Indentification of Over and Under Provision of Liquidity in Real-Time Payment Systems*

Edward Denbee
Bank of England

Rodney J. Garratt
University of California, Santa Barbara

Peter Zimmerman
University of Oxford

August 20, 2015

Abstract

It can be difficult to tell to what extent an individual bank's liquidity provision is intentional and to what extent it results from external factors that are beyond its control. We describe two methods for evaluating liquidity provision in real-time gross settlement payment systems and utilize a recombinant approach to detect instances where observed patterns of liquidity provision are unlikely to have occurred absent of some behavioural or structural factors, such as differences in banks' business models. We apply our techniques to crisis-period data from CHAPS, the UK large-value payment system. We find that smaller banks provided more liquidity to the system than larger banks relative to their payment flows. Moreover, we observe an increase in the degree of inequality of liquidity provision relative to usage across banks following the collapse of Lehman Brothers. Our results suggest that instances of over and under provision of liquidity that appear in our data are more frequent than would be expected from random payment flows.

1 Introduction

A payment system consists of the procedures and associated computer networks used by its participants to transfer money. Sometimes called the 'plumbing' of the financial system, smoothly functioning payments systems are essential to the operation of financial markets. Large-value wholesale payment systems, such as CHAPS in the United Kingdom, are generally considered to be systemically important because of the value and nature of the financial transactions that they facilitate. On a typical business day, transactions with a total value of around £277 billion flow through CHAPS, roughly equivalent to one-sixth of the United Kingdom's annual gross domestic product.[1]

In a real-time gross settlement (RTGS) system, payments settle immediately and with finality in central bank money, providing that the payer has sufficient liquidity to fund the outgoing payment. But the aggregate amount of liquidity needed to fund payment obligations is often much less than

*This paper was completed while RJG was working at the Federal Reserve Bank of New York. The views expressed in this paper are those of the authors, and not necessarily those of the Bank of England, the Federal Reserve Bank of New York or the Federal Reserve System. The authors wish to thank Evangelos Benos, Anton Braun, James Chapman, Roger Clews, Toby Davies, Marco Galbiati, Matti Hellqvist, Matthew Hunt, John Jackson, Will Roberds, Warren Weber, an anonymous referee and seminar participants at the Bank of England and at the Bank of Finland 8th Simulator Seminar and Workshop and the Federal Reserve Bank of Atlanta for useful comments and suggestions. We also thank Lewis Webber for invaluable help with simulations. Any errors are ours. This paper builds on earlier work by the same authors that is available in 'Diagnostics for the financial markets – computational studies of payment system', Bank of Finland Simulator Seminar Proceedings, 2012.

[1]See Bank of England (2014).

gross payment flows. During the course of the day, each payment system participant typically makes and receives thousands of payments. Thus outgoing payments are not only funded from liquidity made available from payment system participants' own resources,[2] but also from liquidity obtained from incoming payments.

If banks were required to process payment requests as soon as they received them, then they would have little discretion over the liquidity they provide to the rest of the payment system. But this is not the case: with the exception of some time-critical payments — or some payments systems designed to process customer payments in real-time — banks do not usually have to process payment requests as soon as they receive them. Rather, banks may choose to delay processing payments in order to conserve liquidity and make use of incoming funds.[3] If too many banks withhold liquidity the payment system can fall into gridlock.[4] Consequently, central banks have an interest in monitoring banks' liquidity provision in order to ensure the continued smooth functioning of the payments system.

In this paper, we measure liquidity provision in two ways. First we look at the maximum net debit position that settlement banks reach in their settlement accounts over the course of each day, during a historic period.[5] The sum of these net debit positions across all banks is the total amount of liquidity that was actually used to make the day's payments. Therefore each bank's own net debit position, divided by the sum of the net debit positions of all banks, gives the share of liquidity provided by each bank. Whenever the value of a bank's payments into the system exceeds that of those it has received, the difference has to be made up either from central bank reserves, or from eligible collateral that a settlement bank pledges intraday in order to obtain liquidity from the central bank.[6] Therefore, we can assume that a net debit position imposes an opportunity cost of using central bank reserves or of pledging eligible collateral, and so our first measure reflects the nominal monetary cost of liquidity provision.

In addition to this monetary cost, making payments earlier can result in a greater exposure to counterparty risk. For example, if the paying bank relies on its own recycled liquidity to fund future payments, then it faces the risk that its counterparty fails to recycle the liquidity back into the payment system in a timely fashion. This may happen, for example, if the counterparty has an operational problem or enters bankruptcy.[7] A settlement bank's *average* net debit position throughout the day is a better proxy for this counterparty risk than the maximum net debit position, so we use this as our second measure of liquidity provision.

Some banks have a higher value of payment activity than others, and hence may reasonably be expected to provide more liquidity, in absolute terms. Therefore we take usage of payments

[2]These are generally either in the form of central bank reserves, or liquidity generated by via repo of eligible assets at the central bank.

[3]The incentive to conserve liquidity arises because funds that banks deposit in their settlement accounts to facilitate payments have an opportunity cost to the bank in terms of foregone investment opportunities, or to mitigate against the possibility of liquidity shocks later in the day.

[4]CPSS (1993) defines a gridlock as a situation where "the failure of some transfer instructions to be executed ... prevents a substantial number of other instructions from other participants from being executed". See Soramäki and Bech (2001) for a discussion of the cost of delay in payment systems.

