

NBER WORKING PAPER SERIES

RELATIONSHIP LENDING AND THE GREAT DEPRESSION:
MEASUREMENT AND NEW IMPLICATIONS

Jon Cohen
Kinda Cheryl Hachem
Gary Richardson

Working Paper 22891
<http://www.nber.org/papers/w22891>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2016

We thank Gabe Chodorow-Reich, Fabrizio Perri, Johannes Wieland, and participants at the NBER SI 2016 Monetary Economics meeting for helpful comments. We also thank Daniel Tracht and Joseph Johnson for invaluable research assistance. Kinda Hachem gratefully acknowledges financial support from Chicago Booth. Part of this research was conducted while Gary Richardson was at the Federal Reserve Bank of Richmond, serving as the official Historian of the Federal Reserve System. The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Richmond, the Federal Reserve System, or the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Jon Cohen, Kinda Cheryl Hachem, and Gary Richardson. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Relationship Lending and the Great Depression: Measurement and New Implications
Jon Cohen, Kinda Cheryl Hachem, and Gary Richardson
NBER Working Paper No. 22891
December 2016
JEL No. E44,G01,G21,L14,N22

ABSTRACT

The Great Depression remains ground zero for studying the non-monetary effects of financial crises. Despite the abundant scholarship on the period, lack of disaggregated data on lending activities has constrained our ability to measure the impact on the real economy of a collapse in long-term lending relationships. We propose here a novel way to extract cross-sectional differences in relationship lending from geographically aggregated financial statements. We find that the banking crises of the early 1930s, by destroying these relationships and the soft yet crucial information garnered from them, explain one-eighth of the economic contraction observed during the Depression. This effect comes specifically from small bank failures which alone explain one-third of the Depression. Large bank failures, on the other hand, were accompanied by a reallocation of deposits towards surviving relationship lenders, leading to economic gains which mitigated the overall negative impact of the banking crises. We show that ignoring cross-sectional differences in continuing relationships on the eve of the Great Depression understates by a factor of 2 the fall in economic activity directly attributable to the banking panics of the early 1930s. We also show that the rebuilding of lending relationships in the mold of those that existed in the 1920s was an important determinant of cross-sectional differences in economic performance during the 1937-38 recession.

Jon Cohen
Department of Economics
University of Toronto
Max Gluskin House
150 St. George Street, 322
Toronto, Ontario M5S 3G7
Canada
jon.cohen@utoronto.ca

Gary Richardson
Department of Economics
University of California, Irvine
3155 Social Sciences Plaza
Irvine, CA 92697-5100
and NBER
garyr@uci.edu

Kinda Cheryl Hachem
University of Chicago
Booth School of Business
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
kinda.hachem@chicagobooth.edu

1. Introduction

Debates among economists about the importance of financial intermediation and the implications of its interruption for business cycle fluctuations, such as those associated with the financial crises in 1997 and 2008, have a long and fractious history. Many of the modern issues in this debate stem from Ben Bernanke's pioneering paper on the Great Depression (Bernanke, 1983), in which he argues that the banking panics of the early 1930s increased the cost of credit intermediation, disrupted lending, and directly deepened the downturn, in addition to reducing the money supply and generating deflation in the way described by Fisher (1933) and Friedman and Schwartz (1963). Since then, a large literature focusing sharply on the Great Depression in the U.S. has emerged and attempted to estimate the size and significance of disruptions to the credit channel during the contraction of the 1930s. Calomiris and Mason (2003), Richardson and Troost (2009), Ziebarth (2013), and Carlson and Rose (2015) all find evidence that the collapse of commercial banking adversely affected the availability of credit and the activities of firms. However, a number of economists, most notably Cole and Ohanian (1999, 2004, 2007), reach exactly the opposite conclusion: the collapse of the United States' commercial banking system had few real economic effects.

Why does the debate continue to rage for an event as large and intensively studied as the Great Depression? The answer, we argue, resides in the challenges associated with data limitations and with issues of identification. An assumption – sometimes articulated, sometimes implicit – that underlies all of the papers on this topic is that banking crises have real effects because they destroy or, at the very least, prevent the immediate redeployment of some critical input into the bank lending process that keeps the cost of credit down. A natural candidate for this input is the soft information that banks acquire about the quality of their borrowers during multi-period lending relationships. Because soft information is difficult to transfer to another bank, failures of lenders who have accumulated knowledge of their borrowers through on-going relationships are more likely to disrupt the flow of funds than failures of lenders who are just starting potential relationships or of those who only provide credit against easy-to-evaluate collateral. The key to unlocking the non-monetary effects of bank failures on economic activity is, therefore, to be found in the proper measurement of continuing relationships on the eve of the Great Depression. However, because we lack the kind of microeconomic data for the earlier

period that economists use in modern analyses of relationship lending, it has been difficult to isolate these relationships.¹ As a result, the majority of papers that attempt to link banking crises to economic activity during the Great Depression have been compelled to assume that most commercial banks were relationship lenders, that most of the loans they made were relationship loans, and that the effect of the failure of any one of these banks was the same as the effect of any other. This, as we show, was not the case.

To be more precise, commercial banks in the 1920s and 1930s made loans to manufacturing and trading firms with which they had, or would potentially have, long-run repeated relationships. They also extended overnight or call loans to brokers and individuals who used the proceeds to purchase securities. The former loans facilitated production and exchange, yielded information that improved the allocation of resources, and were the type of loans whose destruction could have provoked a reduction in potential output. The latter, on the other hand, did very few of these things but constituted roughly 38% of all loans of commercial banks and nearly half of all loans in reserve and central reserve cities which, in addition to being financial centers, were centers of industry and trade and which are often the most influential observations in the panel datasets employed by economic researchers. Banks curtailed both types of lending in the early 1930s when the banking crises hit. By lumping together all loans and ignoring potential differences in their impact on the economy, we run the risk of misstating the effect of a reduction in true relationship loans and hence misstating the non-monetary effects of the banking crises.

To overcome these difficulties, we need to develop a way to extract from aggregated data disaggregate information on loan types and timing – which is precisely the objective of this paper. In short, with theory as our guide, we are able to unbundle the financial statements of banks, make a sharp distinction across jurisdictions between continuing relationship loans and

¹ Modern empirical analyses of relationship lending, typically making use of microdata at the bank-firm level, rely on indicators such as the physical distance between the firm and the bank, the duration of the relationship, and the number of other banks the firm borrows from to measure the strength of a relationship. See, for example, Berger and Udell (1995), Ongena and Smith (2001), Elsas (2005), Chodorow-Reich (2014), Gobbi and Sette (2015), and the references therein. Since this type of matched microdata does not exist for a representative sample of banks and firms in the 1920s and 1930s, it is impossible to use such indicators to capture the decline in relationship lending during the Great Depression and thus to determine its impact on economic activity. An alternative approach that has proved fruitful in a variety of different contexts by Romer and Romer (1989, 2004, 2015) is to make use of historical narratives. While it may be possible to gain some insight into lending practices through the narrative approach, we lack a source which provides narratives of local conditions in the Depression-era U.S. that is consistent across both time and space.

other types of credit activity, and, as a result, demonstrate that (1) the rise and fall of continuing lending relationships contributed significantly to the intensity and duration of the Great Depression and (2) a detailed account of how the revival of these relationships varied across locales post-depression illuminates the pattern and pace of economic recovery.

We start by developing a new measure of continuing lending relationships that can be calculated using only data from geographically aggregated financial statements. In theoretical work on relationship lending and the transmission of monetary policy, Hachem (2011) shows that the loan rates charged by lenders in the middle of a relationship will, on average, be less responsive to changes in bank funding costs than the loan rates charged by other lenders. The new, price-based, measure we develop here builds on this insight. In particular, we use data from the consolidated balance sheets and income statements of national banks in the 1920s and 1930s – reported by the Comptroller of the Currency at a semi-annual frequency for states and reserve cities – to infer the weighted average loan rate in each location at each point in time. We then calculate the elasticity of the loan rate with respect to the discount rate in each location. The twelve Federal Reserve Banks at the time had sufficient latitude to operate largely independent discount windows so there was variation in discount rates across districts, particularly at the onset of the Great Depression (Richardson and Troost, 2009). To control for location-specific differences in rates of return or their responsiveness, we also calculate the elasticity of securities returns with respect to the discount rate in each location and net it out from the loan rate elasticity. We take variations in this net elasticity measure as indicative of differences in the nature and intensity of relationship lending, where areas with relatively less elastic loan rates were those where continuing relationships were most prevalent.

Our first set of results establish the importance of relationship lending for understanding the real effects of banking distress in the early 1930s. We show that the marginal impact of bank suspensions on economic activity was much higher in areas with more continuing relationships.² We then show, through counterfactual analysis, that small bank failures on their own can generate roughly one-third of the economic contraction observed during the Great Depression,

² To put this finding into context, Cole and Ohanian (2007) have argued that the unconditional correlation between bank failures and economic outcomes is weak, even at the state level, thus raising questions about the link between finance and the real economy. Our results essentially show that the correlation becomes much more informative when conditioned on cross-sectional differences in the nature and intensity of lending relationships.

once the negative effects of destroying continuing lending relationships are taken into account. Running the counterfactual with both small and large banks, we find that distress at large banks actually played a mitigating role. In particular, there appear to have been economic gains from the reallocation of deposits toward surviving relationship lenders following suspensions of large banks. On net, we estimate that failures of national banks – small and large – can generate about one-eighth of the economic contraction observed during the Great Depression. We also demonstrate that failure to control for cross-sectional differences in continuing lending relationships on the eve of the crisis would have understated by a factor of 2 the fall in economic activity directly attributable to the banking panics of the early 1930s.

Our second set of results concern the role of relationship lending in recovery. We show that cross-sectional differences in the rebuilding of lending relationships destroyed during the Great Depression are important for understanding cross-sectional differences in economic performance during the 1937-38 recession. This is an interesting and important finding because, very much in the spirit of Rajan and Ramcharan (2016), it suggests that the effects of the Depression persisted long after recovery had set in. While Rajan and Ramcharan (2016) focus on the impact of bank failures in the 1930s on banking concentration in the post-WWII period, we concentrate on the short yet fierce recession of the late 1930s. We find that areas that rebuilt the types of continuing relationships observed in the 1920s fared better during the 1937-38 recession than otherwise similar areas that failed to rebuild. In contrast, areas that rebuilt relationships but, based on their reserve holdings, appear to have been less committed to these relationships, ironically fared worse during this recession than otherwise similar areas that failed to rebuild. These findings would seem to suggest, first, that on-going lending relationships played a significant role in the upswing much as they did in the decline and, second, that the negative consequences of a severe downturn are likely to linger long after the event itself is history.

While, in keeping with Bernanke and others, we focus on the Great Depression – perfectly reasonable given the dimensions of the downturn and the scope of the banking panics – the value of our methodology transcends resolution of data limitations for the 1920s and 1930s. Policymakers working in real time, often in crisis situations, usually have only aggregate data to provide them with information about what is happening at more disaggregated levels. Our method permits them to use these aggregates to extract this more detailed information.

The rest of the paper proceeds as follows. Section 2 discusses the lending activities of Depression-era commercial banks and reviews the theory that leads to our new price-based measure of continuing lending relationships. Section 3 describes the data sources used to construct this measure. Section 4 establishes that the constructed measure is not spurious and does indeed capture variation in the nature and intensity of relationship lending. Sections 5 and 6 present the key empirical results, with Section 5 using cross-sectional differences in continuing relationships on the eve of the Great Depression to pinpoint the real effects of banking distress in the early 1930s and Section 6 demonstrating how relationship rebuilding in the aftermath of the Depression affected economic performance during the 1937-38 recession. Section 7 concludes.

2. Methodology

In this section, we propose two new ways to measure relationship lending in the absence of micro-level data. The first, a quantity-based indicator, permits us to distinguish between different types of bank loans and, at the same time, helps us flesh out the nature of relationship lending. While this indicator is preferable to existing approaches in the literature on the Great Depression, it does not, on its own, permit us to distinguish between new and continuing relationships. As we will explain, the latter are particularly important for picking up the links between banks and the real economy. It is with this in mind that we construct a novel second measure: a price-based indicator of continuing lending relationships.

