Mobile Collateral versus Immobile Collateral*

Gary Gorton, Yale and NBER

Tyler Muir, Yale

June 13, 2016

Abstract

In the face of the Lucas Critique, economic history can be used to evaluate policy. We use the experience of the U.S. National Banking Era to evaluate the most important bank regulation to emerge from the financial crisis, the Bank for International Settlement's liquidity coverage ratio (LCR) which requires that (net) short-term (uninsured) bank debt (e.g. repo) be backed one-for-one with U.S. Treasuries (or other high-quality bonds). The rule is narrow banking. Will this rule reduce fragility in the financial system? The experience of the U.S. National Banking Era, which also required that bank short-term debt be backed by Treasury debt one-for-one, suggests that the LCR is unlikely to reduce financial fragility and may increase it.

*Thanks to Adam Ashcraft, Darrell Duffie, Randy Krozner, Andrei Kirilemko, Arvind Krishnamurthy, Philipp Hildebrand, Manmohan Singh, Paul Tucker, Warren Webber and seminar particpants at the 2015 BIS Annual Conference, the Penn Institute for Economic Research Workshop on Quantitative Tools for Macroeconmic 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 Bruce Champ (deceased) 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 the late Bruce Champ.

1 Introduction

The financial crisis of 2007-2008 emphasized the importance of bank regulation for financial stability and macroeconomic fragility. In response to the financial crisis, many new bank regulations have been implemented. In this paper, we ask: When is a proposed new bank regulation optimal? How can unintended consequences of new regulations be assessed? And, how can we answer these questions in the face of the Lucas (1976) Critique?

We focus on the most important new bank regulation of the post crisis era, the Liquidity Coverage Ratio (LCR).[1] The rule says, in essence, that all (net) short-term debt issued by a bank has to be backed dollar-for-dollar with U.S. Treasuries or similar safe debt, a kind of narrow banking.[2,3] The Bank for International Settlements (2013) labels the LCR as one of the key reforms "to develop a more resilient banking sector" (p. 1). The LCR is a very significant change and thus is expected to have some unintended consequences. The key question is whether this rule will significantly reduce financial fragility, as claimed, or significantly increase financial fragility. In the face of the Lucas (1976) Critique, we argue that economic history can be used to evaluate the LCR. The evaluation is not favorable.

[1] The LCR is a rule first proposed by the Basel Committee on Banking Supervision in December 2010; see BIS (2013). At the end of 2013 the Federal Reserve System issued is own mandatory LCR with slightly different definitions of details. The final rule is 103 pages: https://www.gpo.gov/fdsys/pkg/FR-2014-10-10/pdf/2014-22520.pdf . The European Commission also promulgated the LCR: http://ec.europa.eu/finance/bank/docs/regcapital/acts/delegated/141010_delegated-act-liquidity-coverage_en.pdf

[2] As far as we can tell, little policy evaluation was done before the LCR was adopted. There have 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. See Heller and Vause (2012), Sidanius and Zikes (2012), Fender and Lewrick (2013) and Duffie, Scheicher and Vuillemey (2014).

[3] In fact, since the financial crisis, many 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.

The Lucas Critique is important and is considered a fundamental principle of macroeconomics (e.g., Woodford (2003)). Lucas was responding to the macro econometric practice of the time. Simply put, the critique says that model parameters estimated with past data under a previous policy regime are not invariant to policy changes unless the parameters are "deep" parameters of tastes and technology. The response was to build models on micro foundations that (at least it could be claimed) represented such deep parameters. But, the interpretation and application of the critique has many unresolved issues (e.g., estimated "deep" parameters seem to change over time). These issues however are not our concern here. Whatever one thinks of current DSGE models, it is clear that they have little or nothing to say about financial crises or bank regulation. How then are newly proposed bank regulations to be evaluated? Do they mitigate financial fragility?

The financial crisis of 2007-2008 revealed that the financial system had significantly morphed from a retail insured demand deposit-based banking system into a wholesale banking system. The wholesale banking system relies on collateral to back short-term debt, which is the inside money of the wholesale system, i.e., sale and repurchase agreements (repo). The collateral consists of U.S. Treasuries or AAA asset-backed and mortgage-backed securities ABS/MBS. Our view is that (and we provide evidence for this view below) this system will remain important in the future. ABS/MBS use bank loans as input; that is the bank loans which sat on bank balance sheets as immobile collateral become mobile when securitized, transformed into ABS/MBS.

In this wholesale-based financial system U.S. Treasury debt plays a particularly important role; it has a convenience yield. Krishnamurthy and Vissing-Jørgensen (2012) find that the yield on U.S. Treasuries over 1926-2008 was, on average, 73 basis points lower than it otherwise would have been, due to the "moneyness" and safety of U.S. Treasury securities. In other words, U.S. Treasuries are important as money. Prior to the financial crisis, there was a scarcity of safe debt. When there is a scarcity of safe debt, the private response is to create more privately-produced debt that can act as a substitute (see Krishnamurthy and Vissing-Jørgensen (2012, 2015) and Gorton, Lewellen, and Metrick (2012)). Securitization was the private response, production of ABS/MBS. Using

ABS/MBS as collateral for repo made the financial system fragile.

The LCR aims to make Treasury and other safe debt immobile by requiring that it be used to back short-term bank debt. What will happen under the LCR? The LCR forces one kind of money – Treasuries—to back another type of money—short-term debt such as sale and repurchase agreements (repo). And, the ratio is fixed at one-to-one. If the one-to-one ratio is wrong, and U.S. Treasuries are valuable in an alternative use, rather than as the backing for short-term debt, then the amount of short-term debt issued will be too small. And it would likely, eventually, be produced somewhere else. In other words, it would spur the development of another shadow banking system.

Our main focus is methodological as applied to the LCR. We argue that in the face of the Lucas critique economic history provides some valuable guidance for policy evaluation. Episodes in the past are often similar to current proposed policies and can be a laboratory for studying the effects of proposed policies. History cannot be a perfect guide for policy evaluation and more than a model can. While it is no doubt the case that not every proposed new policy has a parallel in previous history, it may well be that there are close enough parallels to help inform a decision. Think of the historical example as the "model" in which the policy being considered has already been adopted. How far away is the historical episode from the policy to be evaluated? This question is the same when using models to evaluate policy, it's just implicit.

Using economic history to evaluate policy has its roots in Fogel's (1964) study of railroads in the U.S. One of his aims was to evaluate the "take-off" thesis of W.W. Rostow (1960), according to which economic growth needed a central industry, like railroads in the U.S., to achieve "take-off." Fogel (1964) addressed this issue by constructing a counterfactual involving the absence of railroads, replacing them with water transportation along rivers and canals.[4] We do not construct a counterfactual but rather look at a parallel structure in history that can be studied. Closer to this paper are recent examples that use historical parallels to evaluate policy. Bernstein, Hughson and Weidenmier (2015) study the effects of the establishment of the New York Stock Exchange clearinghouse on

[4]Recent work on this include Donaldson and Hornbeck (2015) and Swisher (2014). Also related is Murphy, Shleifer and Vishny (1989).

counterparty risk. Foley-Fisher and McLaughlin (2014) study structural differences between bonds guaranteed by the UK and Irish governments during the period 1920-1938. The events provide a way to think about sovereign debt that is jointly guaranteed by multiple governments, e.g., proposed Euro bonds. Bluedorn and Haelim (2016) study the effects on Pennsylvania banks of the New York Clearing House Association's bailouts of systemically important banks in New York City. They argue that the bailouts "likely short-circuited a full-scale banking panic" (p.1). 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. Bordo and Sinha (2015) study the large Fed bond buying program in 1932 to understand current QE policies.