[5]The term 'settlement bank' refers to a bank that has a settlement account with the central bank for the purposes of making intraday payments through the payments system. Other banks that make payments through CHAPS do so via one or more of these settlement banks.

[6]Banks can also fund payments using liquidity received in interconnected payments systems. For instance, banks which are settlement banks in both CHAPS and CREST — which is the UK securities settlement system — could use a net receiver position in one system to fund payments in the other. We do not consider this in our analysis, partly because in the UK these other interconnected payments systems do not allow settlement banks the same degree of control over intraday timing as CHAPS does.

[7]Manning, Nier and Schanz (2009) examine issues around strategic payment system behaviour.

system liquidity (i.e. gross payment outflows) into account in our two measures of the cost and risk of liquidity provision. We compute our measures for CHAPS settlement banks using data from January 2008 to May 2010, and present aggregated results for groups of banks in two size categories. Although larger banks do provide the bulk of the liquidity in absolute terms, we find that the smaller banks almost invariably provide a larger share of liquidity to the system than their share of payments. This is true under both of our measures.

The ratio of the liquidity provided by a bank to the gross value of payments it makes is the average liquidity cost of the bank's payments. As the above analysis suggests, different banks have different average liquidity costs. We measure inequality in the liquidity cost of payments across CHAPS banks by computing Gini coefficients, and examine these over time. The series shows a significant increase in the period surrounding the collapse of Lehman Brothers. This finding is important because it tells us that CHAPS participants became more dependant on the liquidity provision of others during the crisis.

Unobserved factors may explain some of these differences in liquidity provision.[8] Average payment sizes are likely to be important — a bank which sends and receives a small number of very large payments would be expected to use more liquidity than a bank with a large number of smaller payments. This is because the latter bank would be better able to offset payments and receipts and thus less likely to assume a large net sender position.

Clearing on behalf of clients may have an impact too. There is evidence that this can create efficiencies in liquidity usage (Jackson and Manning (2007)). For larger banks, client clearing tends to form a greater proportion of their payments than it does for smaller banks. Furthermore, it is possible that settlement banks, when making payments on behalf of clients, may on occasion intentionally delay in order to reduce credit exposures to these clients.[9]

These factors could explain the lower levels of liquidity provision, relative to payment values, that we observe for larger banks. In any case, some heterogeneity between individual banks' liquidity provision and usage is inevitable, and does not necessarily imply unfairness. Since the arrival of payment requests from customers is typically outside of the control of the banks, there will be net liquidity providers and users on any given day, even if all banks process payment requests immediately. This means the patterns of liquidity provision that we observe could just be an artefact of the way payment requests happened to arrive. We would like to know when observed differences in liquidity provision are so marked that they are very unlikely to have solely been a result of external factors. We provide a method for identifying when over and under provision of liquidity is unlikely to have occurred by accident. The idea is to ask, given all the different permutations for how payments might have arrived, what would be a very unlikely level of liquidity provision? We answer this question by reshuffling each day's payment schedule many times to generate distributions for our liquidity provision measures. We then check how often actual values of these measures lie in the tails of these distributions. We find that instances where banks are in these tails occur far more frequently than we would expect to see in the absence of behavioural or structural factors.

The analysis presented in this paper can be applied to any payment system, and may be particularly useful to assess liquidity provision in instances where settlement banks have a substantial degree of freedom in choosing when to make payments during the day. To our knowledge, this is the case for most systems used for making wholesale unsecured payments. Care should be taken when comparing results across systems or across time periods that may involve significant reforms

[8] Becher, Galbiati and Tudela (2008) describe CHAPS and the factors that affect the timing of payments in more detail.

[9] According to Valukas (2010), Lehman Brothers' settlement banks tried to reduce their unsecured intraday exposures to the institution once its financial condition began to deteriorate.

to the payment system. In the particular case of CHAPS, it should be noted that the system has undergone several structural changes since the end of our data period in May 2010 that may have led to changes in the patterns of liquidity provision.[10]

2 Data

We use data on payments activity for all CHAPS settlement banks from 2 January 2008 to 28 May 2010. These data are obtained from the payments database maintained by the Bank of England in its role as operator of the RTGS system.

The CHAPS settlement banks during this period were ABN Amro, Bank of England, Bank of Scotland, Barclays, Citibank, Clydesdale, Co-operative Bank, CLS Bank, Danske Bank, Deutsche Bank, Lloyds, HSBC, NatWest, RBS, Santander/Abbey, Standard Chartered and UBS. Bank of England and CLS Bank are excluded from our analysis. Membership is not constant throughout this period: ABN Amro left on 19 September 2008 and Danske Bank joined on 20 April 2009.

We aggregate any figures that are reported separately for NatWest and RBS, since these banks belong to the same group. And a merger with Lloyds meant that the Bank of Scotland reserves account was withdrawn on 5 February 2009, so those two settlement banks effectively operate from a single pool of liquidity after this date.