2.1 Quantity-Based Indicator

In an ideal world, an obvious and straightforward way to construct a quantity-based indicator of relationship lending would be simply to break down all outstanding bank loans into their various components, identify those that we can confidently regard as relationship-dependent, and sum them. The problem, of course, is that the world is far from ideal: loans are reported in coarse categories, categories may change over time, etc. Nevertheless, it is still possible to gain some insights from the quantity data that are available.

The annual report of the Comptroller of the Currency provides information on bank balance sheets in the 1920s and 1930s in a format that is consistent across geographic locations.

Up until 1928, the Comptroller broke down loans into four categories: real estate loans, loans on securities, uncollateralized loans, and loans collateralized by personal security. Based on the Comptroller's description of personal security, the last category consists primarily of loans secured by difficult-to-evaluate collateral such as goods in process of production, warehouse receipts, and, in the case of farm loans, collateralized by future crops.

Since relationship loans are, by definition, those based on a close relationship between the lender and the borrower, it would seem reasonable to consider the last two categories in the Comptroller's report – uncollateralized loans and loans collateralized by personal security – as quintessentially relationship-dependent. Indeed, in addition to being largely uncollateralized, both categories consisted primarily of short-term commercial loans which were often sequenced to finance longer-term projects. Sequencing allowed banks to incorporate information acquired during the initial loan period into future credit terms, including, of course, the possibility of discontinuing the lending relationship. Each short-term loan matured in less than a year, typically less than 180 days, and in the case of outstanding balances drawn from lines of credit, was repayable on demand. Such loans also served as a critical source of working capital for business, financing goods in the process of production, processing, shipment, and sale.

In contrast, real estate loans and loans on securities – the first two categories in the Comptroller's report – are not quintessential relationship lending. Prior to the 1930s, real estate loans had relatively short maturities (3 to 5 years) so refinancing was common. However, unlike the commercial loans described above, mortgages were often not refinanced with the same lender. As for loans on securities, they consisted largely of call loans to brokers and also some call loans directly to firms and individuals, often to facilitate the purchase of stocks, bonds, or other well understood securities. Contemporary observers were unequivocal in their view that loans on securities were information-lite (e.g., Currie, 1931).

Therefore, using the Comptroller's breakdown, we can obtain a reasonable quantity-based indicator of relationship lending activities for the 1920s. The left panel of Figure 1 plots this indicator, defined formally as the fraction of national bank loans that were either unsecured or secured by personal collateral in the 1920s (solid black line). The fraction varies from a high

of 74% in June 1921 to a low of 57% in June 1928 and the split between uncollateralized loans and loans on personal collateral is about 5:1 throughout the period.³

Unfortunately, starting in 1929, the Comptroller reduced the number of loan categories from four to three: loans on securities, real estate loans, and all other loans. All loans eligible to be used as collateral at the Fed's discount window appeared in the "all other loans" category so the Comptroller may have also refined the cutoffs between categories in line with the Fed's 1929 direct action campaign. Despite the changes, the "all other loans" category is still informative about relationship loans much as we regard the sum of unsecured loans and loans on personal collateral as representative of relationship lending. The left panel of Figure 1 shows the fraction of national bank loans that qualified as "all other loans" in the 1930s (dashed black line). For the U.S. as a whole, the fraction of "all other loans" in June 1929 is roughly of the same magnitude as the fraction of loans secured by either no collateral or personal collateral in June 1928. Moreover, and perhaps of greater significance, the cross-sectional correlation between the fraction of "all other loans" in June 1929 and the fraction of loans secured by either no collateral or personal collateral in June 1928 is 0.75, large enough to suggest that the two categories are reasonably similar.

Although the "all other loans" category was not available at a geographically disaggregated level prior to 1929, an aggregate across member banks in leading cities was published through the 1920s and 1930s. The right panel of Figure 1 plots this aggregate. The high correlation between the quantity of relationship loans as measured by this aggregate and the quantity of loans on securities, especially from 1929 to 1933, suggests the need for caution when using these data to estimate the marginal impact of a relationship loan. To understand why, consider a regression of the form $Y = \beta_{s+r}(L_s + L_r)$, where Y is a measure of economic activity, L_s is lending on securities, and L_r is relationship loans. The academic literature on the Great Depression has run versions of this regression using various techniques but, to the best of our knowledge, all have in common the use of the total loan book ($L_s + L_r$) as a regressor, not just the component most closely related to relationship lending (L_r). In the simplest case where relationship loans are homogeneous, the regression that correctly identifies the marginal impact

³ The moderate decline in the ratio of relationship loans to total banking assets between June 1921 and June 1928 reflects an increase in the denominator, not a decrease in the numerator.

of a relationship loan would take the form $Y = \beta_r L_r$. With L_r and L_s highly correlated, β_{s+r} will be a downwardly biased estimate of β_r . This is easy to see in the extreme case where $L_s = L_r$. Then the t-statistics for β_{s+r} and β_r are identical but $\beta_r = 2\beta_{s+r}$.

The preceding discussion demonstrates that failure to distinguish between different types of bank loans will understate the *marginal* impact of relationship lending on the real economy. What about the *aggregate* impact? Returning to the example with $L_s = L_r$, there will be no difference in the aggregate impact because $L_s + L_r$ overstates the quantity of relationship lending by exactly the same factor that β_{s+r} understates the marginal impact of a relationship loan. However, both the marginal impact and the aggregate impact will be inadequately captured by the simple regressions described above if the loans that matter most for the fortunes of local businesses are actually a subset of L_r that moves independently of L_s – that is, if relationship loans are heterogeneous, not homogenous. As we argue in the next subsection, within the set of relationship loans, it is crucial to distinguish between those that are rolled over from one period to the next (i.e., loans made as part of a continuing relationship) and first-time relationship loans. Since there is no such division in the Comptroller’s report, quantity-based measures are unable to capture this distinction and thus necessitate a new approach.

2.2 Price-Based Indicator

Why is it so important to distinguish between new and continuing relationships? The answer is simple – the latter embodies the soft information on borrower quality that is vital to bankers when extending credit to businesses and is accumulated over time through repeated interactions with their clients. Bank failures, such as those that ravaged the U.S. financial system in the early 1930s, destroy this knowledge, disrupt the flow of funds to borrowers and, because it takes time to re-establish these relationships, also impact the pace of recovery. In short, we need to be able to isolate these continuing relationships if we want to capture fully the effects of financial crises on the real economy. As we will now explain, a price-based indicator gives us this ability.

Our indicator builds on a theoretical result in Hachem (2011) that shows how the pricing of a loan changes over the course of a relationship. Appendix A presents a simplified version of

the model. The key insight, for our purposes, is that the returns on continuing relationship loans are less responsive to changes in bank funding costs than are the returns on new relationship loans or the returns on non-relationship (i.e., transactional) loans. While loan returns are not directly reported by the Comptroller, geographically aggregated bank income statements are available. Further discussion of these statements and how we use them to approximate returns is deferred until Section 3.

The intuition behind Hachem's result is as follows. A relationship lender obtains private information about his borrower's abilities during an initial interaction in which credit is extended and, based on this information, determines whether the borrower is good enough to retain. Other banks can infer that retained borrowers are good but, without having participated in the first interaction, they cannot unravel the relationship lender's information set and infer exactly how good a particular borrower is. In other words, competition from these less informed banks constrains, but does not eliminate, the relationship lender's ability to extract rents from the borrowers he retains. A modest risk-shifting problem is then enough to compel the relationship lender to share surplus with some of these retained borrowers in order to incentivize higher repayment rates.⁴ Crucially, the surplus-sharing is shown to take the form of policy-invariant loan rates over intermediate ranges of the policy rate, where the policy rate is the measure of bank funding costs in the model. In this way, the presence of continuing relationships – that is, the financing of retained borrowers by their relationship lenders – lowers the responsiveness of the average loan rate to the policy rate.

Interestingly, this insight garners support from an informal survey of country banking practices conducted by Ford (1928) in Northern Texas. While Ford does not talk about risk-shifting or incentive compatibility, he does observe that, once banks identify good borrowers, they attempt to do everything they can to keep them as customers, a large part of which entails maintaining a constant loan rate.

The environment of the 1920s is consistent with the primitives of Hachem's model, allowing us to use the model's prediction about loan rate responsiveness for identification. First, virtually every county had multiple banks which meant that each lender had to contend with

⁴ Without competition, a more severe risk-shifting problem would be necessary to prompt surplus-sharing, at which point the lender may prefer to demand collateral upfront rather than learn over time.

competitors. Second, commercial loans tended to be uncollateralized and short-term with the possibility of being rolled over (see Subsection 2.1), also a feature of the model. Third, law review articles from the period indicate that taking a borrower to court was very costly in virtually any jurisdiction, leaving incentive compatible contracts as the most effective way to address risk-shifting. As explained above, the key prediction of the model for our purposes is that, in the world just described, the incentive compatible contract offered by a continuing relationship lender reduces the responsiveness of the average loan rate to the policy rate. Appendix A also translates this prediction into a statement about elasticities. In particular, the elasticity of the average loan rate with respect to the policy rate will be lower in an area with more continuing relationships. Accordingly, once we control for regional differences that do not map into different values of the model parameters, we can use cross-sectional differences in loan rate elasticity to identify cross-sectional differences in continuing relationships.

With the theoretical groundwork established, we propose to calculate the elasticity of loan returns with respect to the discount rate, which we take to be the cost of funds to banks during the period we study. From the foundation of the Fed through the 1930s, the discount window was operated in a manner that made the discount rate the marginal cost of funding a commercial loan. More precisely, the Fed allowed, expected, and in fact needed member banks to discount commercial loans with Reserve Banks. It expected banks to do this to accommodate seasonal and cyclical peaks in demands for credit – one of the principle motivations for the creation of the Federal Reserve System. The Fed also needed banks to discount loans because those discounts were its primary source of earned income. From the early 1920s through the early 1930s, the Fed was not allowed to hold government bonds, corporate securities, derivatives, or equities so, to cover its operating costs, the Fed encouraged member banks to use the discount window. The upshot: banks that wanted to expand their commercial loan book knew they could always raise the funds for those loans by discounting them at the Fed, making the discount rate the relevant cost of funds.⁵

⁵ In principle, loans on securities were not discountable at the Fed, making the discount rate the relevant cost for relationship lending alone. In practice, however, the Fed expressed concern about banks using discount loans to invest in securities, suggesting that the discount rate was also a relevant cost of funds for security holdings by banks.

Formally, the elasticity of loan returns with respect to the discount rate comes from estimating equations of the form:

$$\log(\text{ReturnOnLoans}_{i,t}) = \alpha_i^l + \beta_i^l \log(\text{DiscountRate}_{i,t}) + \varepsilon_{i,t}^l$$

where i denotes location, t denotes time, and β_i^l is the relevant elasticity.

Since we ultimately want to compare elasticities across regions, we need to control for the possibility that some differences in loan returns are driven by considerations outside the model. The most important of these considerations is liquidity management. Banks in the model have no need for reserves which is obviously an abstraction from the real world. This abstraction only matters for our analysis insofar as liquidity management differs across regions. The interwar period in the U.S. involved regional differences in interbank networks (Mitchener and Richardson, 2016) as well as differences in lender of last resort policies across Federal Reserve districts (Richardson and Troost, 2009), both of which could lead to cross-sectional differences in liquidity management. It is important to control for these differences because discount rate changes in the 1920s and 1930s usually prompted banks to alter their reserve holdings first and the allocation of funds across other assets – namely, loans and securities – second. The more reserves a bank holds, the lower its interest income unless, of course, additional income is earned by reaching for yield on loans and securities. In this case, we may see higher reserve holdings as well as higher returns on loans and securities in areas with weaker interbank networks and/or less accommodative lender of last resort policies. We may also see more elastic returns on loans and securities in these areas as changes in the discount rate lead to bigger reserve adjustments.