We study the LCR, as an important example of this approach. In particular, the LCR is structurally identical to the U.S. National Banking Era which also required that banks' short-term debt (national bank notes) be backed by U.S. Treasuries.[5] We examine the National Banking Era experience to guide our thinking about the effects of the LCR. Under the National Banking System, national banks could issue distinct "national bank notes" by depositing eligible U.S. Treasury bonds with the U.S. Treasury, which would then print the bank's notes. Originally, the idea was to create a demand for U.S. Treasuries so as to finance the U.S. Civil War. But, it was also believed that backing private money with Treasuries would prevent banking panics. Prior to the National Banking Era, U.S. 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.

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.

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

During a panic, instead of requesting (gold and silver) cash for private bank notes, debt holders demanded national bank notes for their demand deposits. There was another problem with the National Banking System: too little money was issued. 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). The puzzle is that national banks never fully utilized 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.

Because the LCR is structurally the same as the National Banking System, this puzzle is important. In this paper, we show that the one 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. Also, average profit rather than marginal profit was calculated.

For the National Banking 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 historical experience as banking panics occurred frequently throughout the National Banking Era (1873, 1884, 1890, 1893, 1896, and 1907).

Combining the two interrelated steps of our argument in one paper is somewhat unusual, but necessary. We have to establish parallels between the two financial systems. We will argue that during the National Banking Era there was a shortage of U.S. Treasuries, and that it was this shortage that resulted, at least partially, in the under issuance of national bank notes, which accounts, in part, for the growth in (uninsured) demand deposits. To establish the parallel, in order to analyze the LCR, we need to first show that there was, and continues to be, a shortage of U.S. Treasuries today. So, we first, in Section 2, examine the changes in the U.S. financial system over the last thirty or so years. We provide evidence of a scarcity of safe long-term debt prior to the crisis as well as currently. To do this we study the determinants and extent of repo fails. A repo "fail" occurs when one party to the transaction does not perform to the contract at maturity, failing to return the collateral or the cash. We show that repo fails were increasing because of the scarcity of U.S. Treasuries and Agency bonds. Previous research, discussed below, shows that when U.S. Treasuries are scarce, the private sector provides substitutes in the form of asset-backed and mortgage-backed securities.

Having established the shortage of Treasuries in the modern financial system, we turn to the parallel LCR National Banking Era system in Section 3. There we provide evidence of a scarcity of U.S. Treasury debt during the National Banking Era and show that this led to growth in other forms of bank debt, demand deposits, which were vulnerable to runs. We calculate the profitability of national bank note issuance and then show that even in the 19th and early 20th centuries U.S. Treasuries had a convenience yield. We show that proxies for the convenience yield are important in explaining the Treasury convenience yield. Section 4 concludes with implications for the present day and a discussion of using economic history as a way to analyze proposed new policies.

2 Collateral Mobility and Scarcity

In this section we first briefly look at the transformation of the financial system in the thirty years prior to the 2007-2008 financial crisis. Then we provide evidence of the scarcity of safe debt by looking at repo fails.

2.1 The Transformation of the Financial System

Two important changes have occurred in the financial system in the last forty years. First, the system has evolved from one of immobile collateral into a system of mobile collateral. In other words, instead of bank loans remaining on bank balance sheets to provide backing for demand deposits, bank loans were securitized 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. Second, the demand for U.S. Treasuries and other safe debt by the rest of the world grew significantly. A result of these changes has been a shortage of safe debt.

Figure 1, from Gorton, Lewellen, and Metrick (2012), displays the composition of the privately-produced safe debt in the U.S. as a percentage of total privately-produced safe debt in the United States, since 1952 (based on Flow of Funds data). As shown in the figure, there was a very significant transformation of the U.S. financial system starting roughly in the mid-1970s. Demand deposits, which were the dominant form of safe debt for roughly the first 25 to 30 years, constituted nearly 80 percent of the total in 1952 and remained high at 70 percent in the late 1970s. But then demand deposits began a steep decline as the financial architecture changed with the rise of money market mutual funds, money market instruments (e.g., repo and commercial paper), and with securitization. This transformation reflects the changing demands for different types of safe debt, as demands from the wholesale market grew enormously relative to the retail market. The change is the rise of the Shadow Banking System. The figure suggests that this is a permanent change.

Also, in the last forty years or so there has been an enormous demand by foreigners for U.S. Treasury Debt. Bernanke, Bertaut DeMarco, and Kamin (2011): ". . . a large share of the highly rated securities issued by U.S. residents from 2003 to 2007 was sold to foreigners—55 percent. This share was even higher than in the 1998-2002 period—22 percent—even though total net issuance of apparently safe assets rose from \$3.1 trillion in the first period to \$4.5 trillion in the second [period]. (The net issuance of private label AAA-rated asset-backed securities outstanding, including MBS, rose from \$0.7 trillion in the first period to $2 trillion in the second.)" (p. 8). When there is a shortage of public safe debt, the private sector responds by producing substitutes. With shortages developing when the economy transformed from a retail-based banking system to wholesale-based system, two things happened. Commercial banks became much less profitable and there was a need for privately-produced (mobile) safe debt. The conjunction of these two forces led to securitization. Gorton and Metrick (2012b), Gennaioli, Shleifer, and Vishny (2011), Stein (2010), and others, argue that one of the main purposes of securitization is to produce safe assets.

Studies of the private sector issuance of safe debt confirm that issuance responds to widening of the convenience yield spread. Xie (2012) analyzes all private label ABS/MBS issued from 1978 to 2010. His data set is essentially all private label ABS/MBS in the market amounting to 20,000 deals, 300,000 tranches and $11 trillion in issuance. Using daily data Xie finds that more ABS/MBS are issued when the expected convenience yield is high. This phenomenon does not happen in other markets for privately-produced debt, like corporate bond markets. Sunderam (2015) looks at the issuance of asset-backed commercial paper (ABCP) at the weekly frequency. He finds, among other things, that issuance of ABCP also responds to a shortage of T-bills as evidenced by the convenience yield.

Betaut, Tabova and Wong (2014) examine the supply and demand of safe debt since the financial crisis of 2007-2008 and find that the scarcity of safe debt is a continuing problem. They show that post-crisis, the (high grade) foreign financial sector has produced and supplied safe debt to meet U.S. demand for safe assets. And, a large portion of this is in the form of foreign financial wholesale certificates of deposit. In particular, high-grade dollar-denominated debt from Australia and Canada is now 40% of U.S. foreign portfolio of high-grade dollar-denominated bonds, whereas pre-crisis this share was 8% pre-crisis. They also find "a strong negative correlation between the foreign share of the U.S. financial bond portfolio and measures of U.S. safe assets availability; providing evidence on the importance of foreign-issued financial sector debt as a substitute when U.S. issued "safe" assets are scarce."

