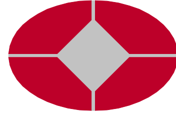
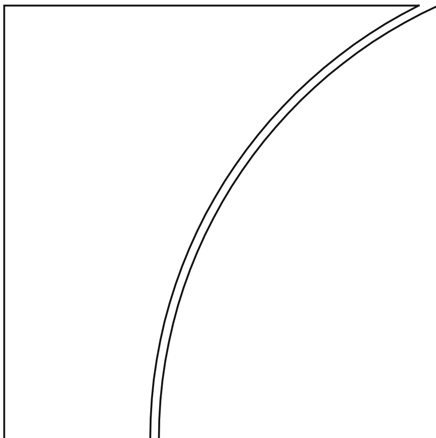


BANK FOR INTERNATIONAL SETTLEMENTS



BIS Working Papers

No 561

Mobile collateral versus immobile collateral

by Gary Gorton and Tyler Muir

Monetary and Economic Department

May 2016

JEL classification: E5, G2

Keywords: Liquidity Regulation, Collateral, Policy Evaluation with Economic History

BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The papers are on subjects of topical interest and are technical in character. The views expressed in them are those of their authors and not necessarily the views of the BIS.

This publication is available on the BIS website (www.bis.org).

© *Bank for International Settlements 2016. All rights reserved. Brief excerpts may be reproduced or translated provided the source is stated.*

ISSN 1020-0959 (print)
ISSN 1682-7678 (online)

Foreword

The 14th BIS Annual Conference took place in Lucerne, Switzerland, on 26 June 2015. The event brought together a distinguished group of central bank Governors, leading academics and former public officials to exchange views on the topic "Towards 'a new normal' in financial markets?". The papers presented at the conference and the discussants' comments are released as *BIS Working Papers* nos 561 to 564.

BIS Papers no 84 contains the opening address by Jaime Caruana (General Manager, BIS), the keynote address by John Kay (London School of Economics) and remarks by Paul Tucker (Harvard Kennedy School).

Mobile collateral versus immobile collateral

Gary Gorton[1] and Tyler Muir[2]

Abstract

The pre-crisis financial architecture was a system of mobile collateral. Safe debt, whether government bonds or privately produced bonds, ie asset-backed securities, could be traded, posted as collateral, and rehypothecated, moving to its highest value use. Since the financial crisis, regulatory changes to the financial architecture have aimed to make collateral immobile, most notably with the BIS "liquidity coverage ratio" for banks. In the face of the Lucas critique, how should these policies be evaluated? We evaluate this immobile capital system with reference to a previous regime, which had this feature: the US National Banks Era.

Keywords: Liquidity Regulation, Collateral, Policy Evaluation with Economic History

JEL classification: E5, G2

[1] Yale and NBER.

[2] Yale.

This paper was prepared for the June 2015 BIS Annual Conference in Lucerne, Switzerland. Thanks to Adam Ashcraft, Darrell Duffie, Randy Krozner, Andrei Kirilemko, Arvind Krishnamurthy, Manmohan Singh, Paul Tucker, Warren Webber and seminar participants at the Penn Institute for Economic Research Workshop on Quantitative Tools for Macroeconomic Policy Analysis, the Stanford Global Crossroads Conference for comments. Thanks to Charles Calomiris, Ben Chabot, Michael Fleming, Ken Garbade, Joe Haubrich, John James, Richard Sylla, Ellis Tallman, Warren Weber and Rosalind Wiggins for answering questions about data. Thanks to Lei Xie and the late Bruce Champ for sharing data. Thanks to Toomas Laarits, Rhona Ceppos, Ashley Garand, Leigh-Anne Clark and Michelle Pavlik for research assistance. Special thanks to the Federal Reserve Bank of Cleveland for sharing the data of Bruce Champ.

1. Introduction

In the 30 years prior to the 2007–08 financial crisis, the global financial system evolved away from a system of immobile collateral into a system of mobile collateral. In this new system, bank loans, instead of remaining immobile on bank balance sheets, were securitised into bonds which could be traded, used as collateral in repo, posted as collateral for derivatives positions, and rehypothecated, moving to the location of their highest value use. In short, US banks' loans, which had been sitting passively on bank balance sheets, were transformed into bonds, making them mobile. Since the financial crisis of 2007–08, regulatory initiatives have been aimed at making collateral once again immobile. These new policies raise a very important Lucas critique problem: How should these policies be evaluated when we do not have adequate models? In this paper, we argue that economic history can be used to evaluate the recent policy of again making collateral immobile. With this approach, we have two main arguments in our evaluation. The first is that there is a cost to making this collateral immobile because it ties up safe debt. The second is that we find weak support for the idea that making collateral immobile has large benefits – namely that it makes the financial system safer and reduces panics. Instead, we find that other forms of bank debt increase when safe collateral becomes immobile, possibly making the system riskier.

The transformation of the US financial system towards mobile collateral is shown in Figure 1 (from Gorton, Lewellen and Metrick (2012)). The figure shows the different forms of privately produced safe debt as a percentage of total privately produced safe debt. In the 1950s and 1960s, demand deposits of banks were 80% of privately produced safe debt. Demand deposits were backed by bank loans, essentially the collateral. Demand deposits then go into a downward trend for the next 30 years. At the same time, the categories of privately produced safe debt that are growing were money market instruments (specifically, repo, commercial paper and money market funds), as well as AAA mortgage-backed securities (MBS) and other AAA asset-backed securities (ABS). By the start of the financial crisis, this "shadow banking system" was as large as the amount of demand deposits (although much of the ABS and MBS ended up as collateral backing repo and asset-backed commercial paper). This shadow banking system is funding the very same bank loans but in a different way.

At the same time that this transformation of the financial system was going on, another important, related, trend was developing. This is shown in Figure 2. Figure 2 shows the holders of outstanding US Treasury debt (excluding Treasuries held by the federal, state and local pensions; social security is also not counted) as a percentage of total outstanding Treasuries. Again, there is a steady downward trend for US banks, as the banks no longer need Treasuries as a component of their portfolios backing demand deposits. Coincidentally, the Rest of the World shows a sharply upward trend in holdings of US Treasury debt. The Rest of the World has a demand for safe, liquid, debt. This is important because of late outstanding US Treasury debt is large, but this should be looked at relative world GDP for perspective. In other words, Treasury debt has a convenience yield. Krishnamurthy and Vissing-Jorgensen (2012a,b) show empirically that investors value the money-like properties of liquidity and safety of these bonds. Also see Duffee (1996). Furthermore, there is a negative correlation between outstanding US Treasury debt and the production of privately produced safe debt (see Krishnamurthy and Vissing-Jorgensen (2012a, b) and Gorton, Lewellen and Metrick (2012)). When Treasuries outstanding (as a percentage of GDP) declines, privately produced safe debt increases to fill the gap. More specifically, Xie (2012)

shows, using daily data, that asset-backed security (ABS) and mortgage-backed security (MBS) issuance occurs when there is a high convenience yield on Treasuries. In other words, when Treasuries are scarce, more ABS and MBS are issued (85% of an issue is AAA). Sunderam (2014) shows that a high convenience yield results in the endogenous response of more asset-backed commercial paper (ABCP) being issued.

The private response to create more privately produced safe debt when there is a scarcity of US Treasuries is partly the focus of the literature on the "global savings glut". This literature argues that increased capital flows into the United States from countries with an excess of savings may have been an important reason that US interest rates were low (see, eg, Bernanke (2005, 2007), Bernanke, Bertaut, De Marco and Kamin (2011)). Foreign investors have a demand for safe assets and focused their investments on US Treasuries (see Caballero and Krishnamurthy (2009)). Maggiori (2013) argues that this as a form of insurance; the US can (almost uniquely) produce safe debt, which is demanded by foreign investors in countries which cannot produce safe debt (China and the oil-producing countries; see Bertaut, De Marco and Kamin (2011)).

The creation of privately produced safe debt is in part a response to a scarcity of government-produced safe debt. Both types of safe debt – government-produced and privately produced – are used as collateral for short-term bank money, ie repo, ABCP and money market funds (MMF). The response to scarcity was not only increased private production of safe debt, but also increased mobility of the debt. In the past, when US Treasuries or bank loans backed demand deposits, the backing loans were immobile. Indeed, since demand deposits were the dominant form of inside money, there was no need for mobility. The bank loans sat on bank balance sheets and were not traded. As the need for demand deposits receded, and foreign investors demanded US Treasuries, the financial system transformed the immobile bank loan collateral into forms of mobile collateral via securitisation. Privately produced safe debt became a product to be used as mobile collateral.

In Section 2 of this paper, we provide evidence of the extent to which the financial system became a system of mobile collateral, indeed, stretching the available mobile debt to meet demands. To do this we study the determinants and extent of repo fails. We show that repo fails were increasing because of the scarcity of US Treasuries and Agency bonds. Another manifestation of system morphing was that repo was significantly expanding beyond the primary dealers prior to the crisis. And, fails were caused in significant part by a demand for liquidity in a world with insufficient safe debt. This was, and remains, a problem. Krishnamurthy and Vissing-Jorgensen (2012b) argue empirically that when the ratio of privately produced safe debt goes up relative to Treasuries, financial crises are more likely. Also see Gourinchas and Jeanne (2012). This scarcity appears to persist today. For example, Bertaut, Tabova and Wong (2014) show that since the financial crisis US investors have invested in the sovereign debt of Australia and Canada because of the shortage of safe debt since securitisation collapsed.

New regulations do not address potential scarcities of safe debt. In fact, since the financial crisis, new regulations aim at returning to a financial system of immobile collateral. For example, under Dodd-Frank and similar European legislation collateral must be posted to central clearing parties (CCPs) (regardless of the private party's net position), while the CCP does not post collateral to participants. CCPs will only accept highly liquid, high-grade collateral. Variation margin has long been part of the bilateral swap market, but importantly, initial margin is new and will increase

substantially the amount of collateral required. Not all swaps trades will be cleared through a CCP. For those that are not, initial and variation margin for each trade must be held by a third party. Further, collateral posted to banks by clients cannot be rehypothecated. And, most importantly, the Basel Committee's Liquidity Coverage Ratio (LCR) requires that (net) short-term bank debt be backed by (essentially) Treasuries ("high-quality liquid assets")[3] – see Basel Committee (2013). In other words, short-term bank money (repo and CP) and Treasuries, both of which have a convenience yield, must be combined. One kind of money must be backed by another kind of money, a kind of narrow banking.

The LCR is the leading example of the move back towards an immobile collateral system. In effect, it attempts to reverse 30 years of change. As such an important change, how should we evaluate the LCR (or any new policy for that matter)? Because of the Lucas (1976) critique, a general equilibrium model would be needed, one which currently does not exist. There have, of course, been numerous more or less ad hoc forecasts of how much collateral the new system will need given the LCR, but these numbers vary a lot and are subject to the Lucas critique, ie there is no accounting for the general equilibrium effects that might occur ("unintended consequences"). In fact, we know from the above studies that privately produced money grows when there is a scarcity of Treasuries.

We are not interested here in forecasting the amount of collateral needed.[4] Rather, in this paper we evaluate the LCR *structure* by analysing a financial system that had the same structure of requiring that US Treasuries back the issuance of privately issued bank money: the US National Banking Era, 1863–1914.[5] Under the National Banking System, national banks could issue distinct "national bank notes" by depositing eligible US Treasury bonds with the US Treasury, which would then print the bank's notes. Originally, the idea was to create a demand for US Treasuries so as to finance the US Civil War. But, it was also believed that backing private money with Treasuries would prevent banking panics. Prior to the National Banking Era, US banks issued their own distinct notes, backed by state bonds (in Free Banking states) or backed by portfolios of bank loans (in chartered banking states). There were systemic banking crises in 1814, 1819, 1837 and 1857. It was expected that the National Banking System would eliminate panics. Similarly, the explicit purpose of the liquidity coverage ratio (LCR) is to make the financial system safer. Basel Committee (2013) labels the LCR as one of the key reforms "to develop a more resilient banking sector" (p 1).

However, this stability did not occur under the National Banking Era. Banking panics were not prevented, but merely shifted from one form of bank money to another. During a panic, instead of requesting cash for private bank notes, debt holders demanded national bank notes for their demand deposits. By the time the National Bank Acts were passed demand deposits had become a sizeable form of privately produced bank money and we show that they continued to grow substantially throughout this period (Figure 9). But, economists did not understand this for decades. Bray Hammond (1957), in his Pulitzer Prize-winning book *Banks and*

[3] Throughout the paper, we will refer to "high-quality liquid assets" as Treasuries and we will refer to net short-term debt as repo.

[4] On this topic see Heller and Vause (2012), Sidanius and Zikes (2012), Fender and Lewrick (2013) and Duffie, Scheicher and Vuillemey (2014).

[5] On the National Banking Era see Noyes (1910), Friedman and Schwartz (1963), and Champ (2011c).

Politics in America, wrote: "… the importance of deposits was not realized by most American economists … till after 1900" (p 80). Hammond goes on to discuss why the growing importance of demand deposits was overlooked. Economist Charles Dunbar (1887) wrote in the inaugural volume of the *Quarterly Journal of Economics*: "The ease with which we ignore deposits as a part of the currency seems the more remarkable, when we consider that few men in business fail to recognize the true meaning of this form of bank liability" (p 402). And Russell C. Leffingwell, the Assistant Secretary of the Treasury wrote as late as 1919: "All of these people who believe in the quantity theory of money … choose to call bank deposits money, but bank deposits are not money" (Leffingwell letters, quoted by Wicker (1966, p 21). Regulators and economists were conceptually confused. What seems so obvious now was not obvious then. Forms of short-term bank debt change and this was not immediately recognised. Worse, the new forms of bank debt and their risks tend to be misunderstood. It is possible that the system even contributed to panics by encouraging a different, less well understood form of bank debt to be created. This is a key point because a main benefit of the proposed system of immobile collateral is that it will make the system safer.

There was another problem with the National Banking System: too little money was issued and the system was inflexible or "inelastic". Too little money was issued even though it was apparently profitable to do so, an apparent riskless arbitrage opportunity. Economists have called this the "under-issuance puzzle" or the "national bank note puzzle." First noticed by Bell (1912), this under-issuance has been a puzzle ever since, for over a century! The puzzle is that national banks never fully utilised their note-issuing powers even though it appears that it was profitable to do so. As Kuehlwein (1992) put it: "… through the turn of the century and into the 1920s banks devoted a significant fraction of their capital to direct loans … despite the fact that national bank notes appeared to be more profitable" (p 111). Friedman and Schwartz (1963, p 23) reached the same conclusion. There is a large literature on this; see Calomiris and Mason (2008) and the citations therein.

Because the LCR is structurally the same as the National Banking System, this puzzle is important. In this paper, we show that the reason that "riskless arbitrage profits" persisted during the National Banking Era was that the calculations of the arbitrage profit done to date ignored the fact that there was a convenience yield to Treasuries and a cost to bank capital. Banks held Treasuries on their balance sheets but, in principle, could have raised capital to buy more Treasuries.[6] But, for the system as a whole, there appears to have been a shortage of safe debt. Simply put, banks had other important uses for Treasuries and bank capital was expensive. We show that the "arbitrage profits" are essentially a proxy for the "convenience yield" on Treasuries or the cost of bank capital or likely both. This suggests that backing one kind of money (National Bank notes) with another kind of money (Treasuries) may not be such a good idea. By linking the two forms of money, another form of private-produced money is likely to appear or grow. This is strongly shown in the data – as the share of Treasuries to GDP declined over this period, deposits grew. And a shortage of safe debt is associated with financial instability. This too is consistent with the data as banking panics occurred frequently throughout the period (1873, 1884, 1890, 1893, 1896, and 1907).

[6] Also, average profit rather than marginal profit was calculated. This distinction is important, as we discuss below.

Of course, it will be objected that the two systems – the current LCR and the historical National Banking Era – are different and that the National Banking Era is not relevant. Indeed, there are important differences. Nevertheless, it seems useful to analyse a system that is structurally identical to the LCR. In terms of policy evaluation, there seems no real alternative.

We summarise with the following conclusions about the LCR or other systems that make collateral immobile. First, these systems have costs because they tie up safe assets and add to issues of scarcity. Second, we cast doubt on the system's supposed benefits – namely preventing crises. Historically, the National Banking Era had many panics. It is likely that a system of immobile collateral that restricts certain forms of bank debt creation simply encourages other forms of bank debt to be created. This is consistent with the growth in deposits during the National Banking Era. There is a remarkably strong correlation between Treasury supply and deposits (Figure 9). These new forms of bank debt are dangerous because they are typically not well understood or acknowledged, as deposits were not well understood in the National Banking Era and as shadow money (repo, ABCP etc) were not understood until the recent crisis.

