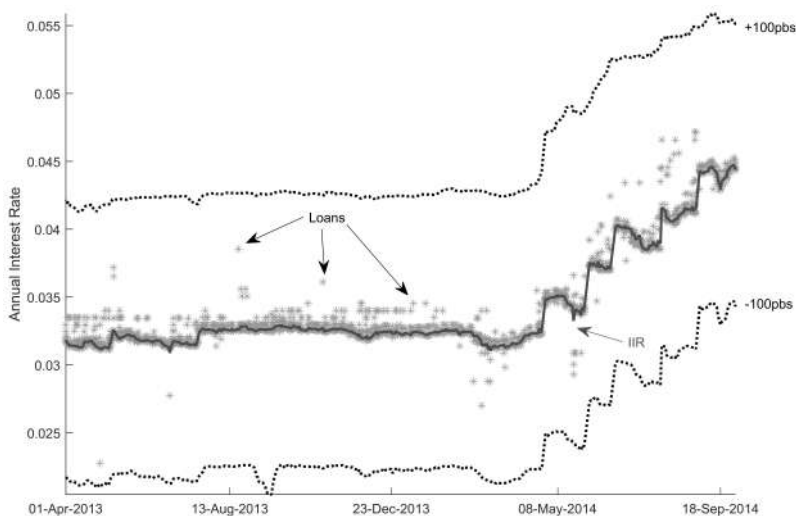
*Source:* authors' calculations.

The graphical outcome of the algorithm for the selected scenario is displayed in Figure 2. Dots correspond to loans identified by the algorithm, whereas the line corresponds to our weighted average implicit interbank interest rate (*IIR*). All time to maturities available are displayed. The corridor corresponds to the selected scenario  $Z_{(100|90)}$ , and it is non-symmetrical due to its design (i.e.,  $\pm 100$  basis points with respect to the maximum and minimum *IBR* for the overnight, one-month, and three-month time to maturities).

Other information valuable for monitoring purposes may be available as well. Figure 3 exhibits a scatter plot that relates identified loans cost and time to maturity; cost is expressed as a spread over the interbank overnight funds average rate (*TIB*), whereas time to maturity corresponds to the number of days at inception. Each axis displays a histogram that corresponds to how the cost and the maturity are distributed in the period under analysis. The average (median) spread over *TIB* is  $-0.0043$  ( $-0.0090$ ) basis points, with a minimum and maximum around 0.81 and 0.59, respectively. The average

(median) maturity of interbank loans is about 2.6 (1) calendar days, with most deviations from the 1-day maturity corresponding to overnight loans coinciding with weekends or holidays. Altogether, Figure 3 depicts that most loans have a low time to maturity at inception (78.86% are overnight loans), and most loans spread over *TIB* tend to be rather small. As expected, there is a positive and non-negligible linear dependence between cost and time to maturity, with a correlation coefficient about 0.34; yet, correlation significantly diverging from unity reveals that –as expected– there is more to cost than time to maturity (e.g., credit risk, market liquidity).

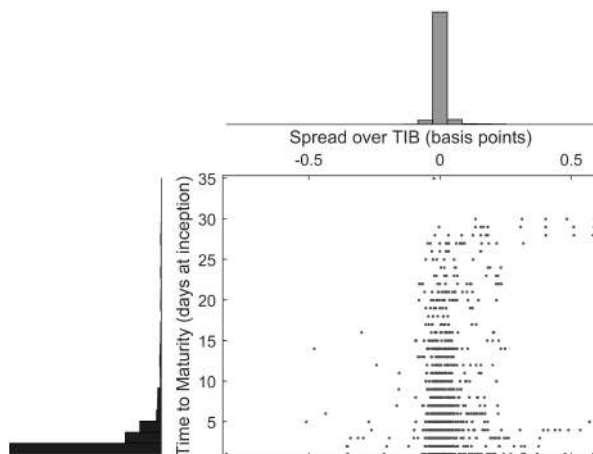
**Figure 2.** *The implicit interbank rate (IIR)*



Source: authors' calculations.