The evidence reviewed so far is very suggestive. We next turn to supplying direct evidence of a shortage of safe debt.

2.2 Evidence of the Scarcity of Treasuries: Repo Fails

Many authors have discussed the shortage of safe debt prior to the financial crisis, e.g., Caballero (2010), Gourinchas and Jeanne (2012), Caballero and Farhi (2014) usually relating it to the global savings glut. But, it has proven difficult to provide evidence for this shortage. In this subsection we provide some evidence for this shortage, which was driving the growth of privately-produced mobile collateral.

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 a bond to return to the borrower, i.e. a fail. There is no direct evidence on this, but we provide a variety of indirect evidence.

We examine data from the New York Federal Reserve Bank on primary dealers' fails.[6] 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: U.S. Treasuries, Agency bonds, and Agency MBS.[7]

Repo fails by asset class are shown in the three panels of Figure 2. From Figure 2 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) in the 1990s compared to the period 2000-2007; also shown are the standard deviation of fails. We formally test for difference between subperiods below.[8]

[6] "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 .

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

[8] The data collected by the New York Fed is very limited. To get some sense of the nar-

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 of time. Secondly, there can be multiple fails due to rehypothecation (the re-use of collateral) chains, i.e., 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 is some survey data from the 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 percent of 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 percent for large firms in 2011. In 2014, ISDA for the first time asked about which bonds

rowness 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. See the DTCC membership list: http://www.dtcc.com/client-center/dtc-directories.aspx In 2011 DTCC settled $1.7 quadrillion in security value. DTCC also has a large repo program. DTCC fails data is for the value of Treasury and Agency fails, that is the amounts that were not delivered to fulfill 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 the larger universe of players. DTCC data show 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 shows. Also, see Gorton and Metrick (2015).

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 3 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 2000-2007, prior to the crisis. Moreover, note the sign difference between Treasuries and MBS during 2000-2007. For Treasuries receive minus fail is very large in 2000-2007, 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.[9]

[9]The fails data on Agency MBS market is 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

2.3 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 sub-panels, 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 September 12, 2001, just after September 11, 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 February 9, 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

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. The TBA market is very large. Average daily fails in this market between December 31, 2009 and December 29, 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).

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, i.e., it is general collateral rather than specific collateral.

The basic idea we explore is whether an increase in the GC repo spread, i.e., 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 as Agency bonds.

We use differences-in-differences in seemingly-unrelated regression on the panel of the Treasuries and Agency bonds, where fails are normalized by fails in 2013. We indicate the three breaks discussed above. The first period is September 12, 2001 going until September 23, 2008 and the second is September 24, 2008 until February 10, 2009, followed by February 11, 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 rate 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 on-line Appendix. In both cases the interaction between the GC repo spread and regime 1 is significant. 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,

[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).

and the interaction terms may be absorbing this effect.[11]

2.4 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 securitization 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 U.S. 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 U.S. 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 U.S. Treasury bonds because without the 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 U.S. Treasury bonds with the U.S. Treasury.[12] In this section we examine the U.S. 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

[11]In the on-line 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 on-line 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, e.g., the rise of large money managers, and foreign investors (see Gorton and Metrick (2015)).

[12]Eligible bonds were U.S. 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 July 12, 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 U.S. bearing interest in cash."

"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 summarizes the results of this section.

3.1 The U.S. National Banking System

During the U.S. National Banking Era banks were required to back their privately-produced money in the form of bank-specific national bank notes with U.S. 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 deposited at the US Treasury went to the banks. With national bank notes backed by U.S. Treasuries there was for the first time in the U.S. a uniform currency. Prior to the Acts, banks issued individual private bank notes which 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 percent of a bank's paid-in capital, but this was changed to 90 percent by the act of March 3, 1865. Also, note issuance was limited to 90 percent of the lower of par or market value. This was changed to 100 percent by an act in 1900. See Noyes (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 utilized. 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 recognize 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).[13]

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

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

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 annualized 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 U.S. 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.[14]

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

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

Cagan (1963) found very high profits rates for the 1870s, 20-30%. More

[13]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.

[14]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).

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 could be used and the bank could earn infinite profits.

Cagan's (1965) formula for calculating the arbitrage profits on note issuance assumes that the notes issued on a bond are used to pay for that very bond. James (1976), however, points out that as a practical matter the bank would see an increase in deposits and would then have a choice of lending out the new deposits directly, or purchasing bonds with the new deposits and then using the bonds as collateral for notes; 90 percent of the bond value would be the amount of notes that could then be lent out (until 1900 when it became 100%). James (1976) argues that if local lending rates are high, it would be more profitable to directly lend the new deposits and avoid losing the interest difference on the market price and 90 percent of the par value. Suppose that the yield on the bond is 3 percent and that the costs of note issue is 1 percent (which James takes as a plausible cost). James use the following condition for the bank to be indifferent between these two alternatives; the local loan rate, r, would have to solve:

$3\% + 0.9r - 1\% = r \Longrightarrow 20\%.(1)$

So, if rates were higher than 20 percent, the bank would not issue notes and would just lend directly, ignoring diversification and the price of risk.

James (1976) argues that local loan opportunities, as reflected in local loan rates, can explain why note issuance varied across the United States, in particular Southern and Western banks issued few notes in excess of the minimum. He regresses the percentage of bonds held in excess of bonds held for circulation above the minimum required on a measure on the difference in a local interest rate and r from equation (1) above. He finds effects of lending rates: "a percentage point decline in local interest rates leading to an increase of more than 5 percentage points in the proportion of notes issued . . . " (p. 366). Calomiris and Mason (2008) refined James' work extensively by studying the cross section of national banks in 1880, 1890, and 1900. They focus on the propensity of banks to issue notes and find that this propensity goes down as the banks have higher loans to total assets ratios (where Treasuries on deposit with the government are not counted in total assets).[15] A higher loans to total assets ratio suggests more lending opportunities and hence lower note issuance, which is what they find. They conclude that lending opportunities varied across the country and this accounts for some of under issuance.

Note that when the Monetary Reform Act of 1900 was passed, national banks received notes equal to the value of the bonds deposited with the Treasury. In that case, the formula does not apply and there is no choice between direct and indirect lending. But, further, the bank does not face the choice of lending directly or depositing bonds with the Treasury and lending the notes unless it is very costly to issue equity. As long as there are arbitrage profits on note issuance as Cagan calculated, then banks should take advantage of it unless it is too costly to issue equity. So, the existence of arbitrage profits suggests that either (1) local lending rates were high prior to 1900; (2) the convenience yield was very high; or (3) equity issuance was very costly, and these are not mutually exclusive.