The paper proceeds as follows. In Section 2, we provide more evidence on the scarcity of safe assets in the period leading up to the financial crisis of 2007–08 by looking at repo fails. Section 3 is devoted to the National Banking Era. We calculate the profitability of national bank note issuance and then show that, even in the 19th and early 20th centuries, US Treasuries had a convenience yield. We then show that this can resolve the century-old national bank note paradox. Section 4 concludes with implications for the present day.

2. Collateral mobility and scarcity

Many authors have discussed the shortage of safe debt prior to the financial crisis, eg Caballero (2010), Gourinchas and Jeanne (2012), usually relating it to the global savings glut. But, it has proven difficult to provide evidence for this shortage. In this section, we provide some evidence for this shortage, which was driving the growth of mobile collateral.

2.1 Repo fails

There is said to be a "repo fail" if one side to the repo transaction does not abide by the contract at maturity, failing to deliver the collateral back (called a "failure to deliver") or failing to repay the loan (called a "failure to receive"). See Fleming and Garbade (2005, 2002) on fails. Repo fails can provide indirect evidence on scarcity and mobility. If collateral is scarce, then it can become more mobile via rehypothecation (re-use) chains, making it more difficult to find the bond to return to the borrower, ie a fail. There is no direct evidence of this, but we provide a variety of indirect evidence.

We examine data from the New York Federal Reserve Bank on primary dealers' fails.[7] The primary dealers are only a subset of all firms involved in the bilateral repo market, as we will see below. But, still it encompasses many large financial firms. The New York Federal Reserve Bank collects data on only three asset classes used as collateral for repo: US Treasuries, Agency bonds, and Agency MBS.[8]

Repo fails by asset class are shown in the three panels of Figure 3. From Figure 3 it is apparent that repo fails were increasing prior to the financial crisis. It is apparent from the figure that the period from January 2000 until January 2010 is more turbulent than the period before and the period after. The turbulence is not just the financial crisis. This is confirmed by Table 1 Panel A which shows the mean dollar amount of fails (in millions of dollars) in the 1990s compared to the period 2000–07; also shown are the standard deviation of fails. We formally test for difference between subperiods below.

The data collected by the New York Fed are very limited. To get some sense of the narrowness of the primary dealer group, we can look at data from the Depository Trust and Clearing Corporation (DTCC) on fails. DTCC has hundreds of members that use DTCC for clearing and settlement.[9] In 2011, DTCC settled $1.7 quadrillion in security value. DTCC also has a large repo programme. DTCC fails data are for the value of Treasury and Agency fails, that is the amounts that were not delivered to fulfil a contract. The DTCC data covers all fails of Treasuries and Agencies, not just repo fails. However, if there is a scarcity of safe debt, then there are likely fails in trades as well as repo. The DTCC series is not as long as the NY Fed's, but it shows a larger universe of players. The data are shown in Figure 4; looking at the scale of the y-axis it is clear that there are many more fails, suggesting that the size of the fails problem is an order of magnitude larger than the NY Fed data show. Also, see Gorton and Metrick (2015).

Aside from operational issues that explain repo fails, there are two other possibilities. First, there is the possibility that a counterparty strategically defaults to retain the bonds or retain the cash, at least for a short period. Secondly, there can be multiple fails due to rehypothecation (the re-use of collateral) chains, ie several transactions are sequentially based on the same collateral. As explained by Fleming and Garbade (2002): "... a seller may be unable to deliver securities because of a failure to receive the same securities in settlement of an unrelated purchase. This can lead to a "daisy chain" of cumulatively additive fails: A's failure to deliver bonds to B causes B to fail on a sale of the same bonds to C, causing C to fail on a similar sale to D, and so on" (p 43). Also, see Singh (2014). We do not have the data, however, to distinguish between fails due to rehypothecation chains from other fails. We cannot distinguish between these possibilities, but the tests below strongly suggest that increasingly fails were not operational errors.

Collateral is mobile if it is in a form that can be traded and posted as collateral in repo or derivatives transactions. Rehypothecation is another form of collateral mobility. What is the extent of rehypothecation? There are some survey data from the

[7] "Primary dealers" are financial firms that are trading counterparties if the New York Fed in its implementation of monetary policy. There are currently 22 primary dealers; see http://www.newyorkfed.org/markets/pridealers_current.html.

[8] "Agency" refers to Fannie Mae, Freddie Mac or Ginnie Mae, government-sponsored enterprises that securitise and guarantee certain types of residential mortgages.

[9] See the DTCC membership list: http://www.dtcc.com/client-center/dtc-directories.aspx.

International Swaps and Derivatives Association (ISDA). ISDA has an annual survey of its members that usually asks about the extent of rehypothecation using collateral received in OTC derivative transactions, in terms of the percentage of institutions that report that they do rehypothecate collateral. In 2001, the first survey, 70% or the respondents reported that they "... actively re-use (or 'rehypothecate') incoming collateral assets in order to satisfy their own outgoing collateral obligations" (p 3). Over the years, the percentage rises to 96% for large firms in 2011. In 2014, ISDA for the first time asked about which bonds were actually used for rehypothecation. Table 1 Panel C shows the results. In 2014, ISDA estimated that total collateral used in non-cleared OTC derivatives to be $3.7 trillion. It would appear that rehypothecation is sizeable. This does not address the question of the length of rehypothecation chains. Singh (2011) estimates that prior to the financial crisis, collateral velocity was three. Also see Singh and Aitken (2010).

This is not the only evidence on scarcity and mobility. The bilateral repo market was expanding significantly beyond the primary dealers in the 2000s. In the New York Fed data, if one dealer fails to deliver to another dealer, then the first dealer records a "fail to deliver" of $N, and the counterparty primary dealer reports a "fail to receive" of $N. So, fails and receives should be equal, unless the primary dealers are trading with firms that are not primary dealers.

To examine whether repo was expanding beyond the primary dealers we look at the difference between receive and fail by asset class. If all the fails are between primary dealers, then this number will be zero. So, if this number is positive, then it means that the party failing to deliver was not a primary dealer, the primary dealer records a "fail to receive". Figure 5 shows failure to receive minus failure to deliver by asset class. Again it is apparent that this number was near zero prior to 2000, meaning that all fails were with another primary dealer. But, after 2000 and prior to the crisis, Receive minus Deliver is clearly not zero. In this period, there are significant fails by non-primary dealer counterparties, suggesting that the bilateral repo market had grown significantly, consistent with collateral being mobile and scarce. (Also see Gorton and Metrick (2015)).

This is confirmed in Table 1 Panel B, where it is clear that failure to receive minus failure to deliver increasingly differs from zero in the period 2000–07, prior to the crisis. Moreover, note the sign difference between Treasuries and MBS during 2000–07. For Treasuries, receive minus fail is very large in 2000–07, again meaning that non-primary dealers are not delivering Treasuries according to their repo contracts. But, in the case of MBS, the number is very negative, meaning that primary dealers are failing to deliver to non-primary dealer counterparties. This is also apparent in the figure.

The fails data on Agency MBS market are very different, likely because the repo fails number includes fails in the "to be announced" (TBA) market, although the data do not allow us to decompose the fails. (See Government Securities Dealers Reports (2015), Board of Governors of the Federal Reserve System). TBA contracts are forward contracts for the purchase of "to be announced" agency MBS. In this market, the MBS to be traded are not specified initially. Rather, the parties agree on six general parameters of the MBS (date, issuer, interest rate, maturity, face amount, price). The contracts involve a delayed delivery, typically an interval of several weeks. In the TBA market transactions are usually "dollar rolls", agreements involving the purchase or sale of an agency MBS with a simultaneous agreement to resell or repurchase MBS at a specified price. In a TBA dollar roll the securities returned need only be "substantially

similar," unlike in a repo transaction. The TBA market is very large. Average daily fails in this market between 31 December 2009 and 29 December 2010, as reported by primary dealers, was $83.3 billion in fails to deliver and $73.8 billion in fails to receive (see Treasury Market Practices Group (TMPG) (2011)). On the TBA market, see Vickery and Wright (2013).

2.2 Fails and the demand for liquidity

We now turn to some formal evidence that collateral became increasingly mobile. We start the analysis by testing to see if there are significant breakpoints in the panel of fails (receive and deliver) data. To do this we follow Bai (2010). Bai (2010) shows how to find breakpoints in panels of data where a breakpoint is in the mean and/or the variance. Assuming a common breakpoint in a panel of data is more restrictive than assuming random breakpoints in the individual different series in the panel, but the method can be used on an individual series as well.

The method can be used to find other breakpoints subsequent to the first. The first breakpoint divides the panel into two subpanels, on each side of the first breakpoint. To find the second breakpoint apply the procedure to each of the two subseries, on the two sides of the first breakpoint. The second breakpoint is the one that gives the larger reduction in the sum of squared residuals, when comparing the break found in each of the two subseries.

We examine a panel of four series: fail to deliver and fail to receive for Treasuries and for Agencies. We omit MBS for reasons discussed above. The sample period of weekly data runs from July 1991 to September 2014. The breakpoints are shown in Table 2. The table shows the 95% confidence intervals in the last two columns in terms of dates. From the figures above it is clear that fails are increasing, starting in the early 2000s. Consistent with this the first breakpoint is 12 September 2001, just after 11 September 2001. This is the start of a different regime and it extends until, not surprisingly, the second break chronologically just after Lehman. The third breakpoint is 9 February 2009.

Why were fails increasing? We will examine the proposition that fails increased as the demand for liquidity increased. We follow Xie (2012) in measuring the convenience yield by the spread between the rate on general collateral (GC) repo and the rate on the Treasury used as collateral for the repo. The maturity is one month. In GC repo, the lenders will accept any of a variety of Treasuries as collateral, ie it is general collateral rather than specific collateral. Xie shows that this spread has historically (1991–2007) been 36 basis points (see Xie (2012)), reflecting the fact that with GC repo the Treasury must be returned and the cash is tied up during the time of the repo. Obtaining a Treasury via reverse repo is not as good as actual ownership of the Treasury.

To be clear, the GC repo to Treasury spread of 36 basis points means that the borrower is *losing* money to the lender for the cash lent. Take the more usual case, say with privately produced collateral. Then, if this underlying collateral earned 6% and the repo rate was 3%, then the borrower would be earning the spread of six minus three. But with Treasuries as collateral, because of their safety and liquidity, the borrower is earning negative 36 basis points. The borrower is paying to own a Treasury financed in the repo market. The amount of repo supplied then is linked to the convenience yield of Treasuries.

The basic idea we explore is whether an increase in the GC repo spread, ie an increase in the convenience yield, is associated with an increase in repo fails. In other words, if there is an increase in the demand for liquidity, then this spread will widen. A widening of the spread corresponds to an increased scarcity of Treasuries, and possibly other safe collateral as well, such as Agency bonds and Agency MBS.

We use differences-in-differences in seemingly-unrelated regression on the panel of the Treasuries and Agency bonds, where fails are normalised by fails in 2013. We indicate the three breaks discussed above. The first period is 12 September 2001 going until 23 September 2008 and the second is 24 September 2008 until 10 February 2009, followed by 11 February 2009 onwards. This means that there are four periods: prior to 2001, break 1, break 2 and break 3. We will look at specifications with and without lags. We also include the change in the one-month T-bill since the level of the interest rate effects the incentive to fail. In a repo fail the implicit penalty is the interest that could have been earned elsewhere, so in a low interest rate environment the penalty is low.[10]

The regression results for fails to receive are shown in Table 3. The table for fails to deliver is in the Appendix. In both cases, the interaction between the GC repo spread and regime 1 is significant. Break 1 is the period prior to the crisis up to 24 September 2008. In both cases, the interaction of the spread with the period of Break 2 is also significant and these coefficients are larger. Break 2 runs from 24 September 2008 to 11 February 2009, the post-Lehman period. Changes in the convenience yield or the demand for liquidity appear to have driven repo fails in the period prior to the financial crisis and during the crisis. The first regime corresponds to the period of the scarcity of safe debt while the second regime corresponds to the flight to quality. This is true for both fails to receive and fails to deliver. The change in the one-month T-bill rate is also significant, suggesting that the incentive to fail is related to the level of the interest rate. Finally, the three dummy variables for the three break regimes are not significant. This may be due to a combination of factors. The break points may be driven by the variance, and the interaction terms may be absorbing this effect.

In the Appendix, we examine the breakpoints for the absolute value of fails to deliver minus fails to receive in the Treasury and Agency MBS repo markets. This is the variable that measures the growth of the repo market beyond the primary dealers. The results in the Appendix show the seemingly unrelated panel regression results for the absolute value of fails to deliver minus fails to receive in the Treasury and Agency MBS repo markets. Not surprisingly this difference in the repo market is not driven by demands for liquidity. Instead, the market is growing for other structural reasons, eg the rise of large money managers, and foreign investors (see Gorton and Metrick (2015)).

The Agency MBS breakpoints for the absolute value of the difference between fails to receive and fails to deliver are shown in Panel B of Table 1. The break points are a bit different from those for Treasuries and Agency MBS. The regressions for Agency MBS fail to receive and fail to deliver are in the Appendix. For both types of fails the liquidity demand, as captured by the GC repo-Treasury spread is significant for the second break period, from 16 September 2009 to 27 December 2011. Throughout the financial crisis, a significant amount of agency MBS continued to be

[10] For this reason, the Treasury Market Practices Group introduced a "dynamic fails charge" to provide an incentive for timely settlement. See Garbade, Keane, Logan Stokes and Wolgemuth (2010).

issued. In 2008 and 2009, $2.89 trillion was issued (see Vickery and Wright (2013)). This was a period of a flight to quality, that is, a desire to hold very liquid securities.

2.3 Summary

The evidence for collateral mobility is indirect because there are no data or limited data on rehypothecation, trading, and collateral posting for derivative positions and for clearing and settlement. Nevertheless, the size of the securitisation market prior to the crisis and the evidence above, indicate the system of mobile collateral that had developed.

3. An immobile collateral system

Bolles (1902) described the US National Bank Act as "... the most important measure ever passed by any government on the subject of banking". The National Bank Acts were passed during the US Civil War; the first Act was passed in 1863 and this law was amended in 1864. The Act created a new national banking system. The Act was intended to create a demand for US Treasury bonds because, without an income tax, it was the only way to finance the North in the war. The Acts established a new category of banks, national banks, which were to coexist with state chartered banks. National banks could issue bank-specific national bank notes by depositing eligible US Treasury bonds with the US Treasury.[11] In this section, we examine the US National Banking Era. In subsection 3.1, we provide a very brief background on the banking system in the era 1863–1914. In subsection 3.2, we introduce the "bank note paradox". We show the "arbitrage profits" that allegedly existed. The analysis of the profitability of note issuance and its relation to the convenience yield on Treasuries is in subsection 3.3. Subsection 3.4 summarises the results of this section.

3.1 The US National Banking System

During the US National Banking Era, banks were required to back their privately produced money in the form of bank-specific national bank notes with US Treasury bonds. One kind of money was required to back another kind of money – narrow banking. So, there was a collateral constraint on the issuance of money by banks. As with repo today, the interest on the bonds went to the banks. With national bank notes backed by US Treasuries there was for the first time in the US a uniform currency. Prior to the Acts, banks issued individual private bank notes that traded at discounts to face value when traded at a distance from the issuing bank. There were hundreds of different banks' notes, making transacting difficult. Initially, national bank note issuance was limited to 100% of a bank's paid-in capital, but this was changed to 90% by the act of March 3, 1865. Also, note issuance was limited to 90% of the lower of par or market value. This was changed to 100% by an act in 1900. See Noyes

[11] Eligible bonds were US Treasury government registered bonds bearing interest in coupons of 5% or more to the amount of at least one third of the bank's capital stock and not less than $30,000. The Act of 12 July 1870 eliminated the requirement that bonds bear interest of 5% or more. After that date eligible bonds were "of any description of bonds of the US bearing interest in cash."

(1910), Friedman and Schwartz (1963), and Champ (2011c) for more information on the National Banking Era.