3 Measures of liquidity provision

3.1 Measuring the cost of liquidity provision

One way to evaluate liquidity provision is to look at the share of total liquidity a bank provides to the system and to relate this to its share of total payments. Suppose there are n banks, which are indexed by $i = 1, ..., n$. Let $x_i^s(t)$ be the amount sent by bank i up to time t on day s, and let $y_i^s(t)$ be the amount received. t lies in the interval $[0, T]$, where $t = 0$ denotes the start of the day and $t = T$ the end. Then the net debit position at time t on day s is:

$$N_i^s(t) = x_i^s(t) - y_i^s(t) \tag{1}$$

The net debit position identifies the liquidity provided by bank i to the rest of the system by time t on day s. The liquidity burden of bank i on day s is determined by the largest net debit position:

$$L_i^s = \max_{t \in [0,T]} N_i^s(t) \tag{2}$$

The largest net debit position incurred by a bank on a given day is the total amount of the bank's own cash and collateral that it actually used to fund its own payments. It is the minimum amount of liquidity that the bank could have held to meet its payment obligations on that day in order to make its payments at the time that they occurred. Note that $L_i^s \geq 0$, since $x_i^s(0) = y_i^s(0) = 0$ for all i, s.

Let $P_i^s = x_i^s(T)$ denote the total value of payments sent by bank i on day s. Our cost-based measure of liquidity provision of bank i on day s is:

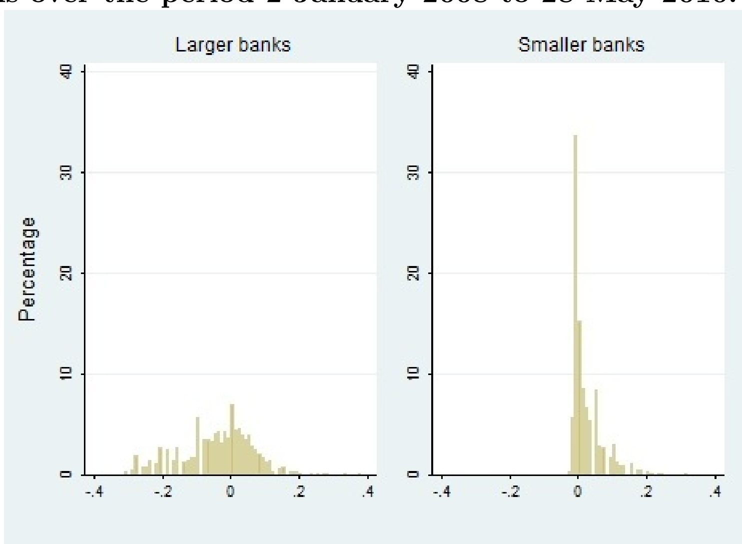
$$c_i^s = \frac{L_i^s}{\sum_{j=1}^n L_j^s} - \frac{P_i^s}{\sum_{j=1}^n P_j^s}. \tag{3}$$

[10] Davey and Gray (2014) analyses changes to banks' liquidity requirements in CHAPS as a result of the introduction of a liquidity saving mechanism in April 2013; and Finan, Lasaosa and Sunderland (2013) describe recent widening in CHAPS settlement bank membership.

This measures the observed difference between bank i's share of liquidity provision and its share of liquidity usage. If the difference is less than 0, then the bank provides less liquidity than might be expected, given its share of payment activity. If it is greater than 0, then it provides more than might be expected, given its share of payment activity.

Figure 1 comprises two histograms showing our measure of the cost of liquidity provision over the period 2 January 2008 to 28 May 2010. We partition the banks into two groups based on their average daily values sent over the period: the 'larger' banks in the left-hand panel and the 'smaller' banks on the right. Larger banks account for over 90% of total payment values sent through CHAPS over this period. The makeup of each group is recalibrated on 5 February 2009.[11] We emphasise that 'larger' and 'smaller' here refer only to average daily values through CHAPS over the period and are not necessarily correlated with other measures of size, such as balance sheet size or payments in non-sterling currencies.

Figure 1: **Frequencies of observations of cost measure of liquidity provision for larger and smaller banks over the period 2 January 2008 to 28 May 2010.**



The bin size on the horizontal axis is 0.01, with the global peaks occurring in the interval containing zero [-0.005, 0.005]. The figure shows that the larger banks tend to have the widest range of values. For smaller banks, the measure over this period takes an average value of 0.03, while for larger banks it is -0.04. This translates into an average provision for each small bank of £600-700 million a day in excess of what might be expected from their share of payment activity. Difference of means tests reveal significance. Table 1 shows that there are differences too in the higher moments of the two distributions: the distribution of scores for the smaller banks has a lower variance, a higher (positive) skewness and a higher kurtosis compared with the distribution for the larger banks.

These differences may not be due to strategic behaviour. It may be that larger banks structurally pay later than the smaller ones, due to the nature of their business. The activity of many of the banks in the 'smaller' group is dominated by retail business, while the payment activity of many of the 'larger' banks may be driven by wholesale business (for example, lending and borrowing in

[11] Before 5 February 2009, there are 7 banks classed as larger and 6 as smaller. From this date, there are 6 in each of the two groups.

Table 1: **Moments of distributions of the cost-based measure of liquidity provision for larger and smaller banks over the period 2 January 2008 to 28 May 2010.**

	Larger banks	Smaller banks
Mean	-0.039	0.029
Variance	0.010	0.002
Skewness	-0.417	1.924
Excess kurtosis	0.093	4.101

the money markets). And many of the larger banks make payments on behalf of other institutions. These factors may lead to structural differences in banks' payment schedules. These results may also suggest that there are liquidity efficiency benefits to being a larger bank.[12]

3.2 Measuring the risk of liquidity provision

Measuring liquidity provision based on largest net debit position has its drawbacks. While this measure addresses the direct cost to a bank of providing liquidity, it does not fully reflect exposure to counterparty risk. Banks rely on each other to recycle payment system liquidity. When a bank assumes a debit position, it has made outward payments before receipt of incoming payments, and so is exposed to risk if a counterparty delays or fails to send its own payments. To reduce this risk, the bank may prefer to delay its own payments.[13] But if all banks delay their payments, then the system would fall into gridlock, so banks must have some willingness to assume net debit positions vis-à-vis their counterparties, at least for some period of time. A measure of the risk of liquidity provision should therefore consider how willing the bank is to hold net debit positions for periods of time during the payment day, as well as the size of those positions.

