Payment Delays and Contagion*

Ben Craig† **Dilyara Salakhova**‡ **Martin Saldias**§

First version: December 29, 2013. This version: November 14, 2014

Abstract

This paper provides a characterization of delays in interbank payments via TARGET2 payment system and conducts a first approximation to the determinants whereby a delay in incoming transactions may cause a delay in other transactions downstream. In contrast to the theoretical literature, our results show that banks do not incur in game-theoretic games but rather set up persistent liquidity strategies. These strategies in turn help identify two types of banks in terms of their initial liquidity allocations and propensity to delay their payments. Results in both cases are robust to alternative environments of high and low market liquidity and distress.

JEL classification: C26, E42, G21.
Keywords: Payment delays, probit models with endogenous regressors, liquidity.

*The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Bank of Cleveland, the Banque de France or the Eurosystem and its staff, the IMF. Martin Saldias worked on the paper while affiliated at Banco of Portugal.

†Cleveland Fed, Deutsche Bundesbank

‡Banque de France, University Paris West Nanterre, dilyara.salakhova@banque-france.fr

§International Monetary Fund

1. Introduction

The smooth functioning of financial infrastructures is crucial for financial stability in a wider sense. Payment systems constitute a core element and play a central role in transmission of liquidity shocks and contagion under market distress that may lead to systemic events. Payment systems also serve as a tool to early identify and then monitor systemic liquidity stress through the behavioral patterns of its participants[1].

Changes in regular patterns of transactions' volumes and values have proven highly informative in the recent crisis period. For instance, during the Lehman Brothers failure, payment markets were among the first to react and show signs of stress, which lasted for several weeks. Banks chose then to delay their payments or reassess their degree of involvement in the market due to concerns of counterparty default risk (Benos *et al.*, 2012).

In this paper, we focus on payment delays –the time passing between when a payment is entered and when it is cleared– as they have highly relevant implications for systemic liquidity risks and also allow to early identify potential changes of behavior of market participants. Delays tend to be reduced as queueing and liquidity-saving mechanisms are put in place to improve overall efficiency in the system design. However, as the potential remaining delays may quickly reflect strategic responses to general or particular shocks of different nature, their determinants and motivations deserve attention.

Our first contribution is a characterization of delays in interbank payments via TARGET2 payment system that is consistent with findings in relevant theoreti-

[1] See Manning *et al.* (2009) for a comprehensive summary of theory and practice of large-value payment systems (LVPS), including delays. Rochet and Tirole (1996) provide additional insights in net settlement systems and payment systems design.

cal and empirical work, including time clustering Armantier *et al.* (2008); Bartolini *et al.* (2010); Proepper *et al.* (2008); Massarenti *et al.* (2013), liquidity management trade-offs (Angelini, 1998; Bech and Garratt, 2003; Bech, 2008), operational risks Merrouche and Schanz (2010), *inter alia*. We then contribute with a first approximation to the determinants whereby a delay in incoming transactions may cause a delay in other transactions downstream under different liquidity and market conditions using probit models with endogenous regressors. Our econometric approach allows us to isolate a contagion phenomenon in normal times and the presence of common factors driving the delays under distressed market conditions.

Our estimation results show that banks do not systematically take strategic bilateral decisions towards other participants in a payment-by-payment basis as each payment is usually too small to induce a strategic game. Conversely, banks take choices on the liquidity to start operating in the system at the beginning of the day and a mechanical process runs on its own throughout the day after the initial decisions. Our estimates are robust and benefit from behavioral patterns of participant banks that we identify in the network structure of TARGET2 as instruments. In particular, we consider temporary losses of liquidity as a result of a large spike in payments needs without an accompanying increase in liquidity, transmission of delay information throughout the network nodes (Blume *et al.*, 2010) and interplay between free riding and reputation risks in bilateral punishment equilibria (Diehl, 2013).

Based on this finding, we then assess the relevance of initial liquidity setup on the strategies of banks throughout the day. Initial liquidity of banks constitutes of the beginning-of-the-day-balance and credit line at the central bank supported by the collateral. Throughout the day, banks make and receive hundreds of payments, and though they may be fairly clear about what payments will come and go out, and the total match between initial liquidity set and in- and out-payments, they face uncertainty about the stream of payments. We find that there are two types of banks in terms of their initial liquidity and their propensity to delay on its payments: banks that set enough initial liquidity to run daily payments (low-α banks) and those which systematically put less than needed initial liquidity (high-α). In particular, high-α banks always rely on payments that come into the system to clear their payments and they account for the majority of the delays observed. In contrast, low-α banks are banks with large initial liquidity and rarely incur in payment delays. Interestingly, they account for the majority of the distribution. Given a low or high α status, this is persistent over time and across the two subsamples that characterize different market conditions. Finally, the division of the sample in two groups allows us to identify four regions of flows, where delays happen more often among high-α banks and when flows go from a high-α to a low-α bank. This result can provide an early warning in terms of market freezes, as delays among low-α banks are rare even under low market liquidity environments.

The remainder of the paper is organized as follows. Section 2 provides a concise review of the literature on delays in payment systems, both from a theoretical and an empirical point of view. The literature suggests that banks are involved in a game with each other, where the strategies are chosen to prevent free riding of liquidity at each payment, where the strategic instruments chosen are bilateral limits on the implicit credit for delay of a particular counter-party and the delay in placing a payment into the system for a counterparty. We show that bilateral limits are rarely chosen strategically, but the delay in placing a payment into the system for a counterparty is difficult to assess empirically. We describe the empirical application in order to obtain and analyze the delays origination and transmission in Section 3.

Results are provided in Section 3.3 where individual payment strategic interactions are shown to be a problematic view of the payments system. This leads us to explore a different approach to delays in intra-day liquidity management which focuses on the initial decisions to set a liquidity reservoir at the beginning of the day. We explore a new way to normalize liquidity decisions to account for the very different payment patterns that face each bank in 4.1. We show that our normalization is bimodal and persistent within a bank in 4.2. We then examine the payment and network patterns that contribute to having a bank to choose low liquidity relative to its payment patterns and find that total value of payments going out or coming in are not important but that network position is in 4.3 and look at the network implications of heterogeneity in α in 4.4. Conclusions and discussion of future research are summarized in Section 5.

2. Literature review

A number of contributions address payment delays from a theoretical perspective and with a game-theoretic approach. They generally link the delays to liquidity as a change of behavior of market participants to shocks or to the payment system structure. For instance, Arjani (2006) conducts a simulation analysis applied to the Canadian LVTS in order to tackle the cost trade-off between settlement delays and intraday liquidity needs in order to improve the efficiency of the system. The author is cautious about the sensitivity of the analysis' results to the assumptions about the behavior of the system participants, which highlights the need for the model to include dynamics and strategic behavior. Along these lines, Schulz (2011) tackles the liquidity-delay trade-off in a simulation exercise calibrated with real world parameters. The author also highlights the sensitivity of the results to behavioral assumptions, pointing out to the need to analyze and characterize them in the data, such as via TARGET2 transactions, in order to obtain a realistic simulations results.

From a game-theory perspective, Bech and Garratt (2003) and Bech (2008) characterize the interaction between intraday liquidity management and payment delays as a coordination game and provide rationale to the payment and delays timing. Buckle and Campbell (2003) show in a theoretical model that delays in a real-time gross settlement (RTGS) systems are likely to occur if banks care about bilateral payment imbalances.

In the empirical literature, Massarenti *et al.* (2013) provide a first and very thorough characterization of the intraday patterns of payments in TARGET2 between 2008 and 2011. The authors identify a number of highly relevant features, including stable and regular payment timing, relationships between interbank transactions and trading activity in financial markets and time clustering for different sub-systems within TARGET2. The authors also tackle delay payments along these lines and find that delays respond to timing clustering, which indicates on one hand strategic liquidity management behavior but also creates foreseeable contexts where payment delays might be more prone to create systemic liquidity distress.

Bartolini *et al.* (2010) match brokered trades and Fedwire payment orders and provide a thorough analysis of payment delays. The authors also find that payment delays can be to some extent predictable due to their time clustering and therefore trigger high-frequency liquidity management decisions to counteract resulting liquidity shortages. They also identify different strategies of market participants, such as the preference of delays of large transactions relative to small trades or, to a lesser extent, delaying settlement when own liquid balances are low. Benos *et al.* (2012) addresses payment delays in the British RTGS or CHAPS[2] in the aftermath of Lehman Brothers bankruptcy as a exposure to counterparty default risk in a con-

[2]The authors also provide an alternative definition of delays to the one studied in this paper.

text of abundant market liquidity. Finally, Heijmans and Heuver (2011) analyze the Dutch part of TARGET2 and find that delays tend to signal changes of behavior of vis--vis its counterparties.

3. Testing payment-by-payment strategic games

3.1. Data and preliminary analysis

Our source dataset comprises all interbank payments at transaction-level settled via TARGET2, the real-time gross settlement (RTGS) system owned and operated by the Eurosystem[3] since November 2007. We chose two days to apply our delay categorization and econometric model. In particular, we studied the delay patterns on 16 July 2010 as our *calm day*, and on 10 March 2009 as our *stress day*. The choice of these two days is based on overall transaction volume in the interbank and bank equity markets and the volatility in the latter relative to the historical trend, i.e. our calm day shows average behavior whereas both volume and transaction volatility are much higher on our stress day.

Based on this rich dataset, we first distinguish between delayed and non-delayed payments. In particular, a payment in our dataset is flagged as delayed if its execution time, i.e. the difference between the point of time that a payment enters the system and the time that it actually clears, exceeds 5 minutes during the last hour[4]. Most of the payments are introduced to the system immediately when pay-

[3]TARGET2 processes payments by three categories of participants: 1) Central banks payments; 2) Ancillary systems; and 3) Commercial institutions. In our analysis, we are interested only in transactions between commercial institutions which make the majority of all the transactions in the system and are also responsible for the most delays. There are 20 classes of transactions, but interbank and customer payments make together 70% of the total number. They are also the ones that are the most delayed. For the current version, we keep all the types of payments between commercial institutions.

[4]This definition excludes delays caused by technical processing time difficulties and we are not interested in the duration of the delay but on its relevance as a determinant of further delays.

ments messages are sent by participants; however, participants have some flexibility to decide on the entry and clearing time by setting earliest debit time. It provides a possibility to send a payment message, e.g., at 7am, to indicate at what time the payment will be introduced to the system, e.g., at 10am. Moreover, payment messages can be sent to the system before the starting of the settlement operations. Therefore, the introduction time is defined as the maximum between the time the payment is sent to the system and earliest debit time if the business date the payment is settled corresponds to the business date the payment is introduced to the system; or 7am otherwise.