3.2 The bank note issuance puzzle

The National Banking Era has been puzzling for economists for well over a century. The puzzle is that there appears to have been high, allegedly sometimes infinite, profits from issuing national bank notes – riskless arbitrage profits – but this capacity to issue notes was never fully utilised. Friedman and Schwartz (1963, p 23): "... despite the failure to use fully the possibilities of note issue, the published market prices of government bonds bearing the circulation privilege were apparently always low enough to make note issue profitable ... The fraction of the maximum issued fluctuated with the profitability of issue, but the fraction was throughout lower than might have been expected. We have no explanation for this puzzle." They go on to write: "Either bankers did not recognise a profitable course of action simply because the net return was expressed as a percentage of the wrong base, which is hard to accept, or we have overlooked some costs of bank note issue that appeared large to them, which seems must more probable" (p 24).[12]

Phillip Cagan (1963, 1965) determined whether it was profitable for banks to issue notes by examining the following formula:

$$r = \begin{cases} \dfrac{r_b p - \tau\alpha \min(p,1)}{p - \alpha \min(p,1)} & if \quad p > \alpha \min(p,1) \\ \infty & if \quad p = \alpha \min(p.1) \end{cases}$$

where: r is the annual rate of return on the issuance of national bank notes; p is the price of the bond held to back the notes (dollars), assuming a par value of one; r_b is the annualised yield to maturity on the bond held as backing; α is the fraction of the value of a given deposit of bonds that could be issued as notes; and τ is the annual expense in dollars of issuing $\alpha \min(p,1)$ in notes. The term $\alpha \min(p,1)$ refers to the amount of notes that are returned to the issuing bank by the US Treasury from the deposit of a bond with price p. The variable τ includes the tax rate on note issuance, which was $0.01 for $1 prior to 1900 and $0.005 on 2% coupon rate bonds after 1900. Also, miscellaneous costs are included here. For example, Cagan used an estimate of these costs of 0.00625 per one-dollar deposit in government bonds.[13]

Champ (2011b) gives the following example. Consider a bank in 1890 (ie α =0.9) that purchased a bond for$1.10, with yield to maturity of 4%. Then, the total cost of note issuance is $\tau = 0.01 + \dfrac{0.00625}{\%0.09} \approx 0.01694$. So, in this case, the rate of profit for issuing notes backed by this bond is:

$$r \approx \frac{(0.04)(1.10) - (0.01694)(0.9)}{1.10 - 0.9} \approx 14.375\%$$

[12] Champ (2011b) also cites these Friedman and Schwartz passages. Friedman and Schwartz's mention of "the wrong base" refers to mistaken calculations by the contemporary Comptroller of the Currency.

[13] The Comptroller used $62.50 for the costs associated with notes issued based on $100,000 of bonds deposited. These costs included the cost of redemption, $45; express charges, $3; engraving plates for the notes, $7.50; and agents' fees, $7. See Champ (2011b).

Cagan (1963) found very high profit rates for the 1870s, 20–30%. More importantly, Cagan and Goodhart (1965) found profit rates of infinity in the early 1900s. An infinite rate of profit occurs when $\alpha = 1$ after 1900 and the bond is selling below par. In that case, the notes the bank could issue based on using that bond as collateral would exactly equal the price paid for the bond, so no capital would be used and the bank could earn infinite profits. Figure 7 shows Cagan's profit series (as computed by Champ (2011b)). The figure shows the conundrum. The gaps in the figure are the instances where the rate of profit was alleged to be infinity; these are all cases where the prices of 2% coupon bonds fell below par. Note that the figure shows calculations using the average rate of profit, not the marginal rate of profit. Yet national banks did not take advantage of apparent profit opportunities.

Figure 8 shows our calculation of the profit series. We used Champ's more accurate representation of the costs of note issuance than the Comptroller of the Currency.[14] We also filled in all the bond prices that were missing from a Bruce Champ spreadsheet (provided to us by the Federal Reserve Bank of Cleveland). We also made further adjustments discussed below.

Why didn't banks take advantage of this arbitrage opportunity? After all, banks held US Treasury bonds on their balance sheets (ie not including Treasuries held to back their notes). See Figure 6. Or, banks could have raised capital and used this to buy bonds for collateral for notes. To explain this puzzle, the literature has focused on hidden transaction costs or the risks of unpredictable redemptions; see Bell (1912), Cagan (1965), Goodhart (1965), Cagan and Schwartz (1991), Duggar and Rost (1969), Champ, Wallace and Weber (1992), and Wallace and Zhu (2004). These explanations are reviewed by Champ (2011b) and Calomiris and Mason (2008).[15] None of these explanations seem particularly persuasive to us, but in any case they are not mutually exclusive with the explanation we propose here.

Prior explanations do not mention that there may be a convenience yield associated with US Treasuries. This is, however, suggested by the work of Duffee (1996) and Krishnamurthy and Vissing-Jorgensen (2012) who look at data over the period 1926–2008, and Krishnamurthy and Vissing-Jorgensen (2013) who analyse the period 1914–2011. Figure 6, showing that national banks held US Treasuries on their balance sheets (ie, the ratio of US government bonds not on deposit at the Treasury to bank loans and discounts) suggests, by revealed preference, that there was a convenience yield associated with Treasuries during the National Banking Era.

At the time, bankers also recognised the convenience yield on Treasuries. For example, *The Financier*, 7 April 1902, Volume LXXIX (The Financier Company): "... banks have always regarded high-class bonds as an offset, so to speak, for risks incurred in discounts yielding a higher rate of interest. In this connection we cannot do better than to quote from a very valuable paper read by A.M. Peabody, of St. Paul, before the St. Paul Bank Clerks' Association, in which this feature is brought prominently forward. After explaining the classification of such investments, Mr. Peabody says: 'They have ever proved themselves the safeguards for banks under pressure of financial panics in times of great stringency, and when it would be impossible to borrow money on any form of security, railroad bonds with government

[14] Based on a spreadsheet of Bruce Champ, provided by the Federal Reserve Bank of Cleveland.

[15] Calomiris and Mason (2008) show that there was ". . . substantial variation in the propensity to issue national bank notes traceable to county, state, and bank-specific characteristics related to the profitability of lending" (p 340).

bonds, are alone available as security for money'" (p 1258). And, *The Bond Buyers' Dictionary* (1907): "... it is possible to say that there is a better market in moments of extreme panic for the Government issues than there is for even the best class railroad bonds. There will not be by any means [be] the same volume of liquidation. For every dollar of Government bonds thrown into a panic market there will be $100 of railroad bonds. ... Government bonds are undoubtedly the safest of all securities ... "(p 73).

But, even if banks wanted to keep these Treasuries on their balance sheets, why didn't they raise bank capital to buy Treasuries to back note issuance. That they did not suggests that bank capital was costly or that banks could not find the bonds. We will argue that bankers did not take advantage of opportunity to issue more national bank notes because it was *not* profitable to do so. We will show that the implicit profit from not issuing notes is driven by measures of convenience yield.

We first return to the calculation of the profit rate from note issuance. As mentioned above, we filled in the missing bonds in Champ's original spreadsheet used for calculating the profit rate to note issuance.[16] We next eliminated bonds that would have been called in the next six months, since then the notes backed by these bonds would have to have been returned, or new bonds would have to have been purchased.[17]

However, there is another issue, namely that it is the *marginal* profit rates that are relevant not the average rate of profit. This is important because in the early 1900s, and possibly before that, US government bonds were hard to find. And, even when banks could find bonds, they had to reverse repo in the bonds at a high cost. Contemporary observers continually wrote about this shortage of safe debt. For example, Morris (1912): "Various reasons have been assigned for the decline in circulation which culminated in 1891, the most probable being the growing scarcity of US bonds and their relatively high premium. It is also alleged that improved banking facilities, allowing a more extensive use of checks, reduced the demand for currency" (p 492). Morris dates the start of the problem as 1891. It is also interesting that Morris points out that the cost of note issuance caused a further development of demand deposits, the shadow banking system of its time.

Borrowing bonds was costly. Francis B Sears, vice president of the National Shawmut Bank of Boston, Mass. (1907–08): "There are two classes of banks – those outside of the large cities, that can get bonds only by buying them, and a few banks in a few large cities that can borrow them. I would like to add that insurance companies and savings banks are large bondholders, and undoubtedly arrangements can be made with them to get bonds for some large banks. The rate is 1½ to 2 percent for borrowing bonds in that way" (p 91). *Bankers Magazine* (March 1908): "Bond borrowings by the national banks have become an important feature of banking in recent years. Where a bank wishes to increase its circulation, or to procure public deposits, and does not happen to have the bonds which must be pledged with the Treasury, and finding the market price of bonds too high to make the transaction profitable if the bonds must be bought, resort is had to borrowing. Bond dealers,

[16] The missing bond prices/amounts were mostly during 1875–1879 plus one bond maturing in 1896. This did not affect the potentially infinite profits, but just added more observations in the earlier period.

[17] Eliminating these bonds removed some spikes in the profit series, one of which was during the period of high profits (1907). But otherwise it has no significant effect on the post-1902 series.

savings banks or private holders may have 'Governments' which they are willing to lend to national banks for a consideration" (p 321).[18] The *Rand-McNally Bankers' Monthly* (September 1902) quoting "a banker": "There is not much profit in issuing circulation on government bonds, but some of the larger banks are willing to take out notes, if they can borrow bonds for that purpose from their friends – not being disposed to buy them for temporary use. ... The real trouble is to find the bonds. Many of them are held by institutions and estates, who cannot legally loan the bonds to National Banks, and as their prices are too high to justify any large purchases of bonds by banks for the purpose of taking out circulation ..." (p 157–58). Gannon (1908), speaking of Treasury bonds: "... such bonds are not easy to buy in quantity, and the greater part of the recent expansion, some $80,000,000 since the panic [of 1907], was accomplished by borrowing bonds" (p 338)

The situation was summarised by *The Financial Encyclopedia* (1911, p 119):

When the banks borrow, either to secure banknote circulation or Government deposits, they make private arrangements with the actual owners of the bonds, including insurance companies, for the use of these securities. The rates banks pay vary, but in general lenders of bonds secure a very substantial profit from this employment of them, in addition to the interest which the bonds themselves carry.

Borrowed bonds were first itemized separately in the national banks' returns under the Comptroller's call of November 25, 1902. At that time the total 'borrowed bonds' reported by national banks of the whole country were $39,254,256 of which New York banks were credited with $21,199,000. In the return of December 3, 1907, the banks of the United States reported bonds of $166,073,021, more than half, or $88,274,330, being held by the forty national banks of this [New York] city. These are by far the largest holdings ever reported by New York banks.

When a bank borrows Government, municipal, or other bonds, from an insurance company, for instance, which are pledged as security for public (Treasury) deposits, it either gives the lender a check for the face value, with a contract stipulating to buy back the bonds at a certain price, or the bank gives the lender other collateral as security for the loan.

In the case of life insurance companies, the collateral offered in exchange for the bonds has often represented bonds in which the lending corporations are allowed to invest, but which were not in the so-called 'savings bank list,' and for that reason were not eligible as security for public deposits. While one or two of the life companies have never consented to lend their bonds, many others, as well as various fire insurance companies, have done so, on the theory that it was a good business transaction, since it yielded them 1 or 1½% in addition to the regular interest return.

The scarcity of bonds meant that the *marginal cost* of conducting the "arbitrage" was higher than the average cost. While the Comptroller started publishing data on bank bond borrowings in 1902, it seems that this problem started earlier. At a meeting of the American Economic Association held in Cleveland, Ohio in December 1897, it

[18] Government deposits in national banks had to be backed by bonds also, but there was a slightly larger list of eligible bonds for this purpose.

was voted to appoint a committee of five economists to consider and report on currency reform in the United States.[19] They turned in a report in December 1898. One point they made was this: "Now it is commonplace that our bank circulation is not a very profitable one." See *The Bankers' Magazine*, February 1899, p 221.

Note issuance profit series that are the average rate of profit are misleading. To adjust the profit calculations to reflect the scarcity and associated high cost of reversing in bonds, we set α in the above calculation of the profit rate to 0.99 instead of 1. Now, there are no instances of infinite profits. We discuss below why this does not greatly affect regression results. Figure 8 shows the series of profit rates in this case.

There is also the issue of the cost of bank capital. This cost is hard to quantify, as it is today. Bank stock during this period was illiquid, trading on the curb market. And there is some evidence that it was held in blocks by insiders. See Gorton (2013). There is no data on bank stock issuance. Contemporaries described the return to bank stock as low, partly due to double liability.[20] For example, Frank Mortimer, cashier of the First National Bank, Berkeley. California, in an address delivered before the San Francisco Chapter of the American Institute of Banking, *American Institute of Banking Bulletin* "When one takes into consideration the risk involved, the capital invested, and the double liability attached to stockholders in national banks, the profit from an investment in bank stock is small, indeed, when compared to the profit accruing from other lines of business." (p 236; reprinted in the *Journal of the American Bankers Association*, vol 6, July 1913–June 1914.).

3.3 The convenience yield on treasuries and the cost of bank capital

In this section, we turn to an analysis of the rate of profit on note issuance. We show that the rate of profit on note issuance is highly related to the convenience yield on Treasuries. We measure convenience yield in two complementary ways. First, we use the supply of Treasuries divided by GDP. Krishnamurthy and Vissing-Jorgensen (2012) show that this measure strongly drives the convenience yield on Treasuries from 1926 to the present. When the supply of those assets is low, that is, safe assets are relatively scarce, then the convenience yield for safe assets increases. Therefore, Treasury supply should be negatively related to the convenience yield. We take two measures of Treasury supply: (US government debt)/GDP (as in Krishnamurthy and Vissing-Jorgensen (2012)), and (available Treasuries)/GDP, where available Treasuries excludes those already held to back bank note issuance and thus captures the remaining supply. Second, we also measure the convenience yield as the spread between high grade municipal bonds from New England and Treasuries. Municipal bond yields are from *Banking and Monetary Statistics* (1976).

Table 4 gives the results of a regression of issuance profits on these measures of the convenience yield from 1880–1913. The results match our intuition. The profit measure is high exactly when the convenience yield to Treasuries is large. We find that a 1% increase in the muni spread is associated with a 15% increase in average profit. As demand for Treasuries increases, the apparent profits also increase. As the supply of available Treasuries decreases, profits also increase. Both the supply

[19] The economists were a very distinguished group: F M Taylor, University of Michigan; F W Taussig, Harvard; J W Jenks, Cornell; Sidney Sherwood, Johns Hopkins; and David Kinley, University of Illinois.

[20] On double liability, see Macey and Miller (1992) and Grossman (2001).

variables and the muni spread are highly significant independently. However, we would suspect that they likely measure similar economic forces, although each is measured with noise. Consistent with this, when we include both the supply of Treasuries and muni spread together, the coefficients on each decrease in absolute value, although they remain statistically significant. This suggests that both are imperfect but overlapping measures of the convenience yield.

We show the results when we use the average profit series as well as the log average profit series. Recall that profit is given by $\frac{r_b p - \tau\alpha \min(p,1)}{p - \alpha \min(p,1)}$. A possible concern is that the profit series is highly non-linear due to the denominator becoming small later in the sample. To mitigate this concern, we also report the results using log profits, which largely alleviates the strong non-linearity in the denominator (see Figure 8 which plots profits on a log scale). Our results do not change drastically with the log transformation, highlighting that non-linearities in the latter half of the sample are not driving the result. In unreported results we also obtain the same basic findings for alternative values of α. Finally, in the Appendix, we show the results when including dummies for pre and post 1900, as the issue with the choice of α is only relevant after 1900. The results, given in Appendix Table 6, show broadly the same pattern in both periods with similar signs, though magnitudes are larger after 1900.

All variables in this regression are persistent which can potentially confound inference. We deal with this in several ways. First, in our main specifications we estimate standard errors using Newey-West with 10-year lags (specifically, we use 10 lags for annual data and 40 lags for quarterly data). Second, we run GLS assuming the error term follows an AR(1). This suggests transforming both our x and y variables by $1-\rho L$ where ρ is the error auto-correlation and L is a lag operator. We find ρ by running OLS as in specification (4) in the Table and computing the sample auto-correlation of the residuals. This does not substantially change the point estimates or inference in terms of what is statically significant. As mentioned by Krishnamurthy and Vissing-Jorgensen (2012), however, coefficients do decrease somewhat in absolute value. A likely reason is that these are noisy measures of convenience yield and measurement error will become more pronounced in the transformed data. This follows from the fact that when x is persistent the variance of x will be dominated by low frequency components. In contrast, in the transformed data measurement error likely accounts for more of the variance of the right hand side variable, resulting in a larger degree of attenuation bias.

Finally, while the results appear fairly strong, we also acknowledge that we are working with a fairly small subsample of data which is a limitation of our analysis. Higher frequency data (eg, monthly data on debt/GDP) won't be particularly helpful here in overcoming the fairly small sample because the variables are highly persistent.

We hypothesise that a third variable – the cost of capital for banks – likely plays a role in explaining the profits on note issuance as well. If the cost of raising capital for banks is high, then banks would find it costly to take advance of note issuance and may leave a puzzlingly large profit on the table. For this conjecture, we can only offer suggestive evidence from Figure 8, which plots the profit series along with NBER recession bars. It is likely that the cost of raising capital for banks increases during recessions, and especially at the onset of recessions, and these are times when we do in fact see increases in the profit series. Thus, there is some suggestive evidence of the cost of capital for banks being positively associated with the profits on note issuance as well.