A bank takes counterparty risk when it is a net sender — that is, its net debit position $N_i^s(t) > 0$. The average risk taken for bank i on day s is:

$$\Lambda_i^s = \frac{1}{T} \int_0^T \max\left[N_i^s(t), 0\right] dt. \tag{4}$$

Using this metric for counterparty risk, we can construct a measure comparing the share of risk borne by each bank to its share of payments, which is given by:

$$\gamma_i^s = \frac{\Lambda_i^s}{\sum_{j=1}^n \Lambda_j^s} - \frac{P_i^s}{\sum_{j=1}^n P_j^s}. \tag{5}$$

If this is less than 0, then the bank takes on less counterparty risk than might be expected, given its share of payment activity. If it is greater than 0, then it takes on more risk than might be expected, given its share of payment activity.

Figure 2 comprises two histograms showing observed values of the measure γ_i^s over the sample period from 2 January 2008 to 28 May 2010. The banks are grouped in the same way as for the measure of cost (see notes to Figure 1). The relationship between size and risk is similar to before. The larger banks bear less risk than their shares of payment activity would suggest. The larger banks have a mean value of -0.05, compared to 0.03 for the smaller banks. Again difference in means tests reveal significance.

[12] Galbiati and Giansante (2010) model liquidity usage as a symmetric random walk. In their model, it can be shown that expected liquidity usage varies asymptotically with the square root of payment volumes.

[13] See Benos, Garratt and Zimmerman (2012).

Figure 2: **Frequencies of observations of risk measure of liquidity provision for larger and smaller banks over the period 2 January 2008 to 28 May 2010.**

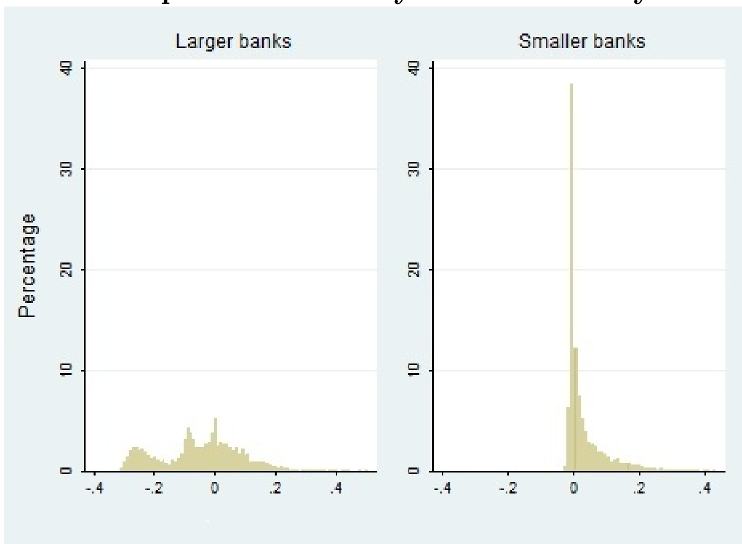


Table 2 shows the higher moments of the two distributions. Again the distribution for smaller banks is more positively skewed and has greater kurtosis.

Table 2: **Moments of distributions of the risk-based measures for larger and smaller banks over the period 2 January 2008 to 28 May 2010.**

	Larger banks	Smaller banks
Mean	-0.045	0.033
Variance	0.019	0.004
Skewness	0.111	2.398
Excess kurtosis	-0.148	6.571

Ball *et al* (2011) explain that many banks have internal schedulers which allow them to place bilateral limits against individual counterparties. By exercising these limits, a bank can directly control its positions against each counterparty, and thus manage its liquidity provision, as captured by the measure of cost. A bank that uses these limits dynamically to control the duration of its bilateral liquidity exposures would be able to manage its counterparty risk, as captured by our measure of risk. The results of the next section suggest that CHAPS settlement banks effectively manage the duration as well as the magnitude of these exposures.[14]

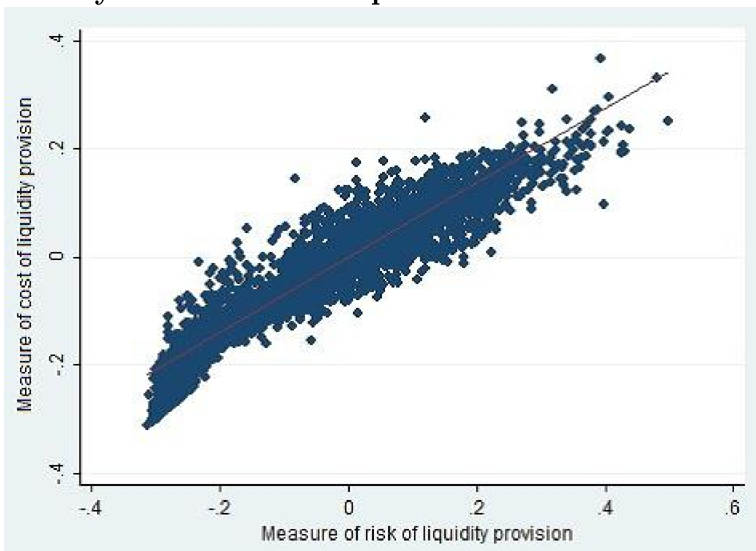
3.3 Comparing the measures of cost and risk of liquidity provision

