NBER WORKING PAPER SERIES

RELATIONSHIP LENDING AND THE GREAT DEPRESSION

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, Revised January 2020

We thank Joseph Johnson, Padma Sharma, and Daniel Tracht for excellent research assistance. We also thank Ben Bernanke, Gabe Chodorow-Reich, Oli Coibion, Jonathan Dingel, Joe Haubrich, Erik Hurst, Chris Koch, Fabrizio Perri, Johannes Wieland, and three anonymous referees, as well as seminar and conference participants at Chicago Booth, UW-Madison, FRB Cleveland, NBER SI 2016 Monetary Economics, the 2018 Texas Monetary Conference, and the University of Michigan for helpful comments. Kinda Hachem thanks Chicago Booth for financial support in earlier stages of this work. 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.

Relationship Lending and the Great Depression
Jon Cohen, Kinda Cheryl Hachem, and Gary Richardson
NBER Working Paper No. 22891
December 2016, Revised January 2020
JEL No. E44,G01,G21,L14,N22

ABSTRACT

The collapse of long-term lending relationships amplified the Great Depression. We demonstrate this by developing a new measure of lending relationships that can be calculated from widely available data at any level of aggregation. Our approach exploits differences in the responsiveness of loan rates to bank funding costs and is supported by historical evidence and theoretical arguments. The new measure reveals that the marginal impact of bank suspensions on economic activity was higher in more relationship-intensive areas, providing the first formal evidence that relationship lending propagated the real effects of banking sector distress in the early 1930s.

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

Kinda Cheryl Hachem
University of Virginia
Darden School of Business
100 Darden Boulevard
Charlottesville, VA 22906
and NBER
hachemk@darden.virginia.edu

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

1 Introduction

The Great Depression was the longest and deepest downturn in U.S. history. Scholars have long debated how the collapse of commercial banking in the early 1930s contributed to the contraction of commerce and industry. The pioneering work of Bernanke (1983) inspired a large literature that seeks to estimate the size and significance of this credit channel. The assumption that underlies almost all of these papers is that bank suspensions had real effects because they destroyed – or at the very least prevented the immediate redeployment of – some critical input into the bank lending process that impacts the cost of credit. Absent this assumption, it is difficult to explain why bank suspensions would contribute to such a deep and protracted decline in economic activity. But what exactly was this critical input and why was it so difficult to redeploy? Despite the abundant scholarship on the Great Depression, data limitations have impeded attempts to answer this question.

In this paper, we demonstrate that the critical input impaired by bank suspensions was the soft information that banks had acquired about the quality of their borrowers over the course of multi-period lending relationships. Soft information cannot be directly observed by the econometrician and the type of microeconomic data on bank-firm interactions used to construct proxies for it in modern analyses of relationship lending are unavailable for the 1920s and 1930s. We resolve this problem by developing a new measure of continuing lending relationships that can be constructed from data aggregated at any level or frequency, enabling comparison to the time-series and panel data typically analyzed by macroeconomists. We then use our measure to show that disruptions to continuing relationships as a result of bank suspensions in the early 1930s contributed to the severity of the Great Depression.

Our measure is based on the idea that loan rates charged in continuing relationships are less responsive to changes in bank funding costs than are those charged on other loans. We start by comparing average interest rates on two types of bank loans: one that historical sources describe as relational (commercial loans) and one that they describe as transactional (brokers' loans). We find that the average interest rate on the relationship loans was less responsive to changes in bank funding costs than was the average interest rate on the transactional loans. We present a model based on Hachem (2011) to help understand why this was the case. We choose this model because its assumptions are consistent with the institutional details of the period we study, namely the short-term and uncollateralized nature of commercial loans, the potential sequencing of these loans as part of longer-term lending relationships, and the competitive nature of bank lending.

At the heart of the model are the asymmetric information problems inherent in financial intermediation. The first is adverse selection, which exists because borrowers have private information about their ability to operate certain projects. The second is moral hazard, where higher loan rates increase a borrower's incentive to undertake projects with high default risk. Moral hazard generates a threshold loan rate above which borrowers shift their unobservable efforts towards riskier projects. Heterogeneous ability implies that this threshold depends on the borrower's type. A continuing relationship lender is able to use information gleaned about his borrower's type during previous interactions to determine the threshold and decide whether or not to pass an increase in bank funding costs through to the borrower. New relationship lenders and/or transactional lenders cannot condition this decision on the borrower's type because they lack information about it.

The model predicts that the incentive compatible contract offered by a continuing relationship lender eliminates risk-shifting by reducing the pass-through from bank funding costs to loan rates for some borrowers. From this, it can be shown that relationship loan rates are on average less elastic with respect to bank funding costs than are transactional loan rates. The difference is driven by continuing rather than new relationships. Moreover, the elasticity of the weighted average of all loan rates in a region is decreasing in the degree to which that region's banks are engaged in continuing relationships. Our new measure of continuing lending relationships is based on this last prediction.

We use data from the consolidated balance sheets and income statements of national banks, as reported by the Comptroller of the Currency in the 1920s, to implement our measure for the eve of the Great Depression. These data allow us to infer the weighted average loan rate at a semi-annual frequency for 82 locations, generating a panel with which we can calculate the elasticity of the loan rate with respect to the marginal cost of funds for banks (the discount rate) in each location. To control for location-specific differences in rates of return and their responsiveness, we also calculate the elasticity of securities returns with respect to the discount rate in each location and use the net elasticity as a gauge of loan rate responsiveness. Regions with relatively less elastic loan rates are those where the theory predicts bank portfolios contained a higher proportion of information-intensive continuing relationship loans. We rule out concentration of the local banking market and differential demand for credit as alternative explanations for the cross-sectional variation in net elasticity. Our measure detects more continuing relationships in areas where historical sources believe relationship lending prevailed, such as rural areas with small banks. That being said, we verify the economic content of net elasticity over and above other local characteristics by controlling for them in our cross-sectional regressions.

Using our new measure, we proceed to study the importance of relationship lending for understanding the real effects of banking distress in the early 1930s. Net elasticity is not correlated with banking distress, indicating that suspensions were not more common in relationship-intensive areas. We estimate that bank suspensions alone explain one-eighth of the national economic contraction observed during the Great Depression. Equally important, we show that the marginal impact of bank suspensions on economic activity was more severe in areas with more continuing relationships, other things the same. In other words, the destruction of continuing relationships amplified the Great Depression. The interaction between banking distress in the early 1930s and our measure of continuing relationships in the 1920s is a statistically significant predictor of cross-sectional differences in economic performance in both 1933 and 1935, suggesting that the real effects of destroying continuing relationships were long-lived.[1] Although Calomiris and Mason (2003), Richardson and Troost

[1] Rajan and Ramcharan (2016) also observe lingering effects of the Depression, but, unlike

(2009), Ziebarth (2013), and Carlson and Rose (2015) all find evidence that the collapse of commercial banks adversely affected the availability of credit and thus the activities of firms during this period, none, due to data limitations, manages to isolate the role of relationship lending in propagating the real effects of banking sector distress.[2]

In keeping with Bernanke (1983) and others, we focus on the Great Depression. No other crisis comes close in terms of the dimensions of the downturn and the largely unfettered nature of the banking panics. However, the value of our methodology transcends resolution of data limitations in the 1920s and 1930s. Policy-makers working in real time, often in crisis situations, usually have only aggregate data to inform them about what is happening at more disaggregated levels. Our method permits the extraction of more detailed information from these aggregates. Our findings also transcend the historical context of the Great Depression. Relationship lending remains a principal source of working capital in many parts of the world, facilitating the allocation of scarce financial resources to the full range of credit-worthy firms. The soft information on which these relationships depend is intrinsically difficult to transfer from one bank to another. Our results suggest that the flow of funds is more likely to be disrupted by failures of lenders who have accumulated knowledge of their borrowers through continuing relationships than by failures of those in the early stages of new relationships or of those who extend credit transactionally. This is a valuable lesson as policy-makers debate the scope for bank bailouts and the parameters of government safety nets.

The rest of the paper proceeds as follows. Section 2 provides historical background on

us, focus on banking concentration post-WWII, not on economic activity in the 1930s and the role of relationship lending.

[2]In contrast to the mainstream view, Cole and Ohanian (2007) argue that the collapse of the U.S. banking system in the early 1930s had few real economic effects, citing weak correlation between bank failures and economic outcomes, even at the state level. On this point, our results suggest that the correlation would be much stronger if calculated using only bank failures that destroyed continuing relationships.

the lending activities of Depression-era commercial banks to motivate our new measure of continuing lending relationships. Section 3 presents theoretical foundations for this measure and Section 4 describes its empirical implementation for the 1920s. Section 5 presents the key empirical results, using cross-sectional differences in continuing relationships in the 1920s to pinpoint the real effects of banking distress in the early 1930s. Section 6 concludes.

2 Historical Background

This section reviews the lending activities of Depression-era commercial banks and furnishes new stylized facts about the responsiveness of interest rates to motivate a new measure of continuing lending relationships that overcomes data limitations.