Based on this rich dataset, we first distinguish between delayed and non-delayed payments. In particular, a payment in our dataset is flagged as delayed if its execution time, i.e. the difference between the point of time that a payment is introduced to the system and the time that it is actually settled, exceeds 5 minutes during the last hour[5]. Most of the payments are introduced to the system immediately when payments messages are sent by participants; however, participants have some flexibility to decide on the introduction and clearing time by setting earliest debit time. It provides a possibility to send a payment message (e.g., at 7am) in order to indicate at what time the payment will be introduced to the system (e.g., at 10am). Moreover, payment messages can be sent to the system before the starting of the settlement operations. Therefore, the introduction time is defined as the maximum between the time the payment is sent to the system and earliest debit time if the business date the payment is settled corresponds to the business date the payment is introduced to the system; or 7am otherwise.

Delayed payments on our calm day represent 17.06% of the total, while on a stress

[5]This definition excludes delays caused by technical processing time difficulties and we are not interested in the duration of the delay but on its relevance as a determinant of further delays.

day, this percentage is slightly smaller (15.87%), although the number is significantly larger (over 50%). Figures 1 and 2 show the distributions of transaction values of delayed and non-delayed payments. Delayed payments are on average larger than non-delayed ones; however there is graphically no significant difference between the two distributions, meaning that banks do not have any strategy of delaying or not a payment based on the transaction value.

Finally, figures 3 and 4 show the kernel distributions of the sum of incoming delayed payments (in log) separately for those payments that in turn produced new delayed payments. For both calm and crisis days, we observe that total values of incoming delayed payments are much higher for the payments that caused delays themselves (red line), in other words, larger delayed incoming payments are in principle associated to a higher probability of a subsequent payment to be delayed. This relationship is especially pronounced on the calm day, which implies that delays are closely related with delay contagion. On the crisis day, there may be also other reasons behind delays as the figures suggest that banks delay more on a crisis day even without having suffered an incoming payment with a delay.

[Insert Figures 3 and 4 here]

3.2. Econometric model of delay drivers

The important behavior we are estimating with the delay data is the extent to which a delay in one transaction causes a delay in another transaction downstream. In other words, when bank **A** receives a payment or series of payments that are delayed, to what extent does this cause it to delay its payments to other banks? The importance of this question is that if the delayed payments cause the bank to delay its payments, then the network, under certain measurable circumstances, can become explosive, or to mix metaphors, to tip, causing the entire connected network to be in delay.

Measuring the effect of a delayed payment on the probability of a delay propagation is essential to understanding whether the network properties of the system intensify the delays, and under what circumstances could the payments system tip and have delays throughout the system. Our measurement, then, could be written as a probit model with endogenous regressors:

$$y^* = X\beta + Z\gamma + \varepsilon$$

$$y = \begin{cases} 1 & \text{if } y^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

Here y^* is an unobserved latent variable, $\mathbf{X}$ is an observed vector of characteristics (including only the intercept in this application), and $\mathbf{Z}$ is some measure of delays coming into the node. $\boldsymbol{\varepsilon}$ has a known distribution, such as a $N(0,1)$ distribution. The parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are to be estimated, and the coefficient of $\mathbf{Z}$, $\boldsymbol{\gamma}$, especially can be used to describe the behavior of the system, i.e. positive statistical significance would indicate that incoming delayed payments are likely to cause a delay in payments of a given node.

Measuring the causality of delays on further delays in our data is complicated by several issues. First, one or a set of common factors could cause delays to be more probable within the entire system. In this case, delays do not propagate themselves, but can look like they do. In our notation, a common shock affects both $\boldsymbol{\varepsilon}$ and $\mathbf{Z}$, in that this shock hits the entire system and thus affects the individual shock $\boldsymbol{\varepsilon}$ and the probability that the delays occur beforehand, $\mathbf{Z}$, in the sense that after the first responder to the common shock, all subsequent responses must follow that first response, leading to a sense that they are responding to the early responses to the common shock rather than to the common shock itself. The delay variable is endogenous, and so the variables $\boldsymbol{\varepsilon}$ and $\mathbf{Z}$ are correlated and a false causality is measured.

Second, a bank that observes a consistent pattern of delays in a single partner can elect to withhold its liquidity from this bilateral partner until this partner clears its own payments. These two nodes can be in a punishment game equilibrium. Because we are interested in how the delays propagate to the rest of the system, we are interested in the effects of $\mathbf{Z}$ on $\mathbf{y}$ which abstract from this. In this scenario, the variables $\boldsymbol{\varepsilon}$ and $\mathbf{Z}$ are correlated because in cases where node **A** delivers $\mathbf{Z}$ of delayed clearing to node **B**, **A** decides as a punishment to delay clearing to node **B**, and so on, and $\boldsymbol{\varepsilon}$ goes up. We handle this latent endogeneity through the use of instrumental variables, using those delayed payments in $\mathbf{Z}$ that do not have a counterpart of a delayed payment from the paying node as instruments, which we denote instrument set Z_1.

The third set of errors is somewhat more subtle. These are the errors mentioned in Blume *et al.* (2010), where a pair of nodes is more likely to delay together. Either because of proximity, or because of similar business models, or exposure of risk from the same funding sources, this pair of nodes share a common unobserved local factor that drives both of them to delay. We solve this problem in an analogous way to their solution, we use the nodes that are not immediately connected to node **A** as instruments for the delays to node **A**. In particular, we look those nodes that are connected to **A** in delays only through an intermediary in a second layer back as instruments for the delays to **A** and that do not go through the self reflecting nodes, described above. This set of instruments are labeled Z_2. Figures 5 to 8 provide an intuitive representation on the estimation strategy and the choice of the instruments.

[Insert Figures 5 to 8 here]

3.3. Results: Delays' propagation

Results of the models described above are reported in Table 1 for a market calm day and in Table 2 for the stress day. In both cases, the benchmark model is presented in columns [1] and corresponds to the case where no instruments are used and we measure the effect of total (log) value of all incoming payments delayed more than 5 minutes during the last hour on the propagation of delays downstream. Columns [2] show the results of the model where incoming delayed payments are instrumented by Z_1, i.e. delayed payments that do not have a counterpart of a delayed payment from the paying node. Finally, columns [3] show the results of the model where the instrument Z_2, where payments are at least one link removed from the node that produces the delay in order to remove the local reflecting, as described in Section 3.2 .

First, we analyze the calm day. Table 1 shows a positive and statistically significant dependence between the delay of a payment and the value of delayed incoming payments to this bank, in other words, large delayed payments to a given bank are associated with a higher probability of subsequent payment delays to other banks. Results of model [2] show a similar picture, after dealing with the potential endogeneity in measuring the causality of the delays using an instrument for mutual delays. The dependence is still positive and significant, therefore the delays between the two nodes are not explained by a common factor, and delay contagion takes place. Finally, in order to assess whether a delay is explained by delay contagion, that is if a delayed payment of bank i is explained by the delayed payment of bank j to bank i, which itself had a delayed payment by bank k. The results in column [3] also show a positive positive and significant effect of large incoming delays. Based on this evidence, we conclude that on the normal day, a payment is delayed with higher probability if other banks delayed a payment to a bank, and they themselves had delayed incoming payments, that is some sort of contagion of delays mainly due to the mechanics of the system.

The estimation results on the crisis day are shown Table 2 and provide different conclusions from what would be expected from theoretical work, where an increase in contagion should be a result of information propagation, punishment mechanisms, end so forth. In the benchmark model [1], a payment is delayed with higher probability the higher is the size of delayed incoming payments: the coefficient is positive and significant. However, when we tackle the potential endogeneity problem with our data and use the instruments (columns [2] and [3]), the coefficients become negative and significant. This result means that delays on the crisis day are largely explained by mutual delaying and second-round effects, which probably means that all the banks are affected by a common factor which makes them delay payments rather than a contagion mechanism. This result means that a closer analysis of the nature of the common unobserved factors is required along the lines of Hsiao *et al.* (2012).

[Insert Tables 1 and 2 here]

This set of results mean that banks do not systematically take strategic bilateral decisions towards other participants in a payment-by-payment basis as each payment is usually too small to induce a strategic game. They rather indicate that banks make their choices on the liquidity to start operating in the system at the beginning of the day and a mechanical process runs on its own throughout the day after the initial decisions, which embeds some persistent prior decisions on liquidity needs and outflows. Finally, as our data and previous empirical work (Diehl, 2013) reveal that bilateral limits are hardly used if at all on a daily basis, which reinforces the empirical evidence against the notion of strategic games from the theoretical literature.

4. Initial liquidity and delays

4.1. Normalization

In all of the discussion below, we test a different hypothesis, namely that delays in the payments system occur due to banks' certain liquidity management behavior. We examine two different months for liquidity management of the banks. The first month, September 2008, was characterized by a scarcity of liquidity brought on by the uncertainty generated by the crisis. Overnight markets were in the process of breaking down and short term liquidity rates were uncertain. We contrast this with the behavior of the banks during a recent month, May 2014 where markets were awash with short term liquidity, available at extremely low rates.

Each bank, when it sets its liquidity, does so in the knowledge of its own particular daily payments patterns. These highly idiosyncratic time series patterns are crucial in influencing the probability that a delay will occur. For example, a bank that knows that its payments are very likely to alternate between payments that arrive and payments that go out with approximately the same amounts can set a liquidity considerably lower than a bank with payments that go out over the beginning of the day and then come in at the end and it still will have a smaller probability of settling at least one of its payments with a delay. Modeling this process parametrically so that the essential part of the process that sets the delays is very difficult. We opt for the more robust non-parametric procedure that block bootstraps the bank's payments, both incoming and outgoing, during the business part of the day, between 8 am and 4 pm, to come up with a distribution on payments patterns for the bank.

The size of the block is set to be the maximum of 10 payments and the value calculated from Politis *et al.* (2009). The minimum payment number allowed for a day is 30 payments. Otherwise α is not calculated. Although this represents an optimal size for variance rather than the minimum, we experimented in our data with different values and found little variation around this value. The bootstrap is then combined with the bank's initial liquidity to calculate a probability that the bank will delay, given the pattern of payments it confronts. This probability, denoted αin this paper, is analogous to the classical Type I error, α, where a high value implies a large error and thus large delay propensity. The bootstrap was sampled 10,000 times, and this value was also examined in a sample of payment-days. We found that increasing the number of bootstrap replications did not change α more than 1%, which was certainly accurate enough for our purposes.