Taken together, our results indicate that the profits to note issuance fluctuate with the convenience yield on Treasuries, and our evidence is consistent with the idea that profits are related to the cost of bank capital.

3.4 Proposals for reform during the National Banking Era

Despite the creation of a uniform currency, the monetary system of the National Banking Era was increasingly unpopular. Reform of the currency system was increasingly discussed because of the problems of inelasticity and banking panics. "Inelasticity" meant that the quantity of money was not sensitive to the business or seasonal cycles. There were spikes in seasonal interest rates and the money supply could not be increased to alleviate bank runs. Following the Panic of 1893, calls for reform of the banking system became louder. All the reforms sought to sever the link between bank notes and Treasuries. This link was viewed as the problem. Replacement proposals were for an "asset-based" system, meaning that currency would be allowed to be backed by bank loans or commercial paper (depending on the party proposing this system).

The first proposal for an asset-backed currency came from the American Bankers' Association in 1894. It was called the "Baltimore Plan" (the Association's meeting was in Baltimore that year). The plan envisioned banks issuing circulating notes under federal supervision, where the notes would be secured by (1) a first lien upon the assets of the issuing bank; (2) the double liability of shareholders; (3) the 5% redemption fund; and (4) a 5% guaranty fund (to be used to repay note holders of a failed bank). Treasury Secretary Carlisle (1894) modified the Baltimore Plan to create more security for the noteholders. He proposed (1) a deposit of 30% of the circulation in legal money with the government; (2), the safety fund of 5%; (3) a requirement upon all other banks to cover any losses beyond the two funds of any failed bank; (4) a lien upon all assets; and (5) a further lien on the shareholders' liability. In addition, both plans wanted to limit note issuance to 50% or 75% of paid-up capital.

There were other subsequent proposals as well, including plans from Eckels (1894a,b) and the Indianapolis Monetary Commission (1900), see Taylor (1898). All of these plans were similar in spirit to the Baltimore Plan. See Wicker (2005), West (1974) and Laughlin (1920) for more detailed discussions of the various plans.

The asset-based monetary proposals sought to separate the two forms of money: Treasuries and bank notes. And they did recognise that bank runs on demand deposits were related to perceptions of the "safeness" of the loans backing the demand deposits. Backing private bank notes with loan portfolios would have created the risk of runs on notes as well as demand deposits, although seniority of notes, guarantee funds, reserves etc were aimed at this problem. But, the proposals embedded the view that an "elastic" currency would alleviate panics. Laughlin (1920): "In all these plans we were really aiming to prevent the difficulties experienced in [the Panic of] 1893 . . ." (p 30–31).

It is interesting that the proposed reforms during the period all aimed to break the link between bank notes and Treasuries.

4. Discussion

Of course, the National Banking System is not exactly like the LCR. There are obvious differences. But, like the National Banking Era, the logic of the LCR is that, if short-term debt is backed by Treasuries, then bank runs will be avoided. Fundamentally, the two systems enforce a correspondence between two types of debt instruments, each with a convenience yield. The input for making one kind of money, bank notes or money market instruments, is required to be Treasuries. Such a system is fragile because, by forcing two kinds of money together, it is likely that there will be a shortage of one kind of money, leading to its private production elsewhere, which creates fragility in the system.

The National Banking System did succeed in introducing a uniform currency where the national bank notes of different banks all traded at par, unlike the pre-Civil War period. This was because banks' national bank notes were backed by US Treasury debt. We have focused on the fact that Treasuries have a convenience yield. In the recent period, a measure of that convenience yield is the GC repo to Treasury spread. This spread is a driver of repo fails when agents want to keep the Treasuries (or the cash) in repo transactions.

This was the same core issue during the National Banking Era. In the National Banking Era, a measure of the convenience yield (and the cost of bank capital) is the implied profitability of issuing bank notes. There is no under-issuance puzzle once this is recognised. If Treasuries have a convenience yield, then they provide safety and liquidity to the agents who demand this. But, these agents or other agents also have a demand for bank notes.

If there were enough Treasuries (high-quality liquid assets) to meet the global demand for safe debt and to back short-term bank debt, then the LCR and related immobile collateral requirements would not be a problem. One potential argument is that in the National Banking Era the supply of Treasuries was low (debt-to-GDP was in the range of 10–30%) so that scarcity was more of an issue in that period then it is today where the supply of government debt is much larger. However, this ignores that the demand for safe US government debt now is also global, which can add to issues of scarcity. The likelihood of such a satiation of the global economy with Treasuries today seems remote. Gorton, Lewellen and Metrick (2012) show that the sum of US government debt outstanding and privately produced safe debt outstanding has been 32% of total assets in the US since 1952. See Figure 10. The figure shows that the bulk of safe debt has never been Treasuries, but has mostly been privately produced debt. (Figure 1 shows the composition of this privately produced safe debt.) There has never remotely been enough US Treasuries to make up the 32% and, given the debt burden of issuing enough to accomplish that, there is never likely to be enough. Furthermore, Treasuries outstanding is a function of fiscal policy not a function of the demands for collateral.

When Treasuries have a convenience yield, and short-term bank debt must be backed by Treasuries, there is a trade-off between the two types of money. More short-term debt means fewer Treasuries for alternative uses. This trade-off is common to the two systems. We saw this in the National Banking Era. The trade-off is evident in the data. Noyes (1910): "A heavy decrease in the outstanding public debt would naturally, at some point, cause a reduction in the bank-note circulation, independently of other influences. A large increase in the government debt would necessarily cause an increase in the supply of bank notes" (p 4). Noyes then traces

this out over the National Banking Era. It is the same statement that was formalised by Krishnamurthy and Vissing-Jorgensen (2012a,b) for the modern era. The convenience yield is negatively related to Treasuries (divided by GDP) outstanding.

One way out of this trade-off, if it is binding, is to privately produce another kind of debt. In the recent period, this was ABS and MBS, which could be used in place of Treasuries to back repo, ABCP, and MMF. In the National Banking Era, it was demand deposits using portfolios of loans as the backing. During the National Banking Era, Treasuries outstanding to GDP fell secularly (see Figure 6, panel C). And, from the start of the National Banking System, the ratio of bank notes to demand deposits fell, from just over 60% in the early part of 1865 to 14% by 1909, as shown in Figure 9. Demand deposits were privately produced safe debt or money, the shadow banking system of its time. So, while the immobile collateral system ended bank runs on bank notes, there were bank runs on demand deposits, another form of bank money. The biggest problem of the National Banking Era was that there were banking panics.

It is difficult to prove causally that demand deposits grew relative to bank notes because of collateral requirements. However, the growth in deposits does line up remarkably well with the decline in Treasury supply and hence the supply of safe assets that could be used to back notes. This accords with the evidence in Krishnamurthy and Vissing-Jorgensen (2012b) that the supply of Treasury crowds out privately produced bank debt. The ratio of national bank notes to demand deposits fell from over 60% to less than 20% over the period and this ratio co-moves strikingly with the debt to GDP ratio as shown in Figure 9. The correlation of the ratio of notes to deposits with the supply of Treasuries to GDP is 0.96. As the supply of Treasuries falls over this period, deposits grow. Demand deposits were conceptually misunderstood during the National Banking Era, although it was clear that demand deposits were the issue in banking panics. This highlights that the system of immobile collateral in the National Banking Era was not successful in mitigating panics and that it likely contributed to the growth of other forms of bank debt.

The Lucas critique seems to be largely ignored by policymakers for the simple reason that the requisite general equilibrium models do not exist.[21] This is a mistake. Economic history provides a laboratory to study large, important, policy changes, like the LCR. There are many other examples where history can be used in this way. Recent examples include Foley-Fisher and McLaughlin (2014) who study structural differences between bonds guaranteed by the UK and Irish governments during the period 1920–38. The events provide a way to think about sovereign debt that is jointly guaranteed by multiple governments, eg, proposed euro bonds. Carlson and Rose (2014) study the run on Continental Illinois in 1984, during which the government provided an extraordinary guarantee of all the bank's liabilities. The authors argue that this example provides insights into the Orderly Liquidation Authority of the Dodd-Frank Act.[22]

[21] A recent example of the unintended consequences of a regulation is provided by Hachem and Song (2015), who show that the restriction of loans to 75% of deposits on Chinese banks led to the development of a shadow banking system.

[22] Bodenhorn (2015) studies the experience of Rhode Island over 1849–1907 to evaluate early release policies as a response to overcrowded prisons.

References

Bai, J (2010): "Common breaks in means and variances for panel data", *Journal of Econometrics* 157, pp 78–92.

Baltimore Plan (1894): American Banking Association, *Proceedings*.

Basel Committee on Banking Supervision (2011): "Basel III Framework for Liquidity – Frequently Asked Questions".

Basel Committee on Banking Supervision (2013): "Basel III: The Liquidity Coverage Ratio and Liquidity Risk Monitoring Tools".

Bell, S (1912): "Profit on National Bank Notes", *American Economic Review*, no 2, pp 38–60.

Bernanke, B (2007): "Global imbalances: recent developments and prospects", speech delivered at the Bundesbank Lecture, Berlin, Germany, 11 September.

——— (2005): "The global savings glut and the U.S. current account deficit", speech delivered at the Sandridge Lecture, Virginia Association of Economists, Richmond, Va., 10 March.

Bernanke, B, C Bertaut, L Pounder DeMarco and S Kamin (2011): "International capital flows and the returns to safe assets in the United States, 2003–2007", Working Paper, Board of Governors of the Federal Reserve System.

Bertaut, Carol, L DeMarco, S Kamin and R Tryon (2012): "ABS inflows to the United States and the global financial crisis", *Journal of International Economics*, no 88, pp 219–34.

Bertaut, C, A Tabova and V Wong (2014): "The replacement of safe assets: evidence from the US bond portfolio", Board of Governors of the Federal Reserve System, working paper.

Board of Governors of the Federal Reserve System (1976): *Banking and Monetary Statistics*.

Bodenhorn, H (2015): "Prison crowding, recidivism, and early release in early Rhode Island", *NBER Working Paper*, no 20837.

Bolles, A (1902): "Responsibility of the National Bank in the present crisis", *Annals of the American Academy of Political and Social Science*, vol 20, pp 1–18.

Bond Buyers' Dictionary, The (1907): edited by S Nelson, S Nelson & Co, New York.

Caballero, R (2010): "The 'other' imbalance and the financial crisis", *NBER Working Paper*, no 15636.

Caballero, R and A Krishnamurthy (2009): "Global imbalances and financial fragility", *American Economic Review: papers & Proceedings*, no 99, pp 584–88.

Cagan, P (1963): "The first fifty years of the national banking system", in *Banking and Monetary Studies*, D Carson (ed), pp 15–42, Irwin.

——— (1965): *Determinants and Effects of Changes in the Stock of Money, 1875–1960*, NBER.

Cagan, P and A Schwartz (1991): "The National Bank note puzzle reinterpreted", *Journal of Money, Credit and Banking*, no 23, pp 293–307.

Calomiris, C and J Mason (2008): "Resolving the puzzle of the underissuance of National Bank notes", *Explorations in Economic History*, no 45, pp 327–55.

Carlisle, J (1894): *Report of the Committee on Banking and Currency together with hearings*, House of Representatives, 8149, 53rd Congress, third session.

Carlson, M and J Rose (2014): "Can a bank run be stopped? Government guarantees and the run on Continental Illinois", Federal Reserve Board, working paper.

Champ, B (2011a): "The National Banking System: The empirical evidence", Federal Reserve Bank of Cleveland, *Working Paper*, no 07-19R.

——— (2011b): "The National Banking System: The National Bank note puzzle", Federal Reserve Bank of Cleveland, *Working Paper*, no 07-22R.

——— (2011c): "The National Banking System: A brief history", Federal Reserve Bank of Cleveland, *Working Paper*, no 07-23R.

Champ, B, N Wallace, and W Weber (1992): "Resolving the National Bank note paradox", Federal Reserve Bank of Minneapolis, *Quarterly Review*, Spring, pp 13–21.

Duffee, G (1996): "Idiosyncratic variations of Treasury bills", *Journal of Finance* , no 51, pp 527–51.

Duffie, D, M Scheicher and G Vuillemey (2014): "Central clearing and collateral demand", *Journal of Financial Economics*, forthcoming.

Duggar, J and R Rost (1969): "National Bank note redemption and treasury cash", *Journal of Economic History*, no 29, pp 512–20.

Dunbar, C (1887): "Deposits as currency", *Quarterly Journal of Economics*, no 1, pp 401–19.

Eckels, J (1894a): *Report of the Committee on Banking and Currency Together with Hearings*, House of Representatives, no 8149, 53rd Congress, third session.

——— (1894b): *Annual Report of the Comptroller of the Currency*, 53rd Congress, third session, 3 December, Government Printing Office, Washington DC.

Fender, I and U Lewrick (2013): "Mind the gap? Sources and implications of supply-demand imbalances in collateral asset markets", *BIS Quarterly Review*, September 2013, pp 67–81.

Fisher, W (1895): "Money and credit paper in the modern market", *Journal of Political Economy*, no 3, pp 391–413.

Fleming, M and K Garbade (2002): "When the back office moved to the front burner: Settlement fails in the treasury market after 9/11", *FRBNY Economic Policy Review*, November, pp 35–57.

Foley-Fisher, N and J McLaughlin (2014): "Sovereign debt guarantees and default: Lessons from the UK and Ireland, 1920–1938", Federal Reserve Board, working paper.

Friedman, M and A Schwartz (1963): *A Monetary History of the United States*, Princeton University Press.

Garbade, K and M Fleming (2005): "Explaining settlement fails", Federal Reserve Bank of New York, *Current Issues in Economics and Finance*, vol 11, no 9.

Garbade, K, F Keane, L Logan, A Stokes and J Wolgemuth (2010): "The introduction of the TMPG fails charge for US Treasury Securities", *FRBNY Economic Policy Review*, October, pp 45–71.

Goodhart, C (1965): "Profit on National Bank notes, 1900–1913", *Journal of Political Economy*, no 73, pp 516–22.

Gorton, G (2013): "The development of opacity in US banking", *Yale Journal of Regulation*, forthcoming.

Gorton, G and A Metrick (2015): "Who ran on repo", Yale working paper.

Gorton, G, S Lewellen, and A Metrick (2012): "The safe-asset share", *American Economic Review: Papers & Proceedings*, no 102, pp 101–6.

Gourinchas, P and O Jeanne (2012): "Global safe assets", paper prepared for the 11th BIS Annual Conference held in Lucerne, 20–21 June 2012.

Grossman, R (2001): "Double liability and bank risk taking", *Journal of Money, Credit and Banking*, vol 33, no 2, pp 143–159.

Hachem, K and Z Song (2015): "The rise of China's shadow banking system", University of Chicago, working paper.

Hammond, B (1957): *Banks and Politics in America from the Revolution to the Civil War*, Princeton University Press.

Heller, D and N Vause (2012): "Collateral requirements for mandatory central clearing of over-the-counter derivatives", *BIS Working Papers*, no 373.

Indianapolis Monetary Commission (1920): *Report of the Monetary Commission of the Indianapolis Convention*, Hollenbeck Press.

James, J (1976): "The conundrum of the low issue of National Bank notes", *Journal of Political Economy*, no 84, pp 359–68.

Kinley, D (1910): *The use of credit instruments in payments in the United States*, National Monetary Commission; Washington DC.

Krishnamurthy, A and A Vissing-Jorgensen (2012a): "The aggregate demand for treasury debt", *Journal of Political Economy*, no 120, pp 233–67.

——— (2012): "Short-term debt and financial crisis: What can we learn from US Treasury supply", *Journal of Financial Economics*, forthcoming.

Kuehlwein, M (1992): "The National Bank note controversy re-examined", *Journal of Money Credit and Banking*, no 24, pp 111–26.

Laughlin, L (1920): *Banking Progress*, Charles Scribners.

Lucas, R (1976): "Econometric policy evaluation: a critique", in K Brunner and A Meltzer (eds), *The Philips Curve and Labor Markets*, Carnegie-Rochester Conference Series on Public Policy, no 1, Elsevier, pp 19–46.

Macaulay, F (1958): *The movements of interest rates, bond yields and stock prices in the United States since 1856*, National Bureau of Economic Research.

Macey, J and G Miller (1992): "Double liability of bank shareholders: History and implications", *Wake Forest Law Review*, no 27, pp 31–62.

Maggiori, M (2013): "Financial intermediation, international risk sharing, and reserve currencies", Harvard University, working paper.

Noyes, A (1910): *A History of the National Bank Currency*, National Monetary Commission, Government Printing Office, Washington DC.

Sidanius, C and F Zikes (2012): "OTC derivatives reform and collateral demand impact", Bank of England, *Financial Stability Paper*, no 18.

