Volatile Times and Persistent Conceptual Errors:

U.S. Monetary Policy 1914-1951

Charles W. Calomiris

October 20, 2010

SECOND DRAFT: NOT FOR QUOTATION

"If stupidity got us into this mess, then why can't it get us out?"

– Will Rogers[1]

I. Introduction

This chapter reviews the history of the early (1914-1951) period of "monetary policy" under the Federal Reserve System (FRS), defined as policies designed to control the overall supply of liquidity in the financial system, as distinct from lender-of-last-resort policies directed toward the liquidity needs of particular financial institutions (which is treated by Bordo and Wheelock 2010 in another chapter of this volume). The history of monetary policy generally focuses on four key sets of questions: what did the monetary authority do, why did it behave the way it did, what effects did its policies have, and what should it have done differently?

In addressing these questions, the monetary history of 1914-1951 can be usefully broken down into three sub-periods – World War I, the interwar period, and World War II and its immediate aftermath. The most interesting monetary policy questions relate to the interwar period, where wartime constraints on monetary policy were absent. The monetary policy of the interwar period has been the subject of voluminous research, and substantial controversy, for decades, both with respect to the intentions and constraints that guided monetary policy during this period and with respect to the effects of monetary policy on the economy. Such studies include those by Chandler (1958, 1971), Friedman and Schwartz (1963), Wicker (1966, 1980, 1982, 1996), Brunner and Meltzer (1964, 1968a), Frost (1971), Meltzer (1976, 2003), Temin (1976), Bernanke (1983, 1995), Miron (1986), Wheelock (1990, 1991, 1992), Eichengreen (1992), Romer (1992), Calomiris and Wheelock (1998), Bordo, Choudri and Schwartz

[1] Cited in Ahamed (2010), p. 347.

(2002), Calomiris and Mason (2003a, 2003b), Hanes (2006), Hsieh and Romer (2006), Bernstein, Hughson and Weidenmier (2010), and Calomiris, Mason and Wheelock (2010).

It would not be possible in this chapter to do justice to these and other contributions to this vast literature, or to describe fully, much less lay to rest, the controversies contained in them. Happily, reviewing that literature in detail is also unnecessary, given the recent comprehensive and authoritative contribution by Meltzer (2003), whose first volume of the history of the Fed not only provides the definitive summary of the actions and intentions of policy makers during this period, but also summarizes, and arguably often resolves, the academic controversies that have surrounded those actions ever since.[2]

My review will focus primarily on a few of the most important and controversial issues surrounding the intentions and consequences of monetary policy during the interwar period, and it does so from the perspective of a larger, historical question: How do monetary policy makers and their critics learn about the proper approach to monetary policy? The obvious part of the answer to that question is that learning must happen over time, largely as a result of trial and error, and subsequent analysis of that trial-and-error policy process. The less obvious part of the answer is that the learning process is not uniform across times and places, and depends on particular historical circumstances, especially the initial conditions that define one's priors, and the specific sequence of shocks that one experiences and from which lessons are supposed to be derived. Historical circumstances were not very favorable for clear and speedy learning about U.S. monetary policy from 1914 to 1951, and that explains why so little seems to have been learned during these years. Indeed, we are still debating some of the fundamental

[2] For similar reasons, I will not review the literature on the transmission mechanism of monetary policy during the interwar period. This literature, especially the contributions of Fisher (1933), Bernanke (1983), Temin (1989), Bernanke and James (1991), are reviewed in Calomiris (1993). See also Calomiris and Hubbard (1989) and Calomiris and Mason (2003b).

questions about what went wrong with monetary policy during this period, what drove those errors, and what effects they had, as my review will show.

The remainder of this chapter is organized as follows: Section II begins with a brief background on the founding of the Fed, its mission, and structure. Section III reviews the perceived failure of the Fed during the Depression, and argues that the Fed's failure, and the protracted delay in learning from that failure reflected the initial conditions in which it was founded – which included the unique circumstances of the U.S. banking system, the guiding conceptual framework of the real bills doctrine, and the dramatic shocks and rapidly changing structure of the economy in its early years.

Section IV considers five key questions in the history of monetary policy during the interwar period: (1)To what extent did the Fed cause the stock market crash of 1929, or alternatively, was the stock market boom of 1928-1929 a source of instability that warranted Fed intervention to prick a ballooning bubble? (2) To what extent did the gold standard limit the Fed's ability to prevent monetary contraction during the Depression? (3) To what extent were the four banking panics identified by Friedman and Schwartz times of unique strain that should have awakened the Fed to the need to prevent a contraction of the money supply? (4) Was the economy in a liquidity trap in the early 1930s, or alternatively, would increased open market operations have prevented the Depression? (5) To what extent were the reserve requirement increases of 1936-1937 responsible for the recession of 1937-1938? Reviewing the changing answers to these questions that have evolved over time illustrates why it was so challenging for Fed officials, and even for subsequent observers of monetary policy, to properly gauge the role that monetary policy played in the Depression, and to identify the particular sources of policy errors. These questions, one can argue, do have answers, but that does not make them easy questions to answer, and the difficulty of answering such questions helps to explain the protracted process of learning about monetary policy that is the central theme of this chapter.

Section V reviews the literature on the Fed's role in smoothing seasonal fluctuations during the interwar period. Section VI describes the role of the Fed during World War II, and the reestablishment of monetary independence from the Treasury in the Accord of 1951. Section VII concludes.

II. Intentions of the Fed's Founders

The original mission of the Fed, as revealed by its structure, by the debates that gave rise to it, and by the statements of its leaders, was quite different from its current focus: "[s]table growth was not part of the Federal Reserve's formal mandate in the early years. Most of the System's leadership would have denied any responsibility for economic activity or employment" (Meltzer 2003, p. 9). The primary concerns that gave rise to the Fed were the so the so-called "inelasticity" of the supply of money and credit, and the peculiar U.S. propensity for banking panics. After 1866, the United States was the only economy in the world that continued to suffer from banking panics: major panics, defined as events in which the New York City Clearing House banks acted cooperatively to deal with liquidity risk occurred in 1873, 1884, 1890, 1893, 1896, and 1907 (Calomiris and Gorton 1991).

The peculiar propensity for panics in the United States reflected its unique "unit" banking structure: unlike other countries, banks in the United States historically (with the exception of the antebellum South, where bank branching was widespread, and a few states in the postbellum North, where branching was permitted on a limited basis) were constrained to operate in one location, and in the few states that permitted branching, branches were generally not allowed throughout the state, and in no case could banks operate branches across state lines. This unit banking structure made the U.S. financial system uniquely risk-prone in its response to real shocks; limitations on branching prevented inter-regional diversification of loan risks ex ante, and hampered the coordination of the banking

system's response to shocks ex post (Calomiris 2010). Other countries had grown out of banking panics, strictly defined, after 1866 (Bordo 1985), but in the United States, panics continued.[3]

Banking panics in the United States were clearly not random events. They occurred at cyclical peaks of economic activity and at seasonal peaks of credit demand in the fall harvesting or spring planting seasons (when the banking system was at its maximum leverage). Calomiris and Gorton (1991) show that, from a cyclical perspective, the panics of the national banking era were quite predictable based on a simple dual-threshold criterion: a banking panic occurred in a particular quarter, if and only if, during the preceding quarter both the liabilities of failed businesses rose by 50% or more (seasonally adjusted), and stock market returns fell by 8 percent or more.

In the minds of the founders of the Fed, the problems of inelasticity and banking panics were closely related. The Fed was founded in the belief that (1) an elastic supply of currency (in the form of reserves supplied by the Fed), would result in an elastic supply of bank credit, and that (2) this would reduce the incidence of banking panics by reducing the exposure of the banking system to liquidity risk. Spikes in the demand for liquidity (at seasonal or cyclical frequency), if not accommodated by increases in currency, produce momentary scrambles for liquidity that would raise interest rates, result in distress sales of assets, and potentially lead, in extreme circumstances, to banking panics. The Fed's job was to accommodate the swings in the demand for liquidity, at seasonal and cyclical frequencies, which would also prevent disruptive interruptions to the availability of credit.

[3] This is not to say that banking systems outside the United States avoided problems during recessions. Australia (in 1893), Argentina (in 1890), Italy (in 1893), and Norway (in 1900) suffered severe bank solvency crises, defined as episodes where the negative net worth of failed banks exceeded one percent of GDP (see Calomiris 2009). Furthermore, banking system shrinkage and distress during recessions, manifested in significant deposit and loan contraction and loan losses, gave rise to credit crunches even when it was not associated with a banking crisis. For a cross-country analysis of the macroeconomic effects of banking system credit contraction, see Bordo and Eichengreen (2003), who study the effects of historical banking distress (broadly defined) on business cycle severity. These episodes, however, are generally not properly regarded as true banking crises, as defined in Calomiris (2009) – that is, episodes either of widespread and sudden withdrawals of deposits due to a panic, or episodes in which losses from bank insolvencies were large as a fraction of GDP.

To help fix ideas, consider the following simple model, illustrated in Table 1 and Figure 1. Table 1 summarizes the balance sheet of a typical bank. Its assets consist of loans and reserves (vault cash plus balances at correspondent banks, and after the founding of the Fed, at Federal Reserve Banks); its liabilities consist of deposits, with the difference between assets and liabilities comprising the bank's equity capital, or net worth. Assume that banks target a low level of default risk on their deposits (for details, see Calomiris and Wilson 2004), which they maintain through a combination of a sufficiently high ratio of reserves to loans, and a sufficiently high ratio of equity to deposits.

In this simple framework, an adverse economic shock causes loan losses, which reduce the ratio of equity to assets. Banks adjust to this shock to restore the low risk of default on deposits by some combination of accumulating capital (which is hard to do at high frequency), and raising their ratio of cash to loans by not renewing some maturing loans (which is easier to do at high frequency).Banks charge for bearing the credit and liquidity risk associated with expanding loans. The short-term loan supply function (Figure 1, Loan Supply without Fed) is upward sloping because short-term increases in lending (holding constant capital and liquid assets) raises the exposure of the bank to both credit risk and liquidity risk. After the founding of the Fed, the new loan supply function is flatter (Figure 1, Loan Supply with Fed). The reason is that, in the presence of the Fed, banks have a new means of dealing with liquidity risk other than re-capitalization or loan liquidation; that third option is borrowing from the Fed against some of their loans. Banks still have to adjust their lending and capital over time to restore their low default risk, as needed, and the Fed does not bear credit significant default risk through discount window lending, but banks do not have to make adjustments suddenly out of fears of illiquidity (during a moment of high seasonal leverage in the banking system). *That means that, in the presence of the Fed, the banking system will avoid magnifying loan loss risk by creating a scramble for liquidating assets at a time of high leverage.* In the presence of the Fed, it is also true that variation in loan demand, seasonally or cyclically, should result in greater variation in the quantity of lending and lower variation in interest

rates, as shown in Figure 1, which is *drawn here assuming* that the primary source of shock in the system is *variation in loan demand* (an assumption which we shall evaluate in more detail below).

Although this model was not articulated in any explicit form by the Fed's founders, it is a reasonable representation of their view that the founding of the Fed would produce greater elasticity of reserves, which would result in greater elasticity of credit, which in turn would reduce the propensity for banking panics. This model also can explain why the United States was uniquely prone to banking panics. Nationwide branching would have produced greater diversification of loan risk (via the law of large numbers), which would have reduced the effect of loan portfolio shocks on banks' risk profiles. Furthermore, branching banks could coordinate better to borrow from each other in response to shocks because they were geographically coincident (see Calomiris and Schweikart 1991 and Calomiris 2000).

Some of the adherents to this view of banking system liquidity risk and the role of the Fed in mitigating it also lamented a closely related structural characteristic of the U.S. banking system that contributed to its liquidity risk: namely the "pyramiding" of reserves in New York City (the tendency for banks in rural areas to park their funds at other banks during the summer and winter seasonal lulls in credit demand). New York City banks reinvested those reserves in the call loan market, resulting in potential linkages between Wall Street securities risks and credit and liquidity risk to the nation's bank deposits. The banking panics of 1857, 1873, and 1907, and arguably others, had their origins in securities market problems concentrated initially in New York. According to those most concerned about this practice (especially, Carter Glass, who chaired the House banking committee that passed the Federal Reserve Act, and later oversaw the reform of banking and Fed practices as a Senator in the 1930s), the Fed would bring an end to this destabilizing tendency by replacing the interbank deposit market based in New York with a decentralized system of deposits by member banks in each Federal Reserve District at their reserve banks.

As envisioned by Glass and other Fed founders, the Federal Reserve Banks would be a repository of excess reserves during times of low demand for their member banks, and a source of additional reserves (via either lending to members or buying assets from them) during high-demand periods. To prevent accommodation of destabilizing increases in demand for reserves (e.g., the fueling of undesirable speculation in stock markets or real estate), the Federal Reserve Banks would focus their activities only on purchasing or lending against "real bills," defined as commercial loans related to the financing of trade. At the Fed's founding, real bills were expected to be the key asset that would pass between member banks and their Reserve Banks.

After an initial struggle for control over the Federal Reserve System between the Reserve Banks and the Board in Washington (Meltzer 2003, pp. 75-82), the twelve Federal Reserve Banks succeeded in maintaining a decentralized system until 1935. The Reserve Banks coordinated their policy decisions through a committee, headed by the New York Fed, but each was free to opt out of any coordinated intervention decided by that committee. That arrangement preserved the decentralized structure originally envisioned by the Fed's founders.[4] The original decentralized structure of the Fed was designed to facilitate a close relationship between the Reserve Banks in each District and their local members, who collectively owned them, and was designed to prevent any capture of the system by Wall Street speculators or Washington politicians.

The early Federal Reserve System's operational structure and behavior reflected its founders' intentions regarding the creation of an elastic supply of reserves, adherence to the real bills doctrine, and the decentralization of decision making, albeit imperfectly. The Fed failed to achieve some

[4] To limit concentration of power, the Fed was divided into twelve quasi-autonomous Districts. But the dollar was a national currency, the banking system was connected through the interbank reserve holdings and the interregional clearing of checks, and the securities markets and real economy were becoming increasingly integrated nationally. The increasing integration of the national economy favored greater coordination of policy. At the same time, sharply divergent regional shocks suggested benefits from preserving District autonomy. The severity of the Great Depression, and the diagnosis that decentralization had contributed to insufficient open market operations in 1931 because of internal dissent (Meltzer 2003, 470-486), brought an end to the era of District autonomy in 1935.

founders' goals. Most obviously, banking system reserve pyramiding in New York City continued after the founding of the Fed, owing to a central flaw in the design of the Fed: Reserve Banks, unlike New York City banks, did not pay interest on interbank deposits; that gave peripheral banks a strong incentive to deposit their reserves in correspondent banks rather than Federal Reserve Banks (White 1983).