4.2. Empirical patterns: High- and low-α Bank

Our main finding from the normalization procedure is that some banks persistently miscalculate the amount of initial liquidity to provide to the system, therefore being high-α banks, whereas the majority feed the system with liquidity, being low-α banks. Figures 9 and 10 show the distributions of average of daily α values (over the respective months) for each bank for two periods, September 2008 and May 2014. We see a clear split: a group of banks with α greater and lower than 0.5, especially so for times of scarce liquidity (September 2008). Moreover, figures of transition probabilities, 11 and 12 demonstrate that banks have very regular behavior by staying in the same category from one day to another. This finding points out to a persistent liquidity management policy over time, therefore confirming our hypothesis that banks do not incur in strategic games but rather make their morning decisions about the intraday liquidity allocation.

[Insert Figures 9 to 12 here]

Before analyzing the behavioral patterns of high- and low- α banks, it is worth recalling that the banks in our whole sample are rather big and active, because in order to compute α values, we keep only banks that make on average more 100 daily transactions during the month. Figures 13, 14, 15 and 16 confirm that since for both periods the two types of banks have similar distributions of number and total value of in- and out- transactions, with the only difference that distributions of high-α banks have a bit fatter right tails. Notwithstanding that, throughout the day, the two groups seem to exhibit different patterns (see Figures 17 and 18): high-α banks tend to receive bigger volumes in the morning and send it in the afternoon, while low-α banks do the opposite.

[Insert Figures 13 to 18 here]

High-α banks also delay more (Figures 19 and 20), and the difference in delayed values is particularly remarkable in September 2008 when liquidity was scarce. During that period, high-α banks transact twice as many payments as low-α banks, whereas they delay ten times bigger volume of payments. It is worth particularly noting that in times of scarce liquidity, high-α banks delay especially to other high-α banks. This finding suggests contagious nature of those delays where potential dynamics may look as follows: a bank with short liquidity relies on the incoming liquidity to clear its payments; meanwhile it delays its payments to other banks downstream. If its counterparts are also high-α banks, they will delay their payments as well since they also count on the incoming liquidity.

[Insert Figures 19 and 20 here]

4.3. What determines a High-α Bank?

We collected a large number of possible determinants of α. Because of data confidentiality issues, we were not able to match the banks with balance sheet determinants. However, the payments data themselves gave us a large number of possible explanatory variables with which to explain the bank's liquidity behavior. We divide the day of the bank into four periods. Two periods, "night" (the hours from 4 pm to 7 am in the morning) and "7-9" (early morning hours) represent the evening settlement hours and the early morning settlement hours where the bank realizes its liquidity for the day, where the evening hours are realized payments in for the next day. Two periods, "morning" (9 am to 12 noon) and "afternoon" (12 noon until 4 pm, just before the large net settlements from other payments systems) represent those times of payments where the daily work day payments are handled, after much of the liquidity reservoir for the payments are set. We tried other ways to denote the daily periods, with little or no changes in our results. These four periods of the day gave times where the different behaviors of the payments give rise to different networks. A bank that is more central in the morning will not be as central in the evening.

The payments network is given by four variables, each measuring a different aspect of the network. While we used other variables to assess the behavior of the payments network, these four seemed to characterize our regressions as well as the other sets of variables. We use the Bonacich centrality measure to focus on the information available to the bank from its payments coming in. Bonacich centrality includes a parameter, β, that must be set which determines how the bank has payments coming from other central nodes. The parameter can be set between the largest eigenvalue of the adjacency matrix which defines the network and the negative of this number. If β is larger, then a higher central node is more likely to be connected (in the sense of payments coming in) with other central nodes. If the parameter is lower, then a more central bank is likely to be connected to more peripheral banks. We report coefficients on the Bonacich centrality for extremely high values of β, following the advice of Bonacich (1987) The Bonacich centrality represents the information coming to the bank. Banks that are more central are able to assess the fact that more or fewer payments of value are coming to it from other central banks, thus providing it with information that is not available to more peripheral banks.

A second parameter, "in degree", simply measures the number of links coming into the bank. It is related to Bonacich in that it represents the total number of paying banks that a bank has access to, and so it tells something about the health of the network available as information to the bank. It also represents the complexity of the network of payments coming into the bank. Banks with low in-degree depend on their payments coming in from only a few banks that can be easily monitored, but then also have access to only information from a few nodes, and so can only see a part of the network, rather than the health of the whole.

The third parameter, "weighted in clustering", represents the riskiness of the position of the payments coming into the bank, following Watts and Strogatz (1998). More formally, it measures the probability that two banks who make payments into a node will make a payment to each other. This measures the riskiness in that while a bank may be able to accurately assess the likelihood that a direct counterparty can make its payment on time, the fact that two counterparties are linked creates a correlation between the two, making one delay more likely to generate a delay on the part of the other, delaying payments from both parties. In this sense, a higher clustering coefficient implies a higher probability of a major delay in incoming payments.

Finally, we choose an obvious measure of the position of the node in the network, the absolute size of the payments coming into the bank, which we calculate as the logarithm of the total incoming payments coming into the bank during a time period. We run two types of probit equations with α as the dependent variable. In one case, we choose α to be a continuous variable, and in the other case, we choose a cutoff value for high and low α to be 0.5, so that high α is set to 1 and low to 0 in the probit estimates. The cutoff value was varied and did not really matter, but reflects the bi-modality of the α values. These are denoted 0-1 in Table 3 where we report our results.

[Insert Table 3 here]

For convenience of interpretation, we divide our results into two times of the day. First, those times which might influence how much liquidity is set. This would be the character of the payment inflows of the evening before and the early morning of the day that the bank observes before its liquidity is topped off at the beginning of the day. Second, the character of payments during the day that might affect the noise of the payments, which would lead, in turn, to a more likely delay. The first we call the "morning liquidity" determinants of α. The second, we call the "day complexity" determinants, because they are aspects of the complexity of the patterns of payments that can lead to delays in payments.

Turning to results in Table 3, we see several patterns. First, the determinants of high α is much more strongly associated with the low liquidity September 2008 month than with the high liquidity May 2014. Coefficients on the each of the variables in the 0-1 regression are all significant for the morning variables which determine the initial liquidity set. High α banks are privy to more information that develops during the evening of the day before through their central network position, although this is mitigated a little bit during the early morning hours (although banks that are central to both evening before and early morning networks are quite likely to be high α banks.) The same could also be said about information obtained from a variety of connections, as evidenced from the in-degree coefficients of the evening, although this, also is mitigated slightly in the morning. Banks that shoulder more risk, as evidenced by the high in-weighted clustering are also more likely to be high α, which is not surprising in that being a high α bank is a risky behavior.

Finally, total payments in has little to do with any α behavior. This is very interesting for several reasons. First, the standard chestnut that banks that process a lot of payments have a riskier liquidity profile because they rely on large numbers just is not borne out in the data. Second, a bank's network position is much more likely to affect its risk attitude to delays. Big payments providers are as likely to be high liquidity providers as not, once network position is taken into account

.

When we examine the high liquidity month, network position has generally the same effect, but much weaker. The coefficients are smaller and of less significance throughout the network variables, or the coefficients have no statistical significance. This is what we might expect from a time period where the costs of mistakes in liquidity can be easily solved by recourse to liquidity cheaply available from a variety of sources. However, the once huge difference between the periods lies in the different importance of size. In highly liquid times, the high α banks are likely to be only those who process more payments in the early morning. This may have to do with the fact that high α banks adjust their liquidity at the latest moment when liquidity is cheap.

A look at the lower portion of Table 3 shows less strong influence of information than in the morning, but a continuing association of high α banks with risk. The morning cluster coefficient associated with riskier network position is also associated with high α. There is also some evidence that more complex morning market positions are associated with high α, although this effect disappears if the complexity continues into the afternoon markets. In the cheap liquidity markets, there are few patterns that can be seen in the estimates. Almost all are insignificant. This suggests that in a regime of cheap liquidity, the only strong influence on whether a bank is a high α bank is the total value of the payments received during the early morning hours.

4.4. Network Implications of High- and Low α Banks

Finally, the two well-defined groups allows us to define a four regions of flows and analyzed their systemic properties. Indeed, flows across these four regions show that delays are more likely to occur between high-α banks and for flow from a high-α bank to a low-α bank. This result could provide an early warning of contagion as delays from high-α to low-α banks and subsequently among low-α banks are signs of market freezes. While this section provides only a hint of the possibilities of a network analysis, it is a focus of our continuing research. The simulation studies of Bech and Soramaki (2005) and others indicate that gridlock can be a common occurrence, particularly in systems which face expensive liquidity shortages. However, these simulation studies take time and focus only on a single array of payments structures. A certain level of heterogeneity is assumed in the behavior of the banks which influence the results. Our focus on the high α banks suggest that the payments system could be even more delicate and subject to gridlock in times of expensive liquidity than previously thought.

Payments gridlock occurs because of delays that are not netted out as pointed out in Bech and Soramaki (2005) and others, including Angelini (2000). Liquidity is needed for the netting out of payments between more than two parties in the Target 2 system because the settlement algorithm will only net out bilateral obligations that do not have enough liquidity to settle. While the system as a whole might have enough liquidity to settle, local areas of the payments network can become sinks where liquidity is unable locally to settle out the obligations. Gridlock is more likely to occur within these low liquidity groups where liquidity does not enter from the outside more liquid areas of the network. Simulation studies in Denmark and Sweden suggest that the probability of gridlock is linearly related to the amount of liquidity available to the network as pointed out in Pettersson (2005). This is more likely to occur in sub networks where there is a lack of liquidity if the subnetwork has little liquidity to begin with, and it does not trade with those groups that have liquidity. The question is whether there is a sub-network in the sense that the lack of liquidity churns within a local area of the network with little entry of liquidity from the outside more liquid banks.

Our evidence from the low liquidity month of September suggests that this may indeed be the case. Figure 21 has several very intriguing features. The figure shows the payments that occur between low and high α banks and between banks that are both high or both low α. If the local areas are well connected to each other, then we should see lots of transactions between them. Figure 21 suggest that not only are the largest numbers of transactions within banks that share α characteristics, but also the lowest number of transactions are from the liquid low α banks to the lower liquidity banks. Further, it is the high α banks that are sparking the delays and our evidence is that the subgroup of high αs tend to pass most of their delays to themselves, making gridlock more likely, although they also pass delays to the other group. In other words, delays propagate most within the high α group leading to concern that gridlock will spark here. Interestingly, the low α group does not pass delays to itself. We can take from this several observations. If gridlock occurs in a network where the total liquidity is "adequate", it will occur within a local low liquidity area of the network that is isolated from the rest of the network. This suggests a stress indicator which measures the degree of connectedness of the local low liquidity high α banks with the low α banks. If this connection is robust, then liquidity flows easily into the high α banks lubricating their transactions. If, as in September 2008, the number of payments from the low α to high α banks are reduced, then the probability of stress increases. A stress indicator could be designed that focused on the liquidity flows from the low α to the high α groups. Without access to the low α liquidity, the probability of gridlock increases within the local high α area of the payments.