Singh, M (2011): "Velocity of pledged collateral", *IMF Working Paper*, no WP/11/256.

——— (2014): *Collateral and Financial Plumbing*, Risk Books.

Singh, M and J Aitken (2010): "The (sizeable) role of rehypothecation in the shadow banking system", *IMF Working Paper*, no WP/10/172.

Taylor, F (1898): "The final report of the Indianapolis Monetary Commission", *Journal of Political Economy*, no 6, pp 293–322.

Treasury Market Practices Group (2011): "Understanding settlement fails in agency mortgage-backed securities."

Vickery, J and J Wright (2013): "TBA trading and liquidity in the agency MBS market", Federal Reserve Bank of New York, *Economic Policy Review*, May, pp 1–18.

Wallace, N and T Zhu (2007): "Float on a note", *Journal of Monetary Economics*, vol 54, pp 229–46.

West, R (1974): *Banking Reform and the Federal Reserve, 1863–1923*, Cornell University Press.

Wicker, E (2005): *The Great Debate in Banking Reform: Nelson Aldrich and the Origins of the Fed*, Ohio State University Press.

——— (1966): *Federal Reserve Monetary Policy 1917–1933*, Random House.

Tables and figures

Table 1: We present summary statistics on fails and rehypothecation.

Panel A: Fails, $ Millions

	Fail to Receive		Fail to Deliver	
	1990-99	2000-07	1990-99	2000-07
Mean	31,676	154,600	11,812	122,363
Std Dev	3,771	6,372	4,105	163,564

Panel B: Receive fails minus deliver fails, $ Millions

	Treasuries		Agencies		MBS	
	1990-99	2000-07	1990-99	2000-07	1990-99	2000-07
Mean	5,169	16,624	-123	338	207	-8,073
Std Dev	3,947	18,268	739	3,696	2,615	22,283

Panel C: Amount of collateral received eligible vs actually rehypothecated (12/31/2013)

	Treasuries	Other
Total Received ($ Millions)	179,366	123,915
Eligible for Rehypothecation	85%	55%
Actually Rehypothecated	55%	30%

Sources: Panels A & B, Federal Reserve Bank of New York. Panel C, ISDA.

Table 2: We report breakpoints for our fails data panel for both failures to receive and deliver along with 95% confidence intervals. The methodology for finding breaks in panels follows Bai (2010).

Panel A: Treasury and Agency Bonds

	Break Date	Lower Bound	Upper Bound
First Break	12-Sep-01	2-May-01	16-Jan-02
Second Break	24-Sep-08	11-Jun-08	31-Dec-08
Third Break	11-Feb-09	14-Jan-09	4-Mar-09

Panel B: Agency MBS

	Break Date	Lower Bound	Upper Bound
First Break	16-Oct-02	18-Sep-02	6-Nov-02
Second Break	16-Sep-09	26-Aug-09	30-Sep-09
Third Break	28-Dec-11	23-Nov-11	25-Jan-12

Table 3: We run seemingly unrelated regressions of Treasury and Agency fails to receive.

	Δ Fails Rec	Δ Fails Rec	Δ Fails Rec	Δ Fails Rec	Δ Fails Rec	Δ Fails Rec
GC Repo-1m T-bill	6.963***	0.695	7.303***	0.640	7.509***	0.620
	(5.57)	(0.41)	(5.78)	(0.38)	(5.91)	(0.36)
L1.GC Repo-1m T-bill			2.609*	0.818	2.951*	0.648
			(2.07)	(0.48)	(2.31)	(0.38)
L2.GC Repo-1m T-bill					2.495*	0.316
					(1.96)	(0.19)
GC Repo-1m T-bill x Break 1		13.35***		13.96***		13.35***
		(5.14)		(5.26)		(5.03)
L1.GC Repo-1m T-bill x Break 1				2.492		1.894
				(0.95)		(0.71)
L2.GC Repo-1m T-bill x Break 1						-2.164
						(-0.82)
GC Repo-1m T-bill x Break 2		39.57***		45.66***		44.08***
		(7.36)		(8.46)		(8.20)
L1.GC Repo-1m T-bill x Break 2				33.27***		37.98***
				(6.55)		(7.46)
L2.GC Repo-1m T-bill x Break 2						32.26***
						(6.43)
GC Repo-1m T-bill x Break 3		-1.485		-1.878		-1.185
		(-0.13)		(-0.16)		(-0.10)
L1.GC Repo-1m T-bill x Break 3				4.103		4.818
				(0.36)		(0.41)
L2.GC Repo-1m T-bill x Break 3						3.120
						(0.27)
Break 1 (9/2001-9/2008)		-8.852		-8.269		-6.430
		(-0.23)		(-0.21)		(-0.17)
Break 2 (9/2008-2/2009)		75.43		153.1		193.5
		(0.57)		(1.16)		(1.48)
Break 3 (2/2009)		11.00		9.994		7.266
		(0.26)		(0.24)		(0.17)
D. 1m T-Bill		-13.08***		-11.39***		-9.288***
		(-6.41)		(-5.53)		(-4.48)
Constant	0.879	-11.81	1.104	-10.05	1.092	-7.817
	(0.05)	(-0.48)	(0.07)	(-0.41)	(0.06)	(-0.32)
Observations	2398	2398	2386	2386	2374	2374
R^2	0.013	0.055	0.015	0.076	0.016	0.095

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: We run regressions of profits on two measures of the convenience yield: Treasury supply and the Municipal bond – Treasury spread. We measure Treasury supply as either debt / GDP or Treasuries available / GDP. Treasury supply variables are annual which reduces the number of observations. The muni spread is quarterly. T-stats are Newey-West with 10 lags for annual regressions and 40 lags for quarterly regressions. The column "GLS" assumes errors follow an AR(1) and hence transforms the x and y variables by $1-\rho L$ where ρ is the autocorrelation of the error term in (4) using OLS and L is a lag operator.

Profit on Convenience Yield, 1880-1913

$$y_t = a + b \times x_t + \varepsilon_t$$

Panel A: y =profit

	(1)	(2)	(3)	(4)	GLS (5)
ln(Debt/GDP)	-43.4				
	[-2.05]				
ln(Avail/GDP)		-31.8		-24.4	-19.3
		[-3.51]		[-4.03]	[-4.33]
Muni spread			54.5	14.9	18.3
			[3.06]	[3.59]	[3.20]
AdjR^2	0.28	0.65	0.50	0.76	0.61
N	34	34	137	34	34

Panel B: $y = \ln(\text{profit})$

	(1)	(2)	(3)	(4)	GLS (5)
ln(Debt/GDP)	-1.78				
	[-2.35]				
ln(Avail/GDP)		-1.18		-1.03	-0.81
		[-4.85]		[-6.81]	[-5.49]
Muni spread			1.81	0.44	0.38
			[3.83]	[4.49]	[2.05]
AdjR^2	0.36	0.67	0.48	0.74	0.54
N	34	34	137	34	34

Figure 1: Composition of Privately Produced Safe Debt (% of Total Privately Produced Safe Debt)

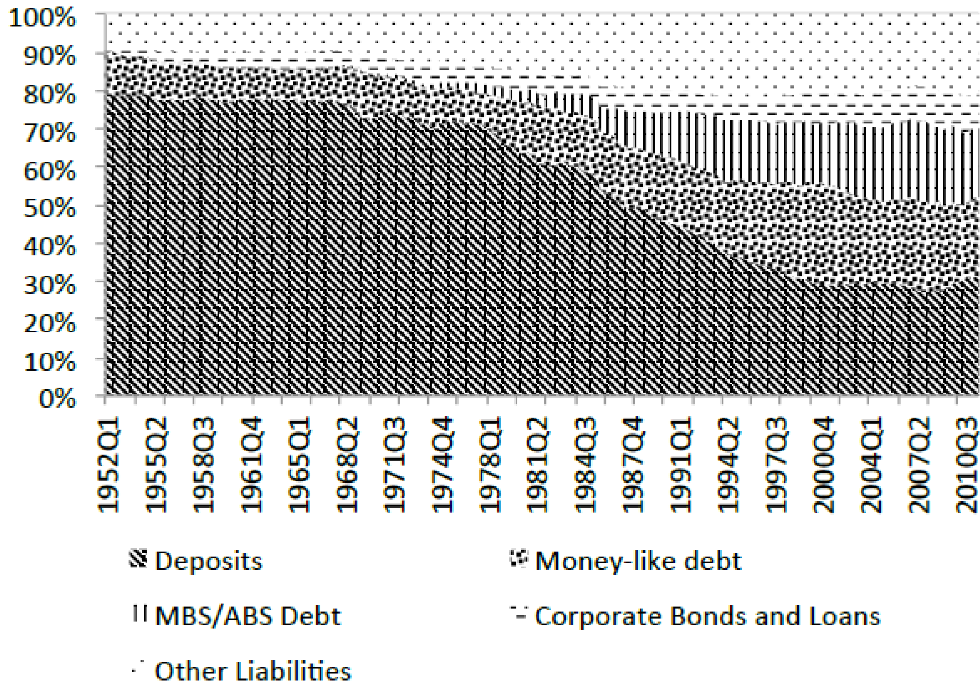


Figure 2: Holders of US Treasury Securities (percent of total outstanding).

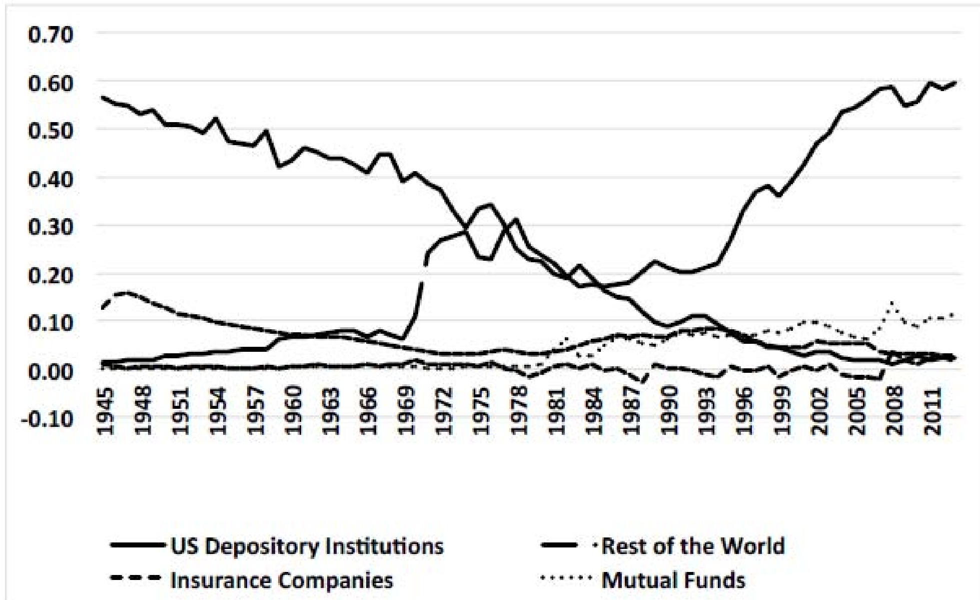


Figure 3: Fails by type.

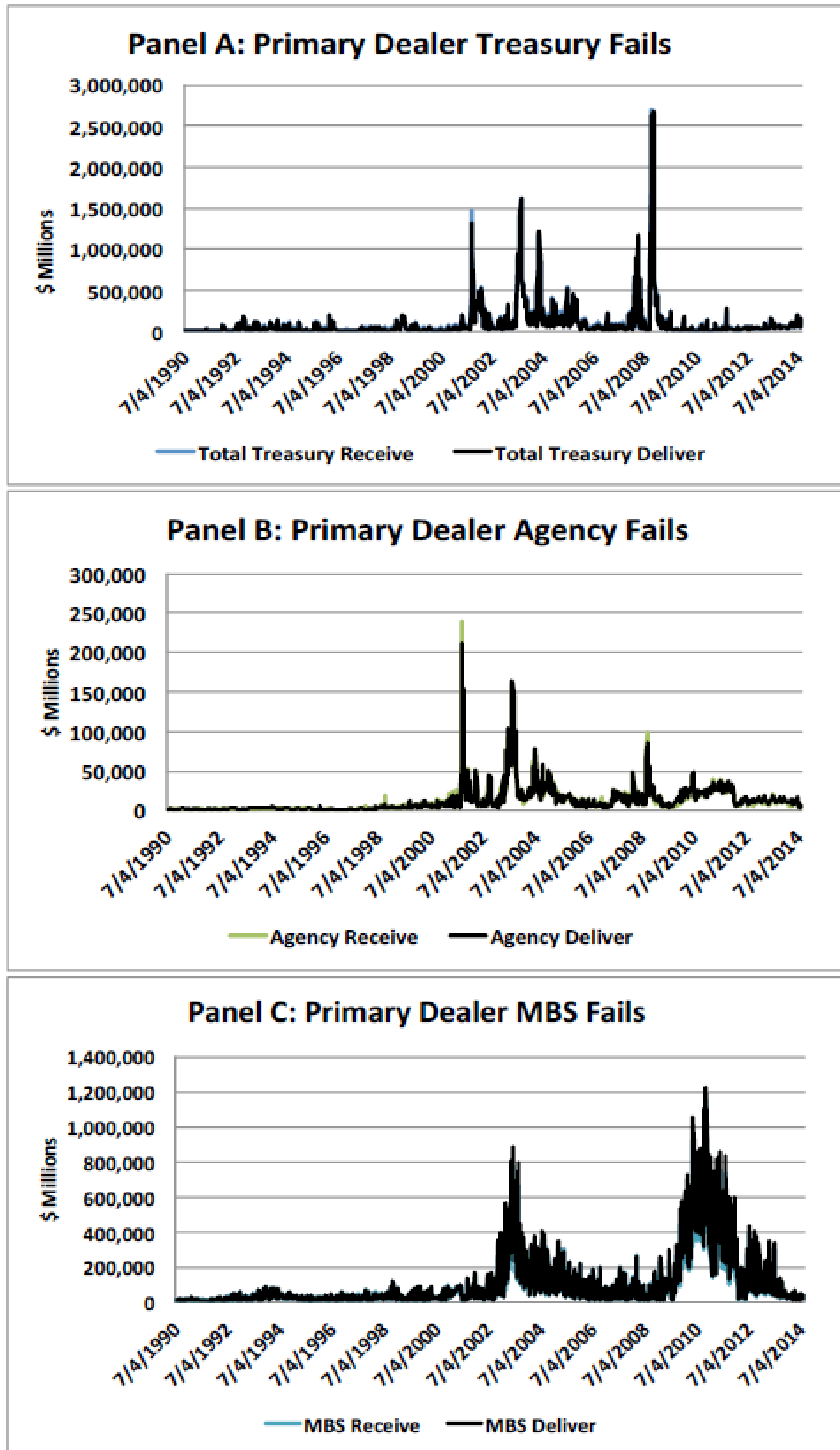


Figure 4: DTCC Fails ($100 million).

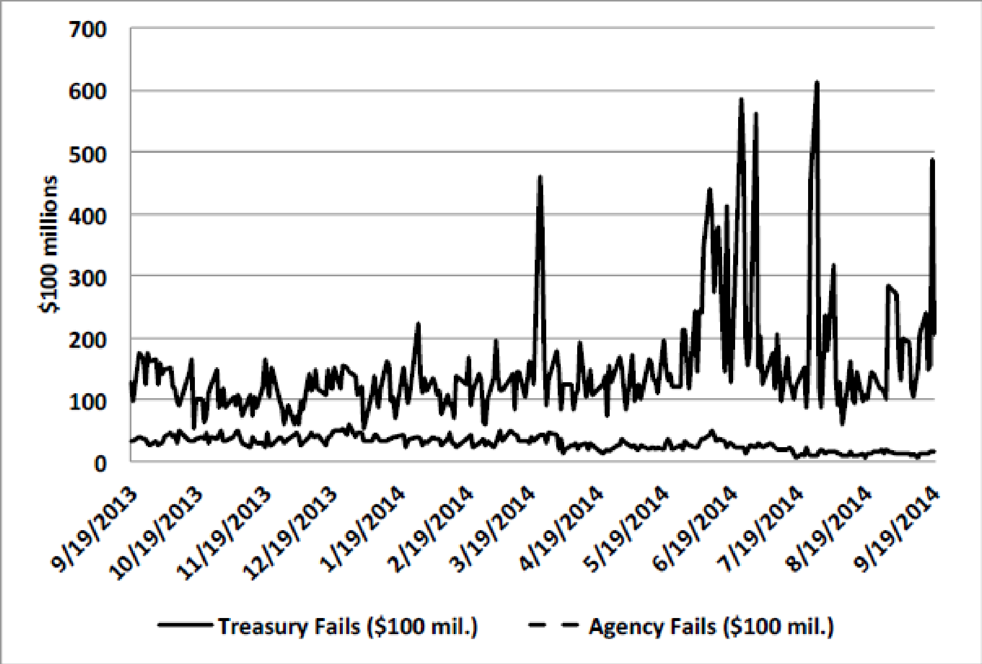


Figure 5: Difference between fail to receive and failure to deliver by type.