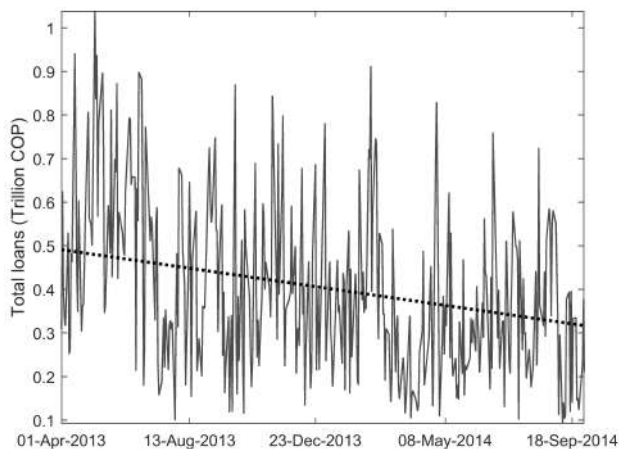
Figure 4 displays the total value of contracted loans for each day in the sample. A decreasing trend is noticeable. This is consistent with two changes in Central Bank's monetary stance depicted in Figure 1. First, in March 2013 the Central Bank halted a 12-month period of rate reductions. Second, in April 2014 the Central Bank began increasing its intervention rate. These two changes may explain to some extent the decreasing trend in interbank loans during the period under analysis.

**Figure 3.** *Distribution of time to maturity of loans (at inception)*



*Source:* authors' calculations.

**Figure 4.** *Value of new loans and linear trend*



*Note:* The average value of loans extended each day is about 0.4 Trillion \$COP, ranging from 0.1 to 1.0 Trillion \$COP.

*Source:* authors' calculations.

### ***B. Contrasting the selected scenario with reported data***

Reports or surveys from financial institutions are the most common source of interbank data. Two types of reports are available from the Colombian Financial Superintendency, both containing different information reported by local financial institutions.

The first type is a report on the interbank funds transactions occurred during the previous business day. It is aggregated by lender, and reports the total value and weighted average interest rate of the loans for three maturities (i.e., overnight, 2-5 days, more than 5 days). After an exhaustive validation process, this report is used by the Central Bank to calculate the interbank overnight funds average rate (*TIB*). However, it does not contain information about the borrower; it is aggregated by maturity, and does not include the pending amount (i.e., the value of the claim). Hence, this type of report is useless for examining and monitoring individual loans and their financial conditions, or for calculating the exposures between financial institutions and building the corresponding interbank claims networks. Consequently, this type of report is discarded for the purpose of this article.

The second type of report discloses each outstanding interbank loan separately, reports the lender, borrower, and pending amount. Each outstanding interbank loan is identified by a unique code, which we use to determine the date in which the loan was contracted. Although the data has a daily frequency, it is made available to the Central Bank in batches transmitted with a lag of between one and two weeks.

As with most reports by financial institutions to financial authorities, the steps involved (i.e., consolidating, processing, transmitting) may result in lagged information and potential errors. Also, as contrasting the reported data is difficult and costly –if possible– for financial authorities, it is uncertain to what extent the information reported is reliable and complete. Furthermore, as the Colombian Financial Superintendency only requires credit institutions to report their interbank funds transactions, other financial institutions that are allowed to participate in the interbank funds market (e.g., brokerage firms) may not be considered.

In contrast, transactional data are available with a minimum lag (i.e., right after the market closes for overnight loans), and it may be monitored in real-time (i.e., as transactions are registered). As the algorithm requires the refunds to match the loans, there is at least a 1-day lag for the overnight loans. Unlike reports required by the Financial Superintendency, due to the non-tiered (i.e., direct) nature of the local large-value payment system, interbank funds transactions with or among non-credit institutions are captured and readily available.

Unfortunately, reliability and completeness of interbank funds transactional data in our case are uncertain as well. The quality of transactional data depends on how careful and truthful financial institutions are when using the codes assigned for registering transactions in the large-value payment system. As with reported data, it is difficult and costly –if possible– for financial authorities to verify that all interbank funds transactions are properly registered.

As the reliability and completeness of both data sources (i.e., transactional and reports) are uncertain, contrasting our results is not straightforward. We are trying to assess the quality of an algorithm run on a dataset whose quality cannot be verified by contrasting its results with financial institutions' reports whose quality has not been verified either. Thus, under the (unverifiable) assumption of reliability and completeness of financial institutions' raw reported data, we expect to find minor and reasonable differences when contrasting transactional data.