2.1 Relationship Versus Transactional Loans

The annual report of the Comptroller of the Currency provides information on U.S. bank balance sheets in the 1920s and 1930s in a format that is consistent across geographic locations. Up until 1928, the Comptroller divided bank loans into four categories: (i) real estate loans, (ii) loans on financial securities, (iii) uncollateralized loans, and (iv) loans collateralized by personal security. The last category consists primarily of loans secured by difficult-to-evaluate collateral such as goods in the process of production and distribution, warehouse receipts, and, in the case of farm loans, future crops.[3]

The sum of (iii) and (iv) constitutes commercial lending and is the largest category overall, accounting for an average of 67% of loans by national banks during the 1920s. In the years just prior to the Great Depression, almost 50% of national bank loans were uncollateralized, with another 10% collateralized by personal security. Category (ii), which consisted largely

[3]After 1928, the Comptroller combined uncollateralized loans and loans collateralized by personal security into one category and refined somewhat the criteria for loans on financial securities in line with the Fed's 1929 direct action campaign.

of loans to brokers, was the second largest category, accounting for an average of 28% of national bank loans during the 1920s. Category (i), real estate loans, was the smallest category, accounting for only 5% of lending by national banks over the same period.

Loans to brokers were short-term (e.g., call loans or 90-day time loans) and facilitated the purchase of stocks and bonds. Historical sources are unequivocal about the lack of relationship lending in brokers' loans. These loans were "usually made on an impersonal basis with the borrower and lender dealing through agents" and "not connected with established customer relationships" (Board of Governors 1943, p. 425). Thomas (1935) further notes that in the U.S., contrary to practices in London, the agents made little inquiry into the borrower's credit-worthiness. Brokers' loans in the U.S. were thus purely transactional, with no decisions made on the basis of soft information gathered in previous periods. The transactional nature of brokers' loans is also emphasized by Currie (1931).

In contrast, commercial lending, which provided working capital to merchants and manufacturers, was a prime venue for relationship building (e.g., Miller (1927), Foulke and Prochnow (1939)). The maturity of bank loans to commercial clients typically ranged from one to six months and never exceeded one year.[4] As noted earlier, commercial loans were either uncollateralized or collateralized by goods in the process of production and distribution. Such goods were difficult to value, costly to repossess, and, if liquidated, could be sold only with a long delay and/or at a considerable loss. For all intents and purposes, then, commercial loans were an unsecured form of lending. As a result, these loans tended to be relationship-based, in that firms borrowed repeatedly from the same bank, often also obtaining deposit, payment, and other intermediation services from this bank. This approach allowed banks to incorporate information acquired during the initial loan period into future credit terms, including the possibility of discontinuing the relationship (Langston (1921)).

[4]To be eligible for use as collateral at a Federal Reserve Bank's discount window, commercial loans had to mature in 180 days or less if collateralized by agricultural commodities and 90 days or less otherwise.

The commercial lending activities of Depression-era banks thus involved the production and use of information. This information was "soft" because it was acquired gradually through repeated personal interactions between banks and their clients. Personal interactions generate assessments of borrower quality that are difficult to standardize, transfer, or replicate without first replicating the entire relationship. Hard information instead involves facts that can be easily obtained prior to origination. By and large, businesses most dependent on banks for commercial loans were small with little or no public record. Surveys by the Department of Commerce and the National Industrial Conference Board determined that 86% of small manufacturing firms in 1929 and 72% of retailers, many of whom sold $12,000 or less of goods annually, depended on banks for working capital (DOC 1935, pp. 65-66; NICB 1932, p. 62).[5] Small, new, or otherwise opaque firms know their intrinsic quality and their motivation to acquire and repay credit; their banks do not but can learn by cultivating a lending relationship.[6]

Within the set of relationship loans, then, it is important to separate loans made as part of a continuing relationship from first-time relationship loans. Since soft information on borrower quality is accumulated over time through repeated interactions, it is embodied in continuing relationships, not new ones. Suspending a bank also suspends its accumulated knowledge and, because it would take time for other banks to replicate the underlying relationships, the flow of funds to firms will be disrupted. We thus need to distinguish be-

[5]Small manufacturers (i.e., those with no more than 250 employees) comprised 97% of all manufacturing firms in the U.S. and employed 48% of all manufacturing workers.

[6]While banks often had boards of local businessmen who could help assess the quality of firms in the community, repeated relationships allowed banks to accumulate information that even the best informed local businessman would not have had (e.g., high-frequency cash flow data obtained by monitoring the firm's deposit accounts, the identities of all parties that the firm paid by check or received checks from, proprietary business plans, etc.). This may have even contributed to local businessmen wanting to be on bank boards in the first place.

tween new and continuing relationships and focus sharply on the latter in order to accurately estimate the real effects of financial crises.

The available data on loan quantities in the 1920s and 1930s do not distinguish between stages of a commercial lending relationship: the Comptroller does not divide commercial loans in this way and we cannot use loan-level data to track interactions between borrower and lender pairs, as is often done in modern analyses of relationship lending, because such data do not exist for a representative sample of banks and firms in our period.[7] To overcome these data limitations, we propose to do with price what we cannot do with quantity. Specifically, we argue that interest income can be used to construct a measure that distinguishes not only between relationship and transactional loans but also between new and continuing relationships. The rest of this section provides some motivating evidence on the information content of interest rates using the distinction between relationship and transactional loans.

2.2 Interest Rate Responsiveness

In the decades surrounding the Great Depression, the Federal Reserve tracked interest rates on several types of bank loans in New York City. Figure 1 plots monthly average interest rates on new 90-day brokers' loans (black dots) and one-to-six month loans to commercial clients (gray dots) for the period January 1919 to December 1938. We also plot the marginal cost of funding a bank loan in New York City, as measured by the discount rate set by the Federal Reserve Bank of New York (black line).[8]

[7]Modern analyses include Berger and Udell (1995), Ongena and Smith (2001), Elsas (2005), and Chodorow-Reich (2014). In other contexts where data are a constraint, the use of historical narratives has often proved fruitful (Romer and Romer (1989, 2004, 2015)). However, we have yet to find a source that provides consistent narratives of local conditions across time and space in the Depression-era U.S.

[8]The discount rate was the marginal cost for commercial loans because banks could fund those loans by discounting them at the Fed. Discounting commercial loans to accommodate

Figure 1 shows that interest rates on brokers' loans rose swiftly and substantially when the discount rate rose. Interest rates on commercial loans moved in the same direction, but slowly and slightly. Even the Federal Reserve noted that "because of a variety of considerations arising out of customer relationships," rates on commercial loans behaved differently than rates on brokers' loans, which were arm's length and impersonal (Board of Governors 1943, p. 426). The Fed did not elaborate on these considerations (we will do so in Section 3) but Figure 1 clearly shows that commercial loan rates and interest rates on brokers' loans differed in their responsiveness to bank funding costs despite having similar maturities.

2.2.1 Contemporaneous Responses

Table 1 presents a more formal treatment of the data plotted in Figure 1. The top panel regresses monthly changes in the commercial loan rate on monthly changes in the discount rate. The middle panel regresses monthly changes in the brokers' loan rate on monthly changes in the discount rate. Column (1) uses all observations from January 1919 to December 1938. On average, the commercial loan rate rose by 19 basis points (bps) while the brokers' loan rate rose by 79 bps in a month when the NY Fed increased the discount rate by 100 bps. These responses differ significantly in statistical terms, as shown in the bottom panel, which rejects the null hypothesis of equality at the 1% significance level. Column (2) adds month fixed effects to control for any seasonal patterns in the data and reaches the same conclusion. Compared to brokers' loan rates, commercial loan rates responded much less substantially within a month to changes in the discount rate.

---

seasonal and cyclical peaks in credit demand was one of the principal motivations for the creation of the Federal Reserve System. The Fed also needed banks to discount loans as they were its primary source of income before it was allowed to hold securities in the 1930s. Loans on financial securities were technically not discountable, but the Fed expressed concern about banks using discount loans to invest in financial securities, suggesting that the discount rate was also the marginal cost for brokers' loans and security holdings by banks more generally.

The rest of Table 1 demonstrates the robustness of this result. Column (3) excludes observations where the NY Fed changed the discount rate by more than 50 bps. There are only 8 such observations in our sample but one may be concerned that they are outliers. Column (4) restricts the sample to months where the NY Fed's discount rate decision is known to have been driven by considerations outside of New York City.[9] Column (5) excludes May to September 1931 to control for the possibility that banks in New York City colluded to keep loan rates constant after the collapse of Creditanstalt in Austria in May 1931. Column (6) excludes observations before 1922 and after 1929 to control for economic turmoil during the Great Depression as well as direct intervention in credit markets by the federal government in the years following WWI and after the Depression. The columns in Table 1 reveal a consistent message: brokers' loan rates responded three to six times as much as commercial loan rates to changes in the discount rate.

2.2.2 Cumulative Responses

Table 2 addresses the possibility that interest rates responded to the discount rate with a lag. We regress changes in loan rates on the contemporaneous discount rate change, six months of lagged discount rate changes, and month fixed effects. Setting the coefficients on the lagged changes to zero would return the specification in the second column of Table 1.

Column (1) in Table 2 shows that, on average, when the NY Fed raised the discount rate by 100 bps, the brokers' loan rate rose by 78 bps in the same month and did not exhibit any statistically significant changes thereafter. The impact of discount rate changes

[9]These months include April 1924 to February 1925 (when the NY Fed lowered rates to help Britain return to the gold standard at its pre-war parity), February 1928 to September 1929 (when the NY Fed kept rates constant because the Federal Reserve Board forbade it from raising rates to control stock market speculation), and October 1931 to January 1932 (when the NY Fed raised rates to stem gold outflows from the U.S. following Britain's departure from the gold standard).

on brokers' loan rates was thus swift and substantial. The response of commercial loan rates was different. Column (2) shows that commercial loan rates in New York City rose by 18 bps in the same month as the discount rate change and continued to rise for at least two and perhaps as long as five months after it. However, the cumulative six-month response of the commercial loan rate (43 bps) remained well below the response of the brokers' loan rate.