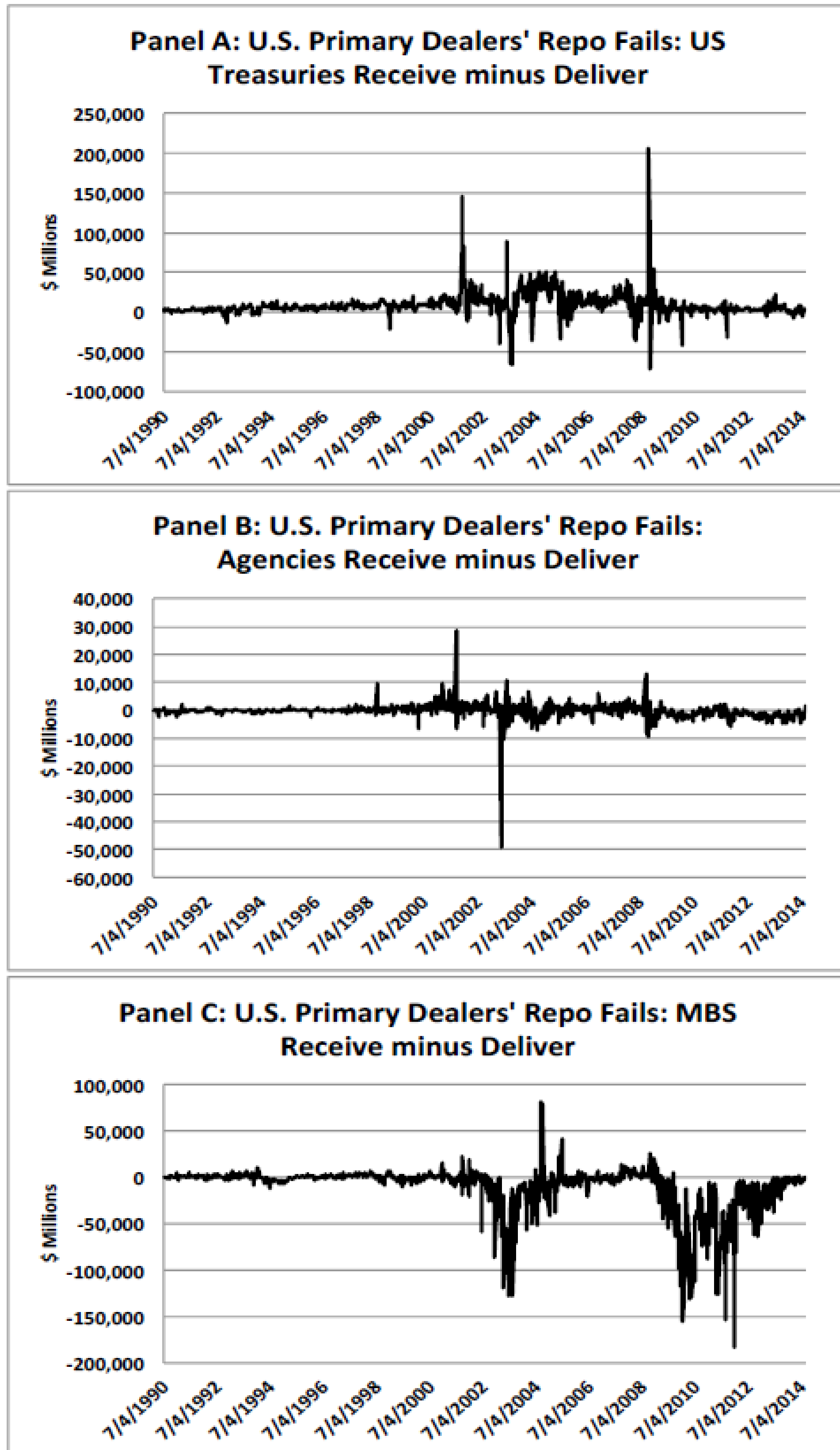


Figure 6: Panel A plots the fraction of Treasuries held to back notes and hence is informative about how aggressively banks were taking advantage of note issuance. Panel B plots the fraction of bonds on hand to loans and discounts. Panel C plots total Debt/GDP outstanding for the US and gives a sense of the total supply of government debt.

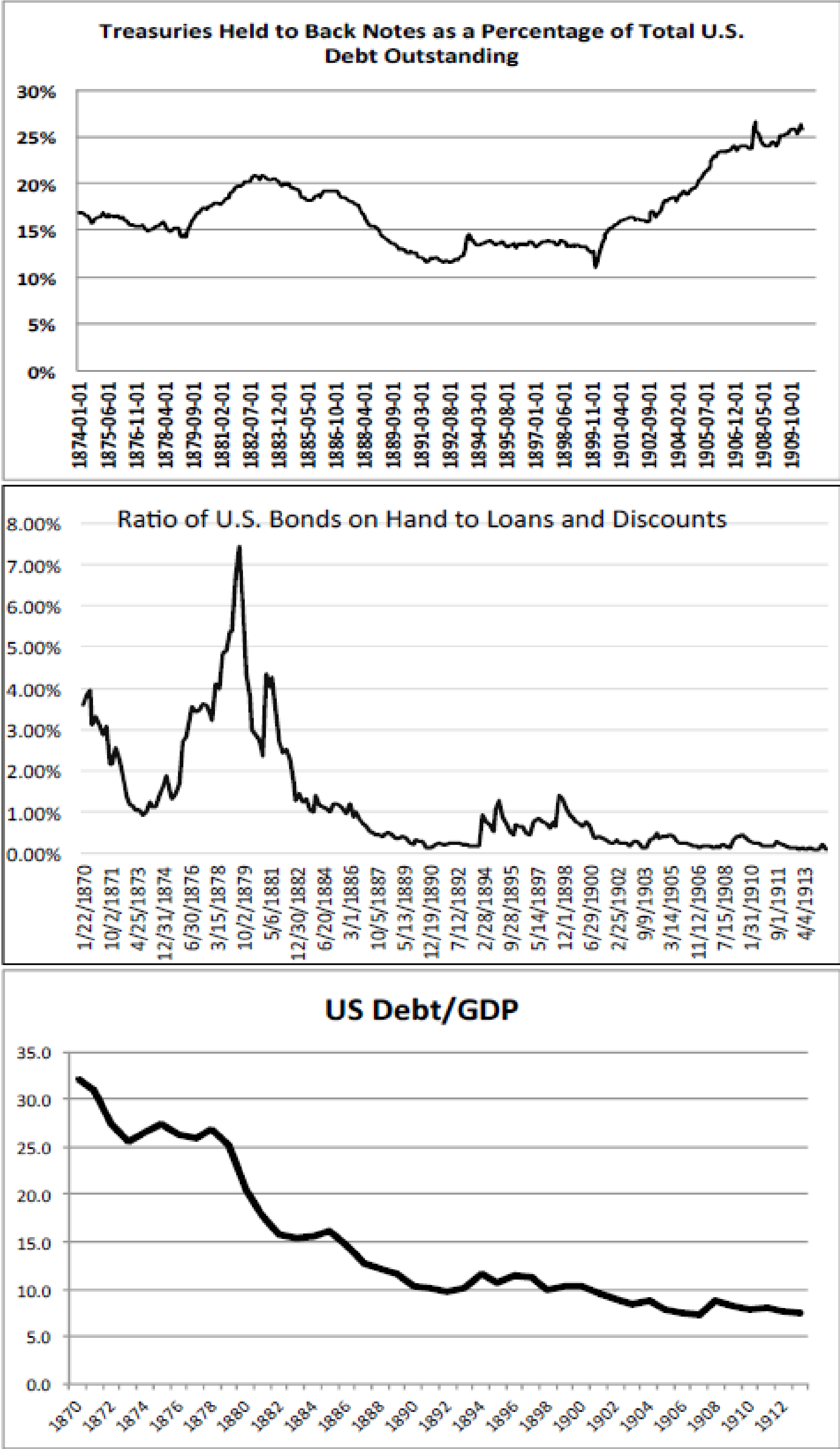


Figure 7: Original profit series as computed by Champ.

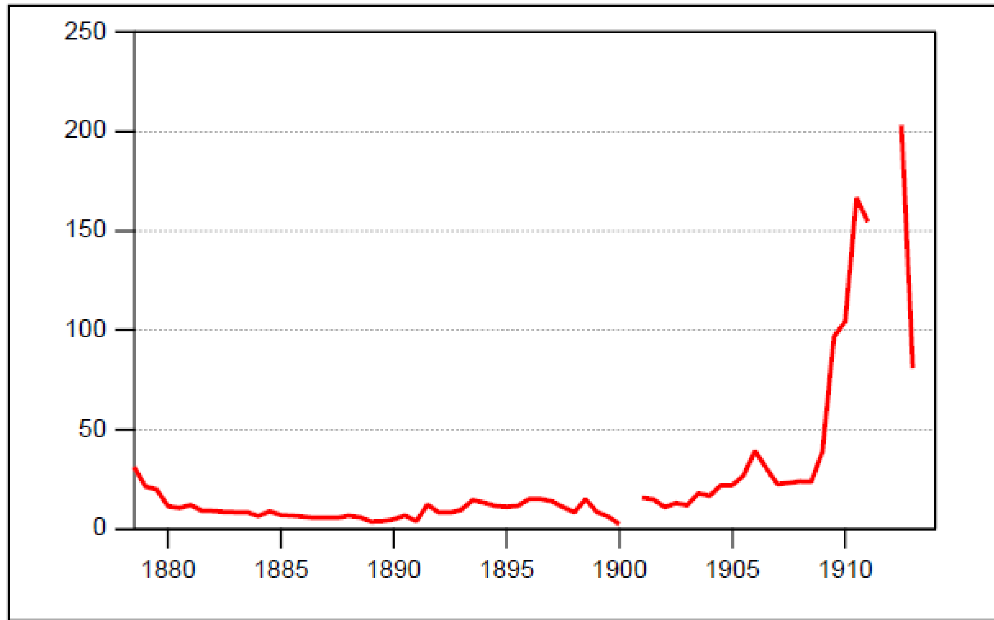


Figure 8: Profit series, plotted in standard (top) and log scale (bottom).

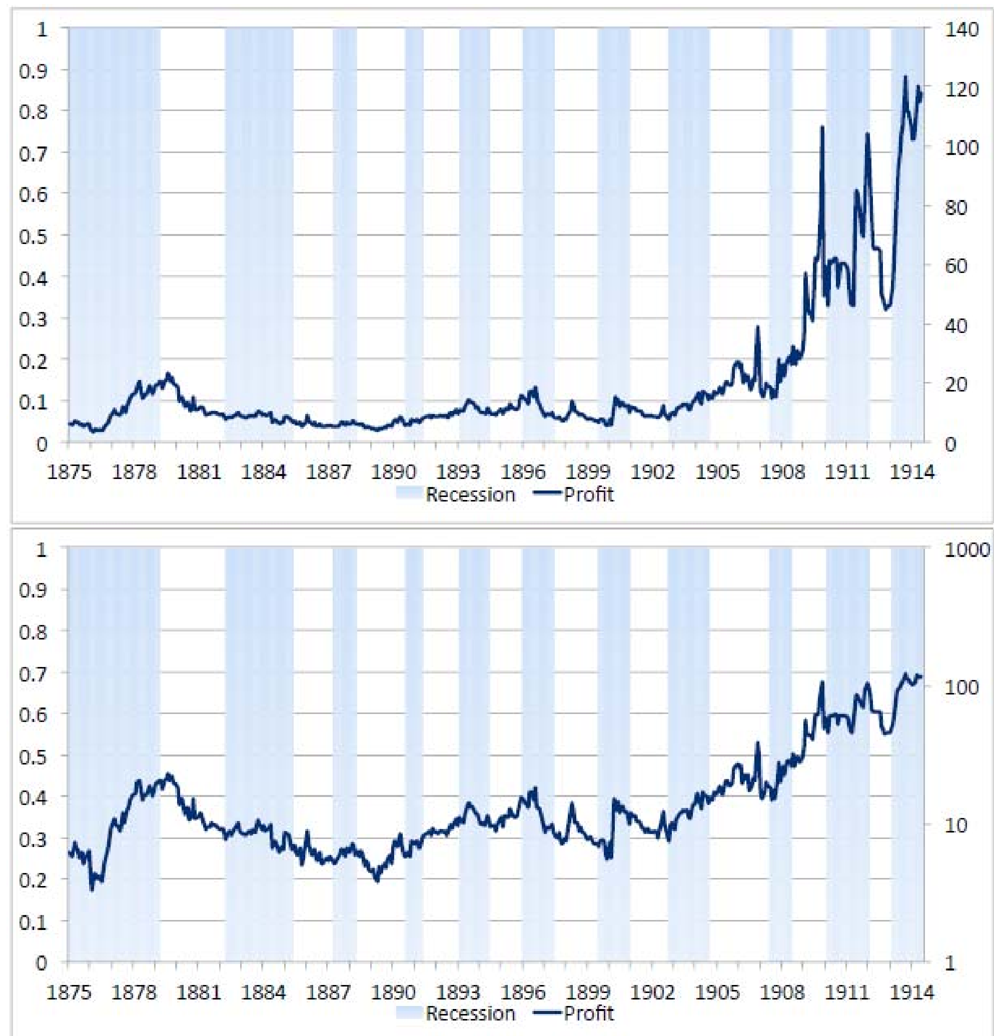


Figure 9: Ratio of notes to deposits. This figure plots the ratio of notes to deposits against US government debt to GDP. It shows that declines in government supply of Treasuries are strongly associated with increases in deposits.

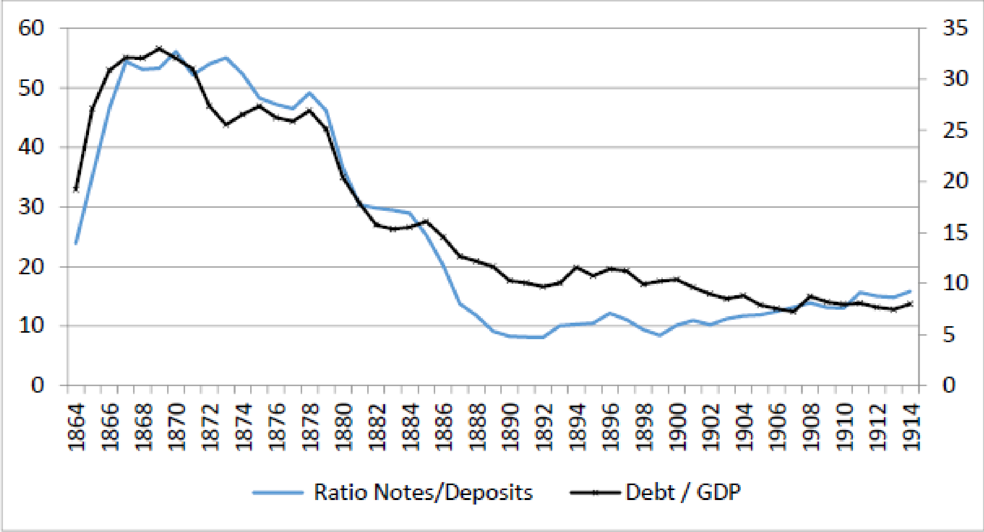
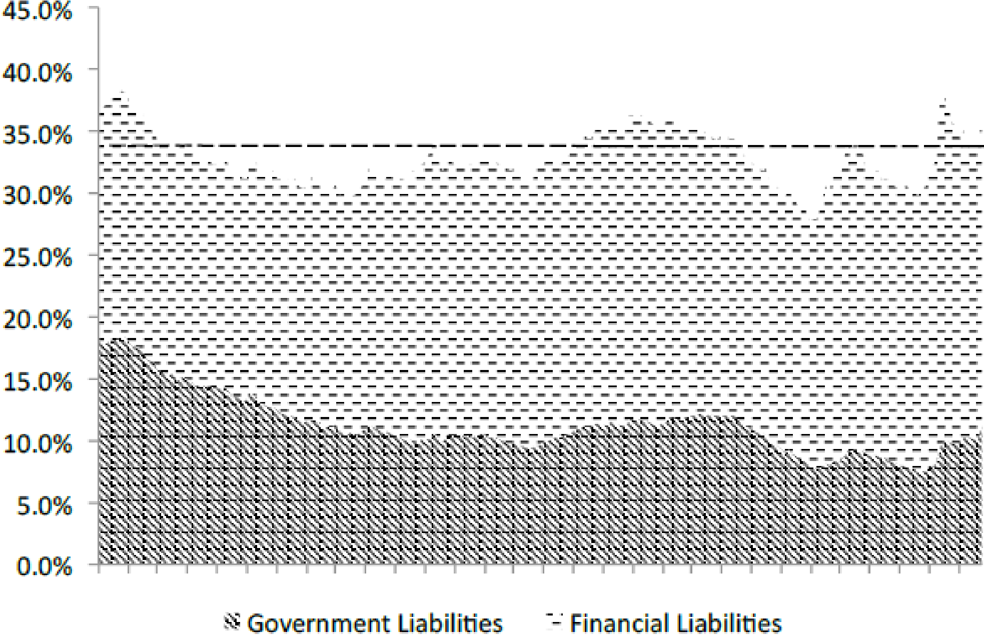


Figure 10: Safe asset share from Gorton Lewellen and Metrick 2012.