The contrast is as follows. We use the second type of report provided by the Colombian Financial Superintendency for the corresponding period (April 1, 2013-September 30, 2014). From this source of raw reported data, we build a consolidated and revised dataset containing the original loan and its main financial features (i.e., contracting date, lender, borrower, loan value, interest rate, and maturity).

Based on the consolidated and revised version of raw reported data we make four contrasts. First, we contrast the number and value of loan transactions identified from both sources; the less dissimilar the datasets, the less disparate the number and value of loan transactions. Second, we calculate the proportion of loan transactions identified from transactional data that coin-

cide with those identified from reported data by date, lender, borrower, and loan value; the higher the proportion, the lower the occurrence of Type 1 errors (i.e., false positives). Third, we discriminate the interest rate matching precision of transactions coinciding from reported and transactional data; we discriminate between exact matches (i.e., zero two-decimal digit absolute difference) and approximate matches (i.e., absolute difference between 1 and 30 basis points). Fourth, we calculate the proportion of loan transactions from reported data that are also identified from transactional data by date, lender, borrower and loan value; the higher the proportion, the lower the occurrence of Type 2 errors (i.e., false negatives).

Table 4 presents the outcome of the contrast. The number and value of identified loans from each data source is rather dissimilar, with reported loans about 1.46 and 1.25 times that from transactional data, respectively. A reasonable explanation for such disparity is related to the loan-by-loan nature of reports required by the Financial Superintendency, even if the conditions (i.e., date, interest rate, maturity) could allow some aggregation. Also, as the large-value payment system charges a per-transaction fee,<sup>10</sup> financial institutions may find optimal aggregating loans into a single registry, and may also aggregate refunds occurring in the same day. An alternative explanation may be the settlement of interbank loans and refunds outside the large-value payment system (e.g., by the delivery of cash or checks). Likewise, the separate settlement of the principal and interest may further complicate the identification of loans and their refunds; yet, the per-transaction fee charged by the large-value payment system should dissuade financial institutions from such separation. Therefore, such disparity may not be surprising, and should also increase the occurrence of Type 2 errors (i.e., false negatives).

The number and value of loans from transactional data that match reported data by date, lender, borrower, loan amount and interest rate is high, 99.70% and 99.83%, respectively; conversely, the number and value of loans from transactional data that do not match reported data is rather low, 0.30% and 0.17%, respectively. Exact interest rate matches (i.e., two-decimal digit

---

<sup>10</sup> The large-value payment system charges \$COP 2,580 (around \$USD 0.86) for every transaction that occurs between 00:00 and 17:00. From 17:00 to 24:00, the per-transaction fee is \$COP 2.5 per million, with a minimum of \$COP 3,730 (around \$USD 1.24).



precision) occur for 91.51% and 88.56% of the number and value of loans from transactional data, respectively. 8.19% and 11.27% correspond to approximate matches, which have an average absolute difference about 2 basis points. All in all, the proportion of loans from transactional data that match reported data suggests that Type 1 errors (i.e., false positives) are rare. Furthermore, as wrong matches can be considered a subset of false positive errors (Arciero et al., 2013), the low occurrence of Type 1 errors and the dominance of exact matches may be interpreted as a signal of a low occurrence of Type 3 errors (i.e., wrong matches).

**Table 4.** *Contrasting transactional and reported data*

| Parameter  |                    | By number of<br>loans <sup>a</sup> | By value of<br>loans <sup>b</sup> |
|--|--------------------|------------------------------------|-----------------------------------|
| Identified loans   | Transactional data | 10.020                             | 142.93                            |
|  | Reported data      | 13.923                             | 177.99                            |
| Proportion of loans from transactional data that match reported data <sup>c</sup>                  |                    | 99.70%                             | 99.83%                            |
| <i>Exact interest rate matches (i.e., two-decimal digit absolute difference precision)</i>         |                    | <i>91.51%</i>                      | <i>88.56%</i>                     |
| <i>Approximate interest rate matches (i.e., absolute difference between 1 and 30 basis points)</i> |                    | <i>8.19%</i>                       | <i>11.27%</i>                     |
| Proportion of loans from transactional data that do not match reported data <sup>c</sup>           |                    | 0.30%                              | 0.17%                             |
| Proportion of loans from reported data that match transactional data <sup>c</sup>                  |                    | 71.81%                             | 80.25%                            |

<sup>a</sup>Number of loans, unless otherwise stated. <sup>b</sup>Trillion \$COP, unless otherwise stated. <sup>c</sup>A match is based on the exact coincidence of date, lender, borrower, and loan amount, whereas the interest rate match considers exact and approximate coincidences.