The above discussion, by outlining the issue to be addressed, also signals the path forward. As long as cross-sectional differences in loan rate elasticity not explained by the model also show up as cross-sectional differences in the elasticity of securities returns, we can net the former from the latter to isolate the effect of relationship lending. To this end, we calculate the elasticity of securities returns with respect to the discount rate:

$$\log(\text{ReturnOnSecurities}_{i,t}) = \alpha_i^s + \beta_i^s \log(\text{DiscountRate}_{i,t}) + \varepsilon_{i,t}^s$$

We then define the net elasticity for location i as:

$$NE_i \equiv \beta_i^s - \beta_i^l$$

A higher value of NE_i means loan rates respond less than interest rates on other financial products when the policy rate changes. We thus take higher values of NE_i to mean that more continuing relationships are present. Section 4 will provide further validation by showing that the only statistically significant predictor of our net elasticity metric is a variable that reasonably proxies a parameter in the model.

3. Data

Our paper analyzes a series of related data panels, all of which were previously constructed from original microdata sources, many of which no longer exist. In this section, we describe the panels and explain how we use them.

Bank Balance Sheets Data on commercial banks come from the Annual Reports of the Comptroller of the Currency. These reports contain tables that indicate the balance sheets of commercial banks with national charters aggregated by Federal Reserve district, state, and major municipalities (principally financial centers then termed reserve cities). National bank assets fall into several key categories. Of particular importance are earning assets which include loans on securities, all other loans (essentially relationship loans as explained in Subsection 2.1), government bonds, and other securities.

Bank Income Statements The Comptroller of the Currency also published earnings and expense statements of nationally chartered banks aggregated at the district, state, and city level twice each year. One table reports data for the months of January through June. The other table reports data for the months of July through December. These tables enable us to calculate average earnings on loans and on securities semi-annually. To the best of our knowledge, the literature has done little with the income statements from this period.

We calculate the lending rate by dividing earnings for a period by the stock of loans at the end of that period. Ideally, we would like to use only the earnings on relationship loans and the stock of relationship loans but the Comptroller's report does not break interest income down

by type of loan for any of the dates in our sample. As a result, we can only use total interest earned on loans relative to total loans to calculate loan returns.⁶

We calculate the securities rate in the same manner as the loan rate (i.e., dividing earnings for a period by securities holdings at the end of that period). Prior to 1926, however, the Comptroller reported only the sum of earnings on loans and securities. Therefore, for 1921 through 1925, we estimate earnings on securities by multiplying market yields on securities (listed in trade publications) by the stock of securities. We also estimate interest earned from balances at other banks. We then subtract these two estimates from the total interest income contained in the Comptroller's report (also removing interest on Fed securities) to get our estimate of loan income. As shown in Figure 2, applying this procedure to data from 1926 to 1929 delivers predictions that line up very well with actuals.

Policy Rates Market interest rates and Federal Reserve policy rates appear in Banking and Monetary Statistics, 1914 to 1941. This tome recapitulates information previously published in annual and monthly reports of the Federal Reserve Board and Fed District Banks as well as trade publications such as the Commercial and Financial Chronicle and the Wall Street Journal. Our principal policy rate is the discount rate charged by each Federal Reserve Bank. These rates often changed during the six-month-periods that make up the time units in our panels. When the rates changed within the period, we calculate an average time-weighted rate in effect for the period. This weighted rate is the sum of the interest rate multiplied by the number of days in which it was in effect divided by the total number of days in the period.

Economic Outcomes We use retail sales as a measure of economic outcomes since it is generally thought of as a contemporaneous indicator (rather than construction contracts or business failures which are leading and lagging indicators respectively). Subsection 5.4 explains how a back-of-the-envelope calculation can be used to scale up the national retail sales effects we estimate into a GNP effect. We obtain retail sales for each location in our sample by aggregating the appropriate counties in the Census of Business. This census is available for the

⁶ This will tend to overstate the elasticity of loan returns and thus bias the net elasticity calculation against us. One very rough way to approximate interest earned from loans on securities is to use the interest rate on New York brokers' loans scaled by the relative interbank deposit rate for each district. Preliminary results suggest that purging loan returns of the resulting approximation and recalculating the net elasticities does not have a major impact on the relative ranking of our locations.

years 1929, 1933, 1935, and 1939. A subsample of the firms from the 1935 census was also surveyed again in 1937. The 1937 survey reports what the 1935 results would have been had they been based on the subsample which, in turn, permits us to scale up the 1937 results and approximate total retail sales for 1937.

Locations We can construct complete time series for 82 locations: 33 reserve cities, 31 states (net of reserve cities) each fully contained within a single Federal Reserve district, 12 parts of states (net of reserve cities) each fully contained within a single district, and 6 district residuals. We want locations that are fully contained in a single district because, as noted earlier, discount rates varied across districts and we need to use the relevant discount rate for each location when calculating net elasticity. The issue arises because some states are split across two districts. For districts that have only one split state, we can start with the district-level data then subtract the city-level data as well as the state-level data for states fully contained in that district to isolate the part of the split state contained in the district. We can then subtract the part we isolated for this state from the state's total in order to ascertain the part of the state that is contained in the other district. As long as this other district does not have more than two split states, we can repeat the process to back out any additional splits in the other district. For some districts, there are too many split states to be fully identified by this iterative procedure so we define district remainders to absorb those cases.

4. Net Elasticity Estimates

Figure 3 plots the net elasticities estimated using data from 1921 to 1929 and compares them to the quantity-based indicator averaged over the 1920s. A majority of locations have positive net elasticity, meaning that their loan rates are less elastic than their returns on securities. This is what the theory predicts when loans are more relationship-based than securities purchases. We can also see that the correlation with the quantity-based measure is positive: areas with more relationship loans as indicated by the quantity-based measure also tended to have more continuing relationships as indicated by the price-based (net elasticity) measure.

There are, however, a few locations with negative net elasticity (i.e., loan rates more elastic than returns on securities). How are they to be interpreted? Recall from Section 3 that we

have to use an approximation to separate interest income into income on loans and income on securities for the pre-1926 period. If we were to dispense with the approximation and calculate net elasticities using only actuals from 1926 to 1937 (stopping in 1929 would be too short), we would find that all net elasticities are positive. However, because we want to separate the 1920s from the 1930s, we cannot use only actuals from 1926 to 1937. The comparison does, nevertheless, shed some light on the question at hand. The market yields that we use to approximate interest from securities in the pre-1926 period are averages over different issues (i.e., the yield on municipal bonds is an average over multiple municipalities) and is, therefore, likely to be too smooth, understating fluctuations in securities income and thus overstating fluctuations in loan income. Indeed, institutional investing advanced during the 1920s as it sought to arbitrage differences in local returns such that, by the end of the decade, there was more co-movement in securities returns across locations than there had been at the beginning of the decade. This means that our approximation procedure will be a bit noisier when applied to the early 1920s than the late 1920s (see Figure 2 for the high accuracy in the late 1920s). However, we are ultimately concerned with the relative not the absolute positions of our various locations (e.g., Dallas is more relationship-intensive than Galveston) and we have no reason to believe that the precision of our approximations will vary systematically across locations. We can therefore interpret the presence of negative net elasticities as a level effect and, as such, it does not compromise our results.

The rest of this section is devoted to providing more formal evidence that our net elasticity variable captures variation in the nature and intensity of relationship lending, not a spurious pattern in the data.

As derived in Subsection 2.2, our price-based approach to measuring continuing relationships is based on the idea that differences in loan rates are driven by inherent differences in lending practices, not by differences in funding practices where deposit rates shape loan rates. We can test this. In particular, using semi-annual data, available from 1925 to 1929 inclusive, we can test the hypothesis that deposit rates did not Granger-cause loan rates in the 1920s for each location in our sample. We find that the average p-value from these tests is 0.23 and the median is 0.11. In other words, for at least half of the locations in our sample, we accept the hypothesis that loan rates were not Granger-caused by deposit rates. Furthermore, there is a mildly positive

cross-sectional correlation between the p-values of the Granger tests and our net elasticity estimates for the 1920s (NE_{20}): 0.2 when all locations are used; 0.35 when only locations with $NE_{20} > 0$ are used. While it would be unwise, given the small sample for the Granger tests, to over-emphasize this correlation, it is at least suggestive that loan rates were less likely to be Granger-caused by deposit rates in precisely the locations where our price-based indicator finds the most evidence of continuing lending relationships.

We can also confirm that our price-based indicator did not somehow become divorced from the theory during the empirical implementation by seeing what, if anything, predicts it. To this end, Table 1 reports the results of a cross-sectional regression of our net elasticity measure in the 1920s on a broad set of controls, namely economic indicators from the 1920 census, banking variables in 1920, district fixed effects, and a dummy variable for reserve cities. Similar regressions are also run using loan rate elasticity and the elasticity of securities returns as the dependent variable instead of net elasticity.

We find that locations where banks were smaller and/or more reliant on deposits for funding in 1920 tended to have less elastic loan rates during the 1920s. Loan rates also tended to be less elastic in locations that registered higher value-added per manufacturing establishment in 1920. It is important to remember that none of these statements are causal. In particular, the negative correlation between productivity and loan rate elasticity could be because long-standing relationships fostered only the most productive firms in the years leading up to the 1920 census. Moreover, none of the variables that are statistically significant predictors of loan rate elasticity retain their significance when we consider net elasticity, our preferred price-based indicator because it controls for cross-sectional differences in returns and/or responsiveness that are most likely unrelated to relationship lending.

Instead, the only statistically significant predictor of net elasticity is the degree of urbanization: the higher the fraction of the population living in urban areas in 1920, the lower the net elasticity during the 1920s. This accords well with the theory. Banks in urban centers were relied upon to be liquidity providers to other financial institutions in emergencies and on short notice, making them more likely to have to suddenly sever relationships with non-financial borrowers for reasons unrelated to the borrower's health. This can be mapped into the model as a

higher probability of exogenous separation in the middle of a potential relationship (see Appendix A). Fewer relationships with policy-invariant loan rates would then be fostered, suggesting that we should indeed observe a lower net elasticity. The first column of Table 1 thus gives us confidence that our net elasticity measure is indeed capturing the incidence of continuing relationships: the only variable that is a statistically significant predictor of $NE20$ is the one that reasonably proxies a “deep parameter” in the model and the correlation between this variable and $NE20$ is as the model would lead us to expect.

5. The Effects of Banking Distress in the Early 1930s

In this section, we show that cross-sectional differences in continuing lending relationships during the 1920s play a fundamental role in determining the real effects of banking distress in the early 1930s. We define two indicators of bank distress for each location i :

$$Suspend_Num32_i = \sum_{t=1930}^{1932} \frac{number_of_suspended_national_banks_{i,t}}{number_of_national_banks_{i,1929}}$$

$$Suspend_Val32_i = \sum_{t=1930}^{1932} \frac{deposits_in_suspended_national_banks_{i,t}}{deposits_in_national_banks_{i,1929}}$$

The fraction of banks suspended in the early 1930s, $Suspend_Num32_i$, captures the dispersion of distress across banks in location i while the share of deposits in suspended banks, $Suspend_Val32_i$, quantifies the size of the shock. We use $Suspend_Num32_i$ as our main indicator of banking distress in location i and $Suspend_Val32_i$ as a control. Our reasoning runs as follows. Suppose bank suspensions in location i amount to 10% of deposits. This could be caused by the suspension of 1 bank with a 10% market share or by the suspension of 10 banks each with a 1% market share. Although the size of the banking shock is the same in both cases – 10% of deposits – the latter more closely approximates the nature of banking panics in the early 1930s than the former. The latter also corresponds to a higher value of $Suspend_Num32_i$ than the former, despite having the same value of $Suspend_Val32_i$.

Our empirical strategy is to regress retail sales in 1933 (measured as a fraction of retail sales in 1929) on our price-based indicator of continuing relationships in the 1920s, the two measures of banking distress defined above, and the interaction between our price-based indicator and each of the banking distress measures. Similar regressions are run using retail sales in 1935 and 1937 instead of 1933. In each regression, the sample consists of the 82 locations discussed at the end of Section 3. We also include Federal Reserve district fixed effects to control for differences in the policy response to the Great Depression as well as differences in retail trade that are demand-driven rather than supply-driven.