Columns (3) and (4) explore the impact of discount rate changes on commercial loan rates outside of New York City. The Fed tracked a sample of these rates and reported one average for Northern and Eastern cities (N/E) and another for Southern and Western cities (S/W). Each average spans several Fed districts, with discount rates varying by district because the twelve Federal Reserve Banks had latitude to operate largely independent discount windows during the period we study (Richardson and Troost (2009)). We report results for the discount rate to which each average appears most responsive. All three commercial loan rates in Table 2 (NYC, N/E, and S/W) respond only partially to changes in the discount rate, reaching an average cumulative response of 45 bps after six months.

Column (5) explores the impact of discount rate changes on the commercial paper rate in New York City. Like commercial loans from banks, commercial paper financed goods in the process of production and sale, in contrast to brokers' loans which financed the purchase of securities. For the period we consider, the Fed kept track of interest rates on commercial paper maturing in four to six months in New York City. This maturity is similar to that of commercial loans. However, unlike the recipients of commercial loans, borrowers in the commercial paper market were large, well-known firms with good credit ratings and a public history of repayment. There was little information left to be acquired about them and hence no need for learning through relationship lending. Investors buying commercial paper and banks making brokers' loans thus shared the feature that neither used soft information acquired as part of a lending relationship.

The two open-market rates in Table 2 (brokers' loan rate and commercial paper rate) behave like each other, illustrating that the difference in the responsiveness of the interest rates plotted in Figure 1 reflects the nature of the loan (relational versus transactional), not the purpose of the loan (financing working capital versus buying securities). The open-market rates respond substantially to discount rate changes in the initial month, less in the next month, and not at all after a few months. Their average cumulative response is 82 bps after six months and the null hypothesis that it equals 100 bps by the second month cannot be rejected. The null hypothesis that the cumulative response of the open-market rates equals that of the commercial loan rates is rejected at the 1% significance level.

2.3 Discussion

The evidence in Tables 1 and 2 makes it clear that, in response to changes in the discount rate, interest rates on bank loans to commercial customers changed less than interest rates on loans that did not occur in the context of a lending relationship. The responsiveness of interest rates can therefore be used to distinguish relationship loans from transactional ones.

Why do differences in responsiveness exist? Before constructing a measure around them, it will be useful to connect the patterns in Section 2.2 to theory. Whatever model one uses, it has to be consistent with the key institutional features of the 1920s and 1930s, namely that the vast majority of bank loans were short-term and lending relationships were effectively uncollateralized. It must also incorporate the fact that the banking sector in the U.S., both nationally and locally, was much less concentrated and therefore much more competitive than it is today. Section 3 presents a model that embraces these institutional features.

The model makes the following predictions. First, the average interest rate on relationship loans will be less responsive to changes in bank funding costs than the interest rate on transactional loans, consistent with the evidence in Tables 1 and 2. Second, the average interest rate charged as part of a continuing relationship will be less responsive to changes in bank funding costs than the interest rate charged at the beginning of a relationship. Continuing relationships are therefore the primary force behind the differential responsiveness of relationship and transactional loan rates found in Section 2.2. Third, the responsiveness of the weighted average of all loan rates in a region to bank funding costs is decreasing in the degree to which that region's banks are engaged in continuing relationships.

This last prediction is important because it allows us to propose a new measure of continuing relationships that exploits changes in loan pricing over time with geographical areas as the unit of observation. We emphasize that our measure is based on differences in the responsiveness of loan rates to bank funding costs over the course of a relationship, not on the level of loan rates at a given point in time. Petersen and Rajan (1994) argue that lending relationships have little effect on the price of credit, but theirs is a result about the level of loan rates in a single cross-section, not a result about changes in loan rates over time. Based on the evidence above, the effects we are interested in are potentially quite large.[10]

3 New Indicator of Continuing Relationships: Theory

To help interpret the evidence in Section 2, we sketch a simple model of relationship lending based on Hachem (2011). We choose this model because its assumptions are consistent with the historical features of national bank lending described above. We then derive a prediction about loan rate elasticities (Section 3.3) that can be used in empirical work.

3.1 Environment

There are two stages and three periods. All agents are risk neutral. There is a continuum of firm types, denoted by ω and distributed uniformly over the unit interval. Types are private information and cannot be credibly communicated by firms. In each stage, firm ω can operate a limited liability production project that generates output $\theta_1 > 0$ with probability $p(\omega)$, where $p'(\omega) > 0$. The project delivers zero output otherwise. For simplicity, we consider

[10] Recent analyses of the transmission of the Lehman shock to Italy also suggest that relationships affect the pricing of credit (e.g., Gambacorta and Mistrulli (2014), Gobbi and Sette (2015), Bolton et al (2016)).

$p(\omega)$ linear. Output is independently distributed across firms and stages.

The firm's project requires one unit of capital input each time it is operated. Capital is available to a mass of ex ante identical banks at an exogenous policy rate r set by the central bank. Firms are not endowed with capital, nor can they store capital or output across stages. The credit contracts that transfer capital from banks to firms are uncollateralized and mature at the end of the stage in which they are signed.

In the first stage, banks are perfectly competitive and offer the same zero-profit interest rate R_1^*. All interest rate decisions are subject to a risk-shifting problem. Specifically, firms can divert capital into speculative projects that deliver $\theta_2 > \theta_1$ with probability $q < p(0)$ and nothing otherwise. We set $q\theta_2 = p(0)\theta_1$ to reduce notation. Banks cannot observe the exact value of output so interest rates cannot be contingent on whether θ_1 or θ_2 is realized. However, the presence of output is detectable so firms with positive output repay their banks.

The loan rate that makes firm ω indifferent between production and speculation is:

$$\overline{R}(\omega) \equiv \frac{p(\omega)\theta_1 - q\theta_2}{p(\omega) - q}$$

where $\overline{R}'(\omega) > 0$.[11] Firms select banks randomly in the first stage then decide which projects to undertake. Firm ω will choose the production project in the first stage if and only if $R_1^* \leq \overline{R}(\omega)$. All agents play the first-stage game in the first period.

At the beginning of the second period, banks and firms face an exogenous separation probability $s \in (0,1)$. A separated firm becomes a new firm, drawing a new type from the uniform distribution and playing the first-stage game again. An unseparated firm continues to the second stage and has its type discovered by its first-stage bank ("insider"). This discovery after one period of lending represents the soft information that a relationship gradually produces. The insider then decides whether to continue the relationship and

[11] At a given loan rate R, firm ω expects payoff $p(\omega)[\theta_1 - R]$ from production and $q[\theta_2 - R]$ from speculation. $\overline{R}(\omega)$ is the loan rate that equates these expected payoffs.

extend another unit of capital to the firm. If the insider wants to keep the firm, the interest rate cannot exceed what other banks ("outsiders") would optimally charge, else the firm will move to an outsider. The firm cannot credibly communicate its type to an outsider, in the same way that it could not credibly communicate its type when it first joined the insider. The relationship thus affords the insider an informational advantage over outsiders.

All interest rate decisions in the second stage (insider and outsider) are also subject to the risk-shifting problem, meaning firm ω chooses the speculative project in the second stage if and only if charged a second-stage interest rate above $\overline{R}(\omega)$. The insider's offer maximizes his profit subject to the risk-shifting problem, which he can now evaluate at the exact ω, and the ability of the firm to move to an outsider. Outsiders are perfectly competitive and offer the same zero-profit interest rate to firms that endogenously separate from their insiders. This rate will differ from the R_1^* offered to new (i.e., exogenously separated) firms because outsiders understand they are being adversely selected by endogenously separated firms.

The game ends at the end of the second stage. To generate the same first-stage interest rate R_1^* in the first and second periods, we model a third period where agents exogenously separated at the beginning of the second period play the second-stage game with probability $1 - s$, after having played the first-stage game during the second period.

3.2 Equilibrium Loan Rates

The equilibrium of the model for a given policy rate r can be described with reference to two cutoff types: $\widetilde{\omega}$ and $\widehat{\omega}$, defined by $p(\widetilde{\omega})\,\overline{R}(\widetilde{\omega}) \equiv r$ and $q\overline{R}(\widehat{\omega}) \equiv r$ respectively, where $\widetilde{\omega} < \widehat{\omega}$. We refer the reader to Hachem (2011) for a detailed proof and only describe here the results we will build on in Section 3.3.

In words, $\widetilde{\omega}$ is the type on which an insider breaks even by charging $\overline{R}(\cdot)$ while $\widehat{\omega}$ is the lowest type that chooses production if charged $\frac{r}{q}$. Types $\omega \leq \widetilde{\omega}$ move to outsiders and are charged a pooled interest rate $\frac{r}{q}$, prompting them to choose the speculative project. Types $\omega > \widetilde{\omega}$ stay with their insiders and are charged interest rates that lead them to choose production. Specifically, insiders match the prevailing outsider rate $\frac{r}{q}$ if $\omega \geq \widehat{\omega}$ but charge $\overline{R}(\cdot)$, which does not vary with r, if $\omega \in (\widetilde{\omega}, \widehat{\omega})$. If types $\omega \in (\widetilde{\omega}, \widehat{\omega})$ were instead charged $\frac{r}{q}$, they would choose the speculative project. Insiders are therefore using their soft information to mitigate risk-shifting, incentivizing higher repayment rates and better project selection by not passing through increases in r to some of the borrowers they retain. The economy thus benefits from continuing relationships.