Intuitively, the two measures of liquidity provision are likely to be correlated. In most cases a bank that builds up a large net debit position will do so gradually over the course of the day, in which case it may have a high score on both metrics. This correlation is shown in Figure 3, which plots the pairs (γ_i^s, c_i^s) for each of the settlement banks on each day over a period extending from

[14]This is confirmed by others too. For example, Jurgilas and Žikeš (2012) find evidence that settlement banks in CHAPS assign a positive and significant interest rate to the intraday timing of their payments.

2 January 2008 to 28 May 2010. Banks that provide more liquidity than their share of payments in both dimensions appear in the top right corner of the figure, whereas banks that provide less in both dimensions appear in the bottom left. This correlation is positive and significant, with an adjusted R^2 value of 0.871.

Figure 3: **The relationship between measures of cost and risk of liquidity provision, 2 January 2008 to 28 May 2010. The line represents the best fit to a linear relationship.**



4 Payment system inequality

In this section we measure payment system inequality in terms of the liquidity cost to the payee of making their daily payments. There are various approaches we could take to measure this inequality. We believe the Gini coefficient, a common measure of income inequality, has desirable properties that make it most suitable for this application.[15] The Gini coefficient has an intuitive interpretation and can be calculated for a sample of entities with different size populations. This is essential to our application, where the number of payments differs both across banks and from day to day.

Entropy measures, such as Theil's T measure, are less desirable for our purposes.[16] The main advantage of Theil's T is that it is decomposable, so that inequality within a group and between groups can be estimated. For instance, economists have compared income inequality within racial groups in a given country; see Bellù and Liberati (2006). We see no immediate use of the decomposition property in our application. Moreover, Theil's T cannot be computed if members of a population sometimes have a 0 value. Again this is unlikely when considering income, but it is feasible that a bank may provide zero liquidity, and hence have a cost per payment equal to zero. Finally Theil's T is not ideal for comparing populations with different sizes. If the number and sizes of groups differ, then the limit of the index will differ.

[15]See Gini (1912), Atkinson (1975) and Kakwani (1977) for the development of this methodology. Barro (2000) and Demirgüç-Kunt and Levine (2009) use inequality to investigate economic phenomena.

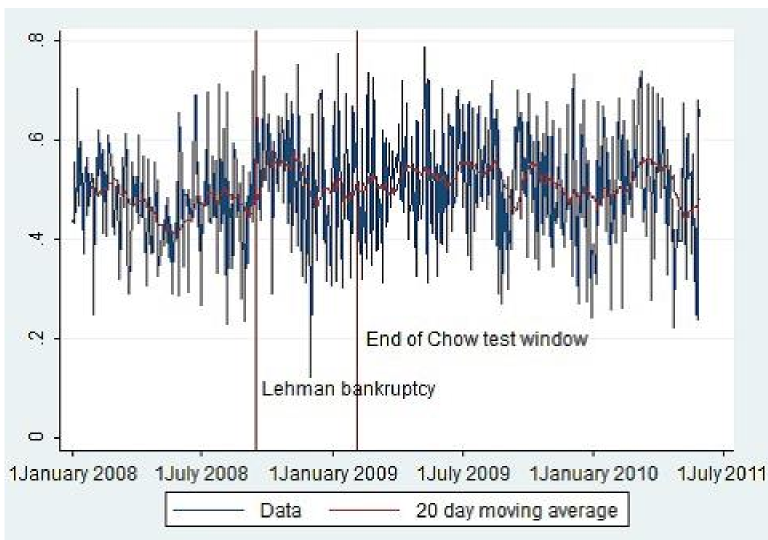
[16]The connection to entropy is very simple. Maximum entropy occurs with a uniform distribution, implying perfect equality. Minimum entropy occurs in a world of perfect certainty: i.e. where one person has all the wealth.

Following Ray (1998), the Gini coefficient for a population with a finite number of groups, each with a finite number of members, on day s is represented by:

$$G^s = \frac{1}{2M^2\mu}\left(\sum_{j=1}^{n}\sum_{k=1}^{n} m_j m_k |\ell_j^s - \ell_k^s|\right) \tag{6}$$

where n is the number of banks, m_j is the number of payments made by bank j, M is the total number of payments made by all banks, μ is the average liquidity cost of all payments and $\ell_j^s = \frac{L_j^s}{P_j^s}$ is the average liquidity cost of payments made by bank j. Rather than looking at differences between individual values and the mean, this measure sums over all the pairwise differences. This is convenient for our purposes since we can weight each settlement bank by the number of payments sent. G^s lies in the interval [0,1]. It takes a value of 0, denoting perfect equality, if all banks make payments using the same amount of liquidity per unit of payment. Conversely, G^s takes a value of 1 if one bank provides all the liquidity with which the other banks make their payments. Figure 4 plots the 20-day moving average value of G^s using data from 2 January 2008 to 28 May 2010.