Since retail sales involve tradeable goods, large regional markets drive both retail prices and the demand for retail items. The Federal Reserve Act of 1913 led to district boundaries that accorded with “the convenience and customary course of business” in 1914, including trade, transportation, and communication links. As a result, Federal Reserve districts, either individually or contiguously, represented large regional markets in the 1920s and 1930s, allowing district-level fixed effects to soak up the demand-side determinants of retail sales. In contrast, the cost of production, which is strongly influenced by the cost and availability of credit, was determined locally during the 1920s and 1930s. This is confirmed by a number of surveys conducted at the time, including the Bureau of Foreign and Domestic Commerce’s *Consumer Debt Study* and the *Survey of Reports of Credit and Capital Difficulties* compiled by the Bureau of the Census. Our empirical strategy thus allows us to identify the supply-side effects of banking distress, namely the idea that economic activity falls because bank failures put an end to existing lending relationships, thereby cutting off the flow of working capital to firms and forcing them to reduce operations.

5.1 Results from Simple OLS

Table 2, estimated using ordinary least squares, shows that the impact of banking distress on retail sales during the Great Depression was most severe in areas with continuing relationships. Conditional on a net elasticity of 0 in the 1920s, suspending 10% of national banks would have led to a 2.01% decline in retail sales between 1929 and 1933 (see the coefficient on *Suspend_Num32* in the first column of Table 2). However, if the jurisdiction had a net elasticity of 0.5 in the 1920s – as shown in Figure 3, there were several such areas – suspending

10% of national banks would have led to an additional 1.63% decline in retail sales between 1929 and 1933 (see the coefficient on the interaction term $Suspend_Num32 \times NE20$, where $NE20$ is our price-based measure). This pushes up the total decline to 3.64%, a roughly 80% jump in the slide of retail activity.⁷

Table 3 confirms that the statistical significance of the results in Table 2 comes from the loan rate part of the net elasticity calculation, not from the return on securities. As explained in Subsection 2.2, low loan rate elasticity is indicative of continuing relationships, all else the same. In the real world, however, all else may not be the same and, to control for this possibility, we calculated the elasticity of loan returns *relative to* the elasticity of securities returns, where a simple difference between these two elasticities provides a compact metric (net elasticity) that is easy to plot. With the same objective in mind, we can include in our regression analysis the elasticity of securities returns as a separate regressor. Table 3 then should be viewed as Table 2 without the restriction that the coefficients on the loan rate elasticity ($LE20$) and the securities rate elasticity ($SE20$) sum to zero. With $SE20$ as a separate control, $LE20$ is the relevant price-based indicator of continuing relationships in Table 3. The results presented in this table confirm the main message of Table 2: banking distress had a more severe effect on retail sales in areas with continuing lending relationships.

This raises an obvious question about the direction of causality. Could we be finding more severe effects in relationship-intensive areas because relationship lenders during this period had weaker borrowers who were more likely to succumb to economic downturns and transmit shocks back to their banks? The theory discussed in Subsection 2.2 suggests otherwise – relationship lenders use their information to retain only sufficiently good borrowers – but, to answer the question empirically, we present two additional analyses which confirm the direction of causality. The first analysis, reported in Subsection 5.2, removes bank suspensions that could plausibly be attributed to bad loans made in the 1920s (the equivalent of loans to weak borrowers) and shows that our findings are robust to such a modification. The second analysis, subsumed in Subsection 5.3, instruments, following Calomiris and Mason (2003), the indicators of banking distress and shows once again that our results remain.

⁷ See Subsection 5.4 for a discussion of the positive (yet statistically insignificant) coefficients on $Suspend_Val32$ and $Suspend_Val32 \times NE20$.

We also rule out the more general possibility of greater inherent sensitivity to shocks in relationship-intensive jurisdictions, of which the weak borrower syndrome is just one manifestation. First of all, this idea is at odds with Hachem (2011) where it is shown that relationship lending tends to smooth the steady state output profile and induce less volatile, not more volatile, responses to certain monetary shocks. Second, as will be seen in Subsection 5.2, bank failures in the early 1930s were not more prevalent in relationship-intensive areas than they were elsewhere. Third, as we will show in Section 6, areas that appear to have been rebuilding relationships in the mold of those destroyed by the Great Depression fared better than others during the 1937-38 recession. Together, these findings indicate that relationship-intensive jurisdictions are not inherently more responsive to shocks unless, of course, the shock is one that forcibly destroys their lending relationships.

5.2 Results from OLS with Purged Suspension

Here we purge the banking distress indicators of suspensions that could have been driven by bad loans made in the 1920s. To this end, define:

$$avgLLpre_i = \frac{1}{4} \sum_{t=1926}^{1929} \frac{loan_losses_{i,t}}{deposits_{i,t}}$$

Table 4(a) shows the results of a Tobit regression of the two distress indicators on *avgLLpre* and some controls. Notice that net elasticity is insignificant in explaining the number of suspended banks. In other words, bank suspensions were not more common in relationship lending areas than in other jurisdictions. Moreover, the correlation between net elasticity and average loan losses in the pre-Depression period (*avgLLpre*) is 0.29 so loan losses were also not much more common in relationship lending areas.⁸ Using the Tobit results, we can remove the effect of average loan losses in the pre-Depression period from the banking distress indicators and re-sensor to get purged indicators of distress. Results based on these purged indicators are shown in Table 4(b) and confirm the findings from the simple OLS regression.

⁸ Some positive correlation here makes sense. Relationship lenders often venture into uncharted territory at the beginning of a relationship (e.g., new entrepreneurs with little history) and not all of these risks will pan out.

5.3 Results from Instrumental Variables Approach

In this subsection, we take a more general approach and instrument the indicators of banking distress. For the instruments to do their job, they need to be able to predict cross-sectional variation in banking distress in the early 1930s but must also be exogenous to other sources of variation in retail sales over the same period.

We start with the instruments proposed by Calomiris and Mason (2003) all measured in 1929: logged banking assets, real estate owned by banks as a fraction of non-cash banking assets, and bank capital as a fraction of banking assets. Calomiris and Mason (2003) argue that differences in bank size across regions are driven by differences in regulation and demography while differences in real estate and capital holdings reflect differences in exposure to agricultural losses in the 1920s. The identifying assumption in Calomiris and Mason (2003) is that shocks in the 1930s were not just a continuation of shocks in the 1920s.

To improve the strength of the first stage – see Table 5(a) – we add cash-to-deposit and loan-to-deposit ratios in 1929 as instruments. These ratios are informative about the ability of banks in different areas to withstand deposit withdrawals (positive for the former, negative for the latter) and should, therefore, have predictive power for the amount of banking distress, at least in the early stages of the crisis. In an alternative specification, we removed the loan-to-deposit ratio as an instrument and replaced it with the ratio of demand deposits to banking assets in 1929. The second stage results, to which we turn below, were quite similar. This is a useful confirmation as one may otherwise have worried that the loan-to-deposit ratio is merely a reflection of variation in loan demand that is also picked up by the error term in the original OLS regression.

The second stage results are reported in Table 5(b). Once banking distress is instrumented, the only significant predictor of the change in retail sales from 1929 to 1933 is the interaction between the fraction of banks that were suspended in the early 1930s and our net elasticity measure from the 1920s. The negative coefficient on this interaction term indicates that retail sales fell in areas that had both a lot of continuing lending relationships and a lot of bank failures. As seen in Table 4(a), these two jurisdictional features are not highly correlated for the

1930s so it seems reasonable to conclude that the destruction of continuing relationships was a major contributor to the fall off in retail sales.

5.4 Direct Effect of Banking Distress on National Retail Sales

We now use our IV results to determine what might have happened to total retail trade in the U.S. had there been no banking distress. In other words, we engage in an exercise in hypothetical history to help us get at the full impact of bank failures on economic activity.

Since distress involves two indicators – the fraction of banks suspended in the early 1930s (*Suspend_Num32*) and the share of deposits in suspended banks (*Suspend_Val32*) – we consider two counterfactuals. The first uses the coefficients from the regression summarized in the first column of Table 5(b) to predict retail sales in 1933 for each location i when $Suspend_Num32_i = 0$. All other regressors, including $Suspend_Val32_i$, are evaluated at their observed values. We obtain a national prediction by summing across the locational predictions. The second counterfactual repeats the exercise but with both $Suspend_Num32_i = 0$ and $Suspend_Val32_i = 0$. To get a benchmark against which to compare the counterfactuals, we use the predictions when all regressors, including $Suspend_Num32_i$ and $Suspend_Val32_i$, are evaluated at their observed values.

In essence, the first counterfactual eliminates distress at small banks: the size of the banking shock in any location i ($Suspend_Val32_i$) is as observed in the data but, instead of being distributed across multiple banks, it is concentrated so that the fraction of banks suspended ($Suspend_Num32_i$) is virtually zero. The second counterfactual, by also setting the size of the shock to zero, eliminates distress at all banks. As a result, the difference between the two counterfactuals isolates the effect of big bank failures while the first counterfactual alone illuminates the effect of small bank failures.

5.4.1 Role of Small Bank Failures

The results from the first counterfactual indicate that national retail sales would have been a statistically significant 9.4% higher in 1933 had the banking distress in each location been extremely concentrated. In other words, 9.4 percentage points of the decline in total retail trade

during the Great Depression are directly attributable to the widespread nature of small bank failures within locations.

Is a 9.4% effect on national retail sales economically significant? Since surveys commissioned during the period suggest that wholesale and manufacturing would have also been directly affected by a decline of working capital through the destruction of existing relationships, we can think of the retail trade sector as a microcosm for the broader economy. If geographically disaggregated measures of GNP were available for the period, our next step would be to redo the analysis in Table 5(b) with GNP as the outcome variable. Such data are, of course, unavailable but it is possible to perform a back-of-the-envelope calculation that translates the total retail sales decline into an aggregate GNP effect. The “multiplier” we use is the coefficient from a simple regression for the period 1920-1929 of GNP growth on retail sales growth, with the latter instrumented by its one period lag. We start in 1920 because that is when the Federal Reserve began publishing its monthly index of retail sales. We seasonally adjust this index using Census software then take annual averages to match the frequency of the GNP estimates available from Romer (1989) and Balke and Gordon (1989). We end in 1929 to ensure that our results are not distorted by the Great Depression, the 1937-38 recession, or WWII. We also eliminate the post-WWII period since the ratio of retail sales to GNP declines markedly after the war. The multipliers from various specifications are reported in the first row of Table 6. Overall, we find that a 9.4% drop in retail sales is consistent with a 15-16% decline in nominal GNP and a 7-8% fall in real GNP. In a nutshell, this means that distress among small banks can account for roughly one-third of the economic contraction experienced during the Great Depression.

To better appreciate the role of relationship lending in generating a result of this magnitude, consider what would happen if cross-sectional differences in continuing relationships were ignored – that is, if the first column of Table 5(b) had been estimated without controlling for our net elasticity measure or its interaction with the banking distress variables. We find that misspecifying the estimating equation in this way would have lowered the 9.4% estimate to 5.0% and produced standard errors which would have prevented us from concluding that the effect is statistically different from zero. Therefore, holding fixed the type of regression we run, the link between the banking crisis and the economy becomes much more apparent when cross-sectional differences in continuing relationships are taken into account.

5.4.2 Role of Big Bank Failures

The results from the second counterfactual indicate that national retail sales would have been 4.1% higher in 1933 had there been no banking distress at all. This is less than half of the 9.4% figure obtained in the first counterfactual. Therefore, conditional on having eliminated distress at small banks, eliminating distress at big banks would have actually reduced economic activity in the early 1930s.

At face value, this is a surprising result. The most plausible interpretation is that deposits in large suspended banks rapidly migrated towards smaller relationship lenders that had weathered the storm, leading to economic gains. The idea is the following. Suppose, for the moment, that depositors of large suspended banks still have access to at least some of their funds. Then the failure of a big bank creates space in the business landscape for smaller banks that survive to increase their deposits and thus their lending activities. This reallocation should be particularly valuable in areas where business is critically dependent on continuing lending relationships, a contention consistent with the positive coefficient on $Suspend_Val32 \times NE20$, over and above the positive coefficient on $Suspend_Val32$, in the first column of Table 5(b). Of course, since both of these coefficients are estimated with large standard errors, particularly in IV estimation, the reallocation effect is also estimated with large standard errors. However, by not dismissing the point estimates, we are providing a conservative estimate of the effect of overall banking distress on retail sales, bringing what could have been a 9.4% effect down to 4.1%.