Subsequently, our focus is on the convenience yield of Treasuries and on bank capital. We observe that banks hold Treasuries on their balance sheets, presumably for risk management reasons. The ratio of Treasuries on the balance sheet to loans and discounts varied from a high of 3.9 percent to a low of 8 bps. In fact, when the arbitrage profits on note issuance were high, banks did respond by reducing their on balance sheet Treasury holdings. Suppose that out of a new deposit of one dollar the bank wants to hold 10 cents of Treasuries on balance sheet. The benefit to the bank from doing this is the "convenience yield," estimated by Krishnamurthy and Vissing-Jorgensen (2012), over the period 1926-2008, to be 73 basis points. Suppose that whether lending directly or with notes, the bank holds 10 percent of the new deposit dollar on balance sheet in the form of Treasuries. So, 0.1*0.73 is earned in either case. Taking this into consideration the indifference condition, ignoring reserve requirements, becomes:

$3\% + 0.9(1 - 0.1)r - 1\% = (1 - 0.1)r \Rightarrow 22.22\%.$

The loan rate rises because less is lent out.

Another explanation for the persistence of the arbitrage opportunity concerns

[15] Banks had to hold a minimum amount of Treasuries with the government as a percentage of paid in capital and the banks issued notes against this minimum. Calomiris and Mason (2008), the define the "propensity to issue" of a bank as: (Actual Notes – Minimum Required Issuance)/(Maximum Permissible Notes – Minimum Required Issuance).

costs of issuance, in particular there may be costs of note redemptions (on this there are a number of papers; see the citations in Calomiris and Mason (2008)). Calomiris and Mason (2008) are skeptical of this argument because of the low volume of redemptions, once worn out notes and the notes of insolvent banks are subtracted.

In this paper, we do not dispute any of the previous explanations, and our explanation is not mutually exclusive. We argue that the persistence of the arbitrage opportunity can be partially explained by the convenience yield on Treasuries and limits to arbitrage in the form of expensive bank capital. The issue of a convenience yield seems very relevant since banks held Treasuries on their balance sheets when they could have simply deposited them with the Treasury. And there is no choice between loan opportunities and profitable note issuance if banks can simply issue new equity. Neither of these two factors have been examined before. Since we do not observe the convenience yield (which also may have varied regionally) or the cost of bank equity, we can only look at the indirect measure of the arbitrage profits as calculated by Cagan. We have a limited goal: to show that the arbitrage profits are, indeed, related to the convenience yield (and bank capital, but we have no measures or proxies for bank capital).

One way for new bank equity to enter the system is via bank entry, a point raised by Hetherington (1990). Hetherington (1990) points out that while note circulation fell 50 percent from \$360 million in 1882 to \$171 million in 1891, the number of national banks doubled. Hetherington (1990) assigns major importance to two regulatory changes during the National Banking Era. The Act of July, 23 1882 caused note issuance to decline. The act reduced the minimum required bond holdings to 25 percent of capital with a ceiling of \$50,000. Because bonds were at a premium, banks made large capital gains by selling bonds and retiring notes. The lower minimum also made it easier for banks to enter. Most of the entry occurred in the South, the Midwest and the Pacific regions. James (1976) argues that the rate of returns on loans was very high in the South and West, explaining the low issuance of national bank notes in these regions. Calomiris and Mason (2008) find that entrants had a low propensity to issue notes. Existing banks, particularly the very large banks in the East could have raised capital. But, Calomiris and Mason (2008) say that "banks seemed loath to increase capital in order to boost note issuance. . . " (p. 346).

Finally, another important issue concerns the fact that the above profit calculations are not marginal calculations, but average profits. We will argue that this distinction is particularly important for the post-1900 period.

Figure 5 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.[16] 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?[17] After all, banks held U.S. Treasury bonds on their balance sheets (i.e., not including Treasuries held to back their notes). See Figure 4.Or, banks could have raised capital and used this to buy bonds for collateral for notes. Prior explanations do not mention that there may be a convenience yield associated with U.S. 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 analyze the period 1914-2011. Figure 4, showing that national banks held U.S. Treasuries on their balance sheets (i.e., the ratio of U.S. 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 recognized the convenience yield on Treasuries. For example, *The Financier*, April 7, 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

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

[17] There is a large literature on the underissuance puzzle that we hae not discussed, e.g. Bell (1912), Cagan (1965), Goodhart (1965), Cagan and Schwartz (1991), Duggar and Rost (1969), Champ, Wallace and Weber (1992), and Wallace and Zhu (2004). None of these explanation are mutually exclusive with those we have discussed.

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 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.[18] 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.[19]

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. U.S. government bonds were hard to find. And, even when banks could find bonds, they had to reverse repo in the bonds

[18]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.

[19]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.

n 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 U.S. 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 $1\frac{1}{2}$ 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, savings banks or private holders may have 'Governments' which they are willing to lend to national banks for a consideration" (p. 321).[20] 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. 157-158).

[20] 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.

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 summarized 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 $1\frac{1}{2}\%$ 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 was voted to appoint a committee of five economists to consider and report on currency reform in the United States.[21] 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 5 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.[22] For example, Frank Mortimer, cashier of the First National Bank, Berkeley. Ca.; 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*,

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

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

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-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 variables and the muni spread are highly significant independently. However, we would suspect that they likely measure similar economic forces though 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 though 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 5 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 aren't 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 (e.g., monthly data on debt/GDP) won't be particularly helpful here in overcoming the fairly small sample because the variables are highly persistent.

We hypothesize 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 5 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 The Rise of Demand Deposits

When Treasuries have a convenience yield, and short-term bank debt must be backed by Treasuries, there is a tradeoff between the two types of money. More short-term debt means fewer Treasuries for alternative uses. This tradeoff is common to the two systems. We saw this in the National Banking Era. The tradeoff 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 formalized by Krishnamurthy and Vissing-Jorgensen (2012, 2015) for the modern era. The convenience yield is negatively related to Treasuries (divided by GDP) outstanding.

One way out of this tradeoff, 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 4, panel C). And, from the start of the National Banking System, the ratio of bank notes to demand deposits fell, from just over 60 percent in the early part of 1865 to 14 percent by 1909, as shown in Figure 6. 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 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 6. 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.

We provide several robustness checks for this result. A potential concern is just that deposits happened to be trending upwards over this time while US debt to GDP is secularly falling. We address this in two ways. First, we run these regressions in changes rather than in levels. Using one year changes the correlation falls to 0.22, using two year changes it increases to 0.58 and using three year changes it becomes 0.69. By using changes, we remove the secular time trend in both series. However, when using shorter horizon changes the results are likely much more affected by measurement error and we are also making an implicit assumption about how quickly deposit growth endogenously responds to scarcity of Treasuries. Thus, more moderate or longer horizons may give more reasonable estimates, though our results indicate that the positive correlation remains for changes rather than levels.

Next, the trend in deposits that we show might be simply a global trend – deposits might be rising in all countries for reasons outside of US policy. To address this, we compute deposits to GDP across 14 developed economies used in Schularick and Taylor (2014). We then regress the ratio of deposits to GDP in the United States on the equal weighted average of the deposits to GDP ratio in all other countries. The coefficient is 1.7, showing that the US had deposits grow much faster over this time period than in other countries. This makes it more plausible that the increase in deposits was due to US specific factors rather than simply a global trend. Taken together, the data suggests that the scarcity of Treasuries played an important role in the growth of deposits over this period.