*Source:* authors' calculations.

Concurrent with the number and value of loans from reported data exceeding those from transactional data, the proportion of loans from reported data that match transactional data is 71.81% and 80.25%, respectively. This suggests that Type 2 errors (i.e., false negatives) are non-negligible. A typical source of Type 2 error is a narrow corridor of plausible rates; however, Tables 1 and 2 show that widening the corridor does not increase the number or value of matched loans notably. Alternatively, as stated before, misusing large-value payment system's registration codes, the aggregation of loans in a single register in transactional data, the separate settlement of principal and interests, and the settlement of loans and refunds outside the large-value payment system may explain differences in the number and value of loans. Still, as the reliability and completeness of reported data is not verifiable, it is uncertain to what extent the excess of reported data and the occurrence of Type 2 errors are related to the quality of both reported and registered information.

### ***C. Contrasting the implicit interbank interest rate with market data***

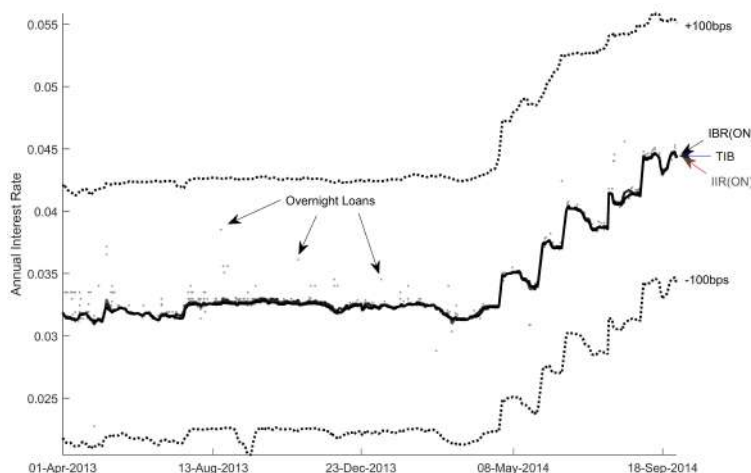
As customary in the literature (see Millard & Polenghi, 2004; Heijmans et al., 2010; Arciero et al., 2013), we contrast our implicit interbank interest rate with the existing interbank interest rate benchmarks. In Colombia, there are two publicly available benchmarks: First, the interbank overnight funds average rate (*Tasa Interbancaria – TIB*), which is calculated and reported by the Central Bank based on financial institutions' reports to the Colombian Financial Superintendency. *TIB* is calculated after an exhaustive validation process of financial institutions' raw reported data, thus it is a reliable and comprehensive benchmark. Second, the interbank overnight reference rate (*Índice Bancario de Referencia Overnight – IBR<sub>ON</sub>*), which is the overnight rate of the interbank reference rate formation program. As *IBR<sub>ON</sub>* results from the interest rate formation program, it is a reliable and comprehensive benchmark as well.

As both benchmarks are overnight rates, we build the corresponding implicit interbank overnight interest rate (*IIR<sub>ON</sub>*). That is, from the selected scenario ( $Z_{(100|90)}$ ) we discard all loans with maturities greater than one business day. Figure 5 displays *IIR<sub>ON</sub>*, *IBR<sub>ON</sub>* and *TIB* for the period under analysis

(1 April, 2013-30 September, 2014). Due to the reliability and comprehensiveness of  $TIB$  and  $IBR_{ON}$ , divergences displayed by the implicit interbank overnight interest rate ( $IIR_{ON}$ ) should correspond to the features and limitations of our algorithm and datasets.

The similarity between the three interbank overnight interest rates is evident. Differences are scarce, and they are not prominent. Concurrent with loans cost distribution (see Figure 3), the implicit overnight loans (the dots) do not deviate too much from the three interest rates in the middle of the corridor.