The question, then, is do we have any evidence to indicate that depositors of large suspended banks were able to access enough of their funds to produce the reallocation effect suggested by our regression results? The answer is yes. Large banks were liquidated quite rapidly, both in absolute terms and relative to smaller banks. In 1931 and 1932, for example, depositors in national banks with more than \$6 million in deposits on the date of failure received an average of 4.2 cents per month for each dollar of deposits during the initial year of liquidation. In comparison, depositors in banks that failed with \$2-6 million in deposits received 3.0 cents per month while depositors in banks that failed with less than \$2 million in deposits received 2.0 cents per month. These figures are all based on receivership data from the annual report of the Comptroller of the Currency. In practice though, the Comptroller's data offers an

extreme lower bound on how quickly suspended deposits in large banks were redeemed. This is because clearinghouses often provided advances to depositors of failed members (which would typically be the largest failures in the municipality) well before the final disbursement recorded by the Comptroller. For example, within days of the failure of the Bank of United States, the *New York Times* reported that the NYC Clearinghouse had arranged for this bank's depositors to borrow 50% of their deposits at a rate of 5% per year from any other member of the clearinghouse.⁹ The *Wall Street Journal* and the *LA Times* described similar arrangements in Columbus, Philadelphia, and Los Angeles, among many other places.

Taken together, our counterfactuals reveal that (1) small bank failures had a very negative effect on economic activity through the destruction of existing lending relationships but (2) this effect may be partly masked by economic gains from the reallocation of deposits toward surviving relationship lenders following large bank failures. On net, we estimate that retail sales in the U.S. would have been 4.1% higher in 1933 had there been no bank suspensions, small or large. Using the multipliers estimated in Table 6, a 4.1% effect on retail sales translates into a 6-7% effect on nominal GNP and a 3-4% effect on real GNP, which is about one-eighth of the economic contraction experienced during the Great Depression. Had Table 5(b) been estimated without controlling for our net elasticity measure or its interaction with the banking distress variables, the 4.1% estimate would have been lowered to 2.0%, understating by a factor of 2 the fall in economic activity that is directly attributable to the banking panics of the early 1930s.

6. Relationship Rebuilding and the 1937-38 Recession

In this section, we show that cross-sectional differences in the rebuilding of lending relationships destroyed during the Great Depression are important for understanding cross-sectional differences in economic performance during the late 1930s recession, thereby shedding light on the heterogeneous nature of recovery from the Depression.

⁹ A year later, the bank superintendent of the state of New York also devised a plan to expedite the return of the remainder of the deposits by reaching a settlement regarding double liability claims.

6.1 Changes in Relationship Lending

Figure 4 reveals how the Great Depression transformed the landscape of relationship lending. Interestingly, the crisis did not lead to a drop in the quantity of relationship loans relative to other assets: as shown in the top right panel of Figure 4, areas that had a high fraction of relationship loans in June 1929 averaged a similarly high fraction during the 1930s.¹⁰ Despite this, the Great Depression did destroy continuing relationships. We can see this in the top left panel of Figure 4: areas that had high net elasticities in the 1920s – that is, areas with a high degree of continuing relationships – suffered the biggest declines in this metric in the 1930s. To this point, the correlation between net elasticity in the 1920s and the change in net elasticity from the 1920s to the 1930s is -0.91. The Great Depression thus had a large negative impact on relationship lenders.

In the theory, a decline in net elasticity can be caused either by a switch from continuing relationships to new relationships (where it is not possible for the lender to offer policy-invariant credit terms because he is still accumulating information on the borrower's type) or by a switch from continuing relationships to no relationships. The former explanation – that is, continuing relationships being replaced by new relationships or, more accurately, relationship rebuilding – finds more support in the data. In particular, there is a negative correlation between the quantity-based indicator in the 1930s and the *change* in the price-based indicator from the 1920s to the 1930s. The correlation is -0.43 if the quantity-based indicator is averaged over the entire decade, -0.48 if averaged over only the Depression years, and -0.36 if averaged over only the post-Depression years. Therefore, areas that had the biggest declines in net elasticity (which we know from above were also areas with a history of relationship lending) kept more of their assets in largely unsecured commercial loans during and after the Great Depression. If this interpretation is correct – that is, if destroyed relationships were being rebuilt – then we would also expect to observe a lower average age of borrowing firms in the 1930s in locations where the decreases in net elasticity were largest, particularly if the net elasticity was high in the 1920s. Appendix B presents some evidence that points in this direction.

¹⁰ We cannot compare the average of the quantity-based measure in the 1930s to its average in the 1920s because of changes in the Comptroller's reporting of geographically disaggregated information (see Subsection 2.1).

6.2 Implications for the Late 1930s Recession

The extent of relationship rebuilding and its implications for recovery in the post-Depression era can also be determined by looking at how different locations fared during the late 1930s recession. As noted earlier, Hachem (2011) predicts that relationship lending makes output less responsive to changes in bank funding costs, all else constant. Therefore, if the recession involved higher funding costs for banks – or at least a perception among banks that such costs were likely to rise – then areas where relationships had been reconstructed should have fared better than otherwise similar areas in the late 1930s. As we will now show, this prediction is borne out by the data. This is an interesting and important result because it reveals a novel link between relationship lending and cross-sectional differences in economic performance during the 1937-38 recession.

The textbook source of higher bank funding costs is contractionary monetary policy. The literature has debated two such policies: the Fed's decision to begin doubling reserve requirements in August 1936 and the Treasury's decision to sterilize gold inflows starting in December 1936. The influential narrative in Friedman and Schwartz (1963) characterizes the Fed's policy as highly contractionary whereas the analysis of interest rates in Hanes (2006) and the simulations of reserve holdings in Calomiris et al (2011) suggest a more important role for the Treasury's policy. On both sides of the debate, however, there is a contractionary policy which we contend relationship lending, once rebuilt, would have mitigated by extending policy-invariant loan rates to some borrowers, enabling those borrowers to continue operations uninterrupted.

We measure the extent of relationship rebuilding prior to the 1937-38 recession using three variables: our quantity-based indicator of relationship lending averaged over the 1930s using only data up to 1936 (λ_{30s}), the change in our price-based indicator of relationship lending from the 1920s to the 1930s (ΔNE) when the 1930s indicator is re-computed using only data up to 1936, and the interaction between λ_{30s} and ΔNE . As we argued in Subsection 6.1, areas with

the highest values of λ_{30s} and the most negative values of ΔNE are the most likely candidates for relationship rebuilding.¹¹

If we were to take this as the end of the story, then the identifying assumption would be that, by 1937, the rebuilding efforts of the mid-1930s had matured into the sort of shock-absorbing relationships consistent with the theory and detected by our price-based indicator in the 1920s. The most direct way to validate this assumption is to compute the price-based indicator using data from the first year of recovery, 1934, to 1936, the last pre-recession year. The problem, aside from the short sample, is that the first real test for the majority of post-Depression relationship lenders would have been the 1937-38 recession, which makes it difficult to know for sure that lenders of rebuilt relationships were as committed to relationship lending on the eve of this recession as they had been in the past.

The most straightforward way to resolve this conundrum is to control for the possibility that the Great Depression created cross-sectional differences in the priorities of banks. The seminal work of Diamond and Dybvig (1983) expounds the fundamental fragility of banks vis-à-vis liquidity so it seems reasonable to focus on how banks may have modified their liquidity management as a result of the Depression. Consider, in particular, a jurisdiction where banks had the following characteristics: (1) they held very thin liquidity cushions on the eve of the Great Depression; (2) they failed en masse during the Depression primarily because of runs on their liabilities, not because of problems with their assets; and (3) they held substantial liquidity cushions on the eve of the 1937-38 recession, largely because of what they had witnessed during the dark days of the Depression. If this jurisdiction had also been a relationship lender in the 1920s, then one might be concerned that its commitment to relationship lending had been tempered by its Depression experience such that the relationships it was rebuilding in the mid-1930s differed in basic ways from those observed in the 1920s. This matters because it may very well have been the case that its banks, at the first sign of trouble, would pre-emptively refuse to continue these new relationships in order to expand their liquidity buffers. Such behavior would

¹¹ Recall from Subsection 2.1 that the quantity-based indicator for the 1930s is based on the “all other loans” category in the Comptroller’s report. The Comptroller separates “all other loans” from “loans on securities” for cities and states but not districts so we cannot use any of the split states or district residuals in the analysis here. As a result, the number of locations will be less than 82.

cause economic activity in the jurisdiction to contract for the very same reason that it did during the Great Depression – the soft information about borrowers cannot be easily redeployed.

We now use the triumvirate of characteristics described above to get at the cross-sectional differences in bank behavior and in economic outcomes during the 1937-38 recession. We quantify item (1) on the above list using the ratio of demand deposits minus cash to total assets in 1929. This captures the vulnerability of each location in our sample once depositors began to panic and withdraw funds in the early 1930s. We quantify (2) using the purged indicator of banking distress in the early 1930s based on the number of suspended banks (see Subsection 5.2). This captures the extent to which good banks succumbed to the panics. Finally, we quantify (3) using the excess reserve ratio in June 1936. We then run a cross-sectional regression of retail sales in 1939 (relative to 1937) on relationship rebuilding as defined above, the excess reserve ratio in June 1936, and the interaction between relationship rebuilding and the excess reserve ratio. We instrument the excess reserve ratio using (1) and (2) so as to isolate the persistent effect of past experience on liquidity management heading into the 1937-38 recession. Areas that had low excess reserve ratios in June 1936 – as predicted by a first stage regression of (3) on (1) and (2) – are areas where we would contend that banks had not developed a preference for liquidity as a result of the Great Depression. Therefore, the relationships rebuilt in these areas are the ones most likely to have been cut from the same cloth as the continuing relationships that existed in the 1920s. In other words, these are precisely the areas where we would expect to observe the greatest positive impact on economic activity during the 1937-38 recession because of the reconstructed lending relationships in the aftermath of the Depression.

The IV results in Table 7 confirm our intuition. They show that areas where relationships were being rebuilt did better than other areas during the late 1930s recession if they had low excess reserve ratios in June 1936 and, by the same token, did worse if they had high excess reserve ratios going into the 1937-38 downturn. In other words, areas that we suspect rebuilt the types of continuing relationships observed in the 1920s did well during the late 1930s recession whereas areas that rebuilt relationships but were, at the same time, more susceptible to fears of yet another liquidity crunch – and thus less committed to maintaining those relationships – did poorly. The past, it would appear, weighed heavily on the present.

7. Conclusion

We proposed, in this paper, a novel measure of continuing lending relationships that resolves the data limitations of the 1920s and 1930s and pinpoints the non-monetary effects of banking distress in a way that the existing literature on the Great Depression has been unable to do. Using our measure, we were able to show that the marginal effect of bank suspensions on economic activity was much higher in areas with more continuing relationships. Taking these cross-sectional differences into account, we then found that small bank failures can generate roughly one-third of the total economic contraction experienced during the Great Depression. Interestingly, however, our analysis also revealed that this effect was partly offset by economic gains from the reallocation of deposits toward surviving relationship lenders following large bank failures. On net, we estimated that national bank suspensions – small and large – generated about one-eighth of the economic contraction during the Great Depression. Counterfactually, we also showed that the fall in economic activity directly attributable to the banking panics of the early 1930s would have been understated by a factor of 2 had we not controlled for cross-sectional differences in continuing lending relationships on the eve of the crisis.

We then used our methodology to show that relationship lending played an important role during the recovery period. In particular, areas that rebuilt the types of continuing relationships that they embraced in the 1920s fared better during the 1937-38 recession than otherwise similar areas that did not rebuild. In contrast, areas that rebuilt relationships but, based on their reserve holdings, appear to have been less committed to these relationships, actually fared worse. In 1937-38, few banks failed and few depositors panicked. This stability largely stemmed from New Deal era reforms of the financial system such as the creation of the Federal Deposit Insurance Corporation. These institutional changes help to explain the different fates of relationship lending regions during the initial contraction of the Great Depression and the double-dip recession nearly a decade later.

In sum, our results indicate that not only does banking distress, if measured properly, have substantial negative consequences for real economic activity but also that the effects of such crises are likely to linger long after the events themselves have passed. While it may appear that modern, highly sophisticated financial systems are immune to such stresses, the Great

Recession and its aftermath would seem to provide a salutary corrective to such overconfidence. And, in many parts of the world where relationship loans constitute a major part of bank assets, financial crises are likely to have serious and potentially long-lasting deleterious effects on the real economy. Fortunately, policymakers, fortified with this knowledge, have the means to mitigate the damage if they act swiftly and strongly enough.