Appendices

Table 5: We report breakpoints for our fails data panel for the absolute value of the difference between fails to receive and fails to deliver along with 95% confidence intervals. The methodology for finding breaks in panels follows Bai (2010).

Panel A: All securities, $|\text{receive} - \text{deliver}|$

	Break Date	Lower Bound	Upper Bound
First Break	12-Sep-01	25-Jul-01	24-Oct-01
Second Break	19-Aug-09	1-Jul-09	30-Sep-09
Third Break	1-Feb-12	23-Nov-11	4-Apr-12

Panel B: Treasury+Agency, $|\text{receive} - \text{deliver}|$

	Break Date	Lower Bound	Upper Bound
First Break	6-Dec-00	31-May-00	6-Jun-01
Second Break	24-Sep-08	25-Jun-08	17-Dec-08
Third Break	11-Feb-09	17-Dec-08	1-Apr-09

Panel C: MBS, $|\text{receive} - \text{deliver}|$

	Break Date	Lower Bound	Upper Bound
First Break	19-Jun-02	1-May-02	31-Jul-02
Second Break	19-Aug-09	1-Jul-09	30-Sep-09
Third Break	1-Feb-12	9-Nov-11	18-Apr-12

Table 6: We recalculate our univariate profit regressions when splitting the sample before and after 1900. This deals with issues of "infinite" profits and our choice of α after 1900.

Profit on Convenience Yield in Subsamples:

$$y_t = a_1 + a_2 \times 1_{t>1900} + b_1 \times x_t + b_2 \times x_t \times 1_{t>1900} + \varepsilon_t$$

Profit on Convenience Yield in Subsamples:

$$y_t = a_1 + a_2 \times 1_{t>1900} + b_1 \times x_t + b_2 \times x_t \times 1_{t>1900} + \varepsilon_t$$

	y =profit		$y = \ln(\text{profit})$	
	(1)	(2)	(1)	(2)
ln(Avail/GDP)	-2.39		-0.28	
	[-1.84]		[-1.82]	
ln(Avail/GDP) $\times 1_{t>1900}$	-37.73		-0.95	
	[-7.29]		[-6.15]	
Muni spread		5.34		0.60
		[3.10]		[2.88]
Muni spread $\times 1_{t>1900}$		56.05		1.16
		[4.40]		[5.23]
AdjR^2	0.90	0.77	0.92	0.83
N	34	137	34	137

Table 7: We run seemingly unrelated regressions of Treasury and Agency fails to deliver.

	Δ Fails Del	Δ Fails Del	Δ Fails Del	Δ Fails Del	Δ Fails Del	Δ Fails Del
GC Repo-1m T-bill	6.727***	0.598	7.002***	0.552	7.207***	0.562
	(5.43)	(0.35)	(5.58)	(0.33)	(5.72)	(0.33)
L1.GC Repo-1m T-bill			2.118	0.770	2.448	0.607
			(1.69)	(0.45)	(1.93)	(0.36)
L2.GC Repo-1m T-bill					2.417	0.406
					(1.92)	(0.24)
GC Repo-1m T-bill x Break 1		12.97***		13.54***		12.96***
		(5.02)		(5.11)		(4.88)
L1.GC Repo-1m T-bill x Break 1				2.172		1.673
				(0.83)		(0.62)
L2.GC Repo-1m T-bill x Break 1						-1.971
						(-0.75)
GC Repo-1m T-bill x Break 2		35.82***		40.87***		39.41***
		(6.69)		(7.57)		(7.32)
L1.GC Repo-1m T-bill x Break 2				27.47***		31.77***
				(5.42)		(6.24)
L2.GC Repo-1m T-bill x Break 2						29.20***
						(5.81)
GC Repo-1m T-bill x Break 3		-0.421		-0.979		-0.152
		(-0.04)		(-0.08)		(-0.01)
L1.GC Repo-1m T-bill x Break 3				2.129		2.970
				(0.19)		(0.25)
L2.GC Repo-1m T-bill x Break 3						3.371
						(0.29)
Break 1 (9/2001-9/2008)		-6.809		-6.450		-4.782
		(-0.18)		(-0.17)		(-0.12)
Break 2 (9/2008-2/2009)		62.38		126.5		163.3
		(0.47)		(0.96)		(1.25)
Break 3 (2/2009)		10.22		9.742		6.906
		(0.24)		(0.23)		(0.16)
D. 1m T-Bill		-11.20***		-9.827***		-7.917***
		(-5.52)		(-4.77)		(-3.81)
Constant	0.966	-10.30	1.273	-8.808	1.145	-6.793
	(0.06)	(-0.42)	(0.08)	(-0.36)	(0.07)	(-0.28)
Observations	2398	2398	2386	2386	2374	2374
R^2	0.012	0.047	0.013	0.062	0.015	0.077

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: We run seemingly unrelated regressions of Treasury and Agency fails to deliver.

	Δ D-R	Δ D-R	Δ D-R	Δ D-R	Δ D-R	Δ D-R
GC Repo-1m T-bill	-0.0555	-0.0353	-0.0555	-0.0274	-0.0457	0.00581
	(-0.25)	(-0.12)	(-0.25)	(-0.09)	(-0.20)	(0.02)
L1.GC Repo-1m T-bill			0.0254	-0.0181	0.0472	-0.0198
			(0.11)	(-0.06)	(0.21)	(-0.07)
L2.GC Repo-1m T-bill					0.185	0.186
					(0.82)	(0.63)
GC Repo-1m T-bill x Break 1		0.118		0.327		0.307
		(0.27)		(0.71)		(0.66)
L1.GC Repo-1m T-bill x Break 1				0.606		0.726
				(1.34)		(1.55)
L2.GC Repo-1m T-bill x Break 1						0.184
						(0.40)
GC Repo-1m T-bill x Break 2		0.351		-0.00556		-0.0348
		(0.38)		(-0.01)		(-0.04)
L1.GC Repo-1m T-bill x Break 2				-1.920*		-1.933*
				(-2.19)		(-2.18)
L2.GC Repo-1m T-bill x Break 2						-0.280
						(-0.32)
GC Repo-1m T-bill x Break 3		0.957		0.618		0.464
		(0.48)		(0.30)		(0.22)
L1.GC Repo-1m T-bill x Break 3				-0.840		-1.049
				(-0.43)		(-0.52)
L2.GC Repo-1m T-bill x Break 3						-0.260
						(-0.13)
Break 1 (9/2001-9/2008)		89.87***		89.50***		89.39***
		(13.55)		(13.47)		(13.40)
Break 2 (9/2008-2/2009)		268.3***		263.4***		263.2***
		(11.88)		(11.61)		(11.56)
Break 3 (2/2009)		21.07**		21.40**		21.69**
		(2.92)		(2.95)		(2.97)
D. 1m T-Bill		-0.373		-0.544		-0.555
		(-1.07)		(-1.53)		(-1.54)
Constant	74.67***	37.40***	74.91***	37.34***	75.11***	37.37***
	(24.98)	(8.81)	(24.95)	(8.77)	(24.89)	(8.73)
Observations	2398	2398	2386	2386	2374	2374
R^2	0.000	0.114	0.000	0.117	0.000	0.117

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: We run seemingly unrelated regressions of MBS fails to receive.

	Δ Fails Rec	Δ Fails Rec	Δ Fails Rec	Δ Fails Rec	Δ Fails Rec	Δ Fails Rec
GC Repo-1m T-bill	1.435	0.164	1.004	0.170	0.942	0.164
	(0.49)	(0.04)	(0.34)	(0.04)	(0.32)	(0.04)
L1.GC Repo-1m T-bill			-4.289	-1.542	-4.208	-1.459
			(-1.45)	(-0.38)	(-1.40)	(-0.35)
L2.GC Repo-1m T-bill					-0.0987	0.0279
					(-0.03)	(0.01)
GC Repo-1m T-bill x Break 1		3.966		2.980		2.937
		(0.63)		(0.46)		(0.45)
L1.GC Repo-1m T-bill x Break 1				-1.898		-2.733
				(-0.30)		(-0.42)
L2.GC Repo-1m T-bill x Break 1						-2.440
						(-0.38)
GC Repo-1m T-bill x Break 2		-3.017		-5.401		-5.440
		(-0.23)		(-0.41)		(-0.41)
L1.GC Repo-1m T-bill x Break 2				-12.09		-12.12
				(-0.98)		(-0.98)
L2.GC Repo-1m T-bill x Break 2						0.613
						(0.05)
GC Repo-1m T-bill x Break 3		41.66		49.92		46.84
		(1.47)		(1.72)		(1.59)
L1.GC Repo-1m T-bill x Break 3				-67.00*		-59.98*
				(-2.43)		(-2.11)
L2.GC Repo-1m T-bill x Break 3						45.72
						(1.62)
Break 1 (10/2002-9/2009)		-4.295		-3.771		-2.076
		(-0.05)		(-0.04)		(-0.02)
Break 2 (9/2009-12/2011)		-95.63		-124.5		-122.9
		(-0.30)		(-0.39)		(-0.38)
Break 3 (12/2011)		22.72		-10.63		-2.786
		(0.22)		(-0.10)		(-0.03)
D. 1m T-Bill		-5.830		-5.987		-6.050
		(-1.18)		(-1.20)		(-1.20)
Constant	5.197	-3.454	-1.352	-3.557	-1.865	-4.992
	(0.13)	(-0.06)	(-0.03)	(-0.06)	(-0.05)	(-0.08)
Observations	1199	1199	1193	1193	1187	1187
R^2	0.000	0.003	0.002	0.012	0.002	0.013

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: We run seemingly unrelated regressions of MBS fails to deliver.

	Δ Fails Del	Δ Fails Del	Δ Fails Del	Δ Fails Del	Δ Fails Del	Δ Fails Del
GC Repo-1m T-bill	1.589	0.167	1.217	0.167	1.128	0.141
	(0.49)	(0.04)	(0.37)	(0.04)	(0.35)	(0.03)
L1.GC Repo-1m T-bill			-4.018	-1.389	-3.961	-1.305
			(-1.24)	(-0.31)	(-1.21)	(-0.29)
L2.GC Repo-1m T-bill					-0.363	-0.120
					(-0.11)	(-0.03)
GC Repo-1m T-bill x Break 1		4.087		3.201		3.229
		(0.60)		(0.46)		(0.46)
L1.GC Repo-1m T-bill x Break 1				-1.654		-2.484
				(-0.24)		(-0.35)
L2.GC Repo-1m T-bill x Break 1						-2.325
						(-0.33)
GC Repo-1m T-bill x Break 2		-2.498		-4.512		-4.395
		(-0.18)		(-0.32)		(-0.31)
L1.GC Repo-1m T-bill x Break 2				-10.19		-10.62
				(-0.76)		(-0.78)
L2.GC Repo-1m T-bill x Break 2						-2.041
						(-0.15)
GC Repo-1m T-bill x Break 3		46.19		56.22		52.77
		(1.49)		(1.77)		(1.64)
L1.GC Repo-1m T-bill x Break 3				-75.50*		-66.92*
				(-2.51)		(-2.15)
L2.GC Repo-1m T-bill x Break 3						47.62
						(1.54)
Break 1 (10/2002-9/2009)		-4.009		-3.488		-2.276
		(-0.04)		(-0.03)		(-0.02)
Break 2 (9/2009-12/2011)		-94.27		-118.5		-120.7
		(-0.27)		(-0.34)		(-0.35)
Break 3 (12/2011)		27.20		-11.20		-3.703
		(0.24)		(-0.10)		(-0.03)
D. 1m T-Bill		-5.665		-5.727		-5.981
		(-1.05)		(-1.05)		(-1.08)
Constant	6.074	-3.539	-1.453	-3.551	-1.983	-4.840
	(0.14)	(-0.05)	(-0.03)	(-0.05)	(-0.05)	(-0.07)
Observations	1199	1199	1193	1193	1187	1187
R^2	0.000	0.003	0.002	0.011	0.001	0.012

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: We run seemingly unrelated regressions of MBS fails to deliver.

	Δ D-R	Δ D-R	Δ D-R	Δ D-R	Δ D-R	Δ D-R
GC Repo-1m T-bill	0.127	0.0590	0.129	0.0633	0.132	0.0517
	(0.33)	(0.12)	(0.34)	(0.13)	(0.34)	(0.11)
L1.GC Repo-1m T-bill			-0.114	-0.0153	-0.0794	-0.00277
			(-0.30)	(-0.03)	(-0.20)	(-0.01)
L2.GC Repo-1m T-bill					0.0672	-0.0434
					(0.17)	(-0.09)
GC Repo-1m T-bill x Break 1		0.0152		0.112		0.118
		(0.02)		(0.15)		(0.16)
L1.GC Repo-1m T-bill x Break 1				0.242		0.246
				(0.33)		(0.32)
L2.GC Repo-1m T-bill x Break 1						0.103
						(0.14)
GC Repo-1m T-bill x Break 2		-0.0356		-0.428		-0.444
		(-0.02)		(-0.28)		(-0.29)
L1.GC Repo-1m T-bill x Break 2				-2.164		-2.104
				(-1.50)		(-1.45)
L2.GC Repo-1m T-bill x Break 2						0.587
						(0.41)
GC Repo-1m T-bill x Break 3		-0.615		-0.463		-0.750
		(-0.19)		(-0.14)		(-0.22)
L1.GC Repo-1m T-bill x Break 3				-6.530*		-5.844
				(-2.03)		(-1.76)
L2.GC Repo-1m T-bill x Break 3						2.303
						(0.70)
Break 1 (10/2002-9/2009)		111.3***		111.1***		111.4***
		(10.20)		(10.17)		(10.17)
Break 2 (9/2009-12/2011)		75.43*		70.27		71.15
		(2.03)		(1.88)		(1.90)
Break 3 (12/2011)		188.8***		187.5***		188.7***
		(15.92)		(15.72)		(15.70)
D. 1m T-Bill		0.289		0.154		0.181
		(0.50)		(0.26)		(0.31)
Constant	108.6***	29.24***	108.1***	29.19***	107.8***	28.97***
	(21.15)	(4.18)	(21.01)	(4.17)	(20.90)	(4.12)
Observations	1199	1199	1193	1193	1187	1187
R^2	0.000	0.188	0.000	0.190	0.000	0.191

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Comments by Randall S Kroszner[1]

Unintended consequences of well motivated regulation: lessons from history for financial regulation today

Since the financial crisis, regulators have raised capital and liquidity requirements for financial institutions substantially. Insufficient capital cushions against losses and liquidity mismatches in which long-term relatively illiquid assets were financed with short-term relatively liquid securities were key sources of the fragility of the financial system (see eg Kroszner and Melick (2011) and Kroszner and Shiller (2011)). The recent crisis made these sources of fragility crystal clear and thus provided a sound basis for increasing capital and liquidity requirements.

Even very well motivated regulations, however, can have unintended consequences. I will consider two historical episodes to draw lessons about the impact, intended and unintended, of financial regulation, with implications for current policy. The first relates to Gorton and Muir's (2015) excellent analysis of the US National Banking Era, from the Civil War to the founding of the Fed (1863–1913). The second relates to close parallels between the international response to the 1912 *Titanic* disaster and the response to the 2008–09 crisis, focusing on the lessons from the long forgotten 1915 *Eastland* tragedy in which more passengers died than in the *Titanic*.

Lessons from the US National Banking Era

Gorton and Muir (2015) draw a fascinating parallel between the regulations of the US National Banking Era and post-crisis bank regulation to learn about the consequences of liquidity regulation. In particular, they develop an analogy between the current "liquidity coverage ratio" (LCR), which requires banks to hold a certain amount of "high-quality liquid assets" such as Treasury securities, and the requirement in the earlier period that notes issued by banks be backed by Treasury securities.