Also, contrary to its founders' vision, the politicization of the system by Washington began almost immediately after the Fed's founding. Under the pressures of World War I's financial challenges, the Fed became an important partner in assisting the U.S. government to market its debts. In 1917, reserve requirements were reduced to permit expanded credit to finance the war (Meltzer 2003, p. 79, footnote 31). And collateral rules for Federal Reserve note issues were relaxed in 1917: the total amount of collateral was reduced, and perhaps more importantly, promissory notes of member banks secured by government bonds could be used as collateral for the notes (Meltzer 2003, p. 89). At the end of World War I, in the interest of boosting demand for outstanding Treasury debts, the Fed also reduced its discount rate for loans collateralized by Treasury securities. This had long-term effects. The discount rate reduction led the Fed to abandon its "penalty" rate policy for targeting the discount rate, which had been one of its core founding principles (Meltzer 2003, pp. 73, 86). This change subverted the Fed founders' intent that the Fed would use a penalty discount rate as its primary tool of managing the cyclical and seasonal availability of credit in the money market. More broadly, the World War I precedent of making the Fed subservient to the interests of marketing Treasury debt not only produced the short-term inflationary binge of 1917-1920 (Meltzer 2003, pp. 90-107), it also set the stage for subsequent changes that eventually made the Federal Reserve a fiscal instrument of the U.S. Treasury. Those changes include the 1932 Glass-Steagall Act, a temporary measure later made permanent, which permitted the use of Treasury securities as collateral for Federal Reserve note issues (Meltzer 2003, pp. 358, 417-418), and the various changes in Treasury monetary powers (discussed below) after 1933,

which gave the Treasury effective control over monetary policy, which it used to target the yields on government debts until the Treasury-Fed Accord of 1951.

III. Initial Conditions

As the Fed began operations in 1914, consider what it had to go on. First, there was no established view of the proper approach to monetary policy, of the precise mission of the Fed, or of how the Fed might accomplish its ill-defined mission. If the point of having a Fed was to increase the elasticity of credit, as distinct from the elasticity of the narrowly defined currency supply (which was the clear purpose of the Aldrich-Vreeland Act of 1908), how would the Fed do so? Was the Fed supposed to focus mainly on seasonal or cyclical elasticity? What sorts of market indicators were likely to give the most reliable signals of loose or tight conditions in the money and credit markets? What were the most effective tools to use to react to those signals, and how much reaction was called for under what circumstances? How would the Fed reconcile its seasonal or cyclical policy objectives with its adherence to the gold standard? Most of these questions had been the subjects of bitter debates leading up to the founding of the Fed, and the debates were not resolved prior to the Fed's founding.

There was no monetary authority track record from the past to guide the new central bank. Furthermore, potentially relevant data were absent, and even some of the simplest conceptual ideas that would be taken for granted today about how to define interesting concepts in order to collect data about them had yet to evolve (e.g., monetary aggregates, according to Friedman and Schwartz 1963, p. 628, are first mentioned by the Federal Reserve in 1948, and are not discussed seriously as a potential criterion for targeting monetary policy until 1952).[5] And, the founding of the Fed was motivated by

[5] For a brief review of the development of theories of central banking, see Meltzer (2003), chapter 2. This review correctly points out that there were early precedents (notably the early 19th c. work of Henry Thornton) to what we would now regards as the "correct" broad view of the role of central banks in stabilizing the economy by targeting the overall supply of liquidity, but these views were not in the mainstream of thinking, inside or outside the United States at the time of the founding of the Fed. The theory of central banking, as it evolved in the 19th and

problems that were recognized at the time as *unique* to the United States' financial system; despite the many volumes of excellent studies of other countries' experiences produced by the National Monetary Commission in 1910, the unique structure of the U.S. financial system (a geographically fragmented unit banking system serving the needs of a vast continent), substantially limited the ability of U.S. policy makers to learn from, say, the Bank of England or any other preexisting central banks.

Thus, the goals, relevant monetary concepts, data, and understanding of policy instruments of monetary policy had to evolve "in real time" alongside policy decisions. U.S. monetary policy makers had no choice but to improvise in the meantime. Improvising meant relying on their preexisting beliefs, and based on those beliefs, arguing with one another about the appropriate goals and indicators on which to focus, and on the appropriate reactions to market indicators.

The dominant monetary policy doctrine at the Fed from the beginning was the "real bills doctrine" – a view of monetary policy that saw trade credit as the desirable focus of commercial bank lending, and that viewed the main role of monetary policy as accommodating shifts in the demand for trade credit. Because trade credit was perceived as closely linked to the real needs of business, and in that respect, distinct from the speculative whims of stock market or real estate speculation, changes in the demand for trade credit were perceived as a more reliable indicator of legitimate shifts in credit demand that were worth accommodating.

By 1923, the Fed had developed a coherent toolkit to implement the real bills doctrine, which was described in its tenth annual report, and which Meltzer (2003, pp. 161-165) refers to as the Riefler-Burgess doctrine. According to that view, the Fed should respond to shifts in the domestic demand for credit by easing through a combination of open market purchases and discount rate reductions

---

early 20th century, focused on two key themes: (1) appropriate mechanisms for ensuring that central banks helped to maintain the long-run commitment to a specie standard, and (2) constructing an appropriate relationship between the central bank and other financial intermediaries to limit liquidity risk and avoid panics.

(typically, implemented together, with open market operations leading subsequent discount window changes) when borrowed reserves were high and nominal interest rates were high.

If it followed the real bills doctrine, the Fed believed that increases in the money supply would be "self liquidating" and, therefore, would pose no risk of fueling speculative bubbles or inflation, nor any threat to the maintenance of the gold standard, so long as the Fed also responded to international shocks in the gold market by tightening credit conditions in response to outflows of gold and loosening in response to gold inflows (the so-called "rules of the game" of the gold standard).[6] The real bills doctrine was an appealing ideology because it offered policy makers a rule that promised to provide short-term flexibility (or "elasticity" of currency supply) but did not disrupt the long-term discipline of adherence to gold.[7] Wicker (1966), Brunner and Meltzer (1964, 1968a), Wheelock (1990, 1991, 1992), and Meltzer (2003, chapter 5) show that – with a few exceptions – the Fed consistently adhered to that real bills approach in the 1920s and the 1930s.

In particular, contrary to Fisher's (1935)[8] and Friedman and Schwartz's (1963) view that Benjamin Strong's death caused the Fed to adopt a new and unwise monetary policy targeting regime after 1928, subsequent econometric work by Wheelock (1900, 1991, 1992) has shown that the same policy reaction function describes Fed actions consistently in the 1920s and 1930s. An analysis of Fed officials' statements by Meltzer (2003) confirms that Fed officials intended to do so. Under Strong's

[6] When the two sets of signals conflicted with one another – especially when gold flows were signaling the need for an expansion of the money supply, as in 1930, when gold flowed in and real bills declined – the Fed favored the real bills doctrine. In 1930, that meant monetary contraction, contrary to the expansion that adherence to the gold standard's rules of the games would have implied (Meltzer 2003, pp. 401-402).

[7] As one Fed official put it: "Probably the most important effect of the Federal Reserve Act was to set up the machinery necessary to provide elastic currency; elastic in that it would be based on self-liquidating credit instruments arising out of the production and distribution of commodities. An obligation of the United States does not represent a transaction of this character…to the extent such obligations back the currency such currency is fiat currency" (statement by John U. Calkins at the Federal Reserve Governors Conference, May 1922, cited in Meltzer 2003, p. 70). This view was widely shared by some of the most prominent monetary economists of the time, including A. Piatt Andrew, H. Parker Willis, J. Laurence Laughlin, and Horace White. For more details, see Mints (1945).

[8] Wheelock (1991, p. 2) cites Fisher's 1935 House Testimony: "…this depression was almost wholly preventable, and …it would have been prevented if Governor Strong had lived."

leadership, the Fed had followed that reaction function, and Strong's less powerful but supposedly enlightened successor at the New York Fed, George Harrison, tended to advocate expansion or contraction in a manner fully consistent with real bills thinking (Meltzer 2003, p. 274). The supposed exceptions to real bills targeting – Strong's successful advocacy of expansionary policy in response to international events in 1924 and 1927 – were not really exceptions, since both occurred at times of high borrowed reserves (when real bills targeting agreed that expansionary policy was called for). Furthermore, as Meltzer (2003, p. 274) shows, the monetary expansion in 1927 was viewed in 1928 by most of the Fed leadership as having been too expansionary, because it had occurred without commensurate growth in real bills; Meltzer argues that even if Strong had lived, he likely would have overseen a tightening of policy in reaction to what was widely perceived in 1928 as an error. Finally, Meltzer (2003, pp. 289-323, 409-410) shows that there is little to the view that there were deep divisions within the Fed or paralysis in the aftermath of Strong's untimely death; most of the disagreements that occurred about policy in 1929 were procedural rather than substantive. The key error of seeing monetary policy as appropriately "easy" despite the worsening contraction in 1930-1932 reflected a common misperception by virtually all Fed officials, which reflected their adherence to the Riefler-Burgess doctrine and their failure to distinguish between nominal and real interest rates.

In the wake of the failure of the Bank of United States in December 1930, and at other times of high bank failure, there was little impetus within the Fed to offset the resulting declines in the supplies of money and credit. As banking system risk increased, and banks responded to the risk with higher reserve demand, Fed officials, including Harrison, followed the Riefler-Burgess doctrine and interpreted rising excess reserves and low borrowings as indicating a lack of demand for money (Meltzer 2003, p. 326-334). As banking failures increased in 1930 and subsequently, Fed officials were aware of failures, but did not see monetary policy as an instrument to be used to prevent bank or commercial failures, which were part of the natural process of maintaining market discipline. Although Fed officials differed

from time to time on the precise levels of purchases or sales of securities, they generally agreed about the direction of policy.

Real bills thinking, as depicted in Figure 1, underlay the primary errors of the Fed in the 1930s: (1) ignoring supply-side shocks in the loan market (which also appeared in the 1930s as shocks to the demand for reserves), and (2) failing to distinguish between nominal and real interest rates.

With respect to the first of these errors, the Fed failed to consider the shift in the demand for reserves by banks in response to recession-induced losses. As banks experienced losses in their loan portfolios and equity capital-to-assets ratio resulting from economic decline, and facing the discipline of the deposit market, they had to either restore lost capital (with either new offerings or increased retentions) or reduce asset risk by increasing the ratio of cash assets relative to loans (Calomiris and Wilson 2004 model that choice). Virtually no banks chose to raise new capital during the 1930s. They chose instead to substantially reduce the ratio of loans to cash assets, and also to cut dividends to boost capital. This inward shift in loan supply is depicted in Figure 2.

As banks were cutting lending, the effect on real interest rates in the loan market was positive. Of course, the recession also caused declines in loan demand, which somewhat offset that effect. As Meltzer (2003) repeatedly emphasizes, however, the Fed did not focus on the real interest rate, which was rising during the early 1930s, in response to increasing deflation; instead, it took the low nominal interest rates in the market as an indication of declining interest rates, which it saw as an indication of a plentiful supply of loans. The result was a sharp contraction in money and credit, which was caused by a contraction in *loan supply*, which substantially aggravated the macroeconomic consequences of deflation (Bernanke 1983, Calomiris and Mason 2003b, Calomiris and Wilson 2004). According to Calomiris and Mason (2003b), after controlling for the effects of the general decline in the supply of money nationally, a contraction in loan supply of one percent (in the state-level cross-section) resulted

in an incremental decline of roughly half a percent in income during the early 1930s. Thus, the combination of adherence to the real bills/Riefler-Burgess doctrine and interest rate money illusion were fundamental conceptual errors that permitted the Fed to oversee the sharp contraction in the supplies of money and credit during the Depression.

The real bills doctrine, and the focus on borrowed reserves and nominal interest rates as reliable signals of market tightness and looseness to which the Fed should respond, have no adherents today. It is now widely recognized that following the Riefler-Burgess framework will tend to aggravate business cycles by expanding the money supply at times of exogenous expansion in aggregate demand. It is also understood today that low nominal interest rates and low levels of borrowed reserves are not reliable indicators of loose money. But that does not mean that Fed officials were stupid. They engaged, unavoidably, in a process of ongoing trial-and-error experimentation, beginning with some simple rules of thumb that seemed to offer a means of balancing short-term flexibility with long-term discipline.

Over time, retrospectively, the Fed would have to draw inferences about past actions and their consequences, and incorporate lessons to guide policy more scientifically. The Fed would have to learn that the real bills doctrine created a pro-cyclical bias in monetary policy (because it was accommodating aggregate demand shocks with increases in money supply). It would have to learn that decreases in borrowed reserves (at a time when banks were scrambling for cash because of system-wide increases in banking risk) could signal tightening credit conditions (as in the early 1930s), or that low nominal interest rates (again, in the early 1930s) could be consistent with high borrowing costs and tight credit conditions (in a deflationary environment). One might have expected that this process would take years.

Why Learning Was So Slow

That expectation would have turned out to be wrong. Despite the huge costs of these conceptual errors in the 1930s, learning did not take years, but rather, several decades (that is, these

lessons were generally, if not universally, understood only by the 1960s). To economists reading this essay in 2010, perhaps the most remarkable single fact to note about monetary policy at the end of the interwar period is that its architects were, for the most part, quite pleased with themselves. Far from learning about the errors of their ways during the interwar period, Fed officials congratulated themselves on having adhered to appropriate principles, and to the extent that they were self-critical, it was because they thought that they had been too expansionary (Meltzer 2003, 410-413).

In part, Fed officials absolved themselves of blame because they believed that monetary policy was of limited use in combating recessions. Monetary policy could only exert proper influence through its effect on bank lending, but Fed officials (who saw low borrowed reserves, high excess reserves, and low nominal interest rates as indicators of loose credit) believed that the economy was in a liquidity trap in the early 1930s. As one official put it, monetary policy did not stop the Depression because "you must have borrowers who are willing and able to borrow" (Meltzer 2003, p. 478). The Fed had not expanded its open market operations because it did not see the point of doing so; more liquidity would not have made a difference. As we shall see below, that view was mistaken, and reflected two errors: (1) the Riefler-Burgess doctrine's approach to measuring the tightness of credit markets, and (2) interpreting rising reserve ratios in the 1930s as indicative of a passive willingness to accumulate reserves by banks, rather than as an active scramble for liquidity.

With some significant exceptions, central bankers from outside the United States, and independent monetary experts – including the famed banking crisis historian Oliver Sprague, who authored (Sprague 1910) one of the most influential reviews of the pre-Fed experience with banking panics as part of the work of the National Monetary Commission – shared the view that the Depression

had not be caused by monetary policy errors (Meltzer 2003, pp. 277-282).[9] Congressional leaders and President Roosevelt (who asked the preeminent advocate of the real bills doctrine in Congress, Senator Carter Glass, to craft regulatory reforms in 1933) seem to have agreed, since the regulatory changes affecting monetary policy during the 1930s were designed in large part to assist Fed monetary policy by making the real bills doctrine work better, based on the view that bank lending for non-real bills purposes had been a contributor to excessive financial expansion (Meltzer 2003, pp. 429-434, Calomiris 2010). This lack of learning had consequences; once the Fed had restored itself to a position of substantial independence from the Treasury after the 1951 Accord, it followed a policy reaction function that was remarkably similar to what it had done in the 1920s and early 1930s (Friedman and Schwartz 1963, pp. 614-32, Brunner and Meltzer 1964, 1968a), and consequently, the Fed made similar "pro-cyclical" errors to those it had made before.