In times of cheap liquidity, this is less likely to occur. Figure 22 has the corresponding numbers for May of 2014. In this scenario, the flow of payments from low to high α is more plentiful than even the flows within the low liquidity banks themselves. Although the system experiences delays, the delays are not focused on the local areas and are less likely to cause gridlock. They are more likely to be one-off occurrences that do not affect the rest of the system and are quickly resolved. Further, a single bank failure is less likely to drag the rest of the system down as the payments to the downstream counterparties from other more liquid banks prevent a gridlock. The degree of connectedness of the low to high α provides a direct indicator of the matter of great concern in a time of crisis: that banks with sufficient liquidity withdraw their liquidity from banks that they deem to have insufficient liquidity. This is clearly a topic of great interest and one of focus in our continuing research.

5. Concluding Remarks

This paper made a characterization of delays in interbank payments via TARGET2 payment system and conducted a first approximation to the determinants whereby a delay in incoming transactions may cause a delay in other transactions downstream under normal liquidity and market conditions and also in distressed times. Taking into account potential endogeneity problems with the data, we estimated probit models with endogenous regressors and found that in normal times the probability of a bank to delay its payments is positively affected by the series and amount of incoming payments that are delayed. This result points out to a contagion phenomenon taking place with important systemic liquidity implications in the network.

This result is robust to the use of alternative instruments that take into account behavioral patterns of nodes/banks. In particular, we look at local liquidity sinks, i.e. a local area of the network that has a temporary loss of liquidity as a result of a large spike in payments needs without an accompanying increase in liquidity. We also consider transmission of delay information throughout the network nodes and free riding bilateral punishment equilibria.

We also find that this contagion phenomenon does not materialize under distressed market conditions and delay contagion is likely to be caused by the presence of common factors, mutual delaying and second-round effects. This result is also robust to the different behavioral patterns of the network nodes described above. These set of results point out however that there is not This conclusion leads future research to conduct a closer analysis of the nature of the common unobserved factors.

A characterization of the banks according to their initial liquidity allocation provides another interesting set of results. First, we identify two types of banks in terms of their initial liquidity and their propensity to delay on its payments: low-α banks start each day with large initial liquidity and rarely incur in payment delays whereas high-α banks are exactly the opposite and account for the majority of delays in both high and low liquidity environments. This bank split is consistent over time and also allows allows identifying four regions of flows, where delays happen more often among high-α banks and when flows go from a high-α to a low-α bank. The latter constitutes in turn an early warning for market freezes.

References

Angelini, P. (1998), "An analysis of competitive externalities in gross settlement systems", *Journal of Banking & Finance*, vol. 22 nº 1: pp. 1 – 18.

Angelini, P. (2000), "Are Banks Risk Averse? Intraday Timing of Operations in the Interbank Market", *Journal of Money, Credit and Banking*, vol. 32 nº 1: pp. 54–73.

Arjani, N. (2006), "Examining the trade-off between settlement delay and intraday liquidity in Canada's LVTS: A simulation approach", Working Papers 2006–20, Bank of Canada.

Armantier, O., Arnold, J. and McAndrews, J. J. (2008), "Changes in the timing distribution of Fedwire funds transfers", *Economic Policy Review*, nº Sep: pp. 83–112.

Bartolini, L., Hilton, S. and McAndrews, J. (2010), "Settlement delays in the money market", *Journal of Banking & Finance*, vol. 34 nº 5: pp. 934–945.

Bech, M. L. (2008), "Intraday liquidity management: a tale of games banks play", *FRBNY Economic Policy Review*, nº Sep: pp. 7–23.

Bech, M. L. and Garratt, R. (2003), "The intraday liquidity management game", *Journal of Economic Theory*, vol. 109 nº 2: pp. 198–219.

Bech, M. L. and Soramaki, K. (2005), "Gridlock Resolution and Bank Failures in Interbank Payment Systems", in Leinonen, H. (editor), *Liquidity, Risks and Speed in Payment and Settlement Systems*, Bank of Finland Studies, pp. 152–177.

Benos, E., Garratt, R. and zimmerman, P. (2012), "Bank behaviour and risks in CHAPS following the collapse of Lehman Brothers", Bank of England working papers 451, Bank of England.

BLUME, L. E., BROCK, W. A., DURLAUF, S. N. and IOANNIDES, Y. M. (2010), "Identification of Social Interactions", Economics Series 260, Institute for Advanced Studies.

BONACICH, P. (1987), "Power and Centrality, A Family of Measures", *American Journal of Sociology*, vol. 92 nº 5: pp. 1170–1182.

BUCKLE, S. and CAMPBELL, E. (2003), "Settlement bank behaviour and throughput rules in an RTGS payment system with collateralised intraday credit", Bank of England working papers 209, Bank of England.

DIEHL, M. (2013), "Measuring free riding in large-value payment systems: the case of TARGET2", *Journal of Financial Market Infrastructures*, vol. 1 nº 3: pp. 1–23.

HEIJMANS, R. and HEUVER, R. (2011), "Is this bank ill? The diagnosis of doctor TARGET2", DNB Working Papers 316, Netherlands Central Bank, Research Department.

HSIAO, C., PESARAN, M. H. and PICK, A. (2012), "Diagnostic Tests of Cross-section Independence for Limited Dependent Variable Panel Data Models*", *Oxford Bulletin of Economics and Statistics*, vol. 74 nº 2: pp. 253–277.

MANNING, M., NIER, E. and SCHANZ, J. (2009), *The Economics of Large-value Payments and Settlement. Theory and Policy Issues for Central Banks*, Oxford University Press, Great Clarendin Street, Oxford OX2 6DP.

MASSARENTI, M., PETRICONI, S. and LINDNER, J. (2013), "Intraday patterns and timing of TARGET2 interbank payments", *Journal of Financial Market Infrastructures*, vol. 1 nº 2: pp. 3–24.

MERROUCHE, O. and SCHANZ, J. (2010), "Banks intraday liquidity management during operational outages: Theory and evidence from the UK payment system", *Journal of Banking & Finance*, vol. 34 nº 2: pp. 314 – 323.

Pettersson, J. (2005), "Simulations of Liquidity Levels and Delays", in Leinonen, H. (editor), *Liquidity, Risks and Speed in Payment and Settlement Systems*, Bank of Finland Studies, pp. 217–226.

Politis, D. N., White, H. and Patton, A. J. (2009), "Correction: Automatic Block-Length Selection for the Dependent Bootstrap", *Econometric Reviews*, vol. 28 nº 4: pp. 372–375.

Proepper, M., van Lelyveld, I. and Heijmans, R. (2008), "Towards a Network Description of Interbank Payment Flows", DNB Working Papers 177, Netherlands Central Bank, Research Department.

Rochet, J.-C. and Tirole, J. (1996), "Controlling Risk in Payment Systems", *Journal of Money, Credit and Banking*, vol. 28 nº 4: pp. 832–862.

Schulz, C. (2011), "Liquidity requirements and payment delays - participant type dependent preferences", Working Paper Series 1291, European Central Bank.

Watts, D. and Strogatz, S. (1998), "Collective Dynamics of Small World Networks", *Nature*, vol. 393 nº 6684: pp. 440–442.

Tables and Figures

Table 1: Probit model with continuous endogenous regressors: **Calm day estimates**

Dependent variable		Model	
Delay propagation	[1]	[2]	[3]
Intercept	-2.343	-1.469	-1.851
	(-229.90)	(-149.22)	(-177.11)
Incoming delays	0.0969	0.0511	0.0619
	(-166.14)	(-95.36)	(-98.10)
Observations	206177	206177	206177

Notes: Incoming delays are transformed in log. All coefficients are statistically significant at 5%, t statistics are presented in parentheses. See Section 3.2 for details of the model specification.

Table 2: Probit model with continuous endogenous regressors: **Stress day estimates**

Dependent variable		Model	
Delay propagation	[1]	[2]	[3]
Intercept	-1.576	-0.394	1.428
	(-213.52)	(-56.31)	(164.49)
Incoming delays	0.0383	-0.0205	-0.1270
	(86.53)	(-54.35)	(-330.68)
Observations	317754	317754	317754

Notes: Incoming delays are transformed in log. All coefficients are statistically significant at 5%, t statistics are presented in parentheses. See Section 3.2 for details of the model specification.

	September 2008	——Low——	—-Liquidity—	————	May 2014	High	Liquidity	
	Morn Liquidity	Morning Liquidity	Day Complexity	Day Complexity	Morning Liquidity	Morning Liquidity	Day Complexity	Day Complexity
		0-1		0-1		0-1		0-1
night_Bon_Centrality	430.5	130.5***			-1.359	-16.51*		
	(1.48)	(3.14)			(-0.40)	(-1.73)		
7-9_Bon_Centrality	0.611	-1.016**			1.270	-0.0992		
	(0.51)	(-2.40)			(1.31)	(-1.35)		
night_in_clust	-1.265	0.891**			-1.450**	-0.0661		
	(-1.50)	(2.09)			(-2.00)	(-0.11)		
7-9_in_clust	3.004**	3.682***			-1.057*	1.162**		
	(2.49)	(4.47)			(-1.68)	(2.01)		
night_in_degree	-0.0150	0.0289***			-0.0144	0.00935		
	(-0.85)	(2.74)			(-0.80)	(0.82)		
7-9_in_degree	0.00135	-0.00831***			-0.00813**	-0.00181		
	(0.24)	(-2.81)			(-2.25)	(-0.74)		
night_in_tot_paym_log	-0.00124	-0.0206			-0.158*	0.000914		
	(-0.02)	(-0.65)			(-1.66)	(0.03)		
7-9_in_tot_paym_log	-0.0678	0.0895			0.492***	0.199**		
	(-0.33)	(1.02)			(3.81)	(2.11)		
aft_Bon_Centrality			0.241	0.0765			0.245	0.113
			(0.89)	(0.52)			(1.05)	(1.50)
morn_Bon_Centrality			-0.222	0.110			0.293	-0.0395
			(-0.58)	(0.61)			(0.85)	(-0.53)
aft_in_clust			-4.822	0.278			-0.577	-1.372
			(-1.30)	(0.12)			(-0.22)	(-0.58)
morn_in_clust			6.923*	5.312**			0.592	2.953
			(1.81)	(2.23)			(0.22)	(1.24)
aft_in_degree			-0.00368	-0.0486**			-0.0222	-0.0376*
			(-0.12)	(-2.41)			(-1.08)	(-1.87)
morn_in_degree			0.00443	0.0529**			0.0160	0.0399*
			(0.12)	(2.46)			(0.75)	(1.87)
aft_in_tot_pay			0.449	0.190			-0.461	0.394
			(0.58)	(0.53)			(-1.17)	(1.20)
morn_in_tot_pay			-0.588	-0.112			0.680*	-0.199
			(-0.76)	(-0.31)			(1.86)	(-0.60)
Constant	2.379	-4.362**	3.896	-5.849***	-3.717	-5.967***	-1.999	-6.392***
	(0.67)	(-2.17)	(1.11)	(-2.90)	(-1.52)	(-2.97)	(-0.67)	(-3.02)
Observations	210	210	210	210	210	210	210	210

Table 3: Regressions of banks' α value on network characteristics of the banks. T statistics are in parentheses. *, **, *** denote significance at 10%, 5% and 1% respectively.