**Figure 5.** *Interbank overnight interest rates comparison*



*Note:* Dots represent the overnight loans identified by the algorithm, whereas the three lines in the middle of the  $\pm 100$  corridor correspond to the implicit interbank overnight interest rate ( $IIR_{ON}$ ), interbank overnight reference rate ( $IBR_{ON}$ ), and the interbank overnight funds average rate ( $TIB$ ). The three interbank overnight interest rates do not deviate significantly from each other. Most loans (i.e., dots) have implicit interest rates that do not deviate too much from the three interbank overnight interest rates.

*Source:* authors' calculations.

Table 5 confirms the linear dependence (i.e., correlation) among the three overnight interest rates level and serial differences, above and below the main

diagonal, respectively. As in Heijmans et al. (2010), this correlation decreases the chance of Type 1 and Type 2 errors.

**Table 5.** *Correlation of the three overnight interest rates<sup>a</sup>*

|            | $IBR_{ON}$ | $TIB$  | $IIR_{ON}$ |
|------------|------------|--------|------------|
| $IBR_{ON}$ | 1          | 0.9997 | 0.9998     |
| $TIB$      | 0.9674     | 1      | 0.9997     |
| $IIR_{ON}$ | 0.9715     | 0.9704 | 1          |

*Note:* <sup>a</sup>Correlation on the level (differences) is above (below) the main diagonal.

*Source:* authors' calculations.

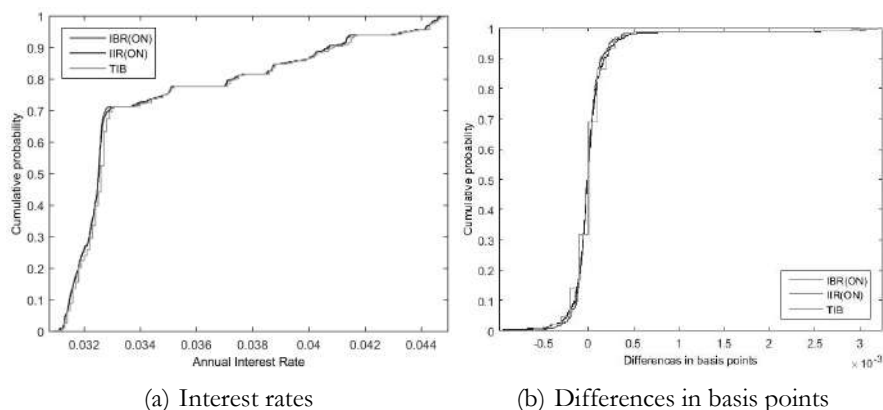
Likewise, Figure 6 confirms the correspondence of the cumulative distributions of interbank overnight interest rates and their differences. The two-sample Kolmogorov-Smirnoff test does not reject the null hypothesis that data in  $IBR_{ON}$  and  $IIR_{ON}$  are from the same continuous distribution at any typical significance level, either for interest rates ( $p$ -value = 0.87) or their differences ( $p$ -value = 0.82). However, this test is rejected when the two samples are  $TIB$  and  $IBR_{ON}$  or  $TIB$  and  $IIR_{ON}$ .<sup>11</sup>

#### **D. Interbank claims networks**

The algorithm does not only allow obtaining interbank interest rates from transactional data, or loans time to maturity and value. It also allows building the corresponding interbank claims networks in a straightforward manner.

<sup>11</sup> As visual inspection of cumulative distribution reveals (Figure 5),  $TIB$ 's level and variation appear to follow discrete changes. Therefore, a preliminary conjecture may relate the rejection of the null hypothesis in the Kolmogorov-Smirnoff test for  $TIB$  to its distributions divergence from a continuous distribution.

**Figure 6.** *Cumulative distributions of interbank overnight interest rates and their differences*



*Note:* visual inspection of the cumulative distributions reveals that the distributions of the three interbank overnight interest rates are rather similar. The two-sample Kolmogorov-Smirnoff test does not reject the null hypothesis that data in  $IBR_{ON}$  and  $IIR_{ON}$  are from the same continuous distribution at the 1% significance level, either for interest rates or their differences.