What explains this lack of learning during the interwar period? It is difficult to answer such a question convincingly, but one thing is for sure: The period 1914-1947 was an unusually unstable three decades in U.S. history. The sequence of large and unique shocks, along with the changing economic structure of the economy, was not conducive to learning about monetary policy; the possibility of meaningful central bank learning about the dynamic structures of the financial and real sectors and their inter-linkages requires a minimum of stability in the basic economic processes that form the backdrop for monetary policy. As Wicker (1966) emphasized, the uniqueness of the events of the interwar period made it difficult for policymakers to react to them properly, especially given the inadequacies and pro-cyclical bias of the policy doctrines the Fed started with.

[9] The most notable exception was Lauchlin Currie (1934), who rejected the real bills doctrine and enjoined the Fed to control the quantity of money, arguing that "there exists no valid theoretical justification for the Commercial Loan Theory of Banking" and he described the Fed's role during the Depression not as a failure of activist policy, but a policy that failed because it "was one of almost complete passivity and quiescence" (cited in Meltzer 2003, p. 474).

Not only was the economy buffeted by large and unique shocks; its structure was also changing over time. No sooner had the Fed been founded, but World War I began. The Fed had been constructed to operate under the international adherence to the gold standard, but World War I saw a collapse of that gold standard. Domestically, World War I saw unprecedented increase in government control of the economy (Rockoff 1998), making it difficult for the nascent central bank to learn about the normal dynamic structure of the financial system and its interactions with the real economy.

The 1920s were also an era of significant structural change and large shocks. The 1920s saw a dramatic mixture of divergent regional and sectoral trends within the United States. Some agricultural areas, for example, those specializing in grains and cotton, were hard hit after World War I by declines in commodity prices. This produced the worst rural mortgage foreclosure outcomes and the highest rural bank failure rates up to that point since the 1830s. Indeed, the high bank failures of the early 1930s are best seen as a continuation of the rural bank failures that had begun in the 1920s (White 1984, Calomiris 1990, 1992, Alston, Grove and Wheelock 1994).

At the same time, however, other sectors of the economy thrived. Not only did much of industry expand, but its structure was changing dramatically. Industrial production was undergoing something of a revolution. New industries and firms based on new products and new production methods were ushering in a new era of dynamic technological progress, which was reflected in the the rapid appreciation of the stock market in the late 1920s (Nicholas 2007, Kabiri 2009). The number of patents was unusually high in the 1920s, but more importantly, of the 19,948 patents granted to firms between 1920 and 1929, an unusually high percentage of them (21%) are cited in patents granted between 1976 and 2002 (Nicholas 1927). Field (2003) shows that the progress of the 1920s was so powerful that it continued into the 1930s, despite the inhospitable macroeconomic environment. The period 1929-1941 saw the “fastest rate of multifactor productivity growth over the last century and a half, and probably

two centuries" (p. 1406), making these years "the most technologically progressive of any comparable period in U.S. economic history" (p. 1399). Mowery and Rosenberg (2000) find that the employment of research scientists and engineers grew by 72.9 percent from 1929 to 1933. Bresnehan and Raff (1991, 1992) examine technological change in the automobile industry, and argue that the Depression hastened the "shake out" of relatively backward firms and the consolidation of market share in technological leaders.

Regional integration through increased communication and transport linkages was accelerating at a rapid rate (automobiles, radios and motion pictures became widespread, and aviation began to spread). In 1919 there were roughly 7 million automobiles in the United States; by 1929, there were more than 23 million. Electricity became widespread in the 1920s, even in rural areas. Annual radio sales grew from $60 million in 1922 to $843 million in 1929.

Integration of communication and transportation affected the nature of consumers' behavior in the 1920s, too, in ways that may have altered aggregate consumption dynamics and financial markets. Electrical appliance use became popular, as did consumer credit for durables purchases. Canned and other prepackaged food purchases grew dramatically. National consumer fads and dramatic news coverage of controversial events (e.g., the Scopes trial in 1925), dramatic increases in attendance at major sporting events, the publication of national magazines with large circulations, and the birth of a national advertising industry all reflected fundamental changes in the 1920s (Allen 1931). Brokerage houses developed the first true network for retail stock investing in the mid-1920s (Ferderer 2007), and consumers' investments in stock became important to an unprecedented degree. Stock prices, especially for high-growth firms and for the large New York banks, grew dramatically during the 1920s, and prompted a growing controversy over whether a bubble had begun in these markets, and whether and how the central bank should respond to it (more on this below).

Banking underwent dramatic changes in the 1920s, too. Rural bank failures prompted widespread bank consolidation and a relaxation of branching restrictions. Many states relaxed their branching laws in the 1920s, and the McFadden Act of 1927 allowed national banks to branch in states that permitted state-chartered banks to branch (Calomiris 2000, chapter 1). From 1920 to 1930, the number of banks operating branches and the number of branches increased from 530 branching banks with 1,281 branches to 751 branching banks with 3,522 branches (Calomiris 2000, p. 57). As banks grew in scale they also grew in scope. U.S. banks entered asset management (trust activities) aggressively during this period. They also participated as underwriters in the growth of the securities markets. And money center banks expanded their operations abroad. The first era of true universal banking in the United States began in the 1920s.

Construction underwent changes, too. Cities saw the spread of skyscrapers of ever-expanding ambition. The first real estate boom and bust in Florida, amid an awakening appreciation of the unique opportunities for development there, occurred from roughly 1920-1926. The population of Miami (only one of the many new resort destinations) rose from 30,000 in 1920 to 75,000 by 1925. A national housing market boom and bust, ending in 1929, was of similar magnitude to the 2000-2007 boom and bust, although the price decline did not have the same impact on the financial sector, owing to the low levels of mortgage leverage in the 1920s (White 2010, Nicholas and Scherbina 2010).

The 1930s, of course, brought even more volatility and shifts in economic and political circumstances than the 1920s. Government policy was a source of substantial volatility, and government grew substantially in scale and scope (Rockoff 1998, Wallis 1998). Many of those shifts affected the structure of the economy, and some of them took the form of institutional changes in the structure of the monetary system and the organizational structure of the Fed.

New Deal industrial policies initially had important microeconomic consequences through the NIRA's attempts to fix prices and wages. Rising unionization and collective bargaining, culminating in the Wagner Act of 1935, altered the behavior of wages, and the macroeconomic consequences of large-scale strikes (like those in the automobile industry in 1937) posed new questions for how monetary policy should react to such shocks.

With respect to banking, failure rates in 1930-1933 were far higher than those of the 1920s, and because failures were concentrated in smaller banks, they produced a significant change in the size distribution of surviving banks. The 1920s were a turbulent decade for agricultural states. The 5,712 banks that failed during the years 1921-1929 had total deposits equal to $1.6 billion at the time of their failure, constituting 3.1% of average total deposits in the banking system from 1921-1929. Losses to depositors for the period 1921-1929 amounted to $565 million, which was 1% of average deposits during the period 1921-1929 and 0.6% of average annual GNP.[10] By way of comparison, during the period 1873-1913, no year had seen losses to depositors in excess of 0.1% of GNP.

Bank failures accelerated in the early 1930s. The number of banks fell 39% from 24,633 in December 1929 to 15,015 in December 1933. The 9,096 banks that failed during the years 1930-1933 tended to be small banks. Failed banks, as defined by the Federal Reserve (1943), represented 37% of the banks in existence at the end of 1929, but the deposits of those failed banks (at their dates of failure) were only 14% of the average level of bank deposits over the years 1930-1933, and losses borne by depositors in failed banks were roughly $1.3 billion, representing 2.7% of the average amount of deposits in the banking system for the years 1930-1933, and 2% of average annual GNP for 1930-1933.

[10] Deposits and failures are from the Federal Reserve Board's data in Banking and Monetary Statistics: 1914-1941 (1943), using suspensions as the measures of failures. Nominal GNP is from the U.S. Department of Commerce's Historical Statistics of the United States, Vol. I (1970).

Consolidation was the twin of bank fragility during the 1920s and 1930s. Thousands of banks were absorbed during the 1920s and early 1930s, accounting for a large share of bank assets, and this process was facilitated by regulatory reforms that expanded permissible bank branching (White 1985, Calomiris 2000, chapter 1).

The 1933 Banking Act, among other things, brought an end to the 1920s merger wave in banking by explicitly limiting consolidation, and by subsidizing small risky banks, who were the primary beneficiaries of the new federal deposit insurance created under the Act. The Act also set demand deposit interest rates at zero, as a means of trying to shrink the interbank deposit market, and restructured the banking system by separating commercial and investment banking. The Reconstruction Finance Corporation, which had begun its lending operations to banks and other firms in 1932, was reformed in 1933 to permit preferred stock investments in banks and other firms, which had important consequences for the banking system (Mason 2001, Calomiris and Mason 2004). All of these changes had potentially large and uncertain effects on the structure of the financial system and the riskiness of financial institutions.

The macroeconomic rules of the game were also in flux. The international return to the gold standard that had been happening during the 1920s came apart in 1931, as Britain and many other countries, but not the United States, left the gold standard. The Fed was enjoined to maintain a commitment to maintaining the gold standard (prior to 1933), and also facilitating the domestic needs of trade. These sometimes conflicted. Furthermore, effective interventions in international currency markets by central banks can require coordinated actions, which further complicated the Fed's responsibilities. In 1933, President Roosevelt left the gold standard and ushered in a new era of gold price targeting, which had dramatic consequences, in particular, by encouraging large gold inflows, which were the main component responsible for expanding the monetary base from 1933 to 1936.

Major fiscal policy changes, especially the large tax increases in 1936 and 1937, also contributed greatly to macroeconomic volatility (Romer 1992, Calomiris and Hubbard 1995, and references therein).

The post-1933 period also changed the structure of monetary policy. Most importantly, it saw substantial increases in the monetary powers of the U.S. Treasury. This occurred via its control of the exchange rate, the control over the sterilization of gold inflows, the Exchange Stabiliziation Fund, established in 1934, and silver-related authority, all of which gave the Treasury the ability to grow the money supply. Because the Treasury wanted to grow the money supply, and because the Fed's balance sheet was too small to stop that growth, even if it had tried, the Fed had no effective tools to use to oppose Treasury expansion (Calomiris and Wheelock 1998). The period of Fed policy from 1933 to 1940 is best captured by Meltzer's (2003) chapter title, "in the back seat."

The back seat also was re-upholstered. The Banking Act of 1935 changed the structure of the Federal Reserve System and its powers, as discussed further below. No sooner had the Fed begun to adjust to those new rules and structure than mobilization for World War II began, with its sectoral shifts, price controls and rationing. Monetary policy did not have to adapt to those changes, since it practically ceased to exist, as the Fed became an instrument for pegging low, constant nominal interest rates on national debt, in support of the Treasury's war financing effort.

In many respects, the America that we have come to know (a technologically advanced, highly mobile society, with nationally integrated financial, product, and factor markets, and eventually, nationwide banks) had its beginnings during the first three decades of the founding of the Fed, but a full understanding of those shifts and their implications for monetary policy could not have been known at the time. Even a central bank equipped with the tools and experience of a 2010 central banker would have found this a challenging environment in which to direct monetary policy. For example, on a

forward-looking basis, it would have been difficult to say what rate of monetary growth, or level of interest rates, was likely to produce price stability in this economy.

Furthermore, unstable times like World War I, the 1920s, and the 1930s, made it hard to identify the contribution of mistaken Fed doctrines and practices (e.g., "real bills" thinking) or the deleterious effects of the Treasury's monetary policies (e.g., the December 1936 decision to sterilize gold flows, which helped to precipitate the recession of 1937-1938, as discussed below). After all, it is difficult to identify monetary policy errors when the effects of other shocks are conflated with monetary policy in their effects on output and employment.

In light of those shocks, perhaps it is not surprising that the Fed found it difficult to identify accurately the negative role that it had played in the economy, or derive lessons from those errors about more appropriate targets and instruments of monetary policy. Indeed, monetary historians and economists are *still* arguing over the nature of monetary policy errors during the interwar period, and are trying to disentangle the effects of monetary policy on the economy from a myriad of other influences.

Friedman and Schwartz (1963) saw the Depression of 1929-1933 and the recession of 1937-1938 largely as consequences of Fed errors that could and should have been avoided. The Friedman-Schwartz narrative of the Great Depression has not always been embraced by others. Temin (1976) argued that monetary policy had largely played a passive role during the Depression. Eichengreen (1992) argued that the Fed was constrained prior to 1932 in its ability to expand the money supply by the gold standard and the rules of the Fed's charter governing its holdings of gold, and others (Bordo, Chaudri and Schwartz 2002, Meltzer 2003, Hsieh and Romer 2006) have disputed Eichengreen's contention. Wicker (1966, 1980, 1996) and Calomiris and Mason (2003a) contested the Friedman and Schwartz view that a national banking panic had occurred in 1930 and early 1931; those alleged panics were a crucial

part of the Friedman-Schwartz argument that the Fed had failed to bolster the money supply at a time when it should have foreseen a panic-induced contraction in the money multiplier. Frost (1971), Calomiris and Wheelock (1998), Hanes (2006), and Calomiris, Mason, and Wheelock (2010) dispute the Friedman-Schwartz account of the role of reserve requirement changes in 1936-1937 in precipitating the recession of 1937-1838.

Overall, recent research has reinforced many aspects of the Friedman-Schwartz view that failures of monetary policy were central to the economic troubles of the 1930s, but researchers disagree about the relative importance of various monetary policy shocks and errors Friedman and Schwartz identify, about the extent to which the Fed, as opposed to the Treasury, was responsible for monetary policy errors in 1936-1937, and about the ideological origins of avoidable mistakes in monetary policy. I review these arguments in more detail in Section IV, which offers a synthesis of the literature on some of these key issues, but I mention them here as a reminder: it has not been easy for monetary economists and historians to identify clearly the nature of monetary policy errors of the interwar period and their effects on the economy, even decades after they allegedly occurred. If it has taken economic historians decades to identify the Fed's contribution to the macroeconomic performance of the Great Depression, perhaps it is not surprising that the Fed did not learn immediately from its alleged mistakes.

Those controversies about the cyclical effects of monetary policy during the interwar period contrast sharply with the literature on the Fed's great success story of the interwar period: the reduction of *seasonal* volatility in the financial system (reviewed in Section V). Unlike the literature on the Fed's contribution to cyclical volatility, the seasonal volatility literature has been remarkably uniform in its conclusions. From its founding, a key part of the Fed's mission (stated clearly by its founders) was to increase the seasonal elasticity of the supply of reserves, to facilitate the seasonal smoothing of the supply of credit. The Fed was given the tools to do so, understood that mission, and seems to have

accomplished it reasonably well (Miron 1986, Bernstein, Hughson and Weidenmier 2009). It is interesting to ask why seasonal smoothing was so much easier than cyclical smoothing, and we shall return to that topic in Section V below.