References

- Balke, N. and R. Gordon. 1989. "The Estimation of Prewar Gross National Product: Methodology and New Evidence." *Journal of Political Economy*, 97(1), pp. 38-92.
- Berger, A. and G. Udell. 1995. "Relationship Lending and Lines of Credit in Small Firm Finance." *Journal of Business*, 68(3), pp. 351-382.
- Bernanke, B. 1983. "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression." *American Economic Review*, 73(3), pp. 257-276.
- Calomiris, C. and J. Mason. 2003. "Consequences of Bank Distress During the Great Depression." *American Economic Review*, 93(3), pp. 937-947.
- Calomiris, C., J. Mason, and D. Wheelock. 2011. "Did Doubling Reserve Requirements Cause the Recession of 1937-1938? A Microeconomic Approach." NBER Working Paper No. 16688.
- Carlson, M. and J. Rose. 2015. "Credit Availability and the Collapse of the Banking Sector in the 1930s." *Journal of Money, Credit, and Banking*, 47(7), pp. 1239-1271.
- Chodorow-Reich, G. 2014. "The Employment Effects of Credit Market Disruptions: Firm-Level Evidence from the 2008-9 Financial Crisis." *Quarterly Journal of Economics*, 129(1), pp. 1-59.
- Cole, H. and L. Ohanian. 1999. "The Great Depression in the United States from a Neoclassical Perspective." *Federal Reserve Bank of Minneapolis Quarterly Review*, 23(1), pp. 2-24.
- Cole, H. and L. Ohanian. 2004. "New Deal Policies and the Persistence of the Great Depression: A General Equilibrium Analysis." *Journal of Political Economy*, 112(4), pp. 779-816.
- Cole, H. and L. Ohanian. 2007. "A Second Look at the U.S. Great Depression from a Neoclassical Perspective." In Kehoe, T. and E. Prescott (eds.), *Great Depressions of the Twentieth Century*, Federal Reserve Bank of Minneapolis, pp. 21-57.
- Currie, L. 1931. "The Decline of the Commercial Loan." *Quarterly Journal of Economics*, 45(4), pp. 698-709.
- Diamond, D. and P. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy*, 91(3), pp. 401-419.
- Elsas, R. 2005. "Empirical Determinants of Relationship Lending." *Journal of Financial Intermediation*, 14(1), pp. 32-57.

- Fisher, I. 1933. "The Debt-Deflation Theory of Great Depressions." *Econometrica*, 1(4), pp. 337-356.
- Ford, A.W. 1928. "Country Banking Practices in Northern Texas." *Journal of Business*, 1(4), pp. 468-484.
- Friedman, M. and A. Schwartz. 1963. *A Monetary History of the United States, 1867– 1960*. Princeton, NJ: Princeton University Press.
- Gobbi, G. and E. Sette. 2015. "Relationship Lending During a Financial Crisis." *Journal of the European Economic Association*, 13(3), pp. 453-481.
- Hachem, K. 2011. "Relationship Lending and the Transmission of Monetary Policy." *Journal of Monetary Economics*, 58(6-8), pp. 590-600.
- Hanes, C. 2006. "The Liquidity Trap and U.S. Interest Rates in the 1930s." *Journal of Money, Credit, and Banking*, 38(1), pp. 163-194.
- Mitchener, K. and G. Richardson. 2016. "Network Contagion and Interbank Amplification during the Great Depression." FRB Richmond Working Paper 16-03.
- Ongena, S. and D. Smith. 2001. "The Duration of Bank Relationships." *Journal of Financial Economics*, 61(3), pp. 449-475.
- Park, H. and G. Richardson. 2012. "Retail Trade by Federal Reserve District, 1919 to 1939: A Statistical History." *Research in Economic History*, 28(1), pp. 151-231.
- Rajan, R. and R. Ramcharan. 2016. "Crises and the Development of Economic Institutions: Some Microeconomic Evidence." Forthcoming in *American Economic Review, Papers & Proceedings*.
- Richardson, G. and W. Troost. 2009. "Monetary Intervention Mitigated Banking Panics During the Great Depression: Quasi-Experimental Evidence from the Federal Reserve District Border in Mississippi, 1929 to 1933." *Journal of Political Economy*, 117(6), pp. 1031-1073.
- Romer, C. 1989. "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869-1908." *Journal of Political Economy*, 97(1), pp. 1-37.

Romer, C. and D. Romer. 1989. "Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz." *NBER Macroeconomics Annual*, Vol. 4, pp. 121-170.

Romer, C. and D. Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review*, 94(4), pp. 1055-1084.

Romer, C. and D. Romer. 2015. "New Evidence on the Impact of Financial Crises in Advanced Countries." NBER Working Paper No. 21021.

Ziebarth, N. 2013. "Identifying the Effects of Bank Failures from a Natural Experiment in Mississippi during the Great Depression." *American Economic Journal: Macroeconomics*, 5(1), pp. 81-101.

Appendix A

This appendix sketches a simplified version of Hachem (2011) with only two periods. The reader is referred to the original paper for more detail.

There is a continuum of borrower types, denoted by ω and distributed uniformly over the unit interval. Borrowers choose between two projects each period. The first is an investment project which produces output θ_1 with probability $p(\omega)$ at the end of the period, where $p'(\omega) > 0$. The second is a speculative project which produces output $\theta_2 > \theta_1$ with probability $q < p(0)$ at the end of the period. Assume $p(0)\theta_1 = q\theta_2$ to reduce notation. The investment and speculative projects produce nothing with probability $1 - p(\omega)$ and $1 - q$ respectively. Note that the investment project second order stochastically dominates the speculative one.

Before a borrower can undertake either project, he must obtain one unit of capital from a lender. Capital is available to lenders at an exogenous policy rate r . The credit contracts that transfer capital from lenders to borrowers are one-period with the possibility of being rolled over. Contracts are characterized by an interest rate which can only be repaid if the borrower's project produces positive output. Lenders can detect the presence of output (but cannot observe its exact value) so firms with positive output repay interest. There is no collateral or quantity rationing.

The loan rate that makes a borrower of type ω indifferent between the two projects is:

$$\bar{R}(\omega) = \frac{p(\omega)\theta_1 - q\theta_2}{p(\omega) - q}$$

The borrower undertakes the speculative project during a particular period if and only if the interest rate he is charged that period exceeds $\bar{R}(\omega)$. All lenders know that this is the best response of a type ω borrower.

At the beginning of the first period, borrower types are private information. Lenders are perfectly competitive and offer the same zero-profit interest rate. Borrowers select their lenders randomly then decide which project to undertake. At the beginning of the second period, each lender learns his borrower's type. Others only observe whether the borrower repaid the interest on his first period loan. Each lender then decides whether to provide another unit of capital to his borrower and, if so, at what interest rate. The interest rate cannot exceed what other lenders would charge, otherwise the borrower will move to one of these other lenders. As shown in Hachem (2011), the second period interest rates that prevail in equilibrium are:

$$R_2 = \begin{cases} r/q & \text{if } \omega \in [0, \tilde{\omega}] \\ \bar{R}(\omega) & \text{if } \omega \in (\tilde{\omega}, \hat{\omega}) \\ r/q & \text{if } \omega \in [\hat{\omega}, 1] \end{cases}$$

where the cutoff types $\tilde{\omega}$ and $\hat{\omega}$ are implicitly defined by:

$$\begin{aligned} p(\tilde{\omega}) \bar{R}(\tilde{\omega}) &= r \\ q \bar{R}(\hat{\omega}) &= r \end{aligned}$$

Types below $\tilde{\omega}$ move to other lenders and are charged r/q , prompting them to undertake the speculative project. Types above $\tilde{\omega}$ stay with their first period lender and are charged interest rates that lead them to choose the investment project. Notice that types between $\tilde{\omega}$ and $\hat{\omega}$ are getting policy-invariant interest rates (i.e., $\bar{R}(\cdot)$ does not depend on r). If these types were instead charged r/q , they would undertake the speculative project.

The average second period loan rate can be expressed as:

$$R_2 = \int_0^{\tilde{\omega}} \frac{r}{q} d\omega + \int_{\tilde{\omega}}^{\hat{\omega}} \bar{R}(\omega) d\omega + \int_{\hat{\omega}}^1 \frac{r}{q} d\omega$$

Its elasticity with respect to the policy rate is then:

$$e_2 \equiv \frac{dR_2}{dr} \frac{r}{R_2} = \frac{1 - \hat{\omega} + \tilde{\omega} + \left[1 - \frac{q}{p(\tilde{\omega})}\right] r \frac{d\tilde{\omega}}{dr}}{1 - \hat{\omega} + \tilde{\omega} + \frac{1}{\bar{R}(\tilde{\omega})} \int_{\tilde{\omega}}^{\hat{\omega}} \bar{R}(\omega) d\omega}$$

Hachem (2011) also shows that first period loan rates are:

$$R_1 = \bar{R}(\xi)$$

where ξ is implicitly defined by:

$$\left[q\xi + \int_{\xi}^1 p(x) dx \right] \bar{R}(\xi) + \beta(1-s) \left[\int_{\tilde{\omega}}^{\hat{\omega}} p(x) \bar{R}(x) dx - \left[1 - \tilde{\omega} - \int_{\tilde{\omega}}^1 \frac{p(x)}{q} dx \right] r \right] = r$$

The second term on the left-hand side is the first period lender's expected continuation value. More precisely, β is a discount factor, s is the probability of exogenous separation, and the term in square brackets that multiplies $\beta(1-s)$ is the lender's expected second period profit from borrowers that he is willing to finance again.

Let ψ denote the fraction of second period borrowers, where $\frac{\partial \psi}{\partial s} < 0$ and $s = 1$ implying $\psi = 0$. The fraction of borrowers in continuing relationships is then $\psi(1 - \tilde{\omega})$. It is trivial to see that lower values of s foster more continuing relationships.

We can now write the average loan rate in the relationship lending model as:

$$R_{rl} = (1 - \psi) \bar{R}(\xi) + \psi \left[\int_{\tilde{\omega}}^{\hat{\omega}} \bar{R}(x) dx + [1 - \hat{\omega} + \tilde{\omega}] \frac{r}{q} \right]$$

Setting $s = 1$ eliminates relationship lending, in which case the average loan rate is simply:

$$R_{no} = \bar{R}(\eta)$$

where η is implicitly defined by:

$$\left[q\eta + \int_{\eta}^1 p(x) dx \right] \bar{R}(\eta) = r$$

The elasticity with respect to the policy rate in the absence of relationship lending is therefore:

$$e_{no} \equiv \frac{dR_{no}}{dr} \frac{r}{R_{no}} = \frac{\left[q\eta + \int_{\eta}^1 p(x) dx \right] \bar{R}'(\eta)}{\left[q\eta + \int_{\eta}^1 p(x) dx \right] \bar{R}'(\eta) - [p(\eta) - q] \bar{R}(\eta)}$$

If s is low, then ψ is high and the elasticity of R_{rl} with respect to the policy rate will be dominated by e_2 . The denominator of e_{no} is positive (see the online appendix of Hachem (2011)) so it remains to show $e_2 < e_{no}$ or, equivalently:

$$\begin{aligned} & \left[q\eta + \int_{\eta}^1 p(x) dx \right] \bar{R}'(\eta) \left[\frac{[p(\tilde{\omega}) - q] \bar{R}(\tilde{\omega})}{p'(\tilde{\omega}) \bar{R}(\tilde{\omega}) + p(\tilde{\omega}) \bar{R}'(\tilde{\omega})} - \frac{1}{\bar{R}(\hat{\omega})} \int_{\tilde{\omega}}^{\hat{\omega}} \bar{R}(\omega) d\omega \right] \quad (1) \\ & < [p(\eta) - q] \bar{R}(\eta) \left[1 - \hat{\omega} + \tilde{\omega} + \frac{[p(\tilde{\omega}) - q] \bar{R}(\tilde{\omega})}{p'(\tilde{\omega}) \bar{R}(\tilde{\omega}) + p(\tilde{\omega}) \bar{R}'(\tilde{\omega})} \right] \end{aligned}$$