For a given r, we can integrate over the type space to get the average interest rate charged in the second stage:

$$R_2^* = \int_0^{\widetilde{\omega}} \frac{r}{q} d\omega + \int_{\widetilde{\omega}}^{\widehat{\omega}} \overline{R}(\omega)\, d\omega + \int_{\widehat{\omega}}^{1} \frac{r}{q} d\omega$$

The region of policy-invariant loan rates is captured by the second integral. While the limits of integration ($\widetilde{\omega}$ and $\widehat{\omega}$) depend on r, the interest rates $\overline{R}(\cdot)$ do not.

Now imagine that fraction φ of the economy follows the model just described while fraction $1 - \varphi$ is characterized by transactional lending. We define a transactional loan to be a one-period contract where the bank and firm exogenously separate with certainty at the end of the period. A transactional loan is similar to a first-stage loan, except that the transactional lender has a continuation value of zero with the firm. Transactional lenders are perfectly competitive and offer the same zero-profit interest rate R_{TL}^* which equates to zero the expected profit from playing the first-stage game with a firm of unknown type (sans second-stage continuation value). In contrast, the first-stage interest rate R_1^* equates to zero the sum of the expected profit from playing the first-stage game with a firm of unknown type and the expected continuation value from playing the second-stage game if the firm's type exceeds $\widetilde{\omega}$. Exact formulas for R_1^* and R_{TL}^* appear in Appendix A.[12]

[12] Our modeling of transactional loans resembles the brokers' loans in Section 2 in that information is not acquired despite an information asymmetry existing between the borrower and lender. In the commercial paper market, information was not acquired because there was no such asymmetry. We model transactional loans more like brokers' loans because any transactional lending by banks beyond brokers' loans would have involved an information

3.3 Elasticity Prediction

The following proposition, proven in Appendix A, shows that the average interest rate on second-stage relationship loans is less elastic with respect to the policy rate than either the interest rate on first-stage relationship loans or the interest rate on transactional loans:

Proposition 1 *Let* $e_i \equiv \frac{dR_i^*}{dr}\frac{r}{R_i^*}$ *denote the elasticity of the interest rate* R_i^* *with respect to the policy rate* r*. The elasticities satisfy* $e_2 < \min\{e_1, e_{TL}\}$.

Consider now the weighted average loan rate in the economy during the second period, which is the main period in the model:

$$\mathcal{R} \equiv \varphi\left[sR_1^* + (1-s)\, R_2^*\right] + (1-\varphi)\, R_{TL}^*$$

Assuming the exogenous separation rate s and the transactional fraction $1-\varphi$ do not vary in a first-order way with the policy rate r, we can derive:

$$\frac{d\mathcal{R}}{dr}\frac{r}{\mathcal{R}} \approx \frac{\varphi s R_1^*}{\mathcal{R}}e_1 + \frac{\varphi\,(1-s)\, R_2^*}{\mathcal{R}}e_2 + \frac{(1-\varphi)\, R_{TL}^*}{\mathcal{R}}e_{TL} \tag{1}$$

where $\frac{d\mathcal{R}}{dr}\frac{r}{\mathcal{R}}$ is the elasticity of $\mathcal{R}$ with respect to the policy rate, $\frac{\varphi s R_1^*}{\mathcal{R}}$ is the fraction of interest income that comes from first-stage relationship loans, $\frac{\varphi(1-s)R_2^*}{\mathcal{R}}$ is the fraction from second-stage relationship loans, and $\frac{(1-\varphi)R_{TL}^*}{\mathcal{R}}$ is the fraction from transactional lending.

Taken together with Proposition 1, Eq. (1) implies that the weighted average loan rate in the economy will be less elastic with respect to the policy rate when banks are more heavily engaged in continuing relationships, as measured by a higher fraction of interest income coming from second-stage relationship loans. For our purposes, this means that areas with less elastic loan rates are areas where continuing relationships are more substantial, other things the same. It is this crucial insight that we exploit next in Section 4.

asymmetry given the small and private nature of bank-dependent borrowers. We refer the reader to Hachem (2011) for a discussion of interest rates when firm types are public.

4 Empirical Implementation for the 1920s

We have demonstrated that the elasticity of the weighted average loan rate with respect to bank funding costs will be lower in areas where continuing relationships are more substantial. We now translate our theoretical prediction into an empirical indicator of continuing lending relationships using the type of data available for the 1920s and 1930s in the U.S., namely the balance sheets and income statements of national banks aggregated by geographic region.[13]

4.1 Methodology

Let β_i^ℓ denote the elasticity of loan returns with respect to the discount rate in location i, calculated prior to the Great Depression. As discussed in Section 2, the discount rate was the marginal cost of funds for banks in the period we study. Discount rates differed across the twelve Federal Reserve districts, with each location i belonging to only one district d. For location i in district d, we estimate β_i^ℓ by running the regression:

$$\log\left(ReturnOnLoans_{i,t}\right) = \alpha_i^\ell + \lambda_i^\ell I_t^{dec} + \beta_i^\ell \log\left(DiscountRate_{d,t}\right) + \varepsilon_{i,t}^\ell \tag{2}$$

where t denotes time. The sample period is 1923 to 1929 and the frequency of observations in the regression is dictated by the availability of commercial bank data in the 1920s, discussed in more detail in Section 4.2. The dependent variable, $ReturnOnLoans_{i,t}$, is calculated by dividing the interest earnings of banks on loans in location i during date t by the stock of bank loans in location i at the end of date t. Appendix B shows that loan returns tended to be higher in the second half of the year. This pattern does not appear for the discount

[13] There is little micro data on how Depression-era banks adjusted loan rates charged to repeat customers. A survey by Ford (1928) of country banking practices in Northern Texas is the only study we found on the issue. Interestingly, Ford did find that relationship lenders often maintained a constant loan rate for borrowers they wanted to keep, even in times when the discount rate was changing.

rate, the interest paid on deposits, or the return on securities discussed below. We therefore include a dummy variable I_t^{dec} in Eq. (2) to control for seasonality in loan returns, with $I_t^{dec} = 1$ if t corresponds to a six-month period ending in December.

The intercepts α_i^ℓ in Eq. (2) control for any heterogeneity across locations that was constant over time. To also control for the possibility of time-varying heterogeneity unrelated to relationship lending, β_i^ℓ can be defined relative to the elasticity of other returns in location i that would have been similarly affected by such heterogeneity. The returns on financial securities held by banks provide a natural candidate. These securities portfolios had a substantial local component, hence their returns would have moved with local conditions but not been reflective of relationship lending as securities were not relationship-based. Accordingly, we estimate the elasticity of securities returns with respect to the discount rate in each location i prior to the Great Depression by running the regression:

$$\log\left(ReturnOnSecurities_{i,t}\right) = \alpha_i^s + \beta_i^s \log\left(DiscountRate_{d,t}\right) + \varepsilon_{i,t}^s \tag{3}$$

The variable $ReturnOnSecurities_{i,t}$ is calculated by dividing the interest earned by banks on securities in location i during date t by the securities holdings of banks in location i at the end of date t. The sample period for the estimation of β_i^s is the same as for β_i^ℓ.[14]

We then define the net elasticity for location i in the 1920s as:

$$NE20_i \equiv \beta_i^s - \beta_i^\ell$$

If $NE20_i > NE20_j$ for two locations i and j, it means that loan rates responded less in location i than in location j, relative to interest rates on other financial products, when the discount rate changed in the 1920s. We thus take $NE20_i > NE20_j$ to indicate the presence of more continuing relationships in location i than in location j on the eve of the Great Depression.

[14]From Appendix B, securities returns may be lower in the second half of the year outside of reserve cities. Our results are robust to adding I_t^{dec} to Eq. (3) when i is not a reserve city.

4.2 Data Sources

The dependent variables in Eq. (2) and (3) are constructed from the annual reports of the Comptroller of the Currency. For June and December of each year, the Comptroller published data on the earning assets of commercial banks aggregated by Fed district, state, and major municipality (reserve cities). Earning assets include loans, government bonds, and other financial securities. Data are reported separately for banks with national charters and banks with state charters. The Comptroller also published earnings and expense tables for nationally-chartered banks aggregated by district, state, and city.

The earnings and expense tables are essentially income statements. They were published as at June 30 of each year until 1925 when the Comptroller also began publishing separate tables for the last six months of the year. Earnings and expense tables were not published for state-chartered banks. Since these tables are necessary to construct the dependent variables in Eq. (2) and (3), we limit our sample to national banks. To the best of our knowledge, the literature has made little use of the income statement data in the Comptroller's reports.

The Comptroller began tabulating earnings on loans separately from earnings on government bonds and other securities in the second half of 1926. Prior to that, we estimate earnings on publicly-traded securities by applying market yields to the securities held on national bank balance sheets. We also estimate interest earned from balances at other banks and interest earned on Fed stock. We then subtract these estimates from the total amount of interest income in the Comptroller's table to isolate loan income.