4 Conclusion

The Lucas critique seems to be largely ignored by bank regulators making new policy for the simple reason that the requisite general equilibrium models do not exist. Proposed new policies can, however, be evaluated with economic history. Economic history provides a laboratory to study large, important, policy changes, like the LCR. When examining the National Banking Era to evaluate the LCR we find a negative experience that does not bode well for the LCR.

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–narrow banking. 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.

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 U.S. government debt outstanding and privately-produced safe debt outstanding has been 32 percent of total assets in the U.S. since 1952. They show further that there has never remotely been enough U.S. Treasuries to make up the 32 percent 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.

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.

References

Bai, Jushan (2010), "Common Breaks in Means and Variances for Panel Data," *Journal of Econometrics* 157, 78-92.

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

Bank for International Settlements (BIS), Bank Commitee on Banking Supervision (2013), "Basel III: The Liquidity Coverage Ratio and Liquidity Risk Monitoring Tools."

Basel Committee on Banking Supervision, Bank for International Settlements (2011), "Basel III Framework for Liquidity–Frequently Asked Questions."

Bell, Spurgeon (1912), "Profit on National Bank Notes," *American Economic Review* 2, p. 38-60.

Bernanke, Ben S. (2007), "Global Imbalances: Recent Developments and Prospects," speech delivered at the Bundesbank Lecture, Berlin, Germany, September 11.

Bernanke, Ben S. (2005), "The Global Savings Glut and the U.S. Current Account Deficit," speech delivered at the Sandridge Lecture, Virginia Association of Economists, Richmond, Va., March 10.

Bernanke, Ben S., Carol Bertaut, Laurie Pounder DeMarco and Steve Kamin (2011), "International Capital Flows and the Returns to Safe Assets in the United States, 2003-2007," *Banque de France Financial Stability Review* (February), 13–26.

Bernstein, Asaf, Eric Hughson and Marc Weidenmier (2015), Counterparty Risk and the Establishment of the New York Stock Exchange," working paper.

Bertaut, Carol, L.P. DeMarco, S. Kamin and R. Tryon (2012), "ABS Inflows to the United States ant the Global Financial Crisis," *Journal of International Economics* 88, 219-234.

Bertaut, Carol, Alexandra Tabova and Vivian Wong (2014), "The Replacement of Safe Assets: Evidence from the U.S. Bond Portfolio," Board of Governors of the Federal Reserve System, International Finance Discussion papers, No. 1123.

Bluedorn, John and Haelim Park (2016), "Stopping Contagion with Bailouts: Microevidence from Pennsylvania Bank Networks during the Panic of 1884," *Journal of Banking and Finance*, forthcoming.

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

Bolles, Albert (1902), "Responsibility of the National Bank in the Present Crisis," *Annals of the American Academy of Political and Social Science*, Vol. 20, 1-18.

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

Bordo, Michael and Arunima Sinha (2015), "A Lesson from the Great Depression that the Fed Might have Learned: A Comparison of the 1932 Open Market Purchases with Quantitative Easing," working paper.

Caballero, Ricardo (2010), "The 'Other' Imbalance and the Financial Crisis," NBER Working Paper 15636.

Caballero, Ricardo, and Emmanuel Farhi (2014), "The Safety Trap," NBER Working Paper No. 20652.

Caballero, Ricardo and Arvind Krishnamurthy (2009), "Global Imbalances and Financial Fragility," *American Economic Review: papers & Proceedings* 99, 584-588.

Cagan, Phillip (1963), "The First Fifty Years of the National Banking System," in *Banking and Monetary Studies,* edited by Deane Carson, 15-42 (Richard D. Irwin; Homewood, IL).

Cagan, Phillip (1965), *Determinants and Effects of Changes in the Stock of Money, 1875-1960* (NBER; New York).

Cagan, Phillip and Anna J. Schwartz (1991), "The National Bank Note Puzzle Reinterpreted," *Journal of Money, Credit and Banking* 23, 293-307

Calomiris, Charles and Joseph Mason (2008), "Resolving the Puzzle of the Underissuance of National Bank Notes," *Explorations in Economic History* 45, 327-355.

Carlisle, John (1894), *Report of the Committee on Banking and Currency Together with Hearings*, House of Representatives, 8149, 53rd Congress, 3rd Session.

Carlson, Mark and Jonathan Rose (2014), "Can a Bank Run Be Stopped? Government Guarantees and the Run on Continental Illinois," Federal Reserve Board, working paper.

Champ, Bruce (2011a), "The National Banking System: The Empirical Evidence," Federal Reserve Bank of Cleveland, Working Paper 07-19R.

Champ, Bruce (2011b), "The National Banking System: The National Bank Note Puzzle," Federal Reserve Bank of Cleveland, Working Paper 07-22R.

Champ, Bruce (2011c), "The National Banking System: A Brief History," Federal Reserve Bank of Cleveland, Working Paper 07-23R.

Champ, Bruce, Neil Wallace, and Warren Weber (1992), "Resolving the National Bank Note Paradox," Federal Reserve Bank of Minneapolis *Quarterly Review* (Spring), 13-21.

Donaldson, Dave and Richard Hornbeck (2015), "Railroads and American Economic Growth: A 'Market Access' Approach," University of Chicago Booth working paper.

Duffee, Gregory (1996), "Idiosyncratic Variations of Tresury Bills," *Journal of Finance* 51, 527-551.

Duffie, Darrell, Martin Scheicher and Guillaume Vuillemey (2014), "Central Clearing and Collateral Demand," *Journal of Financial Economics,* forthcoming.

Duggar, Jan Warren and Ronald Rost (1969), "National Bank Note Redemption and Treasury Cash," *Journal of Economic History 29, 512-520.*

Dunbar, Charles (1887), "Deposits as Currency," *Quarterly Journal of Economics 1, 401-419.*

Eckels, James (1894a), *Report of the Committee on Banking and Currency Together with Hearings*, House of Representatives 8149, 53rd Congress, 3rd Session.

Eckels, James (1894b), *Annual Report of the Comptroller of the Currency*, 3rd Session, 53rd Congress, Dec. 3, Government Printing Office, Washington DC.

Fender, Ingo and Ulf Lewrick (2013), "Mind the Gap? Sources and Implications of Supply-Demand Imbalances in Collateral Asset Markets," *BIS Quarterly Review*, September 2013, 67-81.

Fisher, Willard (1895), "Money and Credit paper in the Modern Market," *Journal of Political Economy* 3, 391-413.

Fleming, Michael and Kenneth 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), 35-57.

Fogel, Robert (1964), *Railroads and American Economic Growth* (Johns Hopkins University Press).

Foley-Fisher, Nathan and Jonathan McLaughlin (2014), "Sovereign Debt Guarantees and Default: Lessons from the UK and Ireland, 1920-1938," Federal Reserve Board, working paper.

Friedman, Milton and Anna Scwartz (1963), *A Monetary History of the United States (Princeton University Press).*

Garbade, Kenneth and Michael Fleming (2005), "Explaining Settlement Fails," Federal Reserve Bank of New York *Current Issues in Economics and Finance*, vol. 11, no. 9 (September).