The same condition is relevant in a model where relational and transactional lending coexist and the weights on R_{rl} and R_{no} in the population average do not vary with r in a first-order way. For ease of exposition, we assume $p(\omega)$ linear in ω and rewrite (1) as:

$$\begin{aligned} & \frac{p(\tilde{\omega}) \tilde{\omega}^2}{[p(1) - p(0)] \tilde{\omega} + p(\tilde{\omega}) \frac{p(0) - q}{p(\tilde{\omega}) - q}} - \frac{1}{p(1) - p(0)} \frac{q}{\theta_1} \int_{\tilde{\omega}}^{\hat{\omega}} [\bar{R}(\omega) - \bar{R}(\tilde{\omega})] d\omega \quad (2) \\ & < \frac{[p(\eta) - q] \eta^2}{p(0) - q} \left[1 - \hat{\omega} + \tilde{\omega} + \frac{[p(\tilde{\omega}) - q] \bar{R}(\tilde{\omega})}{p'(\tilde{\omega}) \bar{R}(\tilde{\omega}) + p(\tilde{\omega}) \bar{R}'(\tilde{\omega})} \right] + \frac{q(\hat{\omega} - \tilde{\omega}) \tilde{\omega}}{p(\tilde{\omega}) - q} \end{aligned}$$

Next, we rearrange $p(\tilde{\omega}) \bar{R}(\tilde{\omega}) = q \bar{R}(\tilde{\omega})$ as defined above to isolate:

$$\hat{\omega} = \frac{p(0) - q}{\frac{q}{p(\tilde{\omega})} \frac{p(\tilde{\omega}) - q}{\tilde{\omega}} - [p(1) - p(0)]}$$

This allows us to rewrite (2) as:

$$\begin{aligned} & -\frac{[p(\tilde{\omega})^2 - q^2][p(1) - p(0)]\tilde{\omega}^2\hat{\omega}}{p(\tilde{\omega})\left[[p(1) - p(0)]\tilde{\omega} + p(\tilde{\omega})\frac{p(0)-q}{p(\tilde{\omega})-q}\right]} - \frac{p(0) - q}{p(1) - p(0)}\frac{q}{\theta_1}\int_{\tilde{\omega}}^{\hat{\omega}}[\bar{R}(\omega) - \bar{R}(\tilde{\omega})]d\omega \\ & < [p(\eta) - q]\eta^2\left[1 - \hat{\omega} + \tilde{\omega} + \frac{[p(\tilde{\omega}) - q]\bar{R}(\tilde{\omega})}{p'(\tilde{\omega})\bar{R}(\tilde{\omega}) + p(\tilde{\omega})\bar{R}'(\tilde{\omega})}\right] \end{aligned}$$

which is true since the left-hand side is negative and the right-hand side is positive. ■

Appendix B

This appendix elaborates on the discussion at the very end of Subsection 6.1. Given the paucity of data on borrower age during this period, we make use instead of readily available information on new business formations since it is perfectly reasonable to assume that areas with a lot of new businesses will also be ones with a high fraction of first-time borrowers. The Statistical Abstract of the United States reports the number of concerns in business (essentially the total number of businesses) and the number of business failures by state for each year. A simple indicator of new business activity in state s during the 1930s can be constructed by taking the total number of businesses at the beginning of 1938, subtracting the total number of businesses at the beginning of 1929, and adding the total number of business failures from 1929 to 1937. To allow comparison across states, the result can be expressed as a fraction of the total number of businesses at the beginning of 1929. The one potential problem with this indicator is that it is likely to understate new business activity because it is unable to account for business consolidation through mergers and acquisitions. To address this, we define $scale_{s,t}$ to be the number of wage earners per manufacturing establishment in state s at time t . These data are also reported in the Statistical Abstract based on the Biennial Census of Manufactures. We then adjust the simple indicator by multiplying the total number of businesses at the beginning of 1938 by $\frac{scale_{s,1937}}{scale_{s,1929}}$ if $scale_{s,1937} > scale_{s,1929}$. For the set of states not split into finer geographic units in the Comptroller's report, the correlation between this adjusted indicator of new business activity and the change in net elasticity from the 1920s to the 1930s is -0.23. However, when we compute the correlation using only states that had high net elasticities in the 1920s, the correlation jumps to -0.48, exactly what our hypothesis would lead us to expect. It may also be useful to note that net elasticity in the 1920s is largely uncorrelated with business formation in the 1920s so our findings for the 1930s cannot be interpreted as relationship-intensive areas simply having more churn.

Tables

Table 1

		NE20		LE20		SE20	
		Coeff	P-value	Coeff	P-value	Coeff	P-value
log_area	log(area in square miles)	-0.175	0.428	0.016	0.906	-0.159	0.464
log_pop	log(population in 1920)	0.088	0.742	0.166	0.529	0.254	0.186
urban	% population urban in 1920	-1.688	0.041	0.698	0.220	-0.990	0.164
nwnp	% population with native white parents in 1920	-0.475	0.455	0.508	0.199	0.033	0.944
age1844	% population aged 18 to 44 in 1920	0.239	0.958	-2.208	0.480	-1.970	0.496
school1620	% of aged 16 to 20 in school in 1920	1.173	0.675	-0.409	0.767	0.764	0.688
log_mfgest	log(number of mfg establishments in 1920)	-0.022	0.936	-0.090	0.560	-0.112	0.542
log_mfgsize	log(workers per mfg establishment in 1920)	-0.103	0.823	0.224	0.232	0.121	0.780
mfgwork	mfg workers as % of population in 1920	-1.589	0.686	2.566	0.230	0.977	0.758
log_mfgva	log(value added per mfg establishment in 1920)	0.356	0.497	-0.407	0.028	-0.051	0.908
log_farms	log(number of farms in 1920)	0.161	0.178	-0.024	0.823	0.137	0.244
acres	farm land as % of area in 1920	0.028	0.966	-0.280	0.413	-0.252	0.685
log_avgacre	log(farm acres / number of farms in 1920)	-0.124	0.637	0.251	0.161	0.127	0.697
log_avgcrop	log(crop value / number of farms in 1920)	-0.244	0.329	0.109	0.273	-0.135	0.538
log_avgvalue	log(value of farm land, equip, etc / no. of farms in 1920)	0.192	0.423	-0.126	0.175	0.066	0.750
ownerop	% of farms owner-operated in 1920	-0.206	0.657	-0.150	0.594	-0.356	0.301
log_banks	log(number of national banks in 1920)	-0.104	0.278	-0.038	0.700	-0.143	0.232
log_banksize	log(banking assets / number of banks in 1920)	-0.020	0.937	0.180	0.037	0.161	0.480
depratio	demand and time deposits / banking assets in 1920	0.133	0.870	-0.763	0.060	-0.629	0.468
ddratio	demand deposits / demand and time deposits in 1920	-0.542	0.436	0.184	0.695	-0.359	0.649
city	dummy variable for reserve cities	-0.237	0.615	-0.171	0.289	-0.409	0.293
Observations		84		84		84	
District Fixed Effects		YES		YES		YES	
R-Squared		0.62		0.68		0.56	

Standard errors clustered at the district level.

Table 2
Results from Simple OLS

	Retail Sales Ratio		
	1933 / 1929	1935 / 1929	1937 / 1929
NE20	0.008 (0.664)	-0.020 (0.155)	-0.431 (0.496)
Suspend_Num32	-0.201 (0.012)	-0.291 (0.031)	-0.817 (0.348)
Suspend_Num32 × NE20	-0.327 (0.068)	-0.227 (0.460)	7.359 (0.451)
Suspend_Val32	0.141 (0.174)	0.250 (0.272)	0.582 (0.599)
Suspend_Val32 × NE20	0.330 (0.246)	0.192 (0.637)	-7.963 (0.382)
Observations	82	82	82
District Fixed Effects	YES	YES	YES
R-Squared	0.50	0.47	0.06

P-Values in parentheses. Standard errors clustered at the district level. All specifications also include a dummy variable equal to one for reserve cities.

Table 3
Results from OLS with Net Elasticity Separated into Components

	Retail Sales Ratio		
	1933 / 1929	1935 / 1929	1937 / 1929
LE20	-0.021 (0.622)	0.022 (0.710)	3.184 (0.275)
SE20	0.000 (0.989)	-0.018 (0.341)	0.432 (0.257)
Suspend_Num32	-0.257 (0.002)	-0.383 (0.052)	-0.970 (0.501)
Suspend_Num32 × LE20	0.433 (0.085)	0.177 (0.567)	-13.402 (0.357)
Suspend_Num32 × SE20	-0.301 (0.148)	-0.116 (0.699)	12.588 (0.386)
Suspend_Val32	0.197 (0.183)	0.206 (0.396)	7.020 (0.281)
Suspend_Val32 × LE20	-0.493 (0.617)	0.489 (0.655)	-12.922 (0.365)
Suspend_Val32 × SE20	0.318 (0.213)	0.075 (0.830)	-16.928 (0.304)
Observations	82	82	82
District Fixed Effects	YES	YES	YES
R-Squared	0.50	0.48	0.08

P-Values in parentheses. Standard errors clustered at the district level. All specifications also include a dummy variable equal to one for reserve cities.

Table 4(a)
Results from Tobit Regression

	Suspend_Num32	Suspend_Val32
NE20	0.010 (0.715)	0.040 (0.107)
avgLLpre	17.287 (0.020)	15.652 (0.032)

P-Values in parentheses. Standard errors clustered at the district level. All specifications also include district fixed effects and a dummy variable equal to one for reserve cities.

Table 4(b)
Results from OLS with Purged Suspensions

	Retail Sales Ratio		
	1933 / 1929	1935 / 1929	1937 / 1929
NE20	0.013 (0.468)	-0.008 (0.566)	-0.384 (0.430)
P_Suspend_Num32	-0.201 (0.007)	-0.366 (0.017)	-0.586 (0.466)
P_Suspend_Num32 × NE20	-0.832 (0.097)	-0.913 (0.287)	11.232 (0.335)
P_Suspend_Val32	0.088 (0.393)	0.131 (0.543)	0.078 (0.985)
P_Suspend_Val32 × NE20	1.032 (0.145)	1.378 (0.268)	-7.669 (0.417)
Observations	82	82	82
District Fixed Effects	YES	YES	YES
R-Squared	0.50	0.49	0.06

P-Values in parentheses. Standard errors clustered at the district level. All specifications also include a dummy variable equal to one for reserve cities.

Table 5(a)
Strength of First Stage for IV

	Suspend_Num32	Suspend_Num32 × NE20	Suspend_Val32	Suspend_Val32 × NE20
SW F-statistic (p-value)	10.99 (0.000)	4.86 (0.010)	32.08 (0.000)	25.65 (0.000)

Table 5(b)
Results from IV

	Retail Sales Ratio		
	1933 / 1929	1935 / 1929	1937 / 1929
NE20	0.025 (0.406)	-0.010 (0.800)	-4.134 (0.406)
Suspend_Num32	-0.326 (0.148)	-0.307 (0.389)	22.606 (0.433)
Suspend_Num32 × NE20	-0.554 (0.058)	-0.267 (0.647)	55.295 (0.391)
Suspend_Val32	0.393 (0.245)	0.447 (0.399)	-19.470 (0.631)
Suspend_Val32 × NE20	0.133 (0.840)	-0.140 (0.842)	8.671 (0.911)
Observations	82	82	82
District Fixed Effects	YES	YES	YES
J-statistic (p-value)	0.830	0.737	0.964
AR Wald Test (p-value)	0.000	0.017	0.918

P-Values in parentheses. Standard errors clustered at the district level. All specifications also include a dummy variable equal to one for reserve cities.

Table 6
Retail Sales and GNP

	Nominal GNP Growth		Real GNP Growth	
	Romer	Balke-Gordon	Romer	Balke-Gordon
Retail_Growth	1.690 (0.000)	1.565 (0.000)	0.696 (0.000)	0.896 (0.000)
Constant	-0.010 (0.305)	-0.006 (0.574)	0.028 (0.000)	0.025 (0.000)

P-values in parentheses. Instrument for Retail_Growth is Retail_Growth_Lag1.