Dividing loan income by the stock of loans at the end of each period and securities income by the stock of securities, we obtain the average returns needed to run Eq. (2) and (3). We then construct the discount rates in effect for each period as time-weighted averages of the discount rates reported in Board of Governors (1943). The final sample that we use in our subsequent analysis contains 82 locations each fully contained within a single Fed district: 33 reserve cities, 31 states (net of any of the 33 reserve cities), 12 split states (2 observations each for 6 states, again net of the reserve cities), and 6 district remainders. We focus on fully contained locations to use the appropriate discount rate when running Eq. (2) and (3).[15]

4.3 Net Elasticity Estimates

The distributions of β_i^ℓ, β_i^s, and $NE20_i$, as estimated from the data, are plotted in Figure 2, with summary statistics reported below the plot. Overall, the distribution of the net elasticities is roughly normal with mean slightly greater than zero. We are interested in the relative ranking of locations (e.g., whether $NE20_i$ is above or below $NE20_j$, as explained in Section 4.1) and do not infer much from the mean of the distribution.[16] The main takeaway from Figure 2 is that there is significant dispersion in net elasticity across locations in the 1920s, which is exactly what we need for our analysis. On average, net elasticity was lower in cities, but dispersion is clearly visible for both cities and states.

Table 3 reports cross-sectional correlations between $NE20_i$ and a variety of demographic and economic indicators from the 1920 population census. Net elasticity tended to be higher in areas that were more rural, as measured by a smaller urban population and/or more

[15] Split states arise because district boundaries run through some states. For a district with only one split state, we subtract from district-level data the fully contained states and reserve cities. Subtracting this result from the split state's total then recovers the part contained in another district. As long as the other district does not have more than two split states, we can repeat the process to back out any additional splits. We define remainders for districts where there are too many split states to be fully identified by this iterative procedure.

[16] If anything, the approximation used to separate loan income from securities income before the Comptroller began tabulating them separately (see Section 4.2) lowers the mean. Market yields average over different issues and thus tend to be smoother than the yields in any individual location. This will lead the approximation to understate (overstate) fluctuations in securities income (loan income), but its precision should not vary systematically across locations, particularly in the early 1920s when institutional investing was less advanced, so comparisons of net elasticity across locations and relative to the mean remain informative.

farming activity, regardless of whether we look at cities, states, or the entire sample. The correlation between the urbanization rate and the log of average bank size is 0.86 in our sample, meaning that more rural areas also tended to have smaller banks.[17] These characteristics arise frequently in historical anecdotes about relationship lenders (e.g., Ford (1928)) and can be intuited from theory. In contrast to big banks in urban centers, small rural banks were not relied upon to be liquidity providers to other financial institutions in emergencies or on short notice, making them less likely to have to suddenly sever relationships with non-financial borrowers for reasons unrelated to the borrower's health. This maps into Section 3 as a lower probability of exogenous separation s in the middle of a potential relationship. More relationships with policy-invariant loan rates would then be fostered in these areas, suggesting that we should indeed observe a higher net elasticity. The regressions we run in Section 5 will control for all of the local characteristics listed in Table 3.

We rule out concentration of the local banking market as an alternative explanation for cross-sectional variation in $NE20_i$. Neumark and Sharpe (1992) have argued that deposit rates are slower to rise in concentrated markets; related arguments also appear in Drechsler et al (2017). One may therefore worry that loan rates are stickier in some areas because market concentration makes deposit rates in those areas stickier, not because there are more continuing relationships. Using deposit shares compiled from the *Rand McNally Bankers' Directory*, we calculate a standard Herfindahl index (HHI) for each county in 1929. We assign to each reserve city in our sample the HHI of its county and to each state, split state, and district remainder the weighted average of its county-level HHIs. The correlation between loan rate elasticity (β_i^ℓ) and HHI is only -0.12 and the correlation between net elasticity ($NE20_i$) and HHI is only 0.16, indicating that our price-based measure of continuing relationships does not confound relationship lending and market power.

The relevance of deposit rates beyond market power can also be tested formally. We

[17] This correlation is around 0.7 when the sample is restricted to only cities or only states. In both subsamples, then, the correlation between urbanization and bank size is high.

use semi-annual data from 1925 to 1929 to explore the dynamics between deposit rates, as imputed from the expense tables published by the Comptroller, and loan returns.[18] For each location in our sample, we calculate the p-value for the null hypothesis that deposit rates did not Granger-cause loan rates. There is zero correlation between that p-value and net elasticity, indicating that the relationship between deposit rates and loan rates did not differ systematically across regions with more or less continuing relationships. We also find little evidence of deposit-driven loan pricing in our sample: we accept the null hypothesis that loan rates were not Granger-caused by deposit rates in almost 50% of our locations, and, of the locations where we reject this hypothesis, 67% are locations where we also reject the hypothesis that deposit rates were not Granger-caused by loan rates.

We also rule out differential demand for credit as a confounding factor. One may worry that loan rates appear stickier in some areas because loan demand moves in the same direction as loan supply following a discount rate change. Graphically, any comovement of demand and supply curves that neutralizes the response of prices amplifies the response of quantities, giving rise to an intuitive test of whether or not such movements in the demand for credit are reflected in our elasticity estimates.[19] For each location in our sample, we run a time series regression of loan growth on the change in the discount rate using semi-annual observations from 1923 to 1929. The coefficient from this regression (negative on average) is negatively correlated with loan rate elasticity (-0.25) and positively correlated with net elasticity (0.31). In other words, locations with stickier loan rates are not locations where lending falls by more in response to higher discount rates, indicating that our price-based measure of continuing relationships does not confound relationship lending and credit demand.

[18]Publication of earnings and expense tables for the second half of each year began in 1925 (see Section 4.2) so we drop 1923 and 1924 here to have equally spaced lags for Granger tests.

[19]While we believe that the netting of β_i^ℓ from β_i^s would absorb this issue if it exists, the test presented here helps evaluate whether the issue exists independently of whether it is or can be controlled for by netting.

We conclude this section with two additional sample statistics. First, the cross-sectional correlation between net elasticity in the 1920s and the average fraction of national bank loans made on little to no collateral over the same period is 0.35. While the sign of this correlation is what one would expect given historical accounts that relationship building rarely ever occurred outside of unsecured lending (see Section 2), the magnitude suggests that not all unsecured loans evolved into continuing relationships. Second, the change in net elasticity from the 1920s to the 1930s is negative for many locations and has a cross-sectional correlation of -0.91 with net elasticity in the 1920s, suggesting that the Great Depression destroyed continuing relationships in areas where such relationships existed.[20]

5 The Effects of Banking Distress in the Early 1930s

We now use our new measure of continuing lending relationships to explore the transmission of banking distress to the real economy during the Great Depression. While many have found that the collapse of commercial banking in the early 1930s adversely impacted the real economy, it remains unexplored whether propagation occurred through relationship lending.

5.1 Empirical Specification

Our first step is to define an indicator of banking distress that provides an accurate representation of its nature in the early 1930s. Conceptually, suspensions capture a broader notion of distress than failures. Many bank suspensions are permanent in that they culminate in failure; other suspensions are only temporary.[21] However, banks that temporarily suspend payments cannot extend loans while suspended and may change their lending activities when

[20]For comparison, the correlation between unsecured lending in the 1920s and the change in unsecured lending from the 1920s to the 1930s (also negative on average) is -0.46.

[21]On average, 25% of bank suspensions during the Great Depression were temporary with a mean suspension length of 102 days.

doors reopen (e.g., cut credit to bolster cash and head off further suspension). Both temporary and permanent suspensions are thus relevant forms of distress. We consider suspensions first then check the robustness of our results to using only failures.

Suppose bank suspensions in location i amounted to 10% of deposits. This could have been caused by the suspension of one bank with a 10% market share or by the suspension of ten banks each with a 1% market share. Although the size of the banking shock is the same in both cases – 10% of deposits – the suspension of many small banks more closely approximates the nature of the banking panics documented by Friedman and Schwartz (1963) and Wicker (1996) for the early 1930s. Accordingly, two separate indicators are needed to accurately describe distress during this period: one that captures the dispersion of distress across banks and another that controls for the size of the shock. We measure the dispersion of banking distress in location i by the fraction of banks suspended in the early 1930s:

$$SuspNum32_i = \sum_{t=1930}^{1932} \frac{number\_of\_suspended\_national\_banks_{i,t}}{number\_of\_national\_banks_{i,1929}}$$

The size of the shock is then measured using the share of deposits in suspended banks:

$$SuspVal32_i = \sum_{t=1930}^{1932} \frac{deposits\_in\_suspended\_national\_banks_{i,t}}{deposits\_in\_national\_banks_{i,1929}}$$

We use $SuspNum32_i$ as our main indicator of banking distress in each location i, with $SuspVal32_i$ as a control. Both variables are constructed from county-level data on bank suspensions compiled retrospectively by the Federal Deposit Insurance Corporation.

Next, we need a measure of the change in economic activity from 1929 to 1933 in each location i. Since our locations include reserve cities, states, split states, and district remainders, we need county-level information that we can aggregate into the appropriate units. This narrows our search to census data. The most precisely measured county-level outcome for the period we study is retail sales, collected as part of the Census of Business.[22] Retail

[22] We refer the reader to the end of Appendix C for a discussion of other indicators.

sales provide a general gauge of economic activity. In the 1920s and 1930s, well before the advent of mainstream consumer credit, the ability of households to purchase retail goods depended on labor income, which itself depended on the ability of firms in the local economy to pay those incomes. The struggles of local firms would thus be reflected in retail sales, regardless of whether or not the median firm was a retailer.