Garbade, Kenneth, Frank Keaane, Lorie Logan, Amanda Stokes and Jennifer Wolgemuth (2010), "The Introduction of the TMPG Fails Charge for U.S. Treasury Securities," *FRBNY Economic Policy Review* (October), 45-71.

Goodhart, Charles A. E. (1965), "Profit on National Bank Notes, 1900-1913," *Journal of Political Economy* 73, 516-522.

Gorton, Gary (2013), "The Development of Opacity in U.S. Banking," *Yale Journal of Regulation*, forthcoming.

Gorton, Gary and Andrew Metrick (2015), "Who Ran on Repo," Yale working paper.

Gorton, Gary, Stefan Lewellen, and Andrew Metrick (2012), "The Safe-Asset Share," *American Economic Review: Papers & Proceedings* 102, 101-106.

Grossman, Richard (2001), "Double Liability and Bank Risk Taking," *Journal of Money, Credit and Banking* 33(2), 143-159.

Hachem, Kinda Cheryl and Zheng Michael Song (2015), "The Rise of China's Shadow Banking System," University of Chicago, working paper.

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

Heller, Daniel and Nicholas 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; Indianapolis).

James, John A. (1976), "The Conundrum of the Low Issue of National Bank Notes," *Journal of Political Economy* 84, 359-368.

Kinley, David (1910), *The Use of Credit Instruments in Payments in the United States* (National Monetary Commission; Washington D.C.).

Krishnamurthy, Arvind and Annette Vissing-Jorgensen (2012a), "The Aggregate Demand for Treasury Debt," *Journal of Political Economy* 120, 233-267.

Krishnamurthy, Arvind and Annette Vissing-Jorgensen (2012), "Short-term Debt and Financial Crisis: What can we learn from U.S. Treasury Supply," *Journal of Financial Economics*, forthcoming..

Kuehlwein, Michael (1992), "The National Bank Note Controversy Reexamined," *Journal of Money Credit and Banking* 24, 111-126.

Laughlin, Lawrence (1920), *Banking Progress* (Charles Scribners; New York).

Lucas, Robert (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 1 (Elsevier; New York), 19-46.

Macaulay, Frederick (1958), *The Movements of Interest Rates, Bond Yields and Stock Prices in the United States Since 1856* (National Bureau of Economic Research; New York, NY).

Macey, Jonathan and Geoffrey Miller (1992), "Double Liability of Bank Shareholders: History and Implications," *Wake Forest Law Review* 27, 31-62.

Maggiori, Matteo (2013), "Financial Intermediation, International Risk Sharing, and Reserve Currencies," Harvard University, working paper.

Murphy, Kevin, Andrei Shleifer and Robert Vishny (1989), "Industrialization and the Big Push," *Journal of Political Economy* 97, 1003-1026.

Noyes, Alexander (1910), *A History of the National Bank Currency*, National Monetary Commission (Government Printing Office; Washington D.C.

Rostow, Walt Whitman (1960), *The Stages of Economic Growth* (Cambridge University Press).

Sidanius, Che and Filip Zikes (2012), "OTC Derivatives Reform and Collateral Demand Impact," Bank of England, Financial Stability Paper No. 18.

Singh, Manmohan (2011), "Velocity of Pledged Collateral," IMF Working Paper WP/11/256.

Singh, Manmohan (2014),, *Collateral and Financial Plumbing* (Risk Books).

Swisher, Scott (2014), "Reassessing Railroads and Growth: Accounting for Transport Network Externality," University of Wisconsin working paper.

Singh, Manmohan and James Aitken (2010), "The (sizeable) Role of Rehypothecation in the Shadow Banking System," IMF Working Paper WP/10/172.

Taylor, F. M. (1898), "The Final Report of the Indianapolis Monetary Commission," *Journal of Political Economy* 6, 293-322.

Treasury Market Practices Group (2011), "Understanding Settlement Fails in Agency Mortgage-Backed Securities."

Vickery, James and Joshua Wright (2013), "TBA Trading and Liquidity in the Agency MBS Market," Federal Reserve Bank of New York *Economic Policy Review* (May), 1-18.

Wallace, Neil and Tao Zhu (2007), "Float on a Note," *Journal of Monetary Economics* 54, 229-246.

West, Robert Craig (1974), *Banking Reform and the Federal Reserve, 1863-1923* (Cornell University Press; Ithaca, NY and London).

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

Wicker, Elmus (1966), *Federal Reserve Monetary Policy 1917-1933* (Random House:New York).

Woodford, Michael (2003), Interest and Prices: Foundation of a Theory of Monetary Policy (Princeton University Press).

5 Tables and Figures

Table 1: We present summary statistics on fails and rehypothecation. Sources: Panels A & B, Federal Reserve Bank of New York. Panel C, ISDA.

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%

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(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: The safe asset share from 1952-2012. Source: Gorton, Lewellen, Metrick (2012).

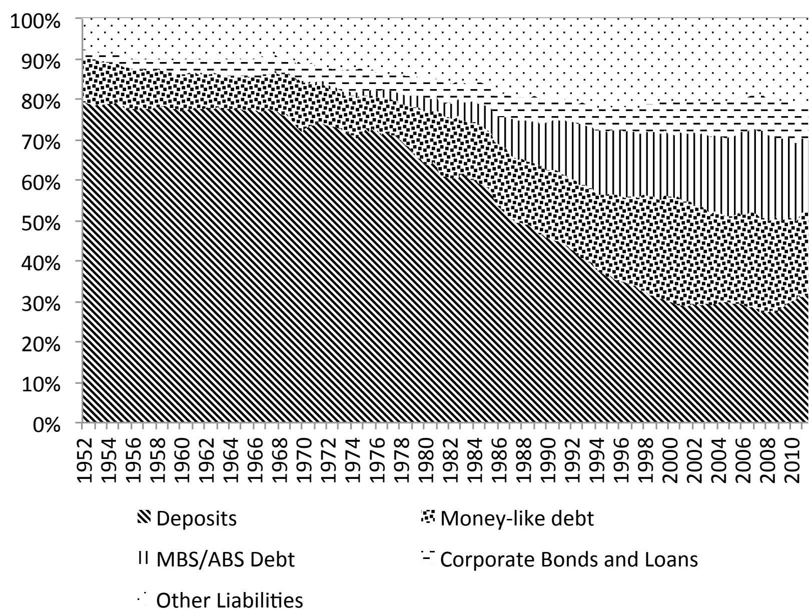


Figure 2: Fails by type.