IV. Identifying Policy Errors Is Challenging, Even with the Benefit of Hindsight

Here I will consider five key questions in the history of monetary policy during the interwar period listed in the Introduction. These questions relate to some of the most egregious alleged errors of the Fed. Our understanding of these errors has evolved over time, and continues to evolve, which is the overarching point of this review. Academics continue to reasonably disagree about the extent to which the Fed can be faulted for the stock market crash of 1929, the failure to sufficiently expand the supply of money in 1930-1933, the extent to which the banking crises of the 1930s were "panics" (which relates to the question of how egregious was the Fed's failure to respond to them), the extent to which monetary policy was limited by a liquidity trap in the 1930s, and the extent to which the reserve requirement increases imposed by the Fed in 1936-1937 helped to precipitate the recession of 1937-1938. The durability of these academic debates itself shows how difficult it was for the Fed to avoid errors in real time, especially given that it was hobbled by conceptual failings (money illusion and real bills thinking). The slow progress toward achieving a consensus on these difficult issues helps to explain the protracted process of learning about monetary policy after the Depression.

The Stock Market Boom and Bust of 1928-1929

From the beginning of 1927 to October 1929, U.S. stocks more than doubled in value (Board of Governors 1943). That impressive average performance masked important differences across types of stocks. High growth stocks grew much faster than average. These consisted of new technology firms like General Motors, RCA and General Electric, and companies in high-growth sectors, like electricity distributers and money center banks. New York City banks were experiencing a boom in revenues as the

result of new services and a growing customer base (trust management, and securities underwriting) and the establishment of global networks.

As discussed above, the United States was experiencing an unparalleled technology boom in the 1920s, one in which many new product and process technologies were developed that would have lasting significance (Field 2003). Nicholas (2007) shows that the number of patents with lasting applications and large numbers of subsequent citations was particularly high during the 1920s. He also shows that the number of patents (weighted by their future citations) is a powerful explanatory variable for the cross-section of stock appreciation during the 1927-1929 boom. The implied value of patents rose over the last years of the 1920s, reflecting the rising perceived real options associated with new technology.

For some sophisticated observers – notably Charles Dice (1929) and Irving Fisher (1930) – productivity growth and other positive long-run trends justified the boom in stock prices. Kabiri (2009) analyses the valuation models that were used by professional investors during the 1920s. He finds that stock prices are consistent with those models. In essence, the high prices of the 1928-1929 boom reflected an expectation of a continuation of the revenue growth of the recent past into the near future. Those expectations also explain the rising leverage that was tolerated by bond market investors during the 1920s (Calomiris 1993). Stock valuations appear not to have been driven by amateur investors; rather, they reflected the consensus among professional investors that growth in earnings for high-growth firms would persist into the future. Calomiris and Wilson (2010) examine the composition of stockholders of Citibank (then First National City Bank) from 1925 to 1931. Large stockholders, especially sophisticated investors who were also bank insiders with very large preexisting stakes, increased their holdings over time; virtually no insiders sold during the boom, and all held on through November 1931.

But for other observers, the rapid rise in prices was seen as an unsustainable and destabilizing frenzy. Some recent academic studies have also agreed with that perspective, emphasizing the new presence of retail investors in the market, the importance of buying on credit, the high mark-ups on closed end mutual funds, and the ex ante perceptions of increasing risk as stock prices rose implicit in securities loan interest rates (DeLong and Shleifer 1991, White 1990, and Rappaport and White 1993, 1994).

For real bills adherents, like Senator Glass and many Fed officials, the major concern was keeping the banking system from becoming entangled in financing the stock market boom. The Fed's actions began as an initiative to prevent bank involvement in the call loan market, which real bills adherents believed would fuel the bubble and leave the money center banks weakened. These beliefs led the Fed to impose limits on member bank lending to the securities market, beginning in 1928 (White 1990, Rappaport and White 1993, 1994). This policy seems to have had little effect either on the amount of credit available for the stock market or on stock prices. Real bills advocates pressed for additional policies to deflate what they saw as a bubble. Ultimately, partly in reaction to the continuing surge in stock prices, the Fed pursued contractionary monetary policy in 1929, which precipitated the recession.

The recession dashed expectations of near-term growth, and resulted in a severe drop in stock prices. However, in 1930, in expectation of a normal economic recovery, stock prices recovered much of the ground they had lost, until the deepening recession caused them to fall once again.

Whether the Fed was prescient about the risks of the stock market, or alternatively, the source of an unnecessary collapse of the stock market, depends on how much weight one attaches to the likelihood of an optimistic growth scenario in the absence of Fed action. For those who equate run-ups in securities prices with bubbles, the Fed was obviously justified, but for those who entertain the

possibility that stock market booms may reflect fundamentals, there is a lot of evidence consistent with that point of view from the 1920s.

Unsurprisingly, the lesson Carter Glass learned was that he had been right all along about the market. That point of view underlay his aggressive and successful actions to limit connections between commercial banking and the securities markets in 1933. The 1933 Act gave the Fed powerful tools to limit bank involvement in securities lending (Meltzer 2003, p. 434). It also separated investment banking and commercial banking and limited interest payments on checking accounts (a convenient means of undermining the pyramiding of reserves by eliminating the interest payments on interbank deposits). Subsequent academic research has seen these as unnecessary and unwise actions, but it took many decades for these regulations to be undone (see Calomiris 2010 for a review).

Did the Gold Standard Constrain Monetary Expansion During the Depression?

Some of the thorniest issues of interpretation about the Fed's actions and intentions during the Great Depression relate to the question of whether the maintenance of the gold standard prevented monetary expansion, especially during the period October 1931-January 1932. Eichengreen (1992) is the primary advocate in favor of the proposition that a lack of "free gold" constrained policy, although support for that point of view can also be found in Temin (1989) and Bernanke (1995). According to Eichengreen (1992), only a coordination of actions by the world's central banks could have freed the Fed from the concerns that inhibited its monetary expansion. Friedman and Schwartz (1963) had taken the opposite view, arguing that the Fed could have expanded the money supply without any risk to its ability to maintain the gold standard.

When the British left gold in September 1931, the United States began to experience a significant drain of gold reserves. By law, the Fed had to maintain the gold standard (the convertibility of its notes into gold), and it also faced statutory requirements regarding the minimum gold reserves it had

to maintain against outstanding Federal Reserve notes.[11] Eichengreen (1992) corroborates the statements of Fed officials (see also, Meltzer 2003, pp. 349-356), who had expressed concern about these risks, both during the 1931-1932 period and subsequently. Eichengreen argues that Fed officials were rightly concerned that an expansion of the money supply in the presence of a reserve outflow would have put the Fed at risk of running out of "free gold" (gold in excess of its 40 percent reserve requirement against notes) and could have prompted a run on the dollar.

Wicker (1966) and Meltzer (2003, pp. 274-276, 347-357, 404-407) review the claims that have been made on both sides of this issue, and document the thinking of the Fed decision makers at the time. They conclude that that Fed officials often felt constrained by the gold standard, but they side with Friedman and Schwartz on the question of whether it would have been possible for the Fed to have expanded the money supply during October 1931-January 1932 without being concerned that doing so would have been difficult or risky. Bordo, Chaudri and Schwartz (2002) argue that the United States was a large country with ample gold reserves, and that it had substantial latitude to expand the money supply without departing from the gold standard. They argue that the United States was not constrained from using expansionary policy to offset banking panics, deflation, and declining economic activity. They perform simulations, based on a model of a large open economy, which indicate that expansionary open market operations by the Federal Reserve at what they regard as two critical junctures (October 1930 to February 1931, and September 1931 through January 1932) would have been successful in avoiding economic contraction without endangering convertibility.

That is also the point of view of Hsieh and Romer (2006), who perform both statistical tests of the effects of monetary expansion on exchange rate devaluation risk, and review market perceptions of

[11] The Fed was required by law to maintain a 40 percent gold reserve against its notes, with the remaining 60 percent in eligible private sector paper. If Reserve Banks held less than the required 40 percent gold reserve, they were subject to a tax (e.g., if the reserve ratio fell to a percentage between 32.5 and 40, the tax rate was 1.5% (Meltzer 2003, p. 356).

that risk, as indicated by market prices, as well as statements by market participants. Hsieh and Romer (2006) show that the expansion of the money supply that occurred after the relaxation of reserve requirements (in February 1932, under the 1932 Glass-Steagall Act) created no risk of devaluation, and they argue that this shows that monetary expansion would not have put the gold standard at risk in late 1931.

There are two separate questions that are both worth addressing, and recognizing them as separate is useful. (1) Did Fed officials hold back in monetary expansion during October 1931-January 1932 that they otherwise would have wanted to undertake because they believed that they were constrained by the gold standard? (2) Irrespective of the beliefs of Fed officials, would it have been possible for the Fed to expand substantially more without threatening the gold standard?

With respect to the intentions and beliefs of Fed officials, it is pretty clear that they were concerned about the gold drain after the British departure from gold, and that this played an important role in delaying expansion during October 1931-January 1932: "Did the free gold problem delay open market purchases? The answer is certainly yes" (Meltzer 2003, p. 357). But Meltzer also shows that Harrison delayed monetary expansion during this period for *several reasons*, and that several other governors opposed purchases for reasons unrelated to gold. One of those reasons was a real-bills-doctrine version of the "liquidity trap" (described below); some thought that monetary expansion would have little effect for that reason (Meltzer 2003, p. 350). Also, many Fed officials believed that the speculative excesses that had produced the Depression had to run their course, and should not be combated with expansion of credit (Meltzer 2003, p. 405).

Contrary to Eichengreen (1992), Meltzer argues that the failure of international coordination was not the binding constraint on Fed actions. International coordination was feasible when the Fed saw it as desirable (Meltzer 2003, p. 344). Meltzer quotes Oliver Sprague, who said in a speech in May 1931

that the lack of monetary expansion by the central banks was not the result of "any difficulty in securing agreement among the three banks (France, U.K., U.S.), but because none of them harbored the belief that [monetary expansion] was the appropriate remedy" (Meltzer 2003, p. 278). Sprague himself agreed with that "liquidationist" position.

Furthermore, the willingness of the Fed to expand after February 1932 (an aggressive but brief expansion that unfortunately came to a halt in June 1932) reflected not only the passage of the 1932 Glass-Steagall Act and the calming of international markets, but also the signals being provided by the real bills indicators:

> The action was consistent with the Riefler-Burgess framework. Member bank borrowing and short-term rates had not declined. Borrowing was well above the $500 million range considered high in an ordinary recession and was almost back to the 1929 peak. A program to reduce the volume of borrowing by undertaking purchases was consistent with the dominant view that credit markets could be eased by forcing a reduction in the System's portfolio of real bills (Meltzer 2003, p. 359).

The fact that the Fed chose to expand the money supply after the passage of the February 1932 Glass-Steagall Act also provides only weak support for the view that the gold reserve was the key limit that prevented policy makers from acting earlier. The 1932 Act did more than just relax the gold reserve requirement; it allowed member bank borrowing at one percent above the discount rate against previously ineligible commercial paper, and it permitted groups of banks (e.g., clearing houses) to borrow collectively from the Fed on the credit of the group (Meltzer 2003, p. 358).

In summary, Fed officials clearly worried about the gold drain of late 1931, and it clearly had a negative effect on their willingness to expand the money supply. After the February 1932 Act, they were willing to expand aggressively. It is clear that free gold and concerns about a potential run on the dollar played an important role in Fed thinking, but not clear precisely how much weight to attach to that factor. Finally, it is important to remember that the period in question during which free gold concerns

mattered as a constraint in the minds of Fed officials was brief (less than six months), and the subsequent period of expansion of the money supply after the passage of the February 1932 Act was also brief (February-June 1932). Free gold was a small issue in the grand scheme of the errors of the Fed during the Depression, and has nothing to say about the failures of Fed policy prior to September 1931 or from June 1932-March 1933.[12]

With respect to the second question – whether, *in fact*, *the Fed was constrained by its gold reserve requirement on notes and by the risk of an attack on the dollar* – there are two aspects to that question (the importance of the physical reserve requirement on notes, and the potential market reaction), and it is important to consider them separately. The primary risk relating to the gold standard did not revolve around the 40 percent gold reserve requirement per se. The possibility that the Fed would be unable to support the dollar because it would keep its gold locked in a vault (as backing for its notes) was remote. Gold inflows in 1930 and early 1931 had been large (raising the gold stock at the Fed 15 percent above its August 1929 value), implying substantial free gold (Meltzer 2003, p. 275). Even after the large $750 million outflow of gold in September and October 1931, there was about $2 billion in surplus gold and gold certificates that the Fed could access, if necessary (Meltzer 2003, p. 276). Harrison himself dismissed the importance of free gold, per se (Meltzer 2003, p. 345).

Furthermore, as Meltzer (2003, p. 356) points out, the Fed enjoyed substantial latitude with respect to its requirement. It could fall below the 40 percent requirement with little cost. A shortfall of reserves would have entailed a small tax payment by the Fed, and even that could have been mitigated by cancelling Federal Reserve notes that were held by the Fed itself, an action that would have freed substantial gold – Meltzer (2003, p. 356). And, of course, it was common for governments to suspend

[12] The possibility of devaluation becomes a policy issue again at the end of 1932, but this has to do with the risk of Roosevelt's leaving the dollar. For conflicting perspectives on that issue, see Wigmore (1987) and Hsieh and Romer (2006).

gold reserve requirements on central banks during crises; the British had suspended the Peel Act of 1844 during all three of Britain's post-1844 banking crises (in 1847, 1857, and 1866).

The more difficult question is whether aggressive monetary expansion in October 1931-January 1932 might have produced an attack on the dollar. This seems to have been more central to the concerns of policy makers. Charles Hamlin, a Federal Reserve Board member, who expressed confidence that free gold, per se, was not a concern, was nonetheless concerned about the escalating and uncertain risks that the Federal Reserve faced in the foreign exchange markets in late 1931:

> The experience of recent weeks brings home to Federal Reserve officials their heavy responsibility, the necessity for keeping their powder dry, so that in these troublous times they may remain the rock that can withstand all storms and upon which world confidence may once more be reconstructed (Meltzer 2003, p. 276).

High interest rates and tight monetary policy were seen as confidence builders, not just by Fed officials (like Harrison), but also by their counterparts abroad. Governor Clement Moret of the Bank of France encouraged Governor Harrison to raise rates in October 1931 for precisely that reason (Meltzer 2003, p. 344-345).

Hsieh and Romer (2006) argue that the lack of any market reaction (increased devaluation risk) to the monetary expansion of 1932 offers evidence against the view that an expansion of the money supply in late 1931 would have produced a run on the dollar. However, one could argue that they overstate the usefulness of their evidence. The market environment of late 1931 was quite different from the one of early 1932. Monetary expansion in an environment that is fearful of collapse and looking for signals of commitment from monetary authorities likely would have punished a move to expand the money supply much more than the market reacted to similar measures in 1932. That is not to say that Hsieh and Romer are wrong in their conclusion, just that it is very hard to have confidence about the counterfactual, as applied to October 1931-January 1932.