Letting $Sales_{i,t}$ denote retail sales in location i in year t, we run regressions of the form:

$$\begin{aligned} \frac{Sales_{i,1933}}{Sales_{i,1929}} &= \gamma_0 + \gamma_1 NE20_i + \gamma_2 SuspNum32_i + \gamma_3 SuspNum32_i * NE20_i \\ &\quad + \gamma_4 SuspVal32_i + \gamma_5 SuspVal32_i * NE20_i + \Gamma X_i + \upsilon_i \end{aligned} \tag{4}$$

where X_i is a vector of controls. Appendix B reports summary statistics for all variables. We first estimate Eq. (4) by ordinary least squares then use a two-stage procedure to address the possibility of endogeneity between economic activity and banking distress.

If banking panics had real effects because they arrested or otherwise impaired soft yet valuable information embodied in continuing lending relationships, we should find $\gamma_3 < 0$. That is, the negative effect of bank suspensions on economic activity should have been more pronounced in locations with more continuing relationships, all else the same. Intuitively, firms in continuing relationships lose a critical source of working capital when their banks are suddenly suspended. If the soft information accumulated over the course of a relationship is either unimportant or easy to redeploy, then the credit challenges faced by these firms should be the same as those faced by firms in otherwise similar locations and we should find no cross-sectional difference in the impact of bank suspensions on economic activity. If instead soft information does matter and is difficult to redeploy across lenders, then it will be harder and/or take longer for these firms to regain access to credit on the terms they had prior to suspension, leading to a more noticeable impact on the local economy.

Our empirical measure of continuing relationships, $NE20_i$, is as derived and implemented in Sections 3 and 4. It is based on the idea that firm funding costs respond less to bank funding costs when loans are made through continuing relationships. From this, one could also derive that continuing relationships reduce the response of economic activity to financial shocks that map broadly into shocks to bank funding costs (Hachem (2011)). However, this insulatory property of continuing relationships would be conditional on banks being operational, as was the case in the 1920s. If banks are instead suspended, their relationships are also suspended and locations dependent on those relationships will suffer to the extent that soft information matters and cannot be redeployed. It does not matter why exactly banks were suspended in the early 1930s as long as the suspensions were not driven by the relationships themselves. The correlation between $NE20_i$ and $SuspNum32_i$ is only 0.15, the correlation between $NE20_i$ and $SuspVal32_i$ is only 0.21, and, conditioned on other explanatory variables, both of these correlations are near zero.[23] In other words, suspensions were not more common in locations with more continuing relationships, removing the concern that banking distress was driven by these relationships.

Setting $NE20_i = 0$ in Eq. (4) returns the type of regression typically run in the literature on the Great Depression to make statements about the effect of banking distress on economic activity. A major shortcoming of the typical regression is its silence on the channels through which banking distress flowed to the real economy. By introducing our new measure of continuing relationships into Eq. (4), along with a battery of controls to absorb other differences across locations, we are able to use γ_3 to study whether the real effects of banking distress were propagated through relationship lending.

5.2 Baseline Results

Results from the estimation of Eq. (4) by ordinary least squares are reported in Table 4. In all columns, the control vector X_i includes district fixed effects, a dummy variable for whether location i is a reserve city, and the 1920 census controls listed in Table 3. Standard errors

[23]The coefficient on $NE20_i$ is near zero and not statistically significant at any reasonable level when added to the first-stage regressions for banking distress in Appendix C.

are bootstrapped, with block bootstrapping at the district level to account for clustering.

Column (1) presents the baseline regression with controls X_i as just described. The estimate of γ_3 is negative and statistically significant. Given the mean (0.057) and standard deviation (0.482) of the net elasticity distribution plotted in Figure 2, the magnitude of γ_3 in the first column of Table 4 indicates that suspending 10% of national banks in the early 1930s would have led to a 4.62% decline in retail sales between 1929 and 1933 in locations where net elasticity in the 1920s was one standard deviation above the mean. This is in contrast to only a 0.49% decline in comparable locations where net elasticity was at the mean. In other words, the marginal impact of banking distress on retail sales during the Great Depression was markedly more severe in areas with more continuing relationships.[24]

We also find that γ_5, the coefficient on the interaction between $SuspVal32_i$ and $NE20_i$ in Eq. (4), is positive and, in the first column of Table 4, statistically significant. Increasing $SuspVal32_i$ without also increasing $SuspNum32_i$ is akin to considering the effects of suspending larger banks. Receivership data for 1931 and 1932 reveal that large banks were liquidated quite rapidly, both in absolute terms and relative to smaller banks.[25] In practice, deposits in large banks could be redeemed even more quickly than the receivership data suggest because clearinghouses often provided advances to depositors of failed members, which would typically be the largest failures in the municipality. The suspension of a large bank

[24]As a placebo test, we replaced $NE20_i$ with the average fraction of national bank loans made on little to no collateral in location i during the 1920s. The results were not statistically different from zero by any reasonable metric, underscoring the importance of extracting continuing relationships from the broader pool of potential relationship loans.

[25]Using the sources described in Richardson (2008), we find that 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. The monthly disbursement was only 3.0 cents for banks that failed with \$2-6 million in deposits and 2.0 cents for those that failed with less than \$2 million in deposits.

would have thus created space in the business landscape for a surviving bank to attract deposits and make loans. The net effect of this reallocation would not exceed zero if the banks were otherwise identical, so we interpret the positive estimate for γ_5 as an indication that large banks tended to be less relationship-intensive than small banks in areas where business was overall more dependent on continuing relationships.[26]

Column (2) in Table 4 adds controls for concentration of the local banking market, specifically HHI as defined in Section 4.3 and its interactions with the banking distress indicators $SuspNum32_i$ and $SuspVal32_i$. The estimate of γ_5 is still positive but no longer statistically significant. The estimate of γ_3, however, remains negative and statistically significant, confirming that the marginal effect of banking distress on retail sales was more negative in areas with more continuing relationships.[27]

Column (3) uses an alternative formulation of $NE20_i$ in which we replace the discount rate on the right-hand side of Eq. (2) with the deposit rate in location i, as imputed from the Comptroller's expense tables, to recalculate β_i^ℓ as the elasticity of loan returns with respect to the average cost of deposits. The estimate of γ_3 in Table 4 is still negative and statistically significant. That said, we prefer the formulation of $NE20_i$ where loan rate elasticity is calculated with respect to the discount rate. In addition to influencing other funding costs through the yield curve, the discount rate was a direct cost of funds for commercial loans

[26]In unreported results, we reran the baseline regression without controlling for $SuspVal32_i$. The coefficient on the interaction between $SuspNum32_i$ and $NE20_i$ was still negative, but its magnitude was much smaller (-0.315 instead of -0.857) and not statistically significant. Running the baseline without $SuspNum32_i$ similarly muted the coefficient on the interaction between $SuspVal32_i$ and $NE20_i$. By including both $SuspNum32_i$ and $SuspVal32_i$ in the same regression, we capture different dimensions of banking distress and obtain sharper and more informative coefficients.

[27]This result is also robust to controlling for the deposit market share of national banks relative to all other banks (e.g., state banks) in 1929.

that banks could discount at their district Fed. The discount rate was also reported directly by the Fed (i.e., we do not need to impute it) and it captures the component of funding costs that commercial banks did not set for themselves, which is also more consistent with the cost of funds in our theoretical model. We therefore revert to the original formulation of $NE20_i$ for the remainder of the paper.

Columns (4) and (5) in Table 4 consider alternative measures of banking distress constructed from original sources. We still separate dispersion of distress from the size of the shock to reflect the nature of the banking panics in the early 1930s. However, instead of using national bank suspensions to calculate $SuspNum32_i$ and $SuspVal32_i$, we use national bank failures in column (4) and national bank mergers under duress in column (5).

The estimate of γ_3 in column (4) is negative and statistically significant. It is somewhat less negative than the estimate in column (2), possibly due to differences in the redeployment of soft information under various forms of distress. As described in Section 5.1, suspensions include both temporary suspensions and permanent suspensions (i.e., failures). Loan officers at failed banks would have had to look for new employment and may have taken their relationship borrowers with them to healthier financial institutions, in contrast with temporary suspensions where loan officers did not have to move but also could not operate as normal. Failures may therefore have been more conducive to the redeployment of soft information. There were, however, major impediments to redeploying soft information from failed banks in the early 1930s. First, new employment was not easy to find. Second, outside of a few big cities, failures were typically resolved through the court system, not through purchase and assumption. Many records were therefore trapped in the courts during the Depression, making it difficult for loan officers to transfer them to new banks. This would have delayed the reformation of old relationships, particularly if managers at the new banks wanted written records to confirm that the relationships were true and sound. On net, then, the estimate of γ_3 is negative and statistically significant, regardless of whether banking distress is measured using failures or suspensions.

The use of mergers under duress in column (5) provides a placebo test. By virtue of being under duress, these mergers indicate banking distress. However, the act of merger, as opposed to suspension or failure, bypasses the labor market frictions that loan officers would face looking for new jobs and avoids the trapping of records that would occur in the courts during bankruptcy proceedings, thus facilitating the transfer of soft information from a distressed bank to a successor. Accordingly, we should not expect a negative and statistically significant estimate of γ_3 when banking distress is measured using mergers under duress. This is confirmed in the fifth column of Table 4.