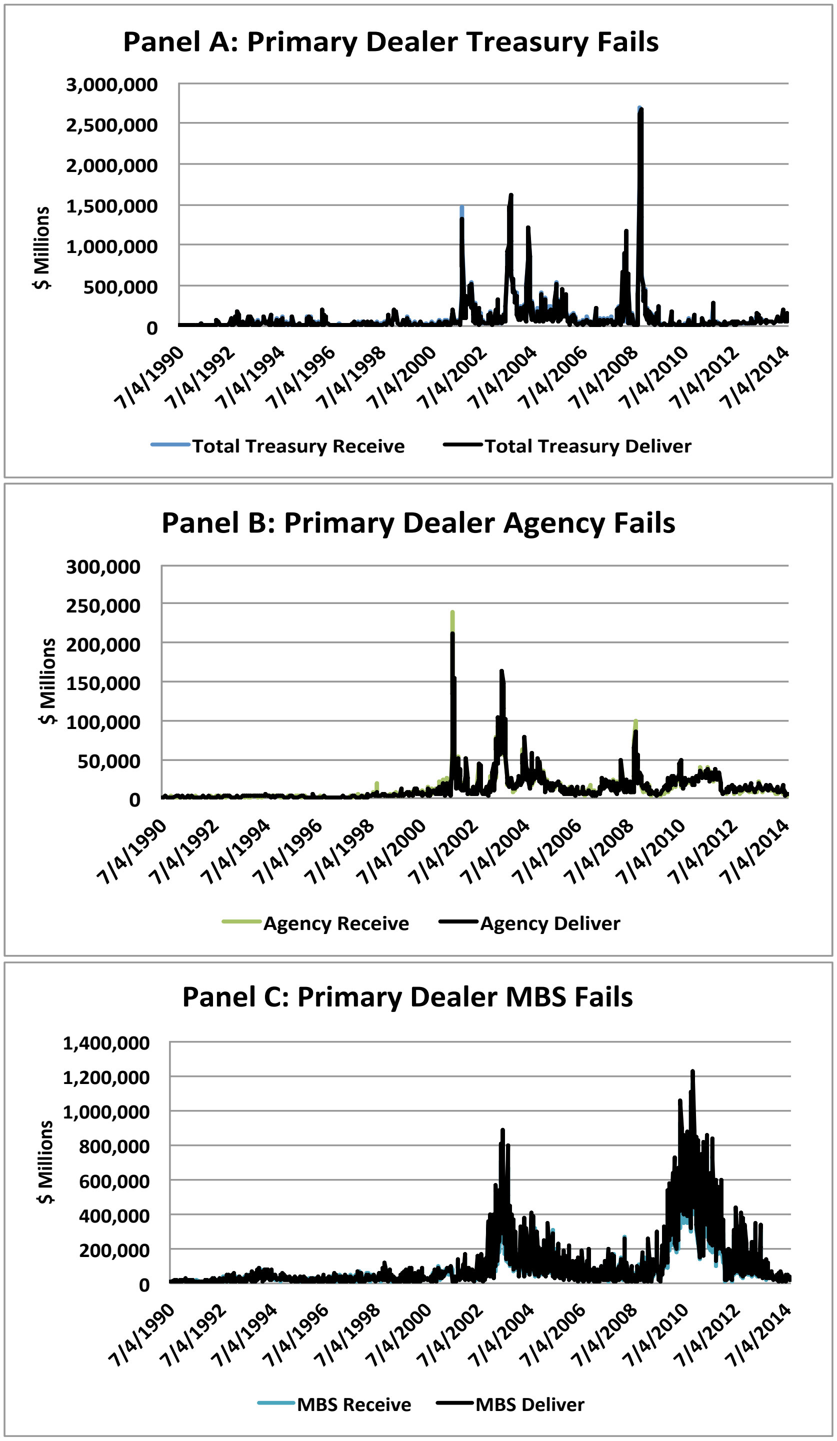


Figure 3: Difference between fail to receive and failure to deliver by type.

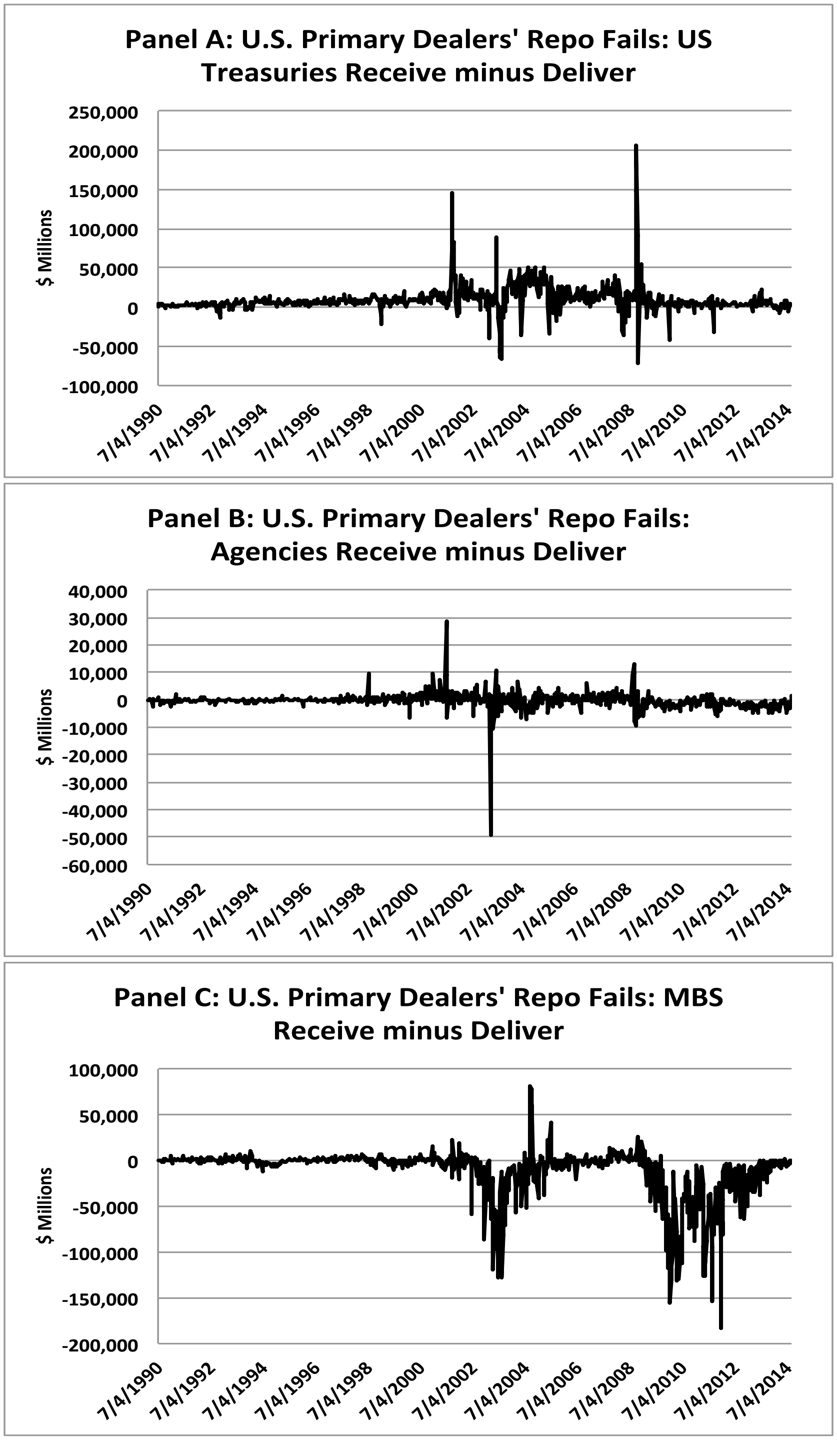


Figure 4: Panel A plots the fraction of Treasuries held to back notes and hence is informative about how aggressively banks were taking advantage of note issuance. Panel B plots the fraction of bonds on hand to loans and discounts. Panel C plots total Debt/GDP outstanding for the US and gives a sense of the total supply of government debt.