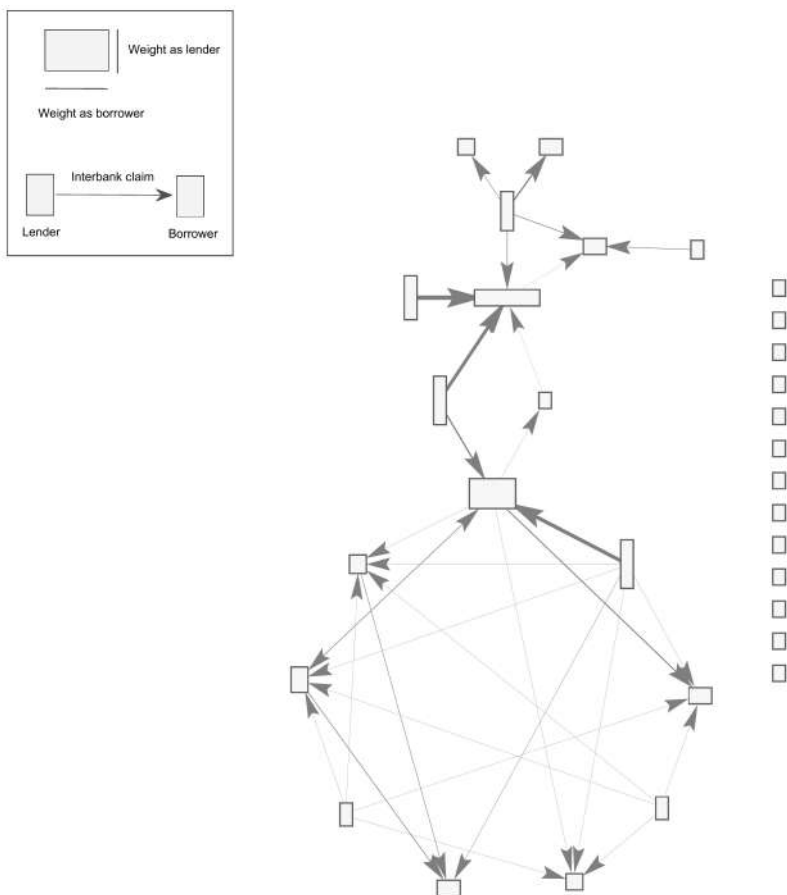
*Source:* authors' calculations.

The procedure added to the typical Furfine's method consists of organizing loans in matrix form, building a hypermatrix (i.e., a cube) of loans, with dimensions  $N \times N \times T$ .  $T$  corresponds to the number of days in the sample ( $t = 1, 2, \dots, T$ ), and  $N$  corresponds to the number of observed participants all over the sample ( $n = 1, 2, \dots, N$ ). Each  $t$ -layer of the hypermatrix will accommodate the cumulated loans granted from financial institution  $i$  to financial institution  $j$  up to day  $t$ . Conversely, all refunds should be organized accordingly. Subtracting each layer in the refunds hypermatrix from the loans hypermatrix will yield the exposure (i.e., outstanding loan) hypermatrix, with each  $t$ -layer containing the interbank claims that  $i$  holds from  $j$  at day  $t$  (i.e.,  $j$  has an outstanding loan from  $i$  at  $t$ ). Hence, each resulting layer is the weighted adjacency matrix of interbank claims.

The graph corresponding to a randomly selected weighted adjacency matrix of interbank claims is portrayed in Figure 7. Nodes or vertexes (in rectangles) correspond to a financial institution that participated at some time  $t$  throughout the sample; for the period under analysis (1 April, 2013-30

September 2014), there are 31 participating financial institutions. The height (width) of each node corresponds to its contribution to the total claims of the market as a lender (borrower). The direction of the arrows (i.e., arcs) represents the existence of an interbank claim (i.e., from the lender to the borrower), whereas their width represents its contribution to the total claims of the system.

**Figure 7.** *Interbank claims network for a randomly selected date*



*Source:* authors' calculations.

## Final remarks

By implementing Furfine's method on a dataset from the Colombian large-value payment system this paper attained three main objectives. First, we filtered out interbank funds transactions in the Colombian market in order to identify interbank loans without relying on data reported by financial institutions. Results suggest that the algorithm is useful and robust as an alternative source of interbank loans data.

Second, we contrasted the loans identified by our algorithm with those identified from reports by financial institutions. Most loans identified by the algorithm coincide with reported data, and most of these coincidences are exact matches. Furthermore, the implicit overnight interest rates successfully match the publicly available overnight rates benchmarks. Such convergence of the implicit overnight interest rate is a fair test of the algorithm's goodness of fit. Interestingly, by running the algorithm backwards we mitigated the over-identification of long-term interbank loans in a straightforward manner.