In summary, the beliefs and actions of Fed officials *regarding the gold standard* were not obviously wrong in late 1931. Fed officials did not regard the physical gold requirement as not much of a barrier to expansion. They were, however, justifiably worried in late 1931 about a potential run on the dollar, and also were loath to loosen out of a misguided philosophical belief in the benefits of allowing the speculative excesses of the Depression to run their course. Worries that led the Fed to project dollar strength in late 1931 receded by early 1932, partly because of several changes introduced in the 1932 Glass-Steagall Act, and partly because of changes in the market's perception of devaluation risk. The reason the Fed was eager to pursue a path of expansion, however, from February-June 1932 was that doing so was fully consistent with the policy rules of the Riefler-Burgess doctrine.

Were the Friedman-Schwartz "Panics" Really Panics?

Friedman and Schwartz (1963) argued that from an early date (December 1930) many bank failures resulted from unwarranted "panic" that occurred in several waves, and that failing banks were in large measure illiquid rather than insolvent. Friedman and Schwartz's emphasis on contagion posited that bank failures, early as well as late in the Depression, mainly reflected a problem of illiquidity rather than insolvency. Illiquid but solvent financial institutions, in their view, failed as the result of withdrawal demands by depositors, particularly during sudden moments of panic. In contrast, an insolvent institution fails to repay depositors as the result of fundamental losses in asset value, rather than the suddenness of depositor withdrawals.

If this account of the banking distress of December 1930-March 1933 is correct, then the failure of the Fed was egregious indeed. In the Friedman-Schwartz view, the Fed not only followed a misguided targeting rule, it failed to recognize and try to address widespread banking panics – the very problem which it was created to prevent. And it failed to do so as early as the end of 1930 and the first half of

1931. That timing is important, since it implies that much of the worst of the Depression might have been avoided if early banking system liquidity risks had been addressed.

Friedman and Schwartz attach great importance to the banking crisis of late 1930, which they attribute to a "contagion of fear" after the failure of a large New York bank, the Bank of United States, which they regard as itself a victim of panic. They also identify two other banking crises in 1931 – from March to August 1931, and from Britain's departure from the gold standard (September 21, 1931) through the end of the year. The fourth and final banking crisis they identify occurred at the end of 1932 and the beginning of 1933, culminating in the nationwide suspension of banks in March. The 1933 crisis and suspension was the beginning of the end of the Depression, but the 1930 and 1931 crises (because they did *not* result in suspension) were, in Friedman and Schwartz's judgment, important sources of shock to the real economy that turned a recession in 1929 into the Great Depression of 1929-1933.

The Friedman and Schwartz argument that these episodes of banking distress reflected liquidity rather than insolvency problems is based upon the suddenness of banking distress and the absence of collapses in relevant macroeconomic time series prior to those banking crises (see Charts 27-30 in Friedman and Schwartz 1963, p. 309). But there are reasons to question Friedman and Schwartz's view of the exogenous origins of the banking crises of the Depression.

As Temin (1976) and others have noted, the bank failures during the Depression mainly marked a continuation of the severe banking distress that had gripped agricultural regions throughout the 1920s. Of the nearly 15,000 bank disappearances between 1920 and 1933, roughly half predate 1930. And massive numbers of bank failures occurred during the Depression era outside the crisis windows identified by Friedman and Schwartz (notably, in 1932). Wicker (1996, p. 1) estimates that "[b]etween 1930 and 1932 of the more than 5,000 banks that closed only 38 percent suspended during the first three banking crisis episodes." Recent studies of the condition of the Bank of United States indicate that it too may have been insolvent, not just illiquid, in December 1930 (Lucia 1985, Wicker 1996). Banks that

considered merging with it determined at the last minute not to do so (Meltzer 2003, pp. 323-4). So there is some prima facie evidence that the banking distress of the Depression era was more than a problem of panic-inspired depositor flight.

Friedman and Schwartz omitted important aggregate measures of the state of the economy relevant for bank solvency, for example, commercial distress and construction activity. Second, aggregation of fundamentals masks important sectoral, local, and regional shocks that buffeted banks with particular credit or market risks. The empirical relevance of these factors has been demonstrated in the work of Wicker (1980, 1996) and Calomiris and Mason (1997, 2003a).

Using a narrative approach similar to that of Friedman and Schwartz, but relying on data disaggregated at the level of Federal Reserve districts, Wicker (1996) argues that it is incorrect to identify the banking crisis of 1930 and the first banking crisis of 1931 as national panics comparable to those of the pre-Fed era. According to Wicker, the proper way to understand the process of banking failure during the Depression is to disaggregate, both by region and by bank, because heterogeneity was very important in determining the incidence of bank failures.[13]

[13] Once one disaggregates, Wicker argues, it becomes apparent that at least the first two of the three banking crises of 1930-1931 identified by Friedman and Schwartz were largely regional affairs. Wicker (1980, 1996) argues that the failures of November 1930 reflected regional shocks and the specific risk exposures of a small subset of banks, linked to Nashville-based Caldwell & Co., the largest investment bank in the South at the time of its failure. Temin (1989, p. 50) reaches a similar conclusion. He argues that the "panic" of 1930 was not really a panic, and that the failure of Caldwell & Co. and the Bank of United States reflected fundamental weakness in those institutions.

Wicker's analysis of the third banking crisis (beginning September 1931) also shows that bank suspensions were concentrated in a very few locales, although he regards the nationwide increase in the tendency to convert deposits into cash as evidence of a possible nationwide banking crisis in September and October 1931. Wicker agrees with Friedman and Schwartz that the final banking crisis (of 1933), which resulted in universal suspension of bank operations, was nationwide in scope. The banking crisis that culminated in the bank holidays of February-March 1933 resulted in the suspension of at least some bank operations (bank "holidays") for nearly all banks in the country by March 6.

From the regionally disaggregated perspective of Wicker's findings, the inability to explain the timing of bank failures using aggregate time series data (which underlay the Friedman Schwartz view that banking failures were an unwarranted and autonomous source of shock) would not be surprising even if bank failures were entirely due to fundamental insolvency. Failures of banks were local phenomena in 1930 and 1931, and so may have had little to do with national shocks to income, the price level, interest rates, and asset prices. The unique industrial organization of the American banking industry plays a central role in both the Wicker view of the process of bank failure during the Depression, and in the ability to detect that process empirically. Because banks in the United

Microeconomic studies of banking distress provide useful evidence on the reactions of individual banks to economic distress. White (1984) shows that bank failures in 1930 were a continuation of the agricultural distress of the 1920s, and are traceable to fundamental disturbances in agricultural markets. Declines in railroad bonds were also significant in some cases (Meltzer 2003, p. 346).

Calomiris and Mason (1997) study the Chicago banking panic of June 1932 (a locally isolated phenomenon). They find that the panic resulted in a temporary contraction of deposits that affected both solvent and insolvent banks. Fundamentals, however, determined which banks survived. Apparently, no solvent banks failed during that panic. Banks that failed during the panic were observably weaker ex ante, judging from their balance sheet and income statements, and from the default risk premia they paid on their debts. Furthermore, the rate of deposit contraction was not identical across banks; deposits declined more in failing banks than in surviving banks.

Calomiris and Wilson (2004) study the behavior of New York City banks during the interwar period, and in particular, analyze the contraction of their lending during the 1930s. They find that banking distress was an informed market response to observable weaknesses in particular banks, traceable to ex ante bank characteristics. It resulted in bank balance sheet contraction, but this varied greatly across banks; banks with higher default risk were disciplined more by the market (that is, experienced greater deposit withdrawals), which encouraged them to target a low-risk of default.

Calomiris and Mason (2003a) construct a survival duration model of Fed member banks throughout the country from 1929 to 1933. This model combines aggregate data at the national, state, and county level with bank-specific data on balance sheets and income statements to identify the key

---

States were smaller, regionally isolated institutions, large region-specific shocks might produce a sudden wave of bank failures in specific regions even though no evidence of a shock was visible in aggregate macroeconomic time series (see the cross-country evidence in Bernanke and James 1991, and Grossman 1994). The regional isolation of banks in the United States, due to prohibitions on nationwide branching or even statewide branching in most states, also makes it possible to identify regional shocks empirically through their observed effects on banks located exclusively in particular regions.

contributors to bank failure risk and to gauge the relative importance of fundamentals and panics as explanations of bank failure. Calomiris and Mason find that a fundamentals-based model can explain most of the failure experience of banks in the U.S. prior to 1933.[14] They identify a significant, but small, national panic effect around September of 1931, and some isolated regional effects that may have been panics, but prior to 1933, panics were not important contributors to bank failures nationally.

The fact that a consistent model of fundamentals can explain the vast majority of bank failures prior to 1933 has important implications. First, the influence of banking panics as an independent source of shock to the economy was not important early in the Depression. Only in 1933, at the trough of the Depression, did failure risk become importantly de-linked from local, regional, and national economic conditions and from fundamentals relating to individual bank structure and performance. Second, the timing of this observed rise in risk unrelated to indicators of credit risk is itself interesting. In late 1932 and early 1933, currency risk became increasingly important; depositors had reason to fear that President Roosevelt would leave the gold standard, which gave them a special reason to want to convert their deposits into (high-valued) dollars before devaluation of the dollar (Wigmore 1987).

As part of their bank-level analysis of survival duration, Calomiris and Mason (2003a) also consider whether, outside the windows of "panics" identified by Friedman and Schwartz, the occurrence of bank failures in close proximity to a bank affects the probability of survival of the bank,

[14] Bordo and Landon-Lane (2010) argue that examiners' reports of failed banks, which sometimes attribute failure to bank illiquidity in the 1930 and early 1931 period, therefore provide evidence of nationwide banking panics. That inference is not warranted, in my view, for two reasons. First, individual bank illiquidity of a failed bank, as inferred by an examiner, is not clearly distinguishable from expected insolvency. After all, suspended banks would have reopened (rather than remained closed) if they did not ultimately become insolvent. A bank whose depositors lose confidence in it, causing insufficiency of reserves, which is also unable to open subsequently, may be deemed to have been "illiquid" by examiners, but may have been made illiquid precisely because it was deemed as insolvent by its depositors (Calomiris and Kahn 1991). Second, even if some banks were driven to failure by their illiquidity (which clearly was the case in some instances) that does not imply the existence of a nationwide panic. Wicker and Calomiris and Mason (2003a) recognize that there were some regional panics during that period, but show that they were not significant at the national level. Calomiris and Mason's (2003a) finding that identifiable fundamental shocks explain individual bank failures as well during the alleged panic episodes of 1930 and early 1931 as during other times in Depression, when Friedman and Schwartz did not allege the existence of a national panic, shows that there was not much additional aggregate importance of non-fundamental factors in explaining bank failures during the early alleged panics.

after taking into account the various fundamental determinants of failure. Calomiris and Mason consider this measure of "contagious failure" an upper bound, since in part it measures unobserved cross-sectional heterogeneity common to banks located in the same area, in addition to true contagion. They find small, but statistically significant, effects associated with this measure. The omission of this variable from the analysis raises forecasted survival duration by an average of 0.2%. They also consider other regional dummy variables associated with Wicker's (1996) instances of identified regional panics, and again find effects on bank failure risk that are small in national importance.

The large number of bank failures in the U.S. during the Great Depression, a phenomenon that was largely confined to small banks, primarily reflected the combination of extremely large fundamental macroeconomic shocks and the vulnerable nature of the country's unit banking system. Panic was not a significant contributor to banking distress on a nationwide basis until near the trough of the Depression, at the end of 1932. For these reasons, the Great Depression bank failure experience has more in common with the bank failures of the 1920s than the panics of the pre-World War I era.

It is probably not correct to argue, then, that the Fed failed to detect avoidable national liquidity crises and prevent waves of bank failures in 1930 and 1931. The Fed properly did not see its role as bailing out failed banks, and thus, it is not surprising that it allowed failed banks to be closed.

That is not to say that there were no bank liquidity problems or panic episodes early in the Depression; there were various local panics at different times during the period 1929-1932 (Florida in 1929, various parts of the South in 1930, Chicago in mid-1932, to name a few), and some were associated with significant bank distress. Furthermore, it would be wrong to presume that the Fed did all it could do address those local problems. While there are inherent limits to what the Fed could accomplish with liquidity assistance through collateralized lending, there is evidence suggesting that

more could have been done.[15] For example, the Atlanta Fed was more activist than the St. Louis Fed, and seems to have been able to prevent liquidity problems from spreading, in Florida in 1929 and in the areas of Mississippi under its control in 1930, while the St. Louis Fed's conservatism may have led to unnecessary bank failures (Richardson and Troost 2006, Carlson, Mitchener and Richardson 2010).[16]

Most importantly, even if the early phases of the Depression lacked nationwide banking panics, that does not mean that the banking distress of the 1930-1931 should have been ignored by the Fed. The Fed should have recognized that the contraction in deposits and increase in the banks' demands for reserves were causing the money multiplier to shrink. The Fed should have responded to this contraction in the supplies of money and credit with expansionary open market operations and discount window lending. In that sense, Friedman and Schwartz are correct to criticize the Fed for its failure to respond. That failure of the Fed did not reflect its inability to perceive financial panic, but rather, adherence to the Riefler-Burgess doctrine, which prevented the Fed from detecting or responding to a contraction in banks' supply of loans.

---

[15] Discount window lending only helps preserve banks that are suffering from illiquidity, which was not the primary problem underlying large depositor withdrawals. Indeed, in 1932, President Hoover created the Reconstruction Finance Corporation (RFC), to enlarge the potential availability of liquidity, but this additional source of liquidity assistance made no difference in helping borrowing banks avoid failure (Mason 2001). As commentators at the time noted, because collateralized RFC and Fed loans were senior to deposits, and because deposit withdrawals from weak banks reflected real concerns about bank insolvency, loans from the Fed and the RFC to banks experiencing withdrawals did not help much, and actually could harm banks, since those senior loans from the Fed and the RFC reduced the amount of high quality assets available to back deposits, which actually increased the riskiness of deposits and created new incentives for deposit withdrawals. In 1933, however, once the RFC was permitted to purchase banks' preferred stock (which was junior to deposits), RFC assistance to troubled banks was effective in reducing the risk of failure (Mason 2001).

[16] Richardson and Troost (2006) show that, despite the limited ability of Fed discount window lending to absorb credit risk, Fed provision of liquidity to member banks mitigated bank failure risk associated with illiquidity somewhat in 1930, and could have played a greater role in stemming illiquidity-induced failures if the Fed had been more willing to relax lending standards to member banks. They study the failure propensities of Mississippi banks. The Federal Reserve Act of 1913 divided Mississippi between the 6th (Atlanta) and 8th (St. Louis) Federal Reserve Districts. The Atlanta Fed championed a more activist role in providing loans to member banks experiencing troubles, while the St. Louis Fed rigidly adhered to the real bills doctrine and eschewed the extension of credit to troubled banks. Mississippi banks in the 6th District failed at lower rates than in the 8th District, particularly during the banking panic in the fall of 1930, suggesting that more aggressive discount window lending reduced failure rates during periods of panic. Carlson, Mitchener and Richardson (2010) show that timely and effective liquidity provision by the Atlanta Fed during the Florida agricultural crisis of 1929 (caused by an insect infestation) was also helpful.