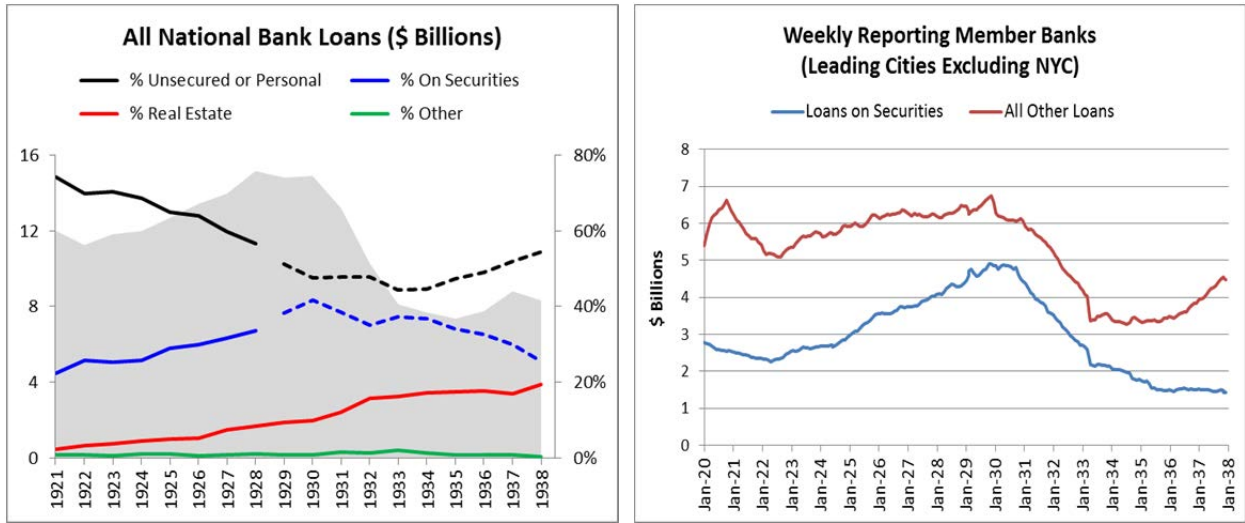
Table 7
Relationship Rebuilding and the Late 1930s

	Retail Sales Ratio: 1939 / 1937		
	OLS	OLS	IV
ΔNE	-0.225 (0.568)	2.001 (0.012)	2.582 (0.017)
λ_{30s}	0.037 (0.944)	0.527 (0.451)	-0.237 (0.860)
$\Delta NE \times \lambda_{30s}$	0.389 (0.547)	-3.305 (0.013)	-5.368 (0.022)
ER_Ratio		4.647 (0.401)	-11.190 (0.445)
$\Delta NE \times ER_Ratio$		-23.510 (0.001)	-29.099 (0.006)
$\lambda_{30s} \times ER_Ratio$		-6.185 (0.462)	14.823 (0.514)
$\Delta NE \times \lambda_{30s} \times ER_Ratio$		38.483 (0.001)	62.788 (0.014)
Observations	63	62	62
District Fixed Effects	YES	YES	YES
R-Squared	0.60	0.67	
J-statistic (p-value)			0.166
AR Wald Test (p-value)			0.000

P-Values in parentheses. Robust standard errors are used. All specifications also include a dummy variable equal to one for reserve cities.

Figures

Figure 1
National Bank Lending



Left panel: Gray area is the dollar value, in billions, of national bank loans plotted against the left axis. Lines are plotted against the right axis. Each year is as of June month-end. The solid black line is based on unsecured loans and loans on personal securities reported by the Comptroller of the Currency at geographically disaggregated levels until 1928. The dashed black line is based on “all other loans” reported by the Comptroller at geographically disaggregated levels starting in 1929.

Figure 2
Predicted versus Actual for State/City Data Prior to December 1929

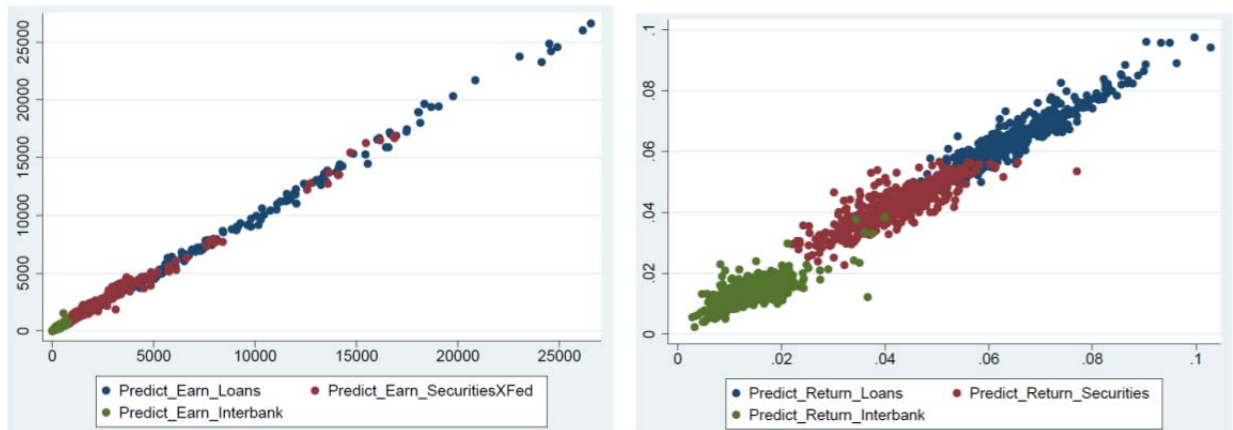
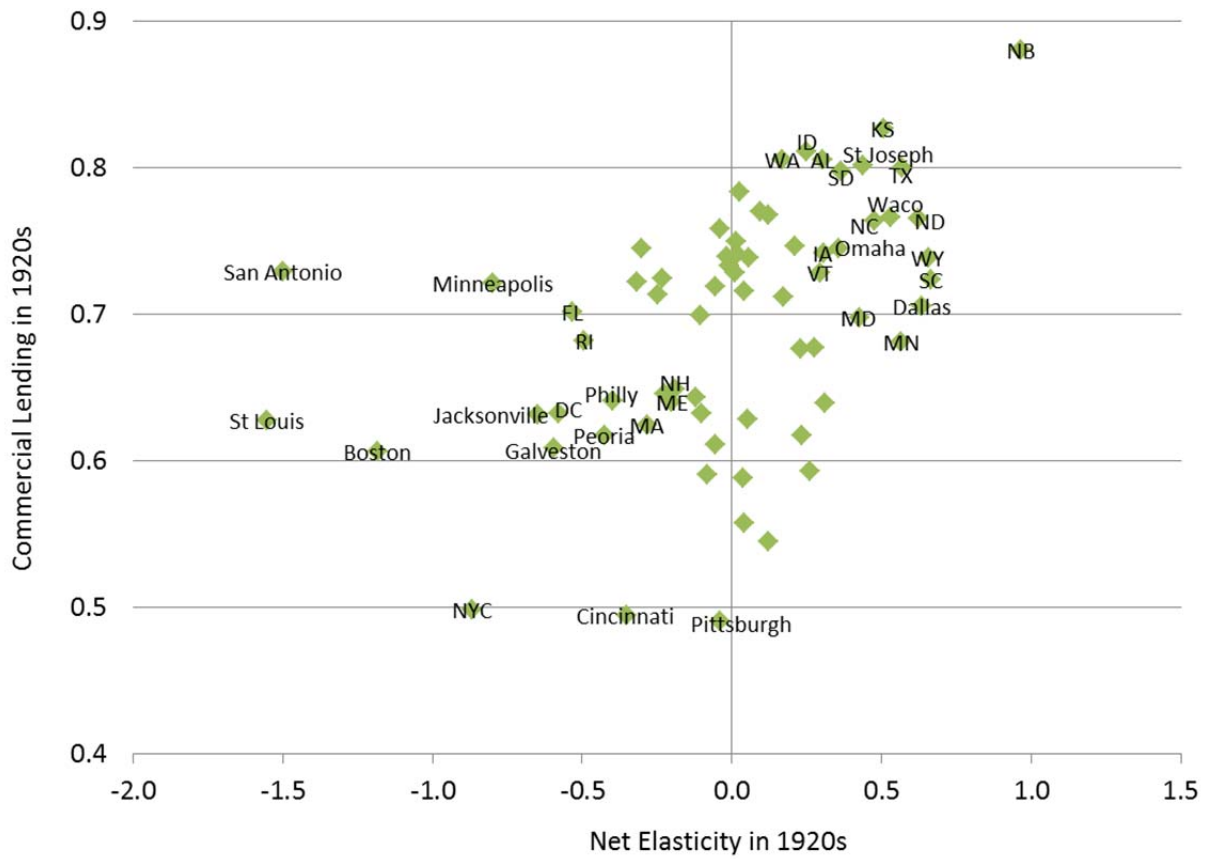
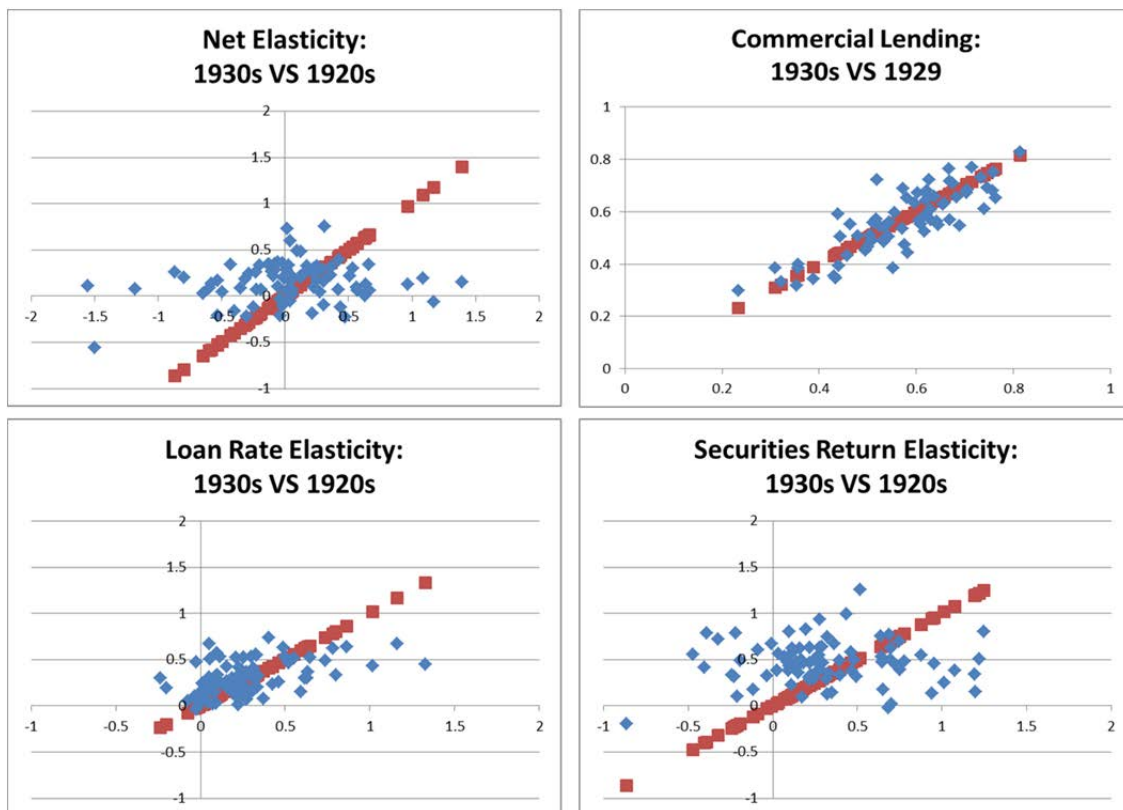


Figure 3
 Relationship Lending in the 1920s



Notes: Net elasticity is as defined in Subsection 2.2. Commercial lending is the fraction of assets in relationship loans, where relationship loans are as defined in Subsection 2.1.

Figure 4
Changes in Relationship Lending



Notes: Elasticities are as defined in Subsection 2.2. Commercial lending in top right panel is the fraction of assets in relationship loans, where relationship loans are as defined in Subsection 2.1.