Figure 4: **Gini coefficient G^s for CHAPS over the period 2 January 2008 to 28 May 2010.**



The figure shows that this measure is highly volatile, changing significantly from day to day. Moveover, the volatility of the series increases throughout 2008, suggesting the distribution of liquidity usage became less consistent. We test for the presence of a structural break in the series with a Quandt-Andrews test for an unknown breakpoint. We use the individual daily data from 2 January 2008 to 4 February 2009, since we might expect the merger of two settlement accounts on 5 February 2009 to cause a structural change to liquidity usage. The test identifies 5 September 2008 as the day when the test statistics were maximised (significant at the 1% level). Table 3 below shows the results. A Chow test for 15 September 2008 rejects at the 1% level the null hypothesis that there is no breakpoint on this date. We conclude that there was a significant increase in payment system inequality across CHAPS members around the time of the Lehman Brothers default on 15 September 2008.[17]

[17]Other studies have observed impacts on participant behaviour in payment systems in response to extraordinary

Table 3: **Quandt-Andrews break test for the Gini coefficient series. The period tested is 2 January 2008 to 4 February 2009. *** denotes significance at the 1% level.**

	Date maximised	F-statistic
Likelihood ratio	05/09/2008	16.48 ***
Wald ratio	05/09/2008	6828.93 ***

The Gini coefficient can also be computed using the measure of the risk of liquidity provision. Since the two measures are highly correlated the results are very similar, and are not included here.

Benos, Garratt and Zimmerman (2012) discuss the changes in CHAPS settlement banks' behaviour that occurred following the Lehman default. Increased concerns about bank-specific and system-wide risks led to slower payment processing, with evidence of targeted delay. This may have caused some banks to use less liquidity as they slowed down their payments, while others may have been forced to use more liquidity as they could not rely on recycling incoming payments.

Central bank actions can also have an effect on the aggregate measure. For example, if one day the Bank of England makes earlier payments in CHAPS, then other settlement banks might end up using less of their own liquidity, as they can recycle central bank liquidity instead. That would result in a decrease in liquidity provision measures. But, if these early payments had taken place in another system and the liquidity then transferred into CHAPS by the settlement banks, then the opposite effect could occur.[18] The actual effect on the Gini coefficient measure will be determined by the degree of heterogeneity in the impact of central bank actions on different settlement banks.

5 A recombinant approach to identifying over and under provision of liquidity

Our measures of liquidity provision capture differences in the amount and duration of liquidity provision by settlement banks. However, payment requests that come to settlement banks from customers are likely to be determined by factors exogenous to the banks, and so can be thought of as random from their perspective. Our liquidity provision measures depend not only on the payments made and received during a day, but also on the order that these payments take place. This means that when we look at a payment file for a given day and compute our measures, the liquidity provision we observe might just be accidental, in the sense that had the same payments arrived in a different order then — assuming all banks process payments immediately — the measures for individual banks would change.

We would like to be able to distinguish between cases where over or under provision is likely to be an artefact of random variations in timing of payments, and cases where they are not. To sort out the random component, we take the observed payments files over a sample period and re-order each day's payments many times to produce thousands of simulated days in which all the same payments are made, but in different orders. We then compute our measures of liquidity provision

events. McAndrews and Potter (2002) showed that, in the immediate aftermath of the events of 11 September 2001, the propensity of banks to send out payments through the Fedwire Funds transfer service declined. Likewise, Bech and Garratt (2012) documented dynamic adjustments in payment processing behaviour by Fedwire participants following the collapse of Lehman Brothers. Neither of these studies considered relative liquidity provision across banks.

[18]For example, a settlement bank taking central bank liquidity in CREST may appear to be using relatively little of its own liquidity in that system, but if the liquidity is moved to CHAPS and used to make payments, then the settlement bank's apparent liquidity provision in CHAPS would increase.

for each of these simulated days and construct a distribution of each measure over each day of the sample.

5.1 Simulation

Since the schedule of payments on each day varies, we need to consider a period of several days. We use 102 days of data (4 January 2010 to 28 May 2010). Each day's simulation involves randomly reordering around 125,000 transactions and so it is too computationally expensive to consider a longer period. Moreover, if we take a longer period there is greater risk of structural changes in banks' liquidity usage (for example, when a settlement bank takes on a large new customer). It is not feasible to consider all possible permutations for every day. Instead we simulate each day 200 times. We find that 200 is sufficient to produce stable empirical results.

We treat payments from or to the Bank of England and CLS Bank as exogenous, since these settlement banks do not have incentives to engage in strategic provision of liquidity. These payments therefore retain their order in the simulations. For example, if such a payment was the tenth payment of the day in reality, then it remains the tenth payment in our recombinant simulations of that day.

5.2 Results

Table 4 provides statistics on the number of occasions when c_i^s, our measure of the cost of liquidity provision of bank i on day s, falls below the 5^{th} or below the 95^{th} percentiles. Table 5 does the same for the measure of risk of liquidity provision. Absent of behavioural or structural factors that influence the timing of payments, we would expect each of these threshold values to be breached on around 5% of occasions. This is not the case — they are breached much more frequently. And there appears to be some heterogeneity across banks: some never breach these thresholds, while others breach them on more than half of occasions. This provides very strong evidence that there are additional structural or behavioural reasons that may cause banks to provide a share of liquidity to payment systems which differs from their share of payment activity.