Was the Economy Stuck in a Liquidity Trap in the 1930s?

The concept of the "liquidity trap," while not consistently defined, has been proposed in various forms to explain the persistence of the Great Depression in the United States as the result of the impotence of monetary policy to promote a recovery. Roughly speaking, there are three versions of the liquidity trap concept: (1) the naïve view of the liquidity trap, based on an assumption of static expectations and a two-asset model of the economy assumptions, and (3) the Fed's own view of the liquidity trap during the Depression, which was based on a particular version of the real bills doctrine.[17]

The naïve view is the one most commonly espoused: When short-term interest rates are at zero, any further expansion of the money supply has no effect on interest rates, as money demand passively accommodates all increases in money supply. Thus, expansion in money supply has no effect on economic activity. According to the naïve view, when short-term interest rates are at zero (implying perfect substitutability between money and Treasury bills), open market operations are a useless tool.

The sophisticated version of the liquidity trap recognizes that there are other assets in the economy, including longer-term Treasury bonds and foreign exchange (or gold), and that expectations of inflation are not constant, but will adapt to changes in policy on a forward-looking basis. Under these assumptions, the liquidity trap is a possibility, but only if the monetary authority acts unwisely. If the short-term nominal interest rate is zero, and the long-term interest rate is positive and above its nominal floor, then a purchase of longer-term bonds can provide monetary stimulus, even before taking into account changes in expected inflation. Furthermore, the central bank can purchase other assets, including foreign exchange, thereby depressing the foreign exchange rate, and boosting the demand for exports. Such expansionary policies, especially if done aggressively, would also lead to expected rises in

[17] For a more detailed treatment of the liquidity trap literature, and the three categories of ideas described here, see Brunner and Meltzer (1968b), Orphanides and Wieland (2000), McCallum (2001), Meltzer (2003, pp. 336-337), 478, Hanes (2006), and Basile, Landon-Lane and Rockoff (2010).

the price level. An expected increase in inflation will lower real interest rates, implying further potency to monetary policy.

If the central bank is unwilling to purchase foreign exchange, long-term bonds or other assets, and if the central bank is not willing to send a signal to the market of continuing expansion in the money supply, then it may be stuck in a liquidity trap. But if the central bank is willing to do these things, then monetary policy can have potent effects, even when short-term nominal interest rates are at their minimum value of zero. In theory, monetary policy never need be stuck in a liquidity trap, if central banks are willing to expand the supply of money on an ongoing basis (so-called "quantitative easing") and purchase assets other than just short-term Treasury bills.

Clearly, during the 1930s, despite episodes when nominal short-term interest rates were at their zero floor, the Fed and the Treasury, through various mechanisms, always had the means at their disposal to expand the money supply, increase inflation, and buy a wide range of assets. Monetary policy, therefore, could have been very effective during the Great Depression. Indeed, Hanes (2006) shows that expansions in the monetary base (e.g., via gold inflows) were effective during the 1930s, in particular, through their effects on long-term interest rates. Basile, Landon-Lane and Rockoff (2010) show that there was substantial ability for the Fed to affect private sector interest rates, too.

The liquidity trap did not constrain the effectiveness of monetary policy. If the Fed and the Treasury had expanded the supply of high-powered money through open market purchases at any time during the period, doing so would have reduced long-term interest rates, expanded money and credit, and resulted in increased economic activity and higher prices.

Nevertheless, Fed officials believed that policy was stuck in a liquidity trap. Their view of that liquidity trap differed from those of economists, and reflected real bills thinking, and a misunderstanding of reserve demand. As Meltzer (2003, p. 336) explains: "Harrison favored delaying

further purchases [in 1931], at first because the international monetary system had deteriorated and he believed the timing was poor, later because the banks held excess reserves. Although he fully discussed the rising rate of failure and insolvency among New York banks, he never mentioned the relation between rising excess reserves and rising failure rates. He believed that open market purchases would be useful only if the banks acquiring reserves used them to acquire low-quality bonds…"

In other words, in mid-1931, and afterward, the reason the Fed chose not to expand the money supply was that *it believed monetary expansion would not result in an expansion of credit*. This was an implication of the Riefler-Burgess doctrine: because loan demand was weak, banks were accumulating excess reserves. Adding more money to the system would just result in greater holdings of bank reserves, not more lending, and thus would have little effect. The Fed failed to see that the expansion in excess reserves reflected banks' increased demand for reserves in the wake of large losses of capital and increased liquidity risk (Calomiris and Wilson 2004, Van Horn 2008, Calomiris, Mason, and Wheelock 2010). The Fed should have interpreted the increase in reserves as indicative of a decline in loan supply (Calomiris and Mason 2003b), as depicted in Figure 2, but instead they saw it as a decline in loan demand, and believed (because of their adherence to the real bills doctrine) that they should accommodate that decline.

Harrison was not alone in that view. Marriner Eccles, in his 1935 testimony before the House Committee on Banking and Currency, agreed with Congressman Goldsborough, who coined the phrase: "one cannot push on a string" (Meltzer 2003, p. 478). Eccles argued that monetary policy was not capable of restoring economic activity, hence Eccles' desire for fiscal activism; for monetary policy to work, "you must have borrowers who are willing and able to borrow" (Ibid.).

The liquidity trap did not constrain monetary policy during the 1930s. False beliefs about the liquidity trap, resulting from the flawed conceptual framework used to define Fed objectives and

interpret economic data, *did* constrain policy. Those false beliefs about the liquidity trap were central in explaining why the Fed failed to expand the money supply during the Depression.

Did the Reserve Requirement Increases of 1936-1937 Cause the Recession of 1937-1938?

The reorganization of the Federal Reserve System in 1935 was presented as a means of centralizing decision making over monetary policy at the Board and giving greater authority to the Federal Reserve System. This reflected the rising influence of Marriner Eccles within the Roosevelt team, and the declining influence of Carter Glass (the architect of the decentralized approach to monetary policy). In the name of coordination, Eccles was especially desirous of diminishing the power of the New York Fed. A modified version of Eccles' plan was adopted, which reorganized the FOMC to put control over monetary policy in the hands of Presidential appointees for the new Board of Governors. The Fed also was granted more authority over the setting of reserve requirements. The Banking Act of 1935 eliminated the so-called Thomas Amendment of 1933, which had required Presidential approval of changes in bank reserve requirements.

On the surface, the Banking Act of 1935 seems like it should have strengthened the power of the Fed to control the money supply. But appearances can be deceiving, especially when the intent is deception. As already noted, after 1933, the Treasury occupied the driver's seat of monetary policy, and the Fed was in the "back seat." The Fed made no change to the size of its open market portfolio from the end of 1933 until 1937. Although monetary policy was highly expansionary from 1933 to the end of 1936, the primary contributor to monetary expansion during that period was gold inflows, which were controlled by the Treasury. Under the Gold Reserve Act of January 30, 1934, the Treasury set the exchange rate, managed the Exchange Stabilization Fund (ESF), and decided on gold sterilization policy. The Silver Purchase Act of June 19, 1934 added further Treasury authority over the creation of money.

The power of the Treasury over monetary affairs was not accidental. There was little the Fed could do to oppose the Treasury's aggressive monetary expansion. That fact reflected Secretary Henry Morgenthau's intentions. Morgenthau's reflections on the limits to Fed power after the Banking Act of 1935 in his diary are particularly telling: "[The Treasury's] power has been the Stabilization Fund plus the many other funds that I have at my disposal and this power has kept the open-market committee in line and afraid of me" (Blum 1959, p. 352). And the appearance of centralized Fed power, alongside the fact of Fed impotence, particularly suited his intentions: "I prophesy that…with the seven members of the Federal Reserve Board and the five governors of the Federal Reserve Banks forming an open market committee, that one group will be fighting the other and that consequently they will not be able to do anything constructive, and that therefore if the financial situation should go sour the chances are that the public will blame them rather than the Treasury" (Blum 1959, p. 352).

The power of the Treasury relative to the Fed was asymmetric; the Fed was not able to challenge the Treasury's desire to *expand* the money supply for the simple reason that the Treasury's capacity to expand high-powered money was greater than the Fed's capacity to contract it, which was limited by the size of the Fed's balance sheet (Calomiris and Wheelock 1998). But the Fed could have forced an increase in the money supply (via open market purchases) even if the Treasury had tried to oppose that increase with a contraction, since in that case, the Fed's power to expand could not have been checked by the Treasury. In effect, the reason the Fed was in the backseat after 1933 was because the Treasury wished to expand as much or more than the Fed would have liked.

In December 1936, fears of increasing inflation risk voiced by Chairman Eccles, as well as by President Roosevelt (Meltzer 2003, p. 502-507, 516), led the Treasury to begin sterilizing gold inflows, which ended the long period of persistent growth of the monetary base. A brief power struggle between

the Fed and the Treasury over who would control sterilization policy was resolved in 1936, when the Treasury won that battle, as Secretary Morgenthau predicted he would (Blum 1959, pp. 360-367).

Although the Fed lacked the balance sheet clout to play chicken with the Treasury credibly over the control of the monetary base, as already noted, the Fed was given new power under the 1935 Act to control reserve requirements. In 1936 and 1937, in a series of actions, the Fed raised the reserve requirements on deposits substantially. The Treasury was sometimes displeased by those actions, and was sometimes successful (in March 1937) in pressuring the Fed to expand the monetary base in order to preempt threatened use of the Exchange Stabilization Fund by Morgenthau to accomplish that same end. On March 15, 1937, under the threat that the Treasury would use the Exchange Stabilization Fund to increase the money supply (described in the minutes of the March 13, 1937 FOMC Executive Committee meeting), the FOMC responded to the Treasury's concerns over a very slight increase in Treasury yields by agreeing to stabilize yields with open market operations, if necessary. As recorded by Morgenthau in his diary, he and President Roosevelt understood that they controlled the money supply, and they were not shy about reminding the Fed of the fact:

> The President suggested that I should say to the Federal Reserve: "Now Congress gave you the job of managing the money market and that is your responsibility. You muffed it. You haven't done it. You have not maintained an orderly market, and this thing is getting steadily worse…Now I, Henry Morgenthau, Jr., talking for the U.S Government, serve notice on the Federal Reserve Board that I ask you to do what Congress has given you the power to do, namely, to increase your portfolio. If you don't do it, the Treasury will step in…We are putting you on notice." (Blum 1959, pp. 373-374, quoted in Calomiris and Wheelock 1998).

This passage shows that the Treasury's definition of the Fed's job during this period was to peg the interest rates on Treasury securities. That was the "orderly market" that Morgenthau was referring to. This passage also shows that, in effect, even the power to change reserve requirements did not give the Fed monetary control. Reserve requirement changes could be offset by an expansion of the base, and the Treasury had enough power to insist on such an expansion, or if necessary, to produce one.

In fact, the reserve requirement increases of 1936-1937 were not intended as a means of contracting the money supply, nor did Fed officials believe that the increases in reserve requirements had done so. Rather, the Fed desired to increase the reserve requirement so that required reserves would be closer in amount to total reserves, to facilitate *future* monetary policy actions, and thus, to forestall long-term inflation risk. This view was part and parcel of adherence to the real bills doctrine, which saw excess reserves as a redundant slack in the system that created an obstacle to the Fed's ability to use its targets predictably (e.g., borrowed reserves).[18] Friedman and Schwartz (1963) recognized that this was the Fed's intent, but they argued that, in fact, the reserve requirement changes had a major contractionary effect on the money supply, over and above the effect of the sterilization of gold inflows (pp. 459, 510-511, 520-522, 544).

More recent scholarship has questioned the importance of the reserve requirement increases for producing the recession of 1937-1938. Meltzer (2003, p. 503) notes that the reserve requirement change of mid-1936 "had no perceptible effect on the economy in 1936," although he is more accepting of the Friedman and Schwartz view for the 1937 reserve requirement increases. Hanes (2006) and Calomiris, Mason and Wheelock (2010) argue that there is no evidence of an effect from reserve requirement changes on the economy, and that one must look elsewhere (gold inflow sterilization, contractionary fiscal policy) to explain the recession of 1937-1938.

Hanes (2006) constructs a model of the effects of monetary policy on long-term interest rates. He shows that Treasury bond yields did not respond to the reserve requirement changes. He concludes that policies affecting the supply of reserves – especially gold inflows before sterilization – had significant effects on interest rates, but reserve requirement changes did not.

[18] See Meltzer (2003, pp. 495-499). Fed officials, particularly Eccles sometimes made statements that seemed to recognize the possibility that the rising reserve holdings of banks may reflect a shift in demand, but this was an aberration from the typical view, and in any case, Eccles, in pushing for the increases in reserve requirements, does not seem to have acted in a manner consistent with that belief (Meltzer 2003, pp. 514-515). Other concerns – especially about inflation risk due to gold inflows – were also relevant (Meltzer 2003, pp. 504-505).

Calomiris, Mason and Wheelock (2010) provide the micro-foundations that explain Hanes's (2006) results. Reserve holdings by member banks, and other banks, reflected risk management (see also Calomiris and Wilson 2004, Van Horn 2008). Given the lost capital of banks in the 1930s, and the liquidity risk banks had experienced, bankers accumulated large reserves, and almost all banks held substantially more reserves than what the Fed required. The Fed reserve requirement increases, although substantial, were still not even close to binding on bank preferences for the most part, since banks wanted to hold substantially more reserves than the Fed ever required them to hold.

To test this theory, Calomiris, Mason and Wheelock (2010) employ microeconomic data on individual bank balance sheets and income statements for 1934 and 1935, and other data, to construct a model that predicts bank reserve ratios in 1935. The coefficients from this model are then used to forecast the ratios of reserves to total assets (or the ratio of reserves to total deposits) for 15 mutually exclusive categories of banks: 12 District-level reserve city bank aggregates, 2 central reserve city aggregates (New York City and Chicago), and a non-reserve city bank aggregate, using various alternative definitions of reserves (including or excluding various categories of liquid assets in the definition of reserves). Calomiris, Mason and Wheelock (2010) then test to see whether the actual reserve ratios in 1936 and 1937 for these 15 bank aggregates were higher than the counterfactual forecasts based on the 1935 estimates. They find no evidence that actual reserve ratios were higher than those forecast based on 1935 estimates. In other words, there is no evidence that reserve requirement changes in 1936-1937 had any effect on the reserve ratios chosen by banks.

Summary

Even with the benefit of seven decades of hindsight, some of the most important facts about the nature of the shocks of the interwar period, and some of the most important alleged sources of errors of monetary policy identified by academics during those seven decades, are not so obvious.

Should the Fed have tried to discourage the run-up in the stock market, or did its attempt to do so create a shock rather than respond to one? The evidence seems at least as compelling in favor of the latter alternative.