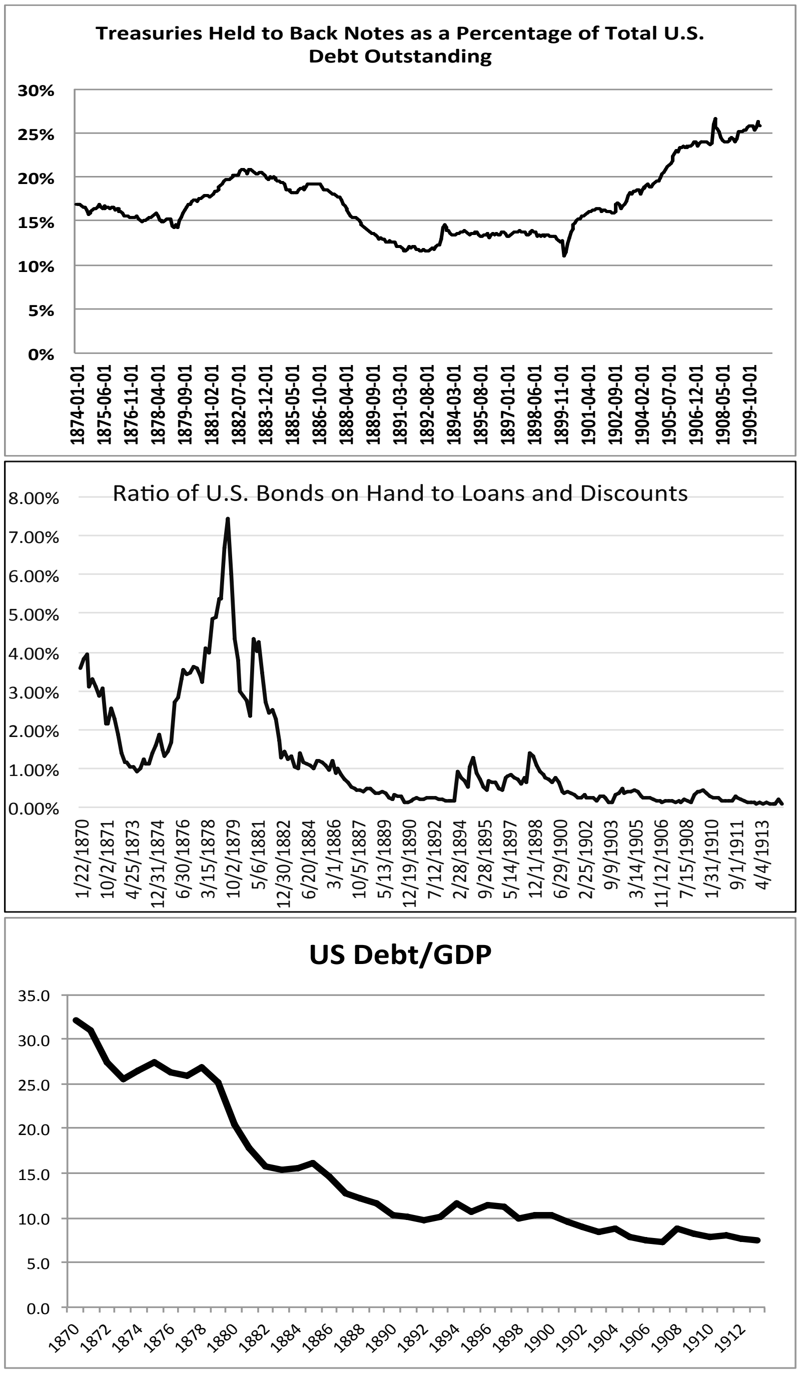


Figure 5: Profit series, plotted in standard (top) and log scale (bottom).

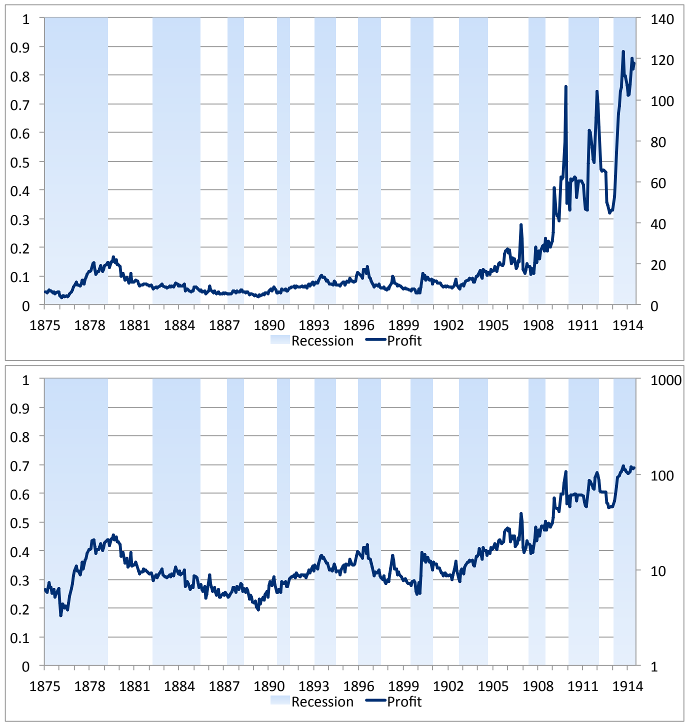
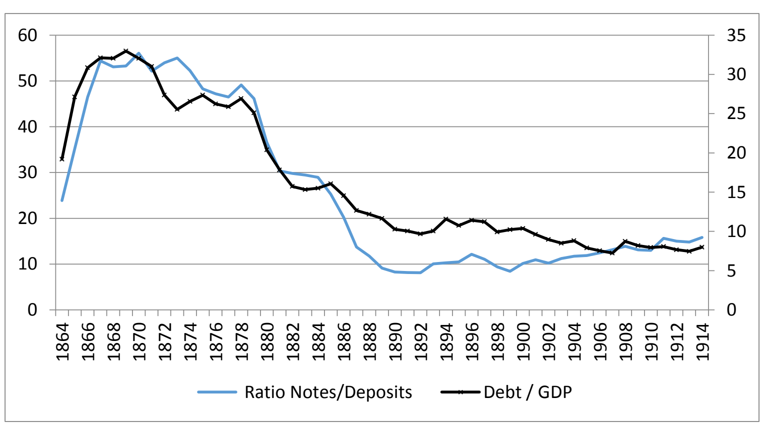


Figure 6: Ratio of notes to deposits. This figure plots the ratio of notes to deposits against US government debt to GDP. It shows that declines in government supply of Treasuries are strongly associated with increases in deposits.



Appendices

ONLINE APPENDIX

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:

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

	y =profit		$y = \ln$(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
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$