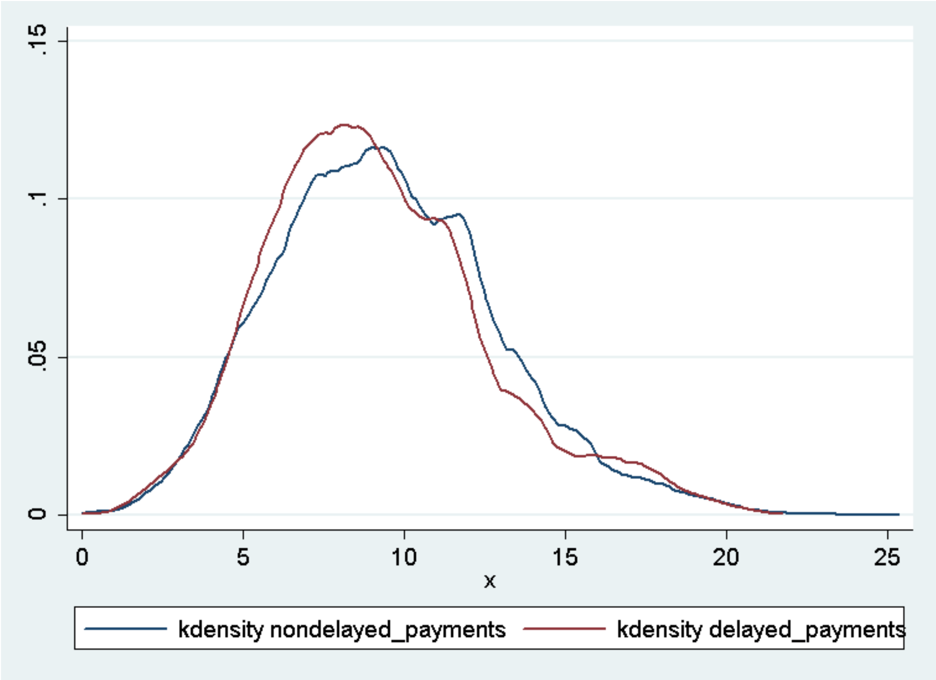

Figure 1: Calm day: Distributions of values of delayed and non-delayed payments

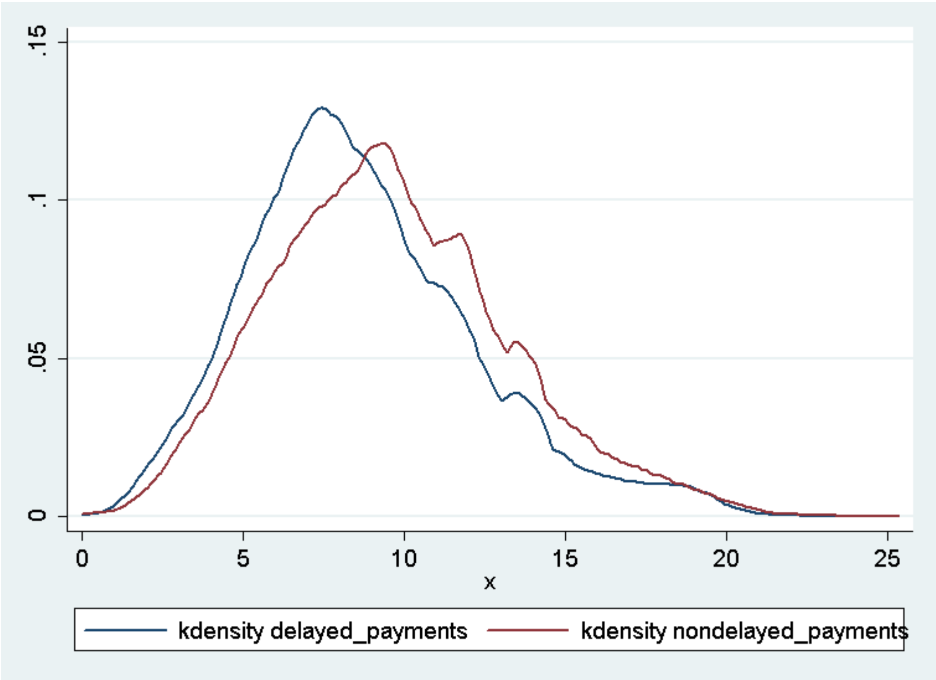

Figure 2: Stress day: Distributions of values of delayed and non-delayed payments

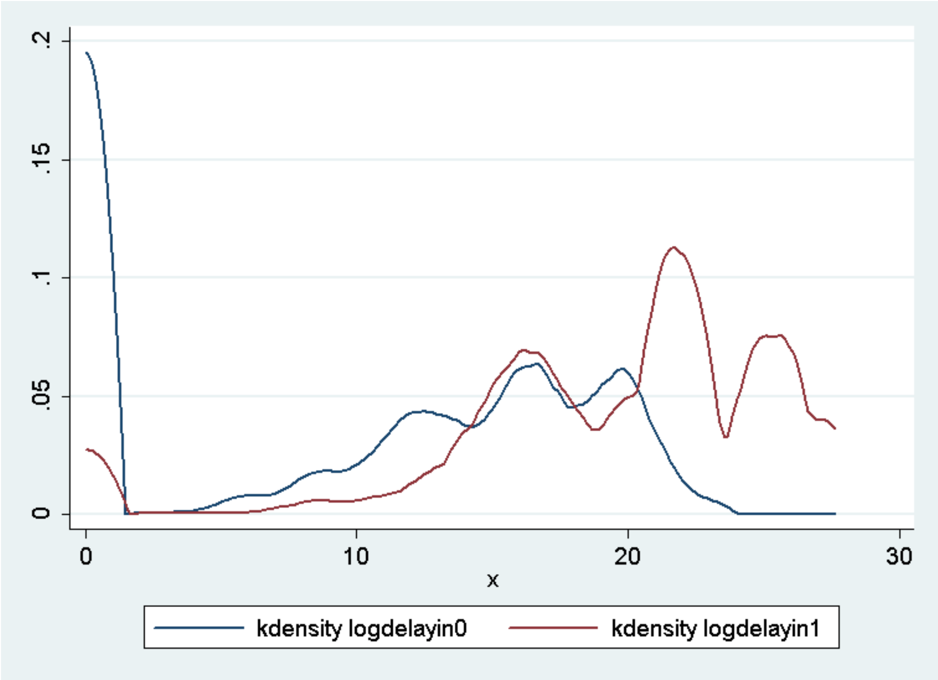

Figure 3: Calm day: Distributions of values of delayed payments that cause and do not cause subsequent delays

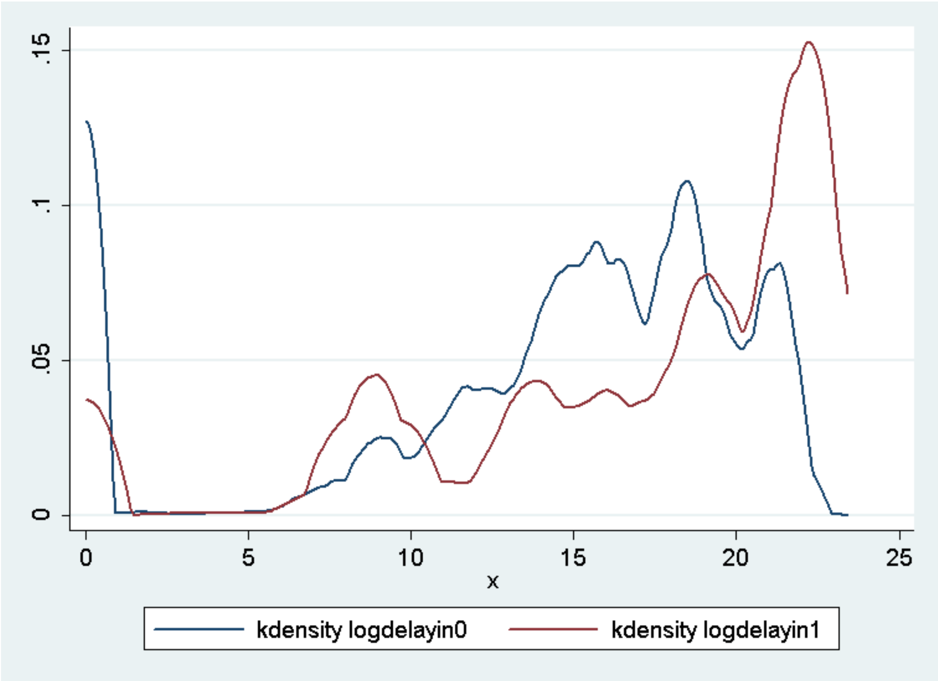

Figure 4: Stress day: Distributions of values of delayed payments that cause and do not cause subsequent delays

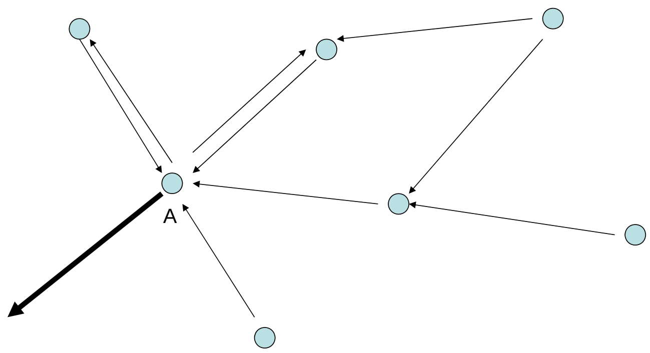

Figure 5: We are interested in measuring the probability of a delay from node **A**, given that node **A** is making a payment. All delays are marked with arrows.