5.3 Additional Results

We conclude this section by discussing the results from a two-stage procedure to address the possibility of endogeneity between retail sales and banking distress in Eq. (4). Specifically, we want to make sure that we have isolated the drop in economic activity caused by the inability of relationship lenders to supply credit, not reverse causality from economic activity to bank distress through decreased demand for credit.

A series of surveys conducted during the Great Depression lend credence to the importance of credit supply shocks. The National Industrial Conference Board found that 22% of industrial firms that normally relied on commercial banks for working capital curtailed operations because of the refusal or restriction of bank credit (NICB 1932, pp. 5, 62).[28] Even firms that "would have readily commanded bank credit in normal times" because their credit ratings were positive, cash flows substantial, and balance sheets liquid and solvent were unable to obtain loans in the early 1930s (NICB 1932, pp. 94, 96, 98-99). Refusals often occurred because lenders' difficulties – typically withdrawals by depositors – forced them to curtail credit or call loans (NICB 1932, p. 111). A survey of small manufacturing

[28]The percentages were 31.1% for firms with 100 or fewer employees, 21.8% for firms with 101 to 250 employees, 19.0% for firms with 251 to 500 employees, 10.2% for firms with 501 to 1,000 employees, and 6.5% for firms with over 1,000 employees (NICB 1932, p. 69).

firms conducted by the Department of Commerce reached similar conclusions: 45% of small manufacturers reported an inability to borrow sufficient funds to finance all of their potential projects (DOC 1935, pp. 16-29) despite a majority of these manufacturers being liquid, solvent, and well-regarded by credit rating agencies (DOC 1935, pp. 41-61).

A simple test based on prices and quantities also indicates that credit supply shocks were of first-order importance in the early 1930s. In contrast to demand-driven decreases in bank lending, which should be accompanied by decreases in loan rates, supply-driven decreases in bank lending should be accompanied by increases in loan rates. Since the income statements described in Section 4.2 allow us to impute average loan returns in various locations during the Great Depression, we overcome the paucity of disaggregated data on loan pricing that would otherwise impede the implementation of this test. For each location in our sample, we compute the correlation between semi-annual loan growth and semi-annual changes in seasonally-adjusted loan returns over the period June 1930 to December 1932. The correlation is negative for 77% of our locations, indicating that loan supply shocks dominated loan demand shocks in the majority of the locations in our sample.

Both the survey evidence and the simple test just described lead us to interpret the results in Table 4 as being largely causal from banking distress to economic activity. However, to provide more formal identification, we use a two-stage strategy similar to the one used by Calomiris and Mason (2003) to identify loan supply shocks in the face of the loan demand critique in Hardy and Viner (1935) and Kimmel (1939). Appendix C provides details on the first-stage estimation and collects the second-stage regression tables. In addition to the specifications in columns (1) and (2) of Table 4, Appendix C runs second-stage regressions that control for interactions between the instrumented indicators of banking distress and variables from the 1920 census. Appendix C also runs regressions that interact $NE20_i$ and the instrumented banking distress indicators with the reserve city dummy to explore whether our results differ across types of locations. Finally, Appendix C runs regressions using the change in retail sales from 1929 to 1935 as the dependent variable, as well as regressions that use alternative measures of economic activity constructed from the Census of Manufactures.

The coefficient on the interaction between the instrumented $SuspNum32_i$ variable and $NE20_i$ is negative and statistically significant in all specifications considered in Appendix C, indicating, as before, that the marginal impact of banking distress on retail sales was more severe in areas with more continuing relationships. The null hypothesis that this coefficient does not differ across the types of locations that comprise our sample cannot be rejected and the coefficient remains negative and statistically significant when the dependent variable is measured in 1935. Thus, the real effects of banking distress were propagated through relationship lending in both cities and states and lingered into the recovery period.

Appendix C also uses the estimated coefficients to make statements about the aggregate implications of banking distress in the early 1930s. We find that retail sales in the U.S. would have been almost 4% higher in 1933 had there been no banking distress. We present aggregate evidence that a 4% drop in retail sales is consistent with a 6-7% drop in nominal GNP and a 3-4% drop in real GNP. This is about one-eighth of the economic contraction experienced during the Great Depression, underscoring the economic significance of our results.

6 Conclusion

We proposed 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. Our measure is based on the idea that longer relationships reveal to banks the highest interest rates they can charge their borrowers without inducing moral hazard, leading to loan rates that are less responsive to changes in bank funding costs. We presented historical and theoretical arguments in support of our measure then implemented it to study the transmission of bank suspensions to economic activity through the disruption of long-term lending relationships.

We found that cross-sectional differences in continuing relationships in the 1920s played a fundamental role in determining the real effects of banking distress in the early 1930s. Our new measure revealed that the marginal impact of bank suspensions on economic activity was much more severe in areas with more continuing relationships, other things the same. In other words, the destruction of continuing relationships amplified the Great Depression, providing the first formal evidence that relationship lending propagated the real effects of banking sector distress in the early 1930s.

In keeping with Bernanke (1983) and others, we focused on the Great Depression. The dimensions of the downturn and the largely unfettered nature of the banking panics make it ground zero for studying the real effects of financial crises. However, our findings are of interest beyond the historical context as relationship lending remains a principal source of working capital in many parts of the world. Our method is also useful for analysis of modern economies when loan-level data are either unavailable or available only with long lags.

By connecting the economic contraction to the collapse of long-term lending relationships, our results also suggest a new avenue for future work on the speed of economic recoveries. When business is dependent on the soft information that banks acquire over time, it is not enough for new banks to replace distressed ones. These banks also need time to interact with firms, learn about them, and rebuild continuing relationships, so, on top of contributing to the severity of the Great Depression, the banking panics of the early 1930s are likely to have contributed to what some believe was a surprisingly slow recovery in the wake of such a catastrophic economic collapse. Relationship rebuilding and the time to recovery from the Great Depression would therefore be an interesting extension for future work.

References

Berger, Allen and Gregory Udell. 1995. "Relationship Lending and Lines of Credit in Small Firm Finance." *Journal of Business*, 68(3), pp. 351-382.

Bernanke, Ben. 1983. "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression." *American Economic Review*, 73(3), pp. 257-276.

Board of Governors. 1943. *Banking and Monetary Statistics, 1914 to 1941.* Washington.

Bolton, Patrick, Xavier Freixas, Leonardo Gambacorta, and Paolo Emilio Mistrulli. 2016. "Relationship and Transaction Lending in a Crisis." *Review of Financial Studies*, 29(10), pp. 2643-2676.

Calomiris, Charles and Joseph Mason. 2003. "Consequences of Bank Distress During the Great Depression." *American Economic Review*, 93(3), pp. 937-947.

Carlson, Mark and Jonathan 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, Gabriel. 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, Harold and Lee 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, Lauchlin. 1931. "The Decline of the Commercial Loan." *Quarterly Journal of Economics*, 45(4), pp. 698-709.

Department of Commerce. 1935. *Survey of Reports of Credit and Capital Difficulties Submitted by Small Manufacturers*. Washington: Government Printing Office.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl. 2017. "The Deposits Channel of Monetary Policy." *Quarterly Journal of Economics*, 132(4), pp. 1819-1876.

Elsas, Ralf. 2005. "Empirical Determinants of Relationship Lending." *Journal of Financial Intermediation*, 14(1), pp. 32-57.

Ford, A.W. 1928. "Country Banking Practices in Northern Texas." *Journal of Business*, 1(4), pp. 468-484.

Foulke, Roy and Herbert Prochnow. 1939. *Practical Bank Credit*. New York: Prentice-Hall.

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

Gambacorta, Leonardo and Paolo Emilio Mistrulli. 2014. "Bank Heterogeneity and Interest Rate Setting: What Lessons Have We Learned since Lehman Brothers." *Journal of Money, Credit and Banking*, 46(4), pp. 753-778.

Gobbi, Giorgio and Enrico Sette. 2015. "Relationship Lending During a Financial Crisis." *Journal of the European Economic Association*, 13(3), pp. 453-481.

Hachem, Kinda. 2011. "Relationship Lending and the Transmission of Monetary Policy." *Journal of Monetary Economics*, 58(6-8), pp. 590-600.

Hardy, Charles and Jacob Viner. 1935. *Report on the Availability of Bank Credit in the Seventh Federal Reserve District*. Washington: Government Printing Office.

Kimmel, Lewis. 1939. *The Availability of Credit, 1933-1938*. New York: Conference Board.

Langston, L.H. 1921. *Practical Bank Operation*. New York: Ronald Press.

Miller, Mahlon. 1927. *Bank Loans on Statement and Character*. New York: Ronald Press.

National Industrial Conference Board. 1932. *The Availability of Bank Credit*. New York.

Neumark, David and Steven Sharpe. 1992. "Market Structure and the Nature of Price Rigidity: Evidence from the Market for Consumer Deposits." *Quarterly Journal of Economics*, 107(2), pp. 657-680.

Ongena, Steven and David Smith. 2001. "The Duration of Bank Relationships." *Journal of Financial Economics*, 61(3), pp. 449-475.

Petersen, Mitchell and Raghuram Rajan. 1994. "The Benefits of Lending Relationships: Evidence from Small Business Data." *Journal of Finance*, 49(1), pp. 3-37.