The motivations for the regulations are similar, namely, to provide greater stability to the system and reduce the likelihood of runs and panics. The LCR can reduce the liquidity mismatch at banks by requiring that a significant fraction of their short-term liabilities, eg commercial paper and repurchase agreements, effectively be backed by short-term liquid Treasuries. In principle, that should reduce the likelihood of "funding runs" where banks cannot roll over their short-term paper to finance their operations.

In the "Free Banking Era" prior to the National Banking Act, relatively illiquid state bonds or portfolios of loans were used to back the issuance of notes, which were the key short-term liquid liabilities of banks. Switching to Treasuries as backing for the bank notes should have reduced the liquidity mismatch and, therefore, enhanced

[1] Booth School of Business, University of Chicago and NBER.

stability. Gorton and Muir (2015) argue, however, that that was not the case and draw lessons for potential unintended consequences of well motivated liquidity regulations such as the LCR today.

In the National Banking period, they argue, the requirement that notes issues be backed by Treasuries caused the collateral to become "immobile" and contributed to a "scarcity" of safe assets. That regulation-induced scarcity led to the private production of assets that appeared to be safe but were vulnerable to runs and involved fragile liquidity mismatches. They argue that, in that period, those privately produced apparently safe assets were bank deposits. They document that bank deposits grew significantly during the National Banking Era and generally inversely with Treasuries outstanding. Instability during the National Banking Era took the form of runs in which depositors demanded Treasury-backed bank notes that the banks were unable to produce.

Gorton and Muir argue that in the 21st century, in response to a scarcity of safe Treasuries, the markets produced apparently safe asset-backed commercial paper (ABCP), mortgage-backed securities (MBS), repurchase agreements (repos) etc that are often considered part of the "shadow" financial sector (see Pozsar et al (2012) and Kroszner (2015a)). Gorton and Muir argue that the LCR may "immobilise" Treasury collateral, contributing to its scarcity, and generating more fragility through forms of shadow banking. Rather than promote stability, regulations like the LCR could have the opposite impact by tying up safe debt and generating incentives for private production of substitutes that may appear safe but are subject to forms of funding runs during crises.

I find the use of historical analogies such as these enlightening and refreshing. Since we cannot run the experiment in advance to assess the costs and benefits of a new rule, we can try to see what evidence history can provide. Certainly, we would not want to draw definitive conclusions from the 19th century for the 21st, but we should not throw up our hands and say that we have no empirical guidance. Gorton and Muir are careful not to make strong predictions about whether the LCR ultimately will undermine stability, but they raise important questions that policymakers should take into account when formulating new rules.

Lessons from the international response to the *Titanic* and the *Eastland* tragedy

I want to generalise from Gorton and Muir's focus on what history can teach us about liquidity regulation to examine more broadly the unintended consequences of well-motivated regulation in response to a key historical crisis and implications for financial regulation today.

Roughly, 100 years ago (24 July 1915), 2,500 passengers in a festive mood boarded the SS *Eastland* for a picnic cruise on Lake Michigan. Just as the ship was about to leave its mooring in the Chicago River, the *Eastland* suddenly capsized and 841 passengers lost their lives – more passengers than in the sinking of the *Titanic* (see Sandburg (1915), Hilton (1995), McCarthy (2014) and Stranahan (2014); this section draws from Kroszner (2015b)).

The fate of the *Eastland* is one of America's great unremembered tragedies, yet it carries important lessons for policymakers around the world today, specifically

about how a regulatory response to one disaster can unleash unintended consequences that could contribute to another.

Following the *Titanic's* sinking, an International Conference on Safety of Life at Sea was convened to develop a global response. Several sensible reforms came from this conference: taking more southerly transatlantic routes to reduce the likelihood of encountering icebergs, and creating an Iceberg Patrol, still run today by the US Coast Guard, to monitor and warn of risks of icebergs.

Similarly, after the 2008–09 financial crisis, the G20, Financial Stability Board, Basel Committee and other international regulatory bodies convened to provide a coordinated global response, promoting rules to reduce risk exposures of banks and to increase macroprudential monitoring by supervisors.

A critical response to the *Titanic* disaster was a call for "lifeboats for all," a measure seen worldwide as necessary to prevent future tragedies at sea. If only the *Titanic* had had enough lifeboats for all its passengers and crew, more lives could have been saved and perhaps no one would have perished. What could be more sensible and obvious?

The measure was adopted in the 1914 International Convention relating to Safety of Life at Sea. It was immediately clear to some, however, that this policy might have unwanted ramifications.

In testimony to Congress, the general manager of the Detroit & Cleveland Navigation Company A A Schantz questioned whether such a requirement should be applied to ships plying the Great Lakes:

> *"The extra weight of the lifeboats and rafts would make them [the ships on the Great Lakes] top-heavy and unseaworthy, and in our judgment, we believe some of them would turn turtle [capsize] if you attempted to navigate them with this additional weight on the upper decks."* (quoted in Hilton (1995), p 11)

While Congress did not mandate "lifeboats for all" for ships on the Great Lakes, the 1915 La Follette Seaman's Act significantly increased the requirements. The *Eastland* added a sizeable number of rafts in 1914, as the federal Bureau of Marine Inspection and Navigation raised licensing requirements even before the Seaman's Act.

Just three weeks prior to the tragedy, and after the passage of the Act, the *Eastland* added more life boats and rafts to boost its licensed capacity to accommodate 2,500 passengers for the ill-fated picnic cruise. The *Eastland* was designed in 1903 with six lifeboats. When the catastrophe occurred, the *Eastland* had 11 lifeboats and 37 rafts. Each life raft weighed some 1,100 pounds.

The additional weight of the lifeboats and rafts may have been only one factor contributing to this calamity, but it illustrates how powerful unintended consequences of even the most sensible-seeming regulatory reforms can be. So what is to be learned from the *Eastland*?

First, even if a rule solves some problems, it doesn't necessarily solve all. In some cases, new regulations can undermine their own goals, creating new sources of instability, whether for financial markets or for ships. New rules can interact with other weaknesses in the system. "Lifeboats for all" can bring a false sense of comfort, and inspectors and supervisors may not look as carefully for other vulnerabilities. What if added lifeboats make suddenly "turning turtle" more likely in the very circumstance

where lifeboats won't help? Certainly, we would not want supervisors to have a false sense of security and overlook particular risks that an institution or activity may pose for the system simply because banks hold more capital (see Kroszner (2012)).

Second, one-size-fits-all regulation may not be appropriate. Higher capital and liquidity requirements for the largest global banks relative to smaller banks may make sense, just as "lifeboats for all" may be appropriate for transoceanic ships but not for steamships in the Great Lakes. This is consistent with regulatory capital and liquidity surcharges, as well as stress tests, applying to "systemically important" financial institutions.

Third, costs and benefits need to be taken seriously in financial regulation (see eg Kroszner and Strahan (2011)). With five or six years since key regulatory reforms like the Dodd-Frank Act have passed, now is an excellent moment to gather data to assess as best we can the benefits and the costs, the interactions of the Act's many components and other regulatory changes, and their potential unintended consequences. Have activities moved into the "shadows", and how has this affected the robustness of the system? What has been the impact on liquidity of markets? As is often said about the risks of icebergs – it's not what you can see but what you can't see below the water that is most dangerous.

Conclusions

The key lesson from the *Eastland* tragedy and the Gorton and Muir (2015) analysis of the National Banking Era is emphatically not that regulation is inevitably counterproductive. Instead, it is that we always need to consider unintended consequences and cost-benefit trade-offs, even for extremely well motivated rules, to protect us from economic turbulence. Policymakers who ignore history do so at their own – and the economy's – peril.

References

Gorton, G and T Muir (2015): "Mobile collateral versus immobile collateral", working paper, 2015.

Hilton, G (1995): *Eastland: Legacy of the Titanic,* Stanford University Press.

Kroszner, R (2012): "Stability, growth, and regulatory reform", *Banque de France Financial Stability Review: Public Debt, Monetary Policy, and Financial Stability*, April, pp 87–93.

——— (2015a): "The future of banks: will commercial banks remain central to the financial system?", working paper, April.

——— (2015b): "A shipping disaster's lessons for Dodd-Frank," *Financial Times*, 30 July.

Kroszner, R and W Melick (2011): "The response of the Federal Reserve to the recent banking and financial crisis", in J Pisani-Ferry, A Posen and F Saccomanni (eds), *An ocean apart? Comparing transatlantic response to the financial crisis*, Bruegel Institute and Peterson Institution for International Economics.

Kroszner, R and R Shiller (2011): *Reforming US financial markets: reflections before and beyond Dodd-Frank*, MIT Press.

Kroszner, R and P Strahan (2011): "Financial regulatory reform: challenges ahead", *American Economic Review*, *Papers and Proceedings*, May, vol 101, no 3, pp 242–6.

McCarthy, M (2014): *Ashes under water: the SS Eastland and the shipwreck that shook America,* Lyons Press.

Pozsar, Z, T Adrian, A Ashcraft and H Boesky (2012): "Shadow banking", Federal Reserve Bank of New York, *Staff Report*.

Sandburg, C (1915): "Looking 'em over," *International Socialist Review*, vol 16, no 3, September, pp 132–5.

Stranahan, S (2014): "The Eastland disaster killed more passengers than the Titanic and the Lusitania. Why has it been forgotten?", *Smithsonian Magazine*, 27 October, www.smithsonianmag.com/history/eastland-disaster-killed-more-passengers-titanic-and-lusitania-why-has-it-been-forgotten-180953146/#Sxbt5gr2Gh6ZKx4e.99

Comments by Andrei Kirilenko[1]

I thank the organisers for the opportunity to discuss this very insightful paper.

The main points of the paper are as follows:

- "In the thirty years prior to the 2007–08 financial crisis, the global financial system evolved away from a system of immobile collateral into a system of mobile collateral."
- "Since the financial crisis, regulatory initiatives have been aimed at making collateral once again immobile."
- "There is a cost to making this collateral immobile because it ties up safe debt."
- Instead of making the system safer, "other forms of bank debt increase when safe collateral becomes immobile, possibly making the system riskier".

In order to make these points, the paper looks to history. The main argument is that the current Basel Committee Liquidity Coverage Ratio (LCR) era is "structurally similar" to the 19th century US National Banking Era.

As a reminder, prior to the US National Banking Acts of 1863 and 1864, banks in the United States were chartered only by individual states. State-chartered banks issued their own bank-specific and location-specific banknotes. Banknotes issued by state-chartered banks were not uniformly accepted across the United States, because they were backed only by the capital of each individual bank. If a state-chartered bank went bankrupt, it was unable to make full payment on the banknotes it had issued.

The US National Banking Era was characterised by a rapid rise in the number of nationally chartered banks, which both issued and accepted national banknotes. This rapid rise was primarily due to two critical advantages that were given to nationally chartered banks vis-à-vis state chartered banks. Firstly, national banks were subject to tax of only 1% (0.5% from 1900) on their national banknotes in circulation. State chartered banks were subject to a 10% tax on their notes in circulation.

Secondly, national banks were able to take the full value of their capital, invest it in government bonds, deposit bonds with the Comptroller of the Currency (a US Federal Government agency created by the National Currency Act of 1863), get 90% (100% after 1900) of that value in national banknotes, and loan the notes out. State-chartered banks could only loan out the notes that they issued backed by their capital.

National banks were given a preferential treatment in order to make the nationwide payment system cheaper and to foster a deeper and safer national banking system. National banknotes were accepted as legal tender in states other than those in which they are issued. National banknotes were designed to be more difficult to counterfeit. Deposits made in national banknotes could be transferred between banks. Thus, national bank notes became a "new technology" that (among other things) helped facilitate a cheap, safe, nationwide payment system via third-party intermediaries – national banks.

The safety and depth of the national banking system depended on the explicit guarantee of the US Treasury. According to the Annual Report of the Comptroller of

[1] Massachusetts Institute of Technology.

the Currency (1874, Section XI): "[T]he United States guarantees the final payment of the [national] notes."

Yet, while number of nationally chartered banks increased rapidly between 1874 and 1913, according to the Annual Report of the Comptroller of the Currency (1914, Table 22), the number of the national notes in circulation did not skyrocket – it kept pace with US GDP growth.

There is a large literature on this so-called "issuance puzzle", ie the amount of notes in circulation seems to be smaller than would be implied by (seigniorage) arbitrage, but perhaps there is yet another explanation: the cost of government insurance of its guarantee on the national notes in the form of regulatory oversight of the national banks' capital adequacy.

In order to protect itself, the US Treasury was entitled to sell bonds it held as collateral to guarantee the issued notes. In addition, the US Treasury had first lien on the assets of a failed national bank. If the bonds and assets were not enough, the US Treasury could use the capital of the national bank shareholders to pay off holders of the national bank notes. Furthermore, national banks were subject to federal supervision and reporting requirements, so that problems could be detected and dealt with quickly.

With regard to the issuance puzzle, Gorton and Muir note, "our results indicate that the profits to note issuance fluctuate with the convenience yield on Treasuries, and our evidence is consistent with the idea that profits are related to the cost of [national] bank capital."

The authors use this result to argue that the National Banking system and the current LCR system "enforce a correspondence between two types of debt instruments, each with a convenience yield". They state: "the input for making one kind of money, bank notes or money market instruments, is required to be Treasuries". They also note that "such a system is fragile because by forcing two kinds of money together it is likely that there will be a shortage of one kind of money, leading to its private production elsewhere, which creates fragility in the system."

This is an important insight, rooted as it is in the history of banking. It provides a framework for understanding the LCR as a "new technology" intended to strengthen the global payment system by mandating global third-party intermediaries to use sovereign bonds issued in "national" currencies.

Yet, in the aftermath of the global financial crisis, it is not entirely obvious whether global third-party intermediaries (banks) offer a good way to make the global payment system cheaper and to foster a deeper and safer global banking system.

Digital currencies offer an alternative for the private production of national currencies for the use in the global payment system. Digital currencies are truly "new technologies" that are being developed to facilitate a cheap, safe, global payment system that does not rely on third-party intermediaries – global banks. One example of this is Bitcoin, which offers a peer-to-peer payment system with a distributed ledger, as well as a digital cryptocurrency. There are other examples of digital currencies – Litecoin, Peercoin, Namecoin, Dogecoin, Primecoin, Mastercoin etc.

If this the way the global system goes, it would need a supervisory regime and reporting requirements that are more similar to the one currently seen in commodity

markets. Regulators had little experience with such a regime a century ago. They do now.

Thank you.

Previous volumes in this series

No	Title	Author
560 May 2016	Has the pricing of stocks become more global?	Ivan Petzev, Andreas Schrimpf and Alexander F Wagner
559 April 2016	A comparative analysis of developments in central bank balance sheet composition	Christiaan Pattipeilohy
558 April 2016	Why bank capital matters for monetary policy	Leonardo Gambacorta and Hyun Song Shin
557 April 2016	How Does Bank Capital Affect the Supply of Mortgages? Evidence from a Randomized Experiment	Valentina Michelangeli and Enrico Sette
556 April 2016	Threat of Entry and Debt Maturity: Evidence from Airlines	Gianpaolo Parise
555 March 2016	The causal effect of house prices on mortgage demand and mortgage supply: evidence from Switzerland	Christoph Basten and Catherine Koch
554 March 2016	Can a Bank Run Be Stopped? Government Guarantees and the Run on Continental Illinois	Mark Carlson and Jonathan Rose
553 March 2016	What drives the short-run costs of fiscal consolidation? Evidence from OECD countries	Ryan Banerjee and Fabrizio Zampolli
552 March 2016	Fiscal sustainability and the financial cycle	Claudio Borio, Marco Lombardi and Fabrizio Zampolli
551 March 2016	When the Walk is not Random: Commodity Prices and Exchange Rates	Emanuel Kohlscheen, Fernando H. Avalos and Andreas Schrimpf
550 March 2016	A new dimension to currency mismatches in the emerging markets: non-financial companies	Michael Chui, Emese Kuruc and Philip Turner
549 March 2016	Monetary policy spillovers and currency networks in cross-border bank lending	Stefan Avdjiev and Előd Takáts
548 March 2016	Moving in tandem: bank provisioning in emerging market economies	Andres Murcia and Emanuel Kohlscheen
547 March 2016	When pegging ties your hands Intertemporal considerations in currency crises	Nikola Tarashev and Anna Zabai
546 March 2016	Financial intermediation and monetary policy transmission in EMEs: What has changed post-2008 crisis?	M.S. Mohanty and Kumar Rishabh
545 February 2016	Booms and banking crises	Frédéric Boissay, Fabrice Collard and Frank Smets
544 February 2016	What drives inflation expectations in Brazil? Public versus private information	Waldyr D Areosa

All volumes are available on our website www.bis.org.