Third, we constructed a hypermatrix containing the interbank claims network for each day in the sample from transactional data. As illustrated during and after the Great Financial Crisis, it is important to highlight that claims networks are vital for examining financial contagion and systemic risk (see Allen & Gale, 2000; Battiston, Delli et al., 2012; Battiston, Puliga et al., 2012; Kambhu, Weidman & Krishnan, 2007; Haldane, 2009; León & Berndsen, 2014).

Three main contributions from our results are worth highlighting. First, consistent with the lack of timely and granular information during the Great Financial Crisis, we provide further evidence on the usefulness of large-value payment systems' data for monitoring, overseeing, and analyzing the interbank funds market.

Second, we emphasize the importance of recent efforts by financial authorities to construct datasets in which unsecured transactions may be filtered out with ease. It is clear that efforts aiming at unusual datasets like ours or at collecting transaction-level data from financial institutions are advantageous. In this vein, as highlighted by Bank of England (2015), enhanced information

on money markets benefits financial authorities' analysis of both monetary policy and financial stability.

Third, by running the algorithm backwards (i.e., starting with the last date of the sample until reaching the first one) we attained a practical method for mitigating the over-identification of long-term interbank loans, an issue also encountered by Guggenheim et al. (2010) and Heijmans et al. (2010). To the best of our knowledge, this backwards-run setup has not been attempted in related literature before, and may be worth exploring in further implementations of Furfine's algorithm.

Challenges arising from this methodological article come in several forms. First, attained numerical results allow for implementing financial contagion and systemic risk models, such as DebtRank (Battiston, Puliga, et al., 2012). Second, consolidating other sources of financial claims (e.g., derivatives, collateralized loans) into a multi-layer network may provide a broader view of financial contagion (see Poledna et al., 2015). Likewise, incorporating non-financial institutions (e.g., households, firms) into a multilayer network may provide a comprehensive view of systemic risk—as in De Castro and Tabak (2013). Finally, comprehensively tracking the dynamics of financial institutions' interbank loans (e.g., their cost, maturity, and counterparties), amid the dynamics of the interbank claims network (e.g., density, and distribution of links), may provide valuable information for authorities contributing to financial stability, and to formulating monetary policy.

## References

- ALLEN, Franklin & GALE, Douglas (2000). "Financial Contagion", *Journal of Political Economy*, Vol. 108, No. 1, pp. 1-33.
- ALLEN, Franklin; CARLETTI, Elena & GALE, Douglas (2009). "Interbank Market Liquidity and Central Bank Intervention", *Journal of Monetary Economics*, Vol. 56, pp. 639-652.
- ARCIERO, Luca; HEIJMANS, Ronald; HEUVER, Richard; MASSARENTI, Marco; PICILLO, Cristina & VACIRCA, Francesco (2013). "How to Measure the



- Unsecured Money Market? The Eurosystem's Implementation and Validation Using TARGET2 Data". *DNB Working Paper 369*. De Nederlandsche Bank.
- ARMANTIER, Olivier & COPELAND, Adam (2012). "Assessing the Quality of Furfine-Based Algorithms", *Staff Report 575*. Federal Reserve Bank of New York.
- BANCO DE LA REPÚBLICA (2015). *Reporte de Sistemas de Pago*. Retrieved from: <http://www.banrep.gov.co/es/reporte-sistemas-pago> (5 August 2015)
- BANK OF ENGLAND (2015). *A New Sterling Money Market Data Collection and the Reform of SONIA: Public Consultation*. Retrieved from: <http://www.bankofengland.co.uk/publications/Pages/news/2015/062.aspx> (12 September 2015)
- BATTISTON, Stefano; DELLI, Domenico; GALLEGATI, Mauro; GREENWALD, Bruce & STIGLITZ, Joseph (2012). "Default Cascades: When Does Risk Diversification Increase Stability?", *Journal of Financial Stability*, Vol. 8, pp. 138-149.
- BATTISTON, Stefano; PULIGA, Michelangelo; KAUSHIK, Rahul; TASCIA, Paolo & CALDARELLI, Guido (2012). "DebtRank: Too Central to Fail? Financial Networks, the FED and Systemic Risk", *Scientific Reports*, Vol. 2, No. 541.
- CASTIGLIONESI, Fabio & WAGNER, Wolf (2013). "On the Efficiency of Bilateral Interbank Insurance", *Journal of Financial Intermediation*, Vol. 22, 177-200.
- DE CASTRO, Rodrigo & TABAK, Benjamin (2013). "Contagion Risk within Firm-Bank Bivariate Networks". *Working Paper Series 322*, Banco Central do Brasil.
- DEMIRALP, Selva; PRESLOPSKY, Brian & WHITESELL, William (2004). "Overnight Interbank Loan Markets". *Manuscript Board of Governors of the Federal Reserve*. Retrieved from: <http://www.federalreserve.gov/pubs/feds/2004/200429/200429pap.pdf> (10 April 2015).