Was the Fed wrong to be concerned about a run on the dollar in October 1931, or was the reaction to gold outflows and perceived risk prudent? Again, it is hard to dismiss out of hand the possibility that the Fed was right to delay monetary expansion until international markets had calmed a bit. (Of course, that possibility in no way absolves the Fed from its disastrous policies before September 1931, or after June 1932.)

Were the alleged Friedman-Schwartz panics of late 1930 and early 1931 national events that should have grabbed the Fed's attention, or were they local affairs, and largely a continuation of the patterns of local bank failures that had occurred for a decade in agricultural areas? The evidence seems to be fairly strong in favor of the latter.

Was there a liquidity trap that rendered monetary policy impotent during the Depression? There are various ways to define the liquidity trap, and none seem applicable to the Depression. Nevertheless, at some points Fed officials thought that there was no point in expanding the money supply because they were observing rises in excess reserves (which indicated no need for additional liquidity, according to the Riefler-Burgess framework).

Were the reserve requirement changes of 1936 and 1937 key sources of the recession of 1937-1938, as Friedman and Schwartz (1963) argued, or were they largely neutral in their effects, as Fed officials had believed? Again, the evidence seems to be fairly strong in favor of the latter.

Even if one disagrees with the conclusions reached here – and there is room for reasonable disagreement – it would be hard to argue that opposite conclusions could be proven with existing data

and analysis even *as of 2010*. That should give those who are puzzled by the slow learning curve of the Fed during the interwar period some pause. If some of the smartest policy analysts in the world, thinking at leisure decades after the Depression was over, were not able to get key aspects of the story about the Depression right in the 1960s, the 1990s or the 2000s, then perhaps it is not so surprising that the Fed took so long to reject the Riefler-Burgess framework.

V. The Fed's Success Story: Stabilizing Seasonal Swings in Interest Rates

There is substantial evidence (Miron 1986, Richardson and Troost 2006, and Bernstein, Hughson, and Weidenmier 2009b) that the founding of the Fed reduced seasonal volatility of interest rates, increased seasonal variation in lending, and reduced liquidity risk in the banking system, which in turn reduced the propensity for bank panics. This must be regarded as the Fed's "success story" of the interwar period. After reviewing the evidence of that success, I consider why the Fed was able to stabilize the financial system at the seasonal frequency so much better than it was able to manage monetary policy over the business cycle.

Miron (1986) showed that the Fed's founding was associated with reduced seasonal variability of interest rates and increased seasonal variability of lending. Why, exactly, did Fed lending practices make the loan supply function more elastic?

Miron's (1986) findings can be explained by a variant of the deposit risk targeting model in Calomiris and Wilson (2004), discussed in Section II above. In that model, the riskiness of deposits is a function of bank asset risk and bank leverage. Because total bank capital and total cash (gold and currency) assets in the economy do not vary much from month to month, a seasonal increase in bank lending (especially to finance crop harvesting and transport in the fall, which Davis, Hanes and Rhode 2007 show was largely driven by the cotton cycle) implies a commensurate increase in bank asset risk and in bank leverage, which unambiguously means an increase in the riskiness of deposits (the actuarily fair default risk premium). This is a source of seasonal variation in the risk of deposit withdrawals, since

market discipline makes the risk of withdrawal in the deposit market sensitive to increases in default risk (i.e., some depositors are intolerant of risk, and will withdraw when risk increases). A bank that increases its lending, ceteris paribus, faces increased deposit withdrawal risk, particularly if an adverse cyclical shock hits during a seasonal lending spike.

As noted before, all six of the major national banking era panics happened at cyclical peaks; they were clearly responses to adverse economic shocks to banks' balance sheets (Calomiris and Gorton 1991). Furthermore, these panics all occurred either during the spring planting season or the fall harvest, at times when lending (and bank liquidity risk) was at a seasonal peak. Liquidity risk in the banking system peaked in the fall and spring. Seasonal liquidity risk, combined with cyclical changes in default risk, can explain why the panics of the national banking era happened at cyclical peaks (defined by GDP, the stock market, and business failures) in months of highest liquidity risk.

From the perspective of this model, the founding of the Fed provided a means of reducing liquidity risk to banks by giving them a source of liquidity, if needed, to stem deposit withdrawals (making them less vulnerable to withdrawal risk at times when seasonal lending peaks coincided with cyclical downturns). The founding of the Fed thus flattened the bank loan supply function, making loans vary more over the seasonal cycle, and interest rates vary less. This seasonal flattening of the loan supply curve is identical to the pattern already illustrated in Figure 1.

Bernstein, Hughson, and Weidenmier (2009b) provide additional evidence consistent with that interpretation of the Fed's effect on seasonal liquidity risk. They compare the standard deviations of stock returns and short-term interest rates over time in the months of September and October (the two months of the year when markets were most vulnerable to a crash because of financial stringency from the harvest season) with the rest of the year before and after the establishment of the Fed. Stock volatility in those two months fell more than 40 percent, and interest rate volatility more than 70 percent, after the founding of the Fed. They also show that this result is driven by years in which

business cycles peaked. In other words, the main risk that the Fed eliminated was associated with combined cyclical peaks in economic activity and seasonal peaks in lending.

What explains the Fed's ability to successfully manage seasonal fluctuations? Most importantly, a largely passive policy of freely discounting bills for short periods, supplemented as need be with short-term season-specific open market operations, should work reasonably well for accommodating seasonal shifts in demand, so long as there are few seasonal shifts in supply (which seems likely). That is, the seasonal variation of the loan market is probably captured reasonably accurately by Figure 1. It also should have been fairly easy for the Fed to learn that its approach to seasonal smoothing of interest rates, and seasonal accommodation of loans, was working well, since the Fed could observe the increased variation in loans over the season and the decreased variation in interest rates. Thus, in contrast to the cyclical frequency – where relying on the Riefler-Burgess framework was disastrous – that framework worked pretty well for managing seasonal fluctuations in the market.

That suggests another possibility. It is possible that seasonal success slowed the Fed's learning process about cyclical failure. Fed officials may have been encouraged by the observable success of the Riefler-Burgess thinking, as applied to the seasonal cycle, and may have wrongly extrapolated that evidence as proof that their policy approach was also useful for managing cyclical variation. Given how hard it was to learn at the cyclical frequency (owing to few observations of cycles, multiple and diverse shocks affecting cycles, and various structural shifts of the economy), the Fed might have given greater weight to its seasonal success story in reinforcing its confidence in the Riefler-Burgess framework.

VI. Toward the Accord

The late 1930s brought World War II, which was associated with continuing fundamental changes in economic structure – mobilization for war, high government spending, sectoral production

shifts, price controls, rationing, and continuing improvements in technology, transportation and communication.

The War saw large increases in government debt, which increased by roughly one quarter of GDP. Fed policy was essentially dormant during the War; the Fed assumed a largely passive role in support of the Treasury's war financing efforts, by placing a ceiling of 0.375 percent on the interest rates for Treasury bills and 2.5 percent for Treasury bonds (Meltzer 2003, p. 580). After the War, the Treasury continued to control Fed policy, in support of Treasury yields, reflecting Treasury and Fed concerns of postwar recession and deflation (Meltzer 2003, chapter 7). The size of Fed holdings of marketable debt grew dramatically during the War and its aftermath. From December 1941 to December 1950, Fed holdings rose from $2.3 billion to $20.8 billion (Meltzer 2003, p. 720).

After some seventeen years of Fed subservience to the Treasury, in 1951, the Fed and the Treasury agreed that the Fed's independence would be restored. This landmark change seems to have reflected two key influences: (1) Rising inflation risk led the Fed to seek and obtain support in Congress for its independence, and for the desirability of shifting its focus from bond yield pegging to controlling inflation (Meltzer 2003, p. 723). (2) The increased size of the Fed's balance sheet increased its clout (Calomiris and Wheelock 1998). Recall that in 1935 Secretary Morgenthau was confident of his ability to force the Fed to do his bidding precisely because his ability to expand (through his various monetary powers) exceeded the Fed's ability to contract. World War II increased the scale of the Fed, substantially enhancing its power to shrink the money supply. This fact gave the Fed a new strategic advantage that it could use to defend itself from Treasury coercion.

The 1951 Treasury-Fed Accord reestablished Fed independence, but the Fed returned to its old ways, adopting once again the Riefler-Burgess framework to guide its policies (Meltzer 2003, p. 723).

VII. Conclusion

The Fed was established to bring stability to the U.S. financial system through the creation of an elastic supply of liquidity. Exactly what that meant and how it would be accomplished was not clear, and given the unique microeconomic (unit) structure of the American banking system, the Fed would not be able to rely much on precedents from abroad to guide it.

The most remarkable aspect of the period 1914-1951 was its volatility, economically, politically and financially. The first four decades of the Fed's existence coincided, practically without interruption, with World War I, the roaring twenties, the Great Depression, World War II, and the Korean War. These episodes saw dramatic restructuring of the U.S. economy, diverse and severe shocks to the economy, and fundamental changes in government powers and Fed structure.

The Fed made numerous errors in cyclical policy, especially during the period 1929-1933, when its actions precipitated the Great Depression. Although there were many contributing sources of error, the key source of the Fed's errors was adherence to the real bills doctrine, implemented through the Riefler-Burgess framework, which failed to distinguish between demand and supply shocks in the loan market, or between real and nominal interest rates. This led the Fed to misinterpret signals about the economy, effectively to mistake declining loan supply for declining loan demand during the Depression.

Interestingly, the Fed did not learn from its mistakes. Fed officials were pleased with their performance during the Depression because they believed that they had followed sound principles, rather than being distracted by ad hoc goals, which they believed would have made matters worse.

The volatility of the period probably explains why learning was so slow. Diverse shocks and structural change made it hard for the Fed to identify its own actions as a source of economic troubles. A review of five major controversies surrounding Fed actions during the Depression – whether the Fed's

intervention to cool the stock market in 1928-1929 was warranted, whether the gold standard was a constraint on monetary expansion from October 1931 to January 1932, whether banking distress of the early Depression years reflected national panics or fundamentals, whether monetary policy was impotent because of a "liquidity trap" during the Depression, and whether reserve requirement increases were a cause of the recession of 1937-1938 – shows that even decades after the end of the Depression, many of the issues at the heart of Fed decision making in the 1930s remain controversial. From that standpoint, the slow pace of Fed learning about its flawed policy framework is not surprising.

The Fed's greatest success during the period was the smoothing of cyclical variations in interest rates and liquidity risk, an achievement that contributed to financial stability and was very much at the heart of the founding of the institution. Ironically, that success may itself have slowed the process of learning about cyclical policy errors, since the demand-side dominated mindset of the Fed about the loan market, embodied in the Riefler-Burgess framework, worked well for purposes of seasonal smoothing; seasonal success may have boosted the Fed's confidence in its cyclical miscalculations.

The structure of the Fed changed dramatically over time. Its founders envisioned a decentralized structure, which would avoid concentration of power, ensure a connection to the local banks' needs, and avoid politicization of Fed actions. The desire for decentralization, however, conflicted somewhat with the need for a national monetary policy. In its early years, the Fed balanced its decentralized structure and its national mission reasonably well. In 1935, in reaction to perceived problems of insufficient coordination, power in the Fed was centralized. However, real power resided at the Treasury until the 1951 Treasury-Fed Accord. These structural shifts are a reminder that the Fed is a creature of Congress, subject to political influence. The recognition of that fact led the Fed to become an active political force in Washington (a process that could be said to have begun in earnest with the lobbying for its own independence in 1951).

References

Ahamed, Liaquat (2010). *Lords of Finance: 1929, The Great Depression, and the Bankers Who Broke the World*, Windmill Books.

Allen, Frederick Lewis (1931). *Only Yesterday: An Informal History of the 1920s*, Harper & Row.

Alston, Lee J., Wayne A. Grove, and David C. Wheelock (1994). "Why Do Banks Fail? Evidence from the 1920s," *Explorations in Economic History* 30, 409-31.

Basile, Peter F., John Landon-Lane, and Hugh Rockoff (2010). "Money and Interest Rates in the United States during the Great Depression," NBER Working Paper No. 16204.

Bernanke, Ben S. (1983). "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression," *American Economic Review* 73, June, 257-76.

Bernanke, Ben S. (1995). "The Macroeconomics of the Great Depression: A Comparative Approach," *Journal of Money, Credit and Banking*, 27, 1-28.

Bernanke, Ben S. and James, Harold (1991). "The Gold Standard, Deflation, and Financial Crisis in the Great Depression: An International Comparison." In R. Glenn Hubbard, ed., *Financial Markets and Financial Crises*. University of Chicago Press, 33-68.

Bernstein, Asaf, Eric Hughson, and Marc D. Weidenmier (2009). "Identifying the Effects of a Lender of Last Resort on Financial Markets: Lessons from the Founding of the Fed," *Journal of Financial Economics*, forthcoming.

Blum, John Morton (1959). *From the Morgenthau Diaries: Years of Crisis, 1928-1938*, Boston: Houghton Mifflin.

Board of Governors of the Federal Reserve System (1943 [1976]). *Banking and Monetary Statistics: 1914-1941.* Washington: Board of Governors.

Bordo, Michael (1985). "The Impact and International Transmission of Financial Crises: Some Historical Evidence, 1870-1933," *Revista di Storia Economica*, 2d ser., v. 2, 41-78.

Bordo, Michael D., Ehsan U. Choudri, and Anna J. Schwartz (2002). "Was Expansionary Monetary Policy Feasible during the Great Contraction? An Examination of the Gold Standard Constraint," *Explorations in Economic History*, Volume 39, January, pages 1-28.

Bordo, Michael D., and Barry Eichengreen (2003). "Crises Now and Then: What Lessons from the Last Era of Financial Globalization?" in *Monetary History, Exchange Rates and Financial Markets: Essays in Honour of Charles Goodhart*, Volume 2, Edward Elgar, 53-84.

Bordo, Michael D., and John Landon-Lane (2010). “The Lessons from the Banking Panics in the United States in the 1930s for the Financial Crisis of 2007-2008,” NBER Working Paper 16365.

Bordo, Michael D., and David C. Wheelock (2010). “The Promise and Performance of the Federal Reserve as Lender of Last Resort 1914-1933,” in this volume.

Bresnahan, Timothy F., and Daniel M.G. Raff (1991). “Intra-Industry Heterogeneity and the Great Depression: The American Motor Vehicle Industry, 1929-1935,” *Journal of Economic History*, 51, June, 317-332.

Bresnahan, Timothy F., and Daniel M.G. Raff (1992). “Technological Heterogeneity, Adjustment Costs, and the Dynamics of Plant Shut-Down Behavior: The American Automobile Industry in the Time of the Great Depression,” Harvard Business School Working Paper #92-097, June.

Brunner, Karl, and Allan H. Meltzer (1964). “The Federal Reserve’s Attachment to the Free Reserves Concept,” Subcommittee on Domestic Finance, Committee on Banking and Currency, Washington: United States House of Representatives, 88th Congress, 2nd Session.

Brunner, Karl, and Allan H. Meltzer (1968a). “What Did We Learn from the Monetary Experience of the United States in the Great Depression?” *Canadian Journal of Economics* 1, May, 334-48.