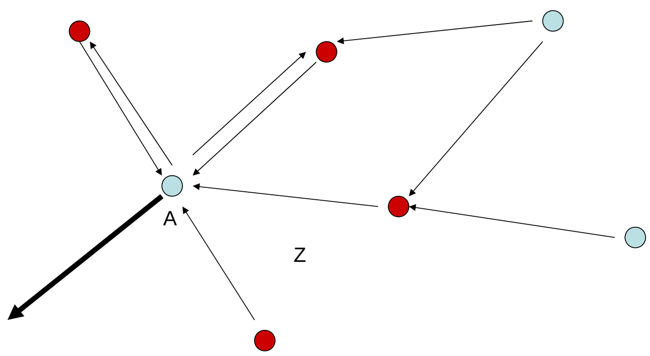

Figure 6: The amount of delays coming in, marked by the sum of the arrows from the red nodes to node **A** are the delays that may be propagated by **A**'s behavior. If all is exogenous, then the instruments would be the same set of arrows.

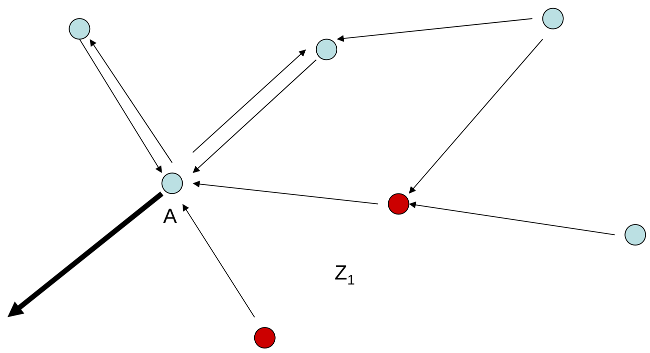

Figure 7: The set of instruments Z_1 are represented by the sum of the arrows going in from the red nodes that do not have a direct reflecting arrow coming back from node **A**.

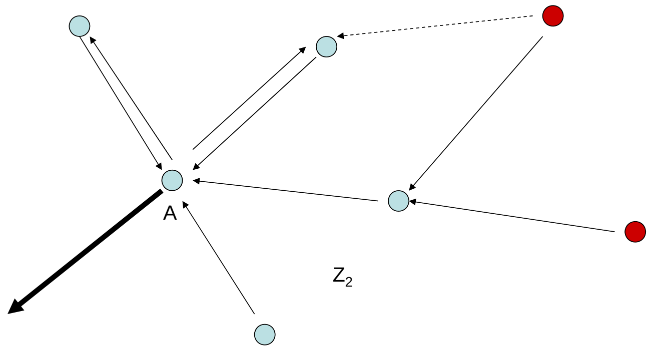

Figure 8: Analogous to Blume *et al.* (2010), we use instruments Z_2, which are at least one link removed from **A**. In order to remove the local reflecting, the instruments do not include those vectors that go through a self reflecting node, so the dotted vector is excluded.

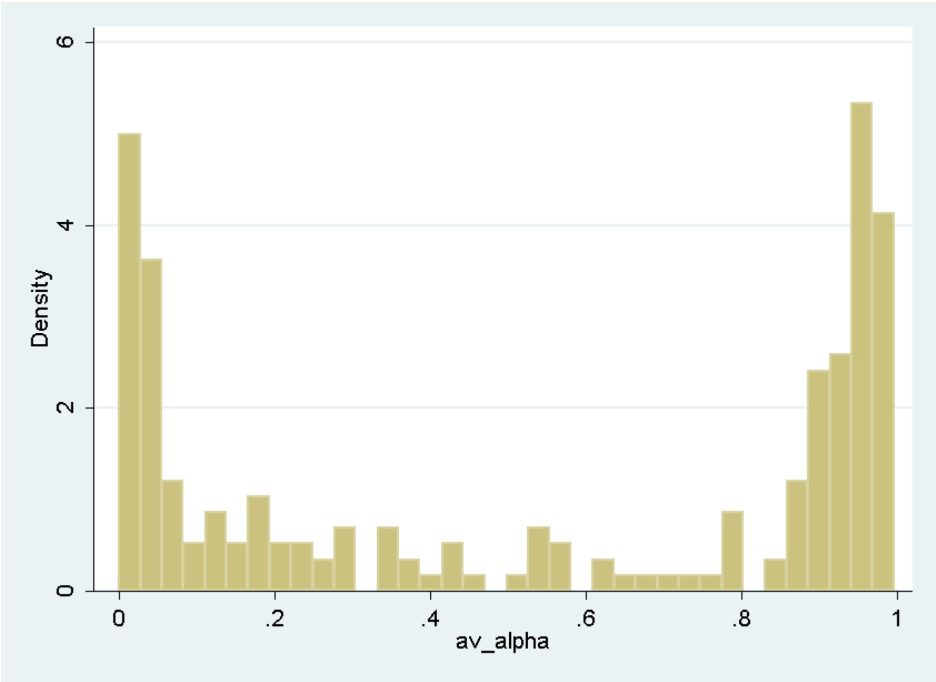

Figure 9: Distribution of alpha values for banks: alpha is an average daily value over September 2008

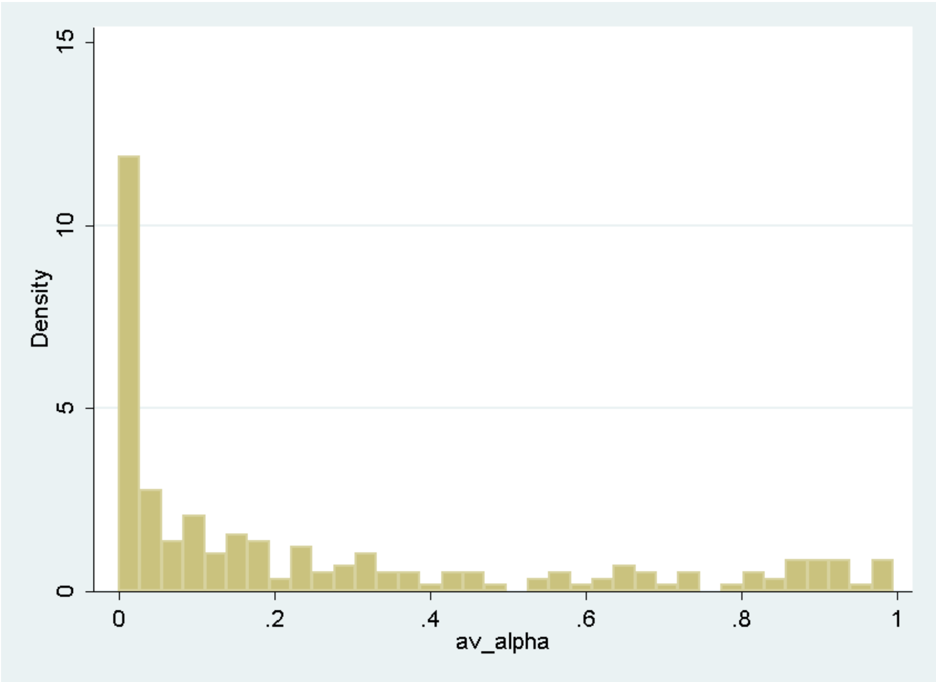

Figure 10: Distribution of alpha values for banks: alpha is an average daily value over May 2014