Table 4: **Cost of liquidity provision: comparison of actual values to threshold values produced from the simulations over the period 4 January to 28 May 2010.**

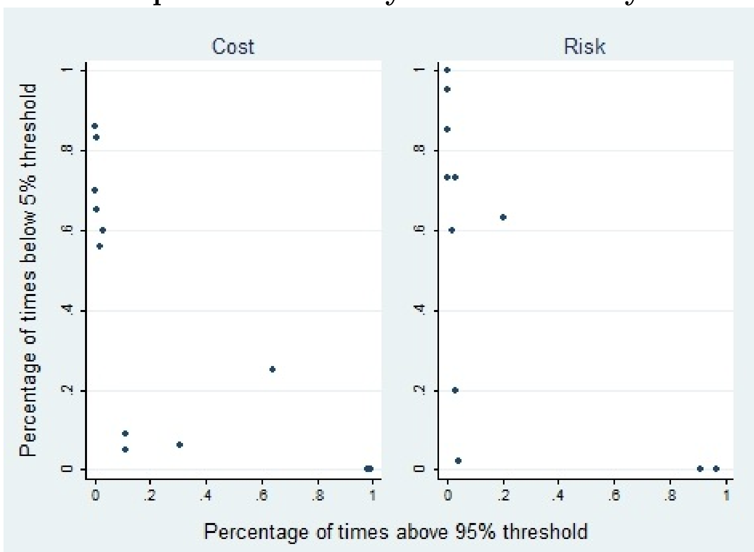
	Threshold	
	5%	95%
Threshold value breached, across all banks	27%	39%
Number of banks which never breach the threshold	2	2
Number of banks that breach the threshold on over half the occasions	3	6

Table 5: **Risk of liquidity provision: comparison of actual values to threshold values produced from the simulations over the period 4 January to 28 May 2010.**

	Threshold	
	5%	95%
Threshold value breached, across all banks	18%	56%
Number of banks which never breach the threshold	5	2
Number of banks that breach the threshold on over half the occasions	2	8

Figure 5 shows the frequency that individual banks are below the 5% threshold (horizontal axis) and above the 95% threshold (vertical) for the two measures of liquidity provision. If liquidity provision were entirely random, then each of these thresholds should be breached around 5% of the time, and we would expect all of the banks to be near the (5%, 5%) point. Instead, they are widely scattered, and there is a clear negative correlation.

Figure 5: **Frequencies of extreme values of liquidity provision measures for 12 settlement banks in CHAPS over the period 2 January 2008 to 28 May 2010.**



5.3 Alternative explanations for simulation results

There is no way to disentangle behavioral and structural causes of over and under provision of liquidity. As discussed above, banks can be under providers if they actively seek to reduce costs of liquidity provision or reduce risk by delaying payments. In what follows, we summarize structural reasons that would lead to over or under provision absent any intentional efforts to manage liquidity usage.

Differences in banks' business models could be a structural factor that affects their liquidity provision. For example, overnight loans involve an advance late in the day with a repayment the next morning. Therefore banks which tend to be net lenders may appear to be providing relatively less liquidity (as they pay late, they can do so from recycled liquidity), while those which are net borrowers may appear to be providing relatively more (as principal repayments are made early, they may need to be funded from reserves).

More generally, some banks may have payments which are time-critical intraday — they need to be made at some point before the end of the day. For example, pay-ins to CLS must be made in the morning UK time. This means that the bank may appear to be generous with its liquidity when in fact the early payment timing is driven by factors outside of the control of the bank's payment system operators.[19]

A bank which sends and receives a small number of very large payments would be expected to use more liquidity than a bank with a large number of smaller payments. This is because the latter

[19] According to Ball *et al* (2011), such time-critical payments comprise around 4% of values in the system.

bank would be better able to net payments and receipts and thus less likely to assume a large net sender position.[20]

Another issue is 'tiering' — that is, the use of the payment system via an account at a settlement bank rather direct membership. One of the reasons for doing so is liquidity pooling: Jackson and Manning (2007) argue that — unless customers' payment flows are highly correlated with those of the settlement bank — the total pool of liquidity needed will be smaller than that required if each customer were to become a settlement bank. This may mean that settlement banks with a large number of clients may need to use less liquidity than those of an equivalent size which do not have a large number of clients.[21] And banks with clients may behave differently, for example to manage their credit risk exposure to their customers.[22]

For banks with a lot of international customers, their location can have an important effect. For example, consider the case of a CHAPS settlement bank with a lot of clients based in the US. These clients might only send same-day instructions in the UK afternoon, due to time zone differences. Such a bank may appear to be hoarding liquidity when in fact it may be making its payments as soon as the instructions arrive. Clients based in Asia may have the opposite effect.

6 Concluding remarks

Our liquidity provision measures allow us to see whether banks provide amounts of liquidity that are comparable to their share of payment flows, and whether this changes over time. We provide evidence that small banks tend to provide more liquidity than larger banks in relative terms (though not necessarily in absolute terms). We also find evidence that the patterns of liquidity provision by some participants are unlikely to be explained solely by random variations in the timing of payment requests.

We cannot distinguish whether observed differences in liquidity provision result from banks intentionally conserving liquidity[23] or whether they result from structural factors. This distinction would be important if the goal was to make a normative assessment of a particular bank's behaviour. And it is possible that a regulator, operator or other participants in the system might want to do this, and perhaps take action in response. CHAPS has throughput guidelines designed to encourage banks to complete a certain percentage of their day's payments by a particular point in time during the day, and banks that fail to do this over an extended period of time have to explain themselves to a panel of their peers. Their arguments could relate to structural factors such as those discussed in Section 5.3. An individual investigation of a bank and its business model could be useful for shedding light on the extent to which it may be intentionally providing an amount of liquidity different to that which might be expected, given its payment activity.