Brunner, Karl, and Allan H. Meltzer (1968b). “Liquidity Traps for Money, Bank Credit and Interest Rates,” *Journal of Political Economy* 76, 1-37.

Calomiris, Charles W. (1989). “Deposit Insurance: Lessons from the Record,” *Economic Perspectives*, Federal Reserve Bank of Chicago, May/June, 10-30.

Calomiris, Charles W. (1990). "Is Deposit Insurance Necessary? A Historical Perspective," *Journal of Economic History*, 50, 283-95.

Calomiris, Charles W. (1992). “Do Vulnerable Economies Need Deposit Insurance? Lessons from U.S. Agriculture in the 1920s.” In Philip L. Brock, ed., *If Texas Were Chile: A Primer on Bank Regulation*. San Francisco: The Sequoia Institute, 237-349, 450-458.

Calomiris, Charles W. (1993). "Financial Factors in the Great Depression,” *Journal of Economic Perspectives*, Spring, 61-85.

Calomiris, Charles W. (2000). *U.S. Bank Deregulation in Historical Perspective*, Cambridge University Press.

Calomiris, Charles W. (2009). “Banking Crises and the Rules of the Game,” Working Paper, Columbia Business School.

Calomiris, Charles W. (2010). “The Political Lessons of Depression-Era Banking Reform,” *Oxford Review of Economic Policy,* forthcoming.

Calomiris, Charles W., and Gary Gorton (1991). “The Origins of Banking Panics: Models, Facts, and Bank Regulation,” in R. G. Hubbard, ed., *Financial Markets and Financial Crises*, University of Chicago, 107-73.

Calomiris, Charles W., and R. Glenn Hubbard (1989). “Price Flexibility, Credit Availability, and Economic Fluctuations: Evidence from the U.S., 1894-1909,” *Quarterly Journal of Economics*, August, 429-52.

Calomiris, Charles W., and R. Glenn Hubbard (1995). "Internal Finance and Investment: Evidence from the Undistributed Profits Tax of 1936-1937,” *Journal of Business* , October 1995, 443-482.

Calomiris, Charles W., and Charles M. Kahn (1991). "The Role of Demandable Debt in Structuring Optimal Banking Arrangements," *American Economic Review* 81, 497-513.

Calomiris, Charles W., and Joseph R. Mason (1997). "Contagion and Bank Failures During the Great Depression: The June 1932 Chicago Banking Panic," *American Economic Review* 87, 863-83.

Calomiris, Charles W., and Joseph R. Mason (2003a). “Fundamentals, Panics and Bank Distress During the Depression,” (with Joseph Mason) *American Economic Review* 93, 1615-47.

Calomiris, Charles W., and Joseph R. Mason (2003b). “Consequences of Bank Distress During the Great Depression,” *American Economic Review* 93, 937-47.

Calomiris, Charles W., and Joseph R. Mason (2004). “How to Restructure Failed Banking Systems: Lessons from the U.S. in the 1930s and Japan in the 1990s” in *Governance, Regulation, and Privatization in the Asia-Pacific Region*, Takatoshi Ito and Anne Krueger, eds., University of Chicago Press, 2004, 375-420.

Calomiris, Charles W., Joseph R. Mason, and David C. Wheelock (2010). “The Reserve Requirement Increases of 1936-1937 and the Recession of 1937-1938,” Work in progress.

Calomiris, Charles W., and Andrew Powell (2001). “Can Emerging Market Bank Regulators Establish Credible Discipline: The Case of Argentina, 1992-99,” in *Prudential Supervision: What Works and What Doesn’t*, edited by Frederic S. Mishkin, 147-96.

Calomiris, Charles W., and Larry Schweikart (1991). "The Panic of 1857: Origins, Transmission, and Containment," *Journal of Economic History*, 51, 807-34.

Calomiris, Charles W., and David C. Wheelock (1998). “Was the Great Depression a Watershed for American Monetary Policy?” in M.D. Bordo, C. Goldin, and E.N. White, editors, *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century*, Chicago: University of Chicago Press, pp. 23-66.

Calomiris, Charles W., and Eugene N. White (1994). “The Origins of Federal Deposit Insurance,” in Claudia Goldin and Gary Libecap, eds., *The Regulated Economy: A Historical Approach to Political Economy*, University of Chicago, 145-88.

Calomiris, Charles W., and Berry Wilson (2004). "Bank Capital and Portfolio Management: The 1930s 'Capital Crunch' and Scramble to Shed Risk," *Journal of Business* 77, 421-55.

Calomiris, Charles W., and Berry Wilson (2010). "New York City Banks' Policies Affecting Ownership Structure, Capital Raising, and Listing in 1927-1931," Working Paper, Columbia University.

Carlson, Mark, Kris James Mitchener, and Gary Richardson (2010). "The Florida Banking Panic of 1929: Real Shocks, Runs and Intervention," Working Paper, University of California, Irvine.

Currie, Lauchlin (1934). "The Failure of Monetary Policy to Prevent the Depression of 1929-32," *Journal of Political Economy*, 42, April, 145-177.

Davis, Joseph, Christopher Hanes, and Paul Rhode (2007). "Harvests and Business Cycles in Nineteenth Century America," Working Paper.

DeLong, J.Bradford, and Andrei Shleifer (1991). "The Stock Market Bubble of 1929: Evidence from Closed-End Mutual Funds," *Journal of Economic History*, 51, September.

Dice, Charles Amos (1929). *New Levels in the Stock Market*, McGraw-Hill.
Demirguc-Kunt, Asli, Edward Kane, and Luc Laeven, eds.(2009). *Deposit Insurance Around the World,* Cambridge, Mass., MIT Press.

Ferderer, Peter (2007). The Rise of the American Over-the-Counters Markets, 1876-1924, Working paper, Macalaster College.

Field, Alexander J. (2003). "The Most Technologically Progressive Decade of the Century," *American Economic Review* 93, September, 1399-1413.

Fisher, Irving (1930). *The Stock Market Crash – And After*, Macmillan.

Fisher, Irving (1933). "The Debt-Deflation Theory of Great Depressions," *Econometrica* 1, pp. 337-57.

Fisher, Irving (1935). Testimony on the Banking Act of 1935, U.S. House of Representatives Committee on Banking and Currency, 74th Congress, 1st Session.

Flood, Mark (1991). "The Great Deposit Insurance Debate," *Federal Reserve Bank of St. Louis Review* 74, July-August, pp. 51-77.

Friedman, Milton, and Anna J. Schwartz (1963). *A Monetary History of the United States, 1867-1960*, Princeton University Press.

Gorton, Gary (1985). "Clearing Houses and the Origin of Central Banking in the United States," *Journal of Economic History* 45, 277-83.

Grossman, Richard S. (1994). "The Shoe That Didn't Drop: Explaining Banking Stability During the Great Depression." *Journal of Economic History,* 54, 654-82.

Hsieh, Chang-Tai, and Christina Romer (2006). "Was the Federal Reserve Constrained by the Gold Standard During the Great Depression? Evidence from the 1932 Open Market Purchase Program," *Journal of Economic History* 66, March, 140-176.

Kabiri Ali (2009). "Wall St. 1929: Modern valuation models, Money Market Arbitrage and Liquidity Black Holes," PhD Dissertation, Cass Business School, City University, July.

Lucia, Joseph L. (1985). "The Failure of the Bank of United States: A Reappraisal," *Explorations in Economic History* 22, 402-16.

Mason, Joseph R. (2001). "Do Lender of Last Resort Policies Matter? The Effects of Reconstruction Finance Corporation Assistance to Banks During the Great Depression," *Journal of Financial Services Research*, September 20, 77-95.

McCallum, Bennett T. (2001). "Inflation Targeting and the Liquidity Trap," NBER Working Paper 8225.

Meltzer, Allan H. (2003). *A History of the Federal Reserve, 1913-1951*, University of Chicago Press.

Miron, Jeffrey (1986). "Financial Panics, the Seasonality of the Nominal Interest Rate, and the Founding of the Fed," *American Economic Review* 76, 125-140.

Mowery, David, and Nathan Rosenberg (2000). "Twentieth Century Technological Change," in Stanley Engerman and Robert Gallman, eds., *The Cambridge Economic History of the United States*, 693-742.

Nicholas, Tom (2007). "Stock Market Swings and the Value of Innovation, 1908-1929," in *Financing Innovation in the United States, 1870 to the Present*, N.R. Lamoreaux and K.L. Sokoloff, eds., 217-45.

Nicholas, Tom, and Anna Scherbina (2010). "Real Estate Prices During the Roaring Twenties and the Great Depression," Working paper, Harvard Business School.

Orphanides, Athanasios, and Volker Wieland (2000). "Efficient Monetary Policy Design Near Price Stability," *Journal of Japanese and International Economics* 14, 327-365.

Rappaport, Peter, and Eugene White (1993). "Was There a Bubble in the 1929 Stock Market," *Journal of Economic History*, 53, September, pp. 549-74.

Rappaport, Peter, and Eugene White (1994). "Was the Crash of 1929 Expected?" *American Economic Review*, 84, March, pp. 271-81.

Richardson, Gary, and William Troost (2006). "Monetary Intervention Mitigated Banking Panics During the Great Depression? Quasi-Experimental Evidence from the Federal Reserve District Border in Mississippi, 1929 to 1933," NBER Working Paper No. 12591.

Sprague, Oliver M. W. (1910). *History of Crises under the National Banking System*, National Monetary Commission.

Temin, Peter (1976). *Did Monetary Forces Cause the Great Depression?* W.W. Norton.

Temin, Peter (1989). *Lessons from the Great Depression,* MIT Press.

U.S. Department of Commerce (1970). *Historical Statistics of the United States.*

Van Horn, Patrick (2008). "Excess Reserves During the Great Depression: Evidence from the Central Money Market of New York City, 1929 to 1932," Working paper.

Wallis, John Joseph, and W. E. Oates (1998). "The Impact of the New Deal on American Federalism," in M.D. Bordo, E. Goldin, and E.N. White (Eds.) *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century*. Chicago: University of Chicago Press.

Wheelock, David C. (1990). "Member Bank Borrowing and the Fed's Contractionary Monetary Policy during the Great Depression," *Journal of Money, Credit and Banking*, 22, November, 409-426.

Wheelock, David C. (1991). *The Strategy and Consistency of Federal Reserve Monetary Policy, 1924-1933*, Cambridge: Cambridge University Press.

Wheelock, David C. (1992). "Monetary Policy in the Great Depression: What the Fed Did and Why," *Federal Reserve Bank of St. Louis Review*, 74, 3-28.

White, Eugene N. (1983). *The Regulation and Reform of the National Banking System, 1900-1929*, Princeton University Press.

White, Eugene N. (1984). "A Reinterpretation of the Banking Crisis of 1930," *Journal of Economic History,* 44, 119-38.

White, Eugene N. (1985). "The Merger Movement in Banking, 1919-1933," *Journal of Economic History*, 45, June, 285-291.

White, Eugene N. (1990). "The Stock Market Boom and Crash of 1929 Revisited," *Journal of Economic Perspectives*, 4, Spring, 67-83.

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

Wicker, Elmus (1980). "A Reconsideration of the Causes of the Banking Panic of 1930." *Journal of Economic History,* 40, 571-83.

Wicker, Elmus (1996). *The Banking Panics of the Great Depression.* Cambridge University Press.

Wigmore, Barrie A. (1987). "Was the Bank Holiday of 1933 a Run on the Dollar Rather than the Banks? *Journal of Economic History* 47, 739-56.

Figure 1

The Loan Market with and without the Fed

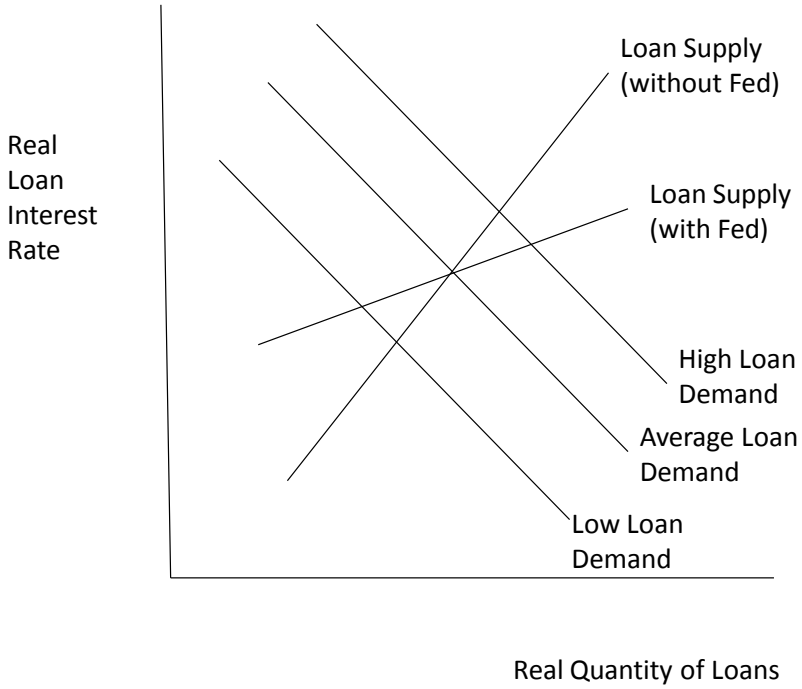

Figure 2

Loan Supply Contraction and Rising Real Loan Rates

During the Depression

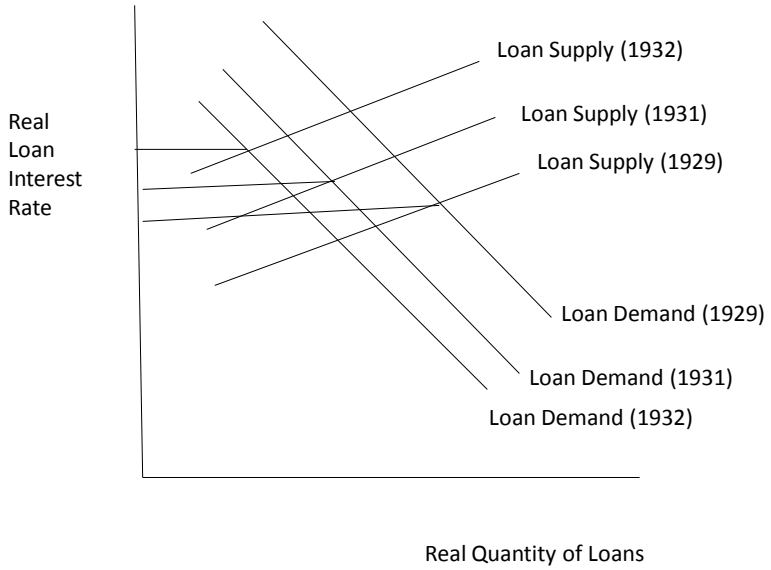

Table 1

A Representative Bank's Balance Sheet

Assets		Liabilities and Net Worth	
80	Loans	70	Deposits
20	Liquid Reserves	30	Equity Capital
100	Total Assets	100	Total Liabilities and Net Worth