- FRICKE, Daniel & LUX, Thomas (2014). “Core-Periphery Structure in the Overnight Money Market: Evidence From the e-MID Trading Platform”, *Computational Economics*, Vol. 45, No. 3, pp.359-395. doi:10.1007/s10614-014-9427-x.
- FURFINE, Craig (1999). “The Microstructure of the Federal Funds Market”, *Financial Markets, Institutions and Instruments*, Vol. 8, pp. 24-44.
- FURFINE, Craig (2001). “Banks Monitoring Banks: Evidence From the Overnight Federal Funds Market”, *Journal of Business*, Vol. 74, No. 1, pp. 33-58.
- GUGGENHEIM, Basil; KRAENZLIN, Sébastien & SCHUMACHER, Silvio (2011). “Exploring an Uncharted Market: Evidence on the Unsecured Swiss Franc Money Market”. *SNB Working Papers 2011-5*. Swiss National Bank.
- HALDANE, Andrew (2009, April). “Rethinking the Financial Network”. Speech Delivered at the Financial Student Association, Amsterdam, Netherlands.
- HEIJMANS, Ronald; HEUVER, Richard & WALRAVEN, Danielle (2010). “Monitoring the Unsecured Interbank Money Market Using TARGET2 Data”, *DNB Working Paper 369*. De Nederlandsche Bank.
- HENDRY, Scott & KAMHI, Nadja (2007). “Uncollateralized Overnight Loans Settled in LVTS”. *Bank of Canada Working Paper 07-11*. Bank of Canada.
- KAMBHU, John; WEIDMAN, Scott & KRISHNAN, Neel (2007). “New Directions for Understanding Systemic Risk”, *Economic Policy Review*, No.13, No. 2.
- KYRIAKOPOULOS, Franziskos; THURNER, Stefan; PUHR, Claus & SCHMITZ, Stefan (2009). “Network and Eigenvalue Analysis of Financial Transaction Networks”, *The European Physical Journal B*, Vol. 71, pp. 523-531.
- LEÓN, Carlos & BERNDSEN, Ron (2014). “Rethinking Financial Stability: Challenges Arising From Financial Networks’ Modular Scale-Free Architecture”, *Journal of Financial Stability*, Vol. 15, pp. 241-256. doi:10.1016/j.jfs.2014.10.006.

- MARTINEZ, Constanza & LEÓN, Carlos (2015). “The Cost of Collateralized Borrowing in the Colombian Money Market: Does Connectedness Matter?”, *Journal of Financial Stability* (accepted paper). doi:10.1016/j.jfs.2015.10.003.
- MILLARD, Stephen & POLENGHI, Marco (2004). “The Relationship Between the Overnight Interbank Unsecured Loan Market and the CHAPS Sterling System”, *Bank of England Quarterly Bulletin, Spring*. Bank of England.
- POLEDNA, Sebastián; MOLINA-BORBOA, José Luís; MARTÍNEZ-JARAMILLO, Serafín; VAN DER LEIJ, Marco & THURNER, Stefan (2015) “The Multi-Layer Network Nature of systemic Risk and its Implications for the Costs of Financial Crises”, *Journal of Financial Stability*, Vol. 20, pp.70-81.
- ROCHET, Jean-Charles & TIROLE, Jean (1996). “Interbank Lending and Systemic Risk”, *Journal of Money, Credit and Banking*, Vol. 4, No. 28, pp. 733-762.