The Gini coefficient we compute provides a measure of how the average liquidity cost of payments is distributed across the system. High values of the Gini coefficient indicate that some banks are much larger providers of liquidity relative to payments than others, and that some are effectively free-riding on the liquidity provision of others. The implication is that disruptions in liquidity provision due to failures or operational outages could have significant effects, because some banks

[20] This pooling effect is described in Galbiati and Giansante (2010).

[21] We cannot observe the internalisation effect of tiering. Flows between a settlement bank and its customers do not need to be submitted to CHAPS at all, but can simply be cleared across the settlement bank's own accounts without any central bank liquidity required.

[22] The Bank of England has identified tiering in UK payment systems, particularly CHAPS, as a source of risk to financial stability, as it leads to intraday credit and liquidity exposures between members and indirect participants. See Bank of England (2014).

[23] See Norman (2010) or Ball *et al* (2011).

are dependant on the liquidity provision of others. We found that the Gini coefficient in CHAPS increased around the time of the Lehman Brothers failure, increasing this concern.

An interesting by-product of our simulations is that we obtain a measurement of system-wide liquidity usage that abstracts from intraday payment arrival timing and processing decisions. Total liquidity actually used by these twelve banks averaged £18.7bn each day over the sample period, while in the simulations the system required £17.9bn on an average day. These results suggest that strategic choice in payment submission does not result in a substantially less efficient system, in aggregate, compared with our model where payments are simply made randomly.

References

Atkinson, A B (1975), 'The economics of inequality', *Oxford University Press.*

Ball, A, Denbee, E, Manning, M and Wetherilt, A (2011), 'Intraday liquidity: risk and regulation', *Bank of England Financial Stability Paper*, No. 11.

Bank of England (2014), 'The Bank of England's supervision of financial market infrastructures', *Annual Report*, March.

Barro, R J (2000), 'Inequality and growth in a panel of countries', *Journal of Economic Growth*, Vol. 32, pages 5-32.

Bech, M L and Garratt, R J (2012), 'Illiquidity in the interbank payment system following wide-scale disruptions', *Journal of Money, Credit and Banking*, Vol. 44, No. 5, pages 903-29.

Becher, C, Galbiati, M and Tudela, M (2008), 'The timing and funding of CHAPS Sterling payments', *Economic Policy Review*, Federal Reserve Bank of New York, Vol. 14, No. 2, pages 113-33.

Bellù, L G and Liberati, P (2006), 'Inequality analysis: the Gini Index', Food and Agriculture Organization, United Nations.

Benos, E, Garratt, R J and Zimmerman, P (2012), 'Bank behaviour and risks in CHAPS following the collapse of Lehman Brothers', *Bank of England Working Paper*, No. 451.

Committee on Payment and Settlement Systems (1993), 'Payment systems in the Group of Ten countries', *Bank for International Settlements*, December.

Davey, N and Gray, D (2014), 'How has the Liquidity Saving Mechanism reduced banks' intraday liquidity costs in CHAPS?', *Bank of England Quarterly Bulletin*, Vol. 54, No. 2, pages 180-9.

Demirgüç-Kunt, A and Levine, R (2009), 'Finance and inequality: theory and evidence', *World Bank policy research working paper*, No. 4330.

Finan, K, Lasaosa, A and Sunderland, J (2013), 'Tiering in CHAPS', *Bank of England Quarterly Bulletin*, Vol. 53, No. 4, pages 371-8.

Galbiati, M and Giansante, S (2010), 'Emergence of networks in large value payment systems (LVPSs)', *DEPFID Working Paper No. 01/2010.*

Gini, C (1912), 'Variabilità e mutabilità', *reprinted in 'Memorie di metodologica statistica', eds. Pizetti, E and Salvemini, T (1955).*

Jackson, J P and Manning, M J (2007), 'Central bank intraday collateral policy and implications for tiering in RTGS payment systems', *The Future of Payment Systems, Routledge.*

Jurgilas, M and Žikeš, F (2012), 'Implicit intraday interest rate in the UK unsecured overnight money market', *Bank of England Working Paper*, No. 447.

Kakwani, N C (1977), 'Applications of Lorenz curves in economic analysis', *Econometrica*, Vol. 45, No. 3, pages 719-727.

Manning, M, Nier, E and Schanz, J, eds. (2009), 'The economics of large-value payments and settlement', *Oxford University Press.*

McAndrews, J J and Potter, S M (2002), 'Liquidity effects of the events of September 11, 2001', *Economic Policy Review*, Federal Reserve Bank of New York, Vol. 8, No. 2, pages 59-79.

Norman, B (2010), 'Liquidity saving in real-time gross settlement systems – an overview', *Bank of England Financial Stability Paper*, No. 7.

Ray, D (1998), *Development Economics*, Princeton, NJ: Princeton University Press.

Soramäki, K and Bech, M L (2001), 'Gridlock resolution in interbank payment systems', *Bank of Finland Working Paper*, No. 9.

Valukas, A R (2010), 'Lehman Brothers Holdings Inc. Chapter 11 Proceedings examiner's report', U.S. Bankruptcy Court Southern District of New York, *Jenner and Block LLP*, Section III.A.5.