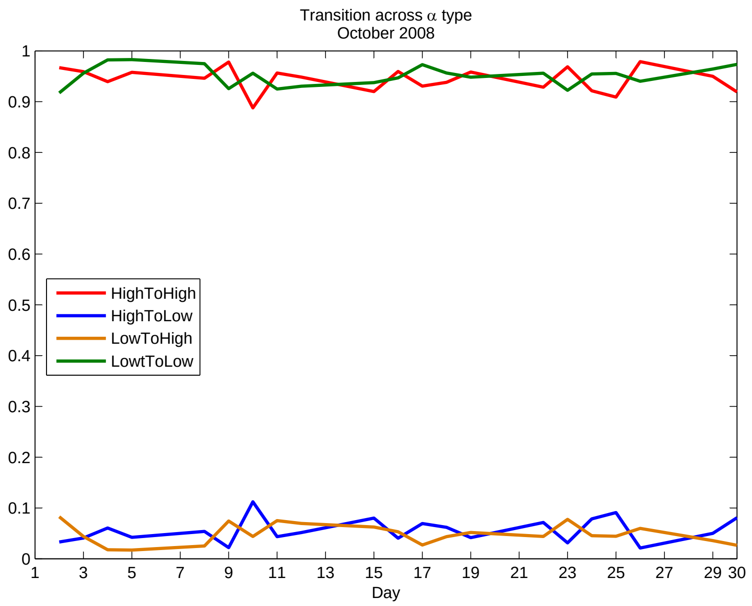

Figure 11: Transition probabilities across *alpha* status

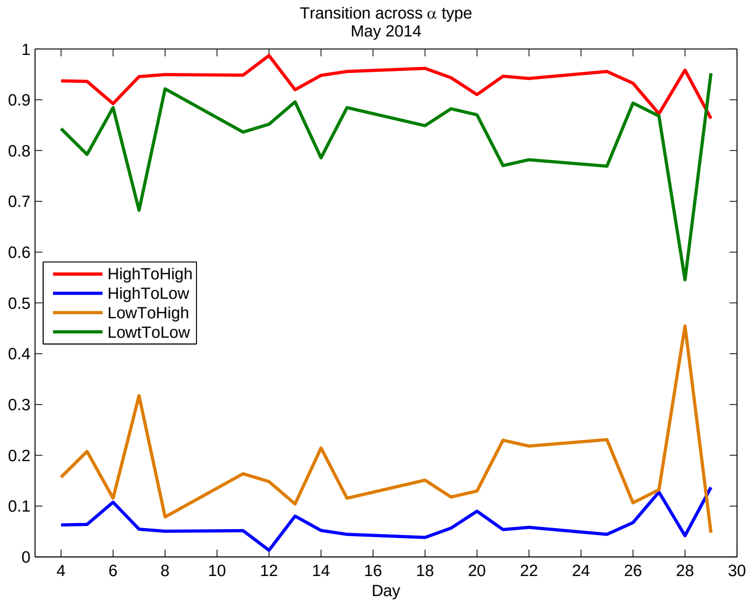

Figure 12: Transition probabilities across *alpha* status

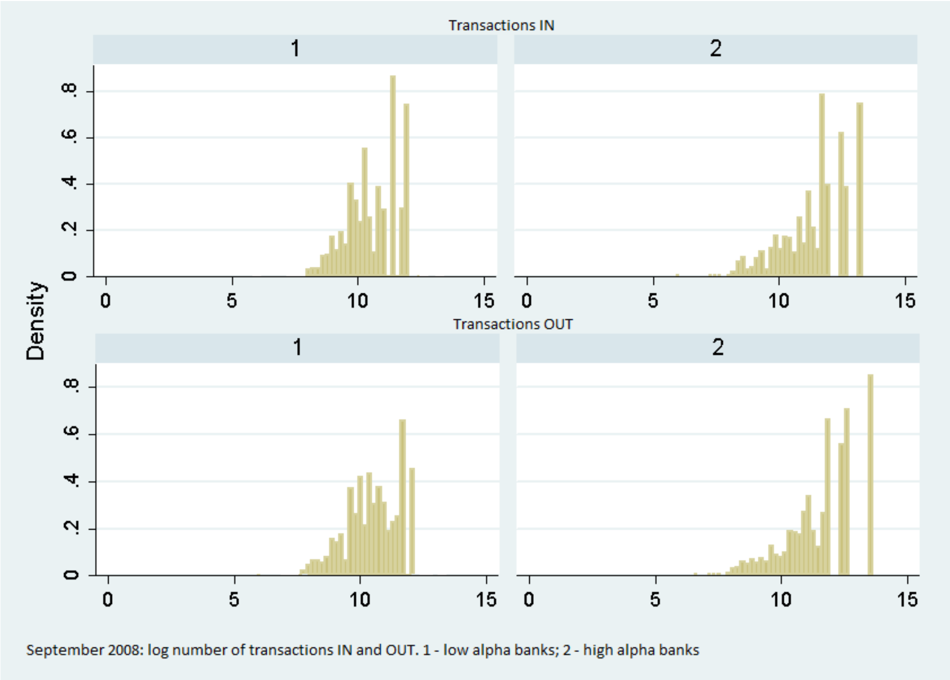

Figure 13: Log number of IN and OUT transactions by two types of banks in September 2008

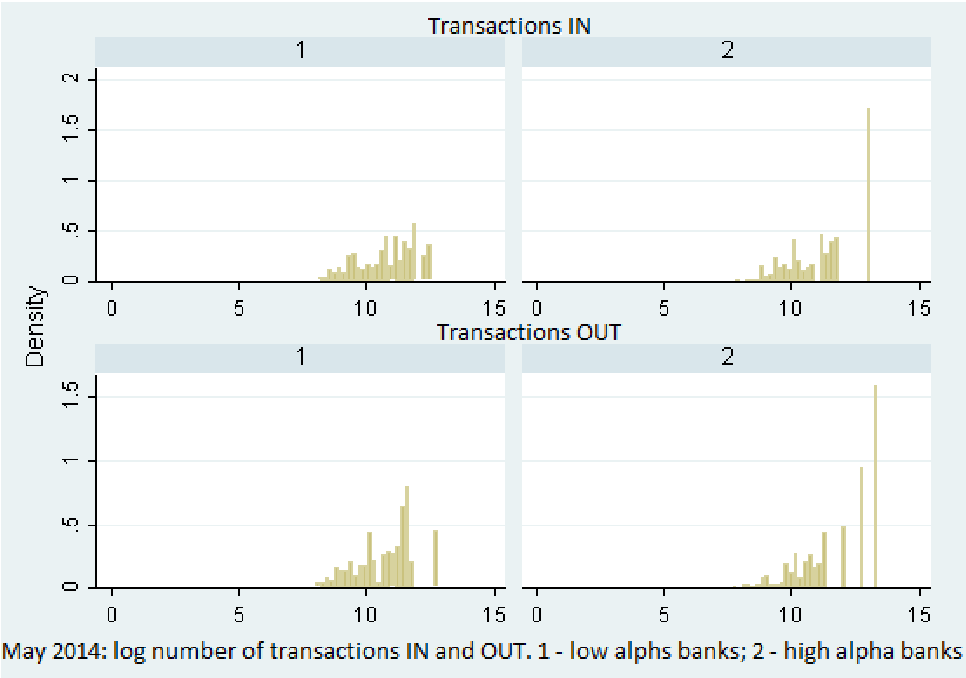

Figure 14: Log number of IN and OUT transactions by two types of banks in May 2014

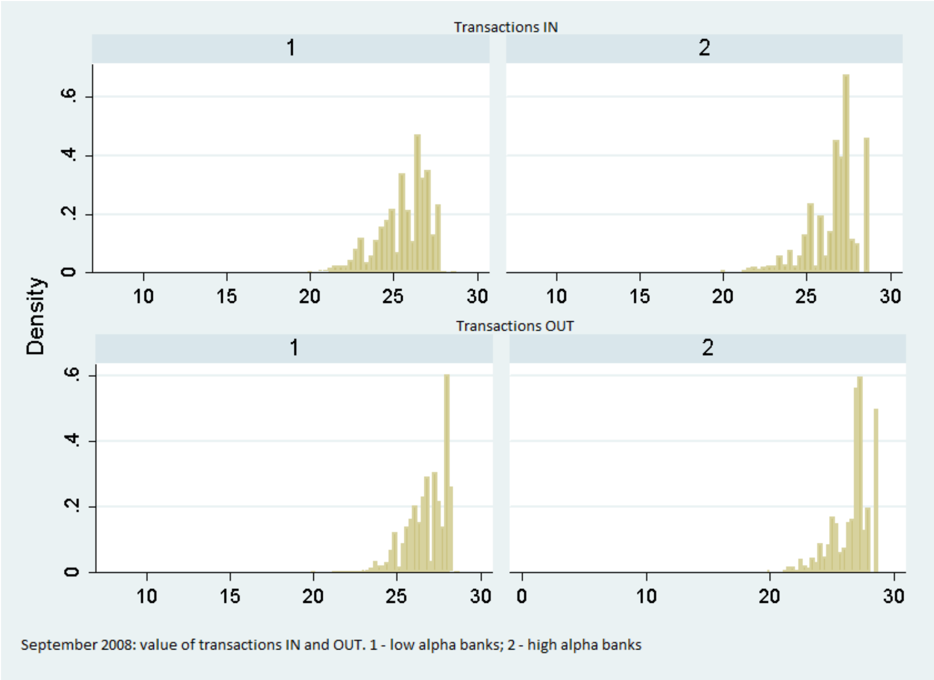

Figure 15: Log value of IN and OUT transactions by two types of banks in September 2008

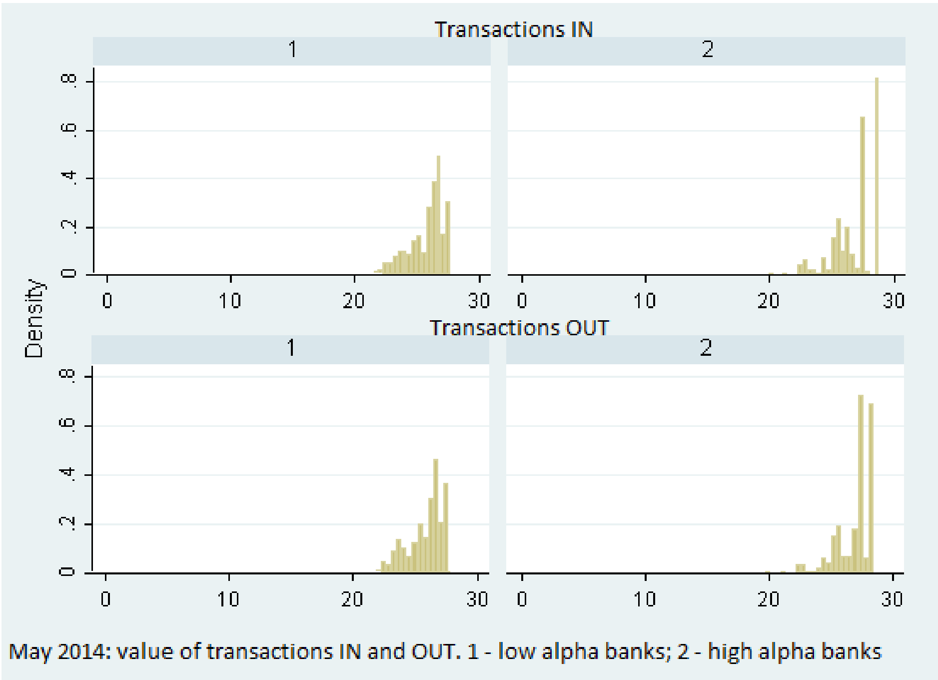

Figure 16: Log value of IN and OUT transactions by two types of banks in May 2014

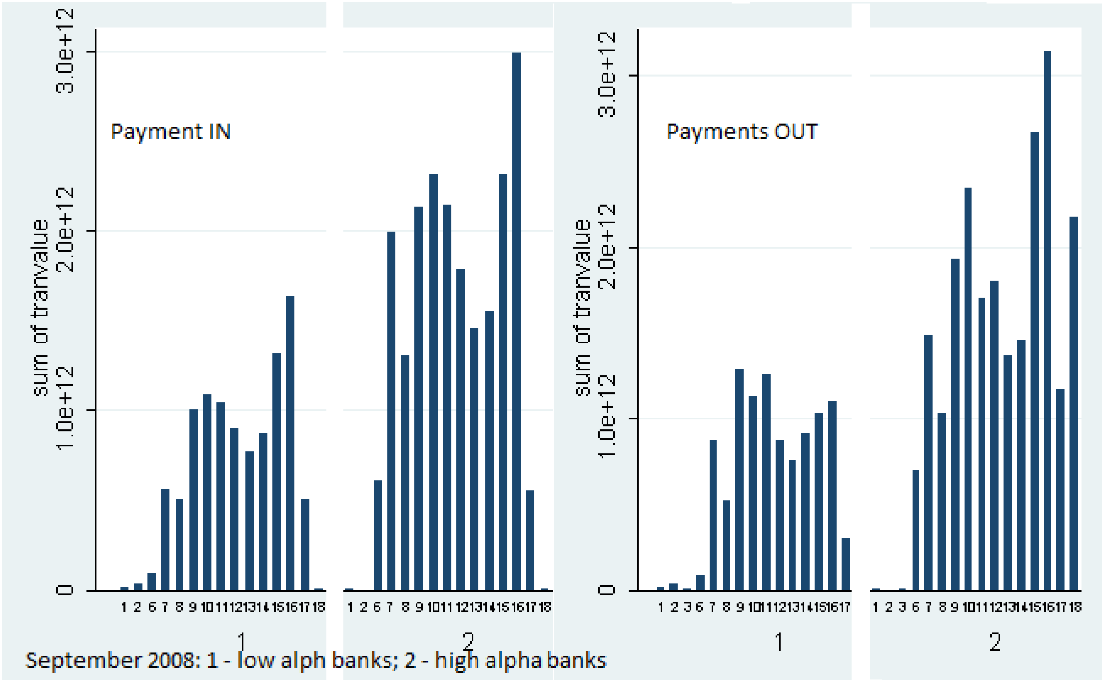

Figure 17: Value of IN and OUT payments by two types of banks throughout the day in September 2008

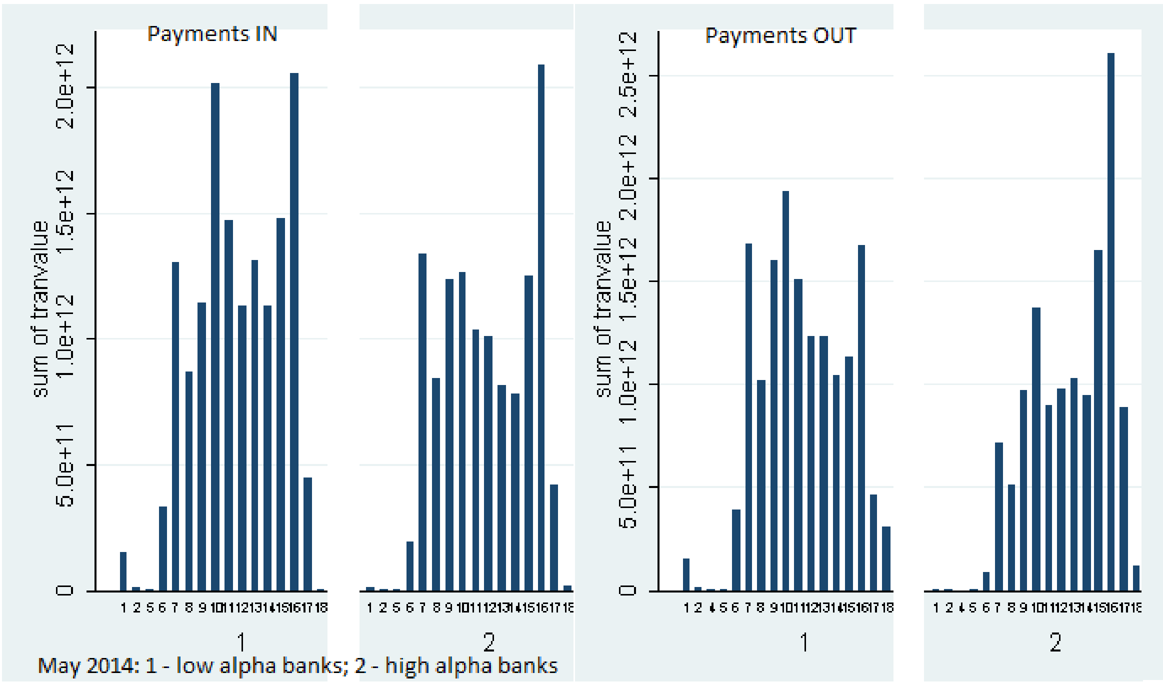

Figure 18: Value of IN and OUT payments by two types of banks throughout the day in May 2014

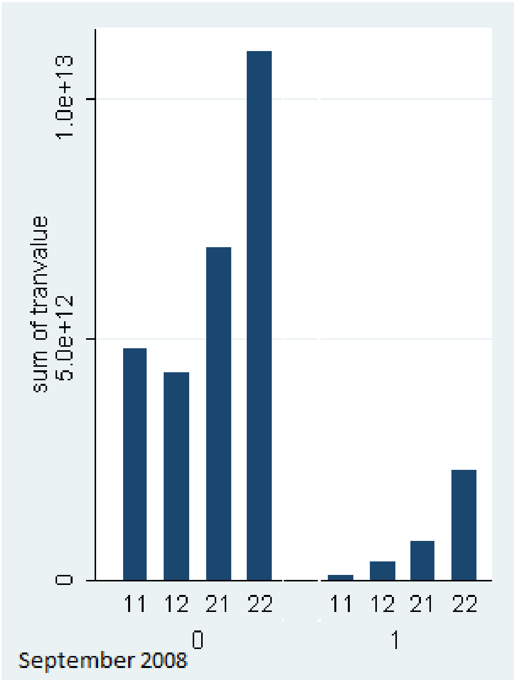

Figure 19: Total value of transacted (0) and delayed (1) payments in September 2008 by transaction category: 11 - from low to low alpha banks; 12 - from low to high; 21 - from high to low; 22 - from high to high

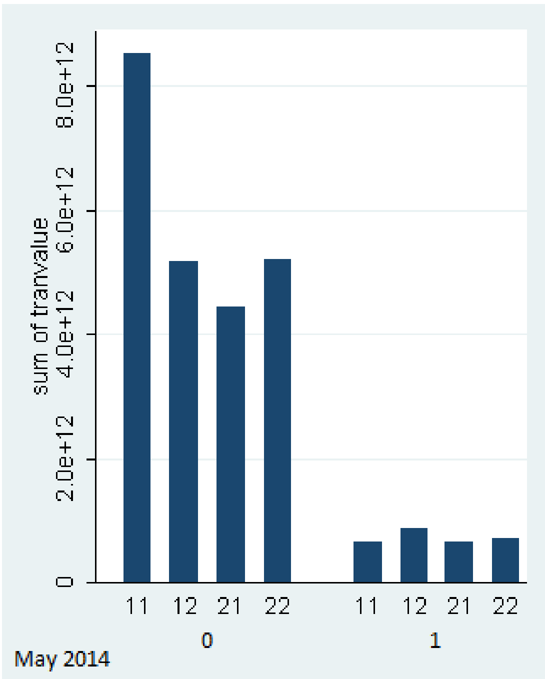

Figure 20: Total value of transacted (0) and delayed (1) payments in May 2014 by transaction category: 11 - from low to low alpha banks; 12 - from low to high; 21 - from high to low; 22 - from high to high

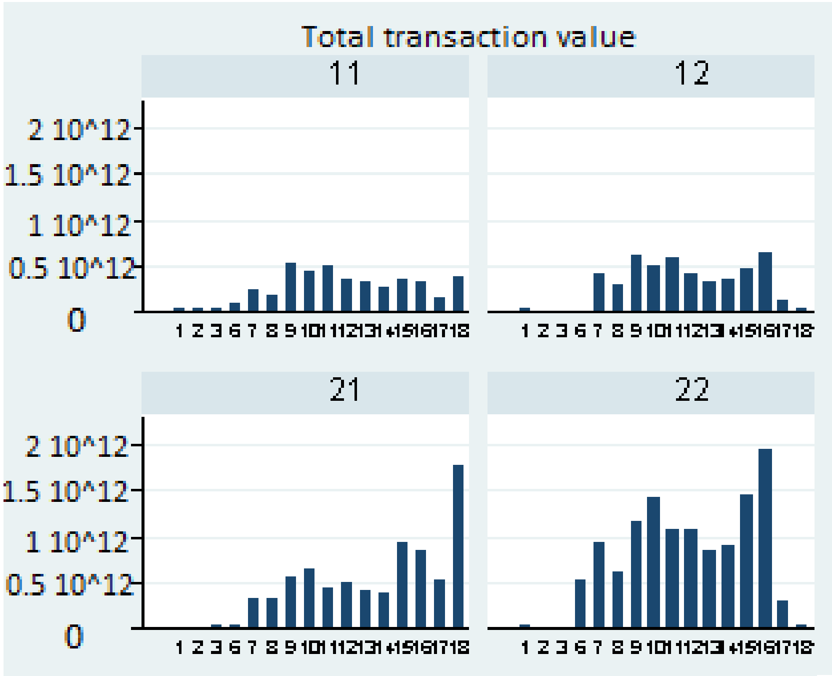

Figure 21: Total transaction value of payments by category throughout the day in May 2014 (hourly averages): 11 - from low to low alpha banks; 12 - from low to high; 21 - from high to high; 22 - from high to high

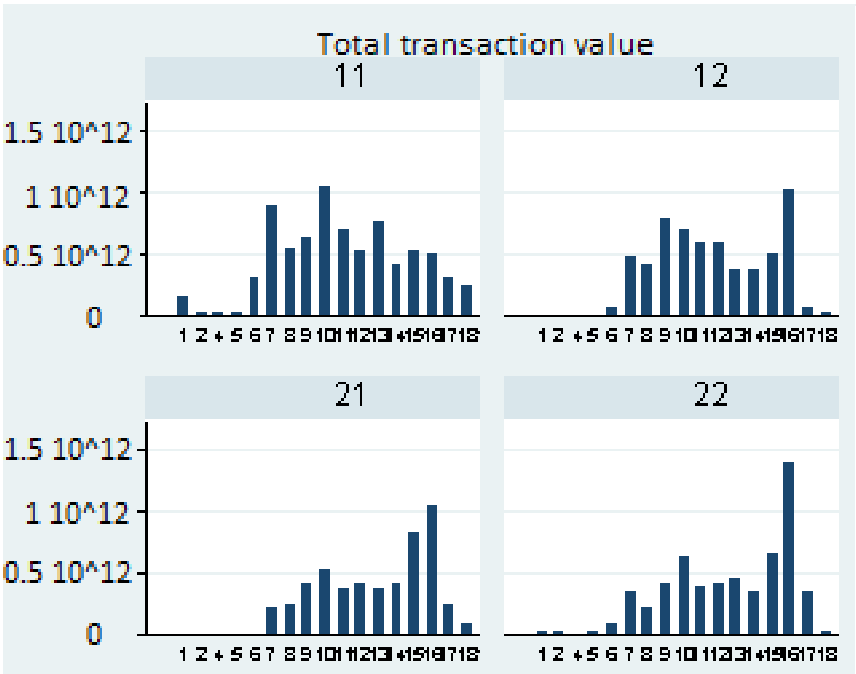

Figure 22: Total transaction value of payments by category throughout the day in September 2008 (hourly averages): 11 - from low to low alpha banks; 12 - from low to high; 21 - from high to high; 22 - from high to high