Rajan, Raghuram and Rodney Ramcharan. 2016. "Crises and the Development of Economic Institutions: Some Microeconomic Evidence." *American Economic Review P&P*, 106(5), pp. 524-527.

Richardson, Gary. 2008. "Quarterly Data on the Categories and Causes of Bank Distress during the Great Depression." *Research in Economic History*, Vol. 25, pp. 37-115.

Richardson, Gary and William Troost. 2009. "Monetary Intervention Mitigated Banking Panics During the Great Depression: Quasi-Experimental Evidence from the Federal Reserve District Border in Mississippi." *Journal of Political Economy*, 117(6), pp. 1031-1073.

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

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

Thomas, Woodlief. 1935. "Use of Credit in Security Speculation." *American Economic Review*, 25(1), pp. 21-30.

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

Ziebarth, Nicolas. 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.

Table 1: Contemporaneous Interest Rate Responses to Discount Rate Changes

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	± 0.5	External	No Collude	1920s
Dependent Variable: Δ Comm. Loan Rate, NYC						
Δ Discount Rate	0.185***	0.181***	0.187***	0.180***	0.182***	0.141***
	(0.041)	(0.039)	(0.060)	(0.042)	(0.039)	(0.044)
Dependent Variable: Δ Brokers' Loan Rate						
Δ Discount Rate	0.789***	0.777***	0.530***	0.742***	0.777***	0.864***
	(0.135)	(0.128)	(0.125)	(0.060)	(0.129)	(0.186)
H_0: Equality of Coefficients						
Chi-squared	26.51	27.76	10.40	53.64	27.04	13.30
P-value	0.0000	0.0000	0.0013	0.0000	0.0000	0.0003
Observations	239	239	231	34	234	96
Month FE	NO	YES	YES	YES	YES	YES

Notes: Top and middle panels regress monthly changes in the indicated interest rate on contemporaneous changes in the NY Fed discount rate. Bottom panel tests whether the coefficients in the top and middle panels are equal. Column (1) uses all observations from January 1919 to December 1938. Column (2) adds dummies for month of the year. Column (3) excludes observations where the discount rate changed by more than 50 bps. Column (4) uses only months where NY Fed decisions are known to have been driven by external considerations. Column (5) excludes the period of possible rate collusion after the collapse of Creditanstalt in Austria. Column (6) uses only observations from the 1920s. Robust standard errors are in brackets. ***p<0.01, **p<0.05, *p<0.1

Table 2: Lagged Interest Rate Responses to Discount Rate Changes

	(1)	(2)	(3)	(4)	(5)
Months Since Change	Δ Brokers'	Δ Commercial Loan Rate			Δ Comm.
in Discount Rate	Loan Rate	NYC	N/E Cities	S/W Cities	Paper Rate
0	0.781***	0.177***	0.093***	0.086***	0.507***
1	0.108	0.118**	0.098***	0.072***	0.260***
2	0.041	0.072**	0.070**	0.059***	0.067**
3	-0.078	0.028	0.082**	0.057***	0.022
4	-0.039	-0.011	0.093***	0.117***	0.001
5	-0.023	0.078***	0.059**	0.030	-0.013
6	0.010	-0.036	-0.017	0.019	0.003
Cumulative	0.800***	0.425***	0.479***	0.440***	0.847***
Discount Rate	New York	New York	Boston	Richmond	New York

Notes: Sample period is January 1919 to December 1938. Each column regresses monthly changes in the indicated interest rate on contemporaneous and lagged changes in the indicated discount rate. All columns include six months of lagged changes and dummies for month of the year. N/E and S/W are averages constructed by the Fed. N/E includes Chicago, Boston, Philadelphia, Pittsburgh, Buffalo, Cleveland, Detroit; S/W includes San Francisco, Los Angeles, St. Louis, Dallas, Minneapolis, Kansas City, New Orleans, Seattle, Atlanta, Baltimore, Richmond (Board of Governors 1943, p. 427). ***$p<0.01$, **$p<0.05$, *$p<0.1$

Table 3: Net Elasticity and 1920 Census Variables, Correlations

		Correlation with NE20		
		All	Cities	States
LogArea	log(area in square miles)	0.426	0.237	0.220
LogPop	log(population in 1920)	0.132	-0.278	0.024
Urban	% population urban in 1920	-0.521	-0.374	-0.449
Nwnp	% population with native white parents in 1920	0.302	0.467	0.128
Age1844	% population aged 18 to 44 in 1920	-0.414	-0.051	-0.351
School1620	% of aged 16 to 20 in school in 1920	0.284	0.377	0.155
LogMfgEstPc	log(no. of mfg establishments per capita in 1920)	-0.410	-0.217	-0.419
MfgWork	mfg workers as % of population in 1920	-0.365	-0.241	-0.438
LogMfgVa	log(value added per mfg establishment in 1920)	-0.268	0.061	-0.336
LogFarmsPc	log(no. of farms per capita in 1920)	0.521	0.401	0.431
Acres	farmland as % of area in 1920	0.251	0.402	0.182
LogAvgCrop	log(crop value per farm in 1920)	0.175	0.271	0.064
LogAvgValue	log(value of farmland, equip, etc per farm in 1920)	-0.099	0.054	0.040
HomeOwnClr	home ownership rate in 1920	0.431	0.290	0.307

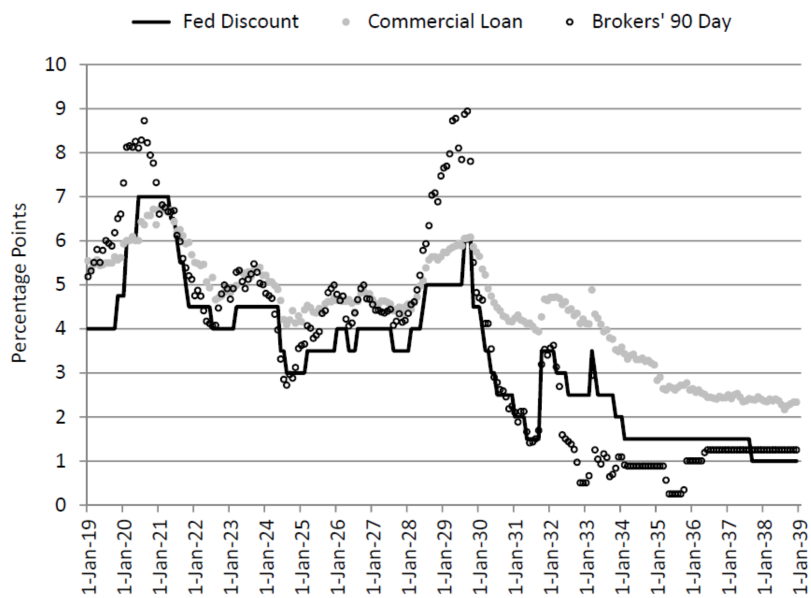
Notes: Manufacturing is abbreviated mfg. "All" calculates the cross-sectional correlation between the indicated variable and NE20 using the full sample (82 locations), "Cities" uses only the 33 reserve cities, and "States" uses only the 49 states, split states, and district remainders.

Table 4: OLS Estimates, Retail Sales Ratio, 1933/29

VARIABLES	(1) Baseline	(2) Control HHI	(3) Alt. NE20	(4) Using Failures	(5) Using Mergers
NE20	0.024	0.029	0.008	0.031	0.001
	(0.024)	(0.028)	(0.021)	(0.029)	(0.023)
SuspNum32	-0.170	-0.116	0.021	-0.213	-0.240
	(0.105)	(0.229)	(0.261)	(0.296)	(0.895)
SuspNum32 x NE20	-0.857***	-0.804**	-0.544**	-0.530*	0.848
	(0.304)	(0.350)	(0.242)	(0.287)	(1.608)
SuspVal32	0.107	-0.170	-0.229	-0.044	-0.721
	(0.125)	(0.279)	(0.332)	(0.363)	(0.929)
SuspVal32 x NE20	0.883*	0.609	0.461	0.096	-1.667
	(0.453)	(0.525)	(0.346)	(0.378)	(1.551)
HHI		0.005	0.037	0.019	0.034
		(0.124)	(0.128)	(0.125)	(0.126)
SuspNum32 x HHI		-0.297	-0.828	-0.062	-0.394
		(0.928)	(0.980)	(1.127)	(4.092)
SuspVal32 x HHI		1.249	1.457	0.997	6.866
		(1.028)	(1.136)	(1.253)	(5.300)
Observations	82	82	82	82	82
Adjusted R-squared	0.596	0.621	0.608	0.582	0.459
District FE	YES	YES	YES	YES	YES
Census Controls	YES	YES	YES	YES	YES

Notes: In all columns, the dependent variable is retail sales in 1933 as a fraction of retail sales in 1929. NE20 is the measure of continuing relationships on the eve of the Great Depression. With the exception of column (3), where the loan rate elasticity component of NE20 is calculated with respect to the average cost of deposits, NE20 is always calculated with respect to the discount rate. SuspNum32 and SuspVal32 measure respectively the number and deposits of distressed national banks from 1930 to 1932 relative to the number and deposits of all national banks in 1929. Distress is defined using suspensions in columns (1)-(3), failures in column (4), and mergers under duress in column (5). HHI measures local banking market concentration in 1929. Bootstrapped standard errors clustered at the district level are in brackets. ***$p<0.01$, **$p<0.05$, *$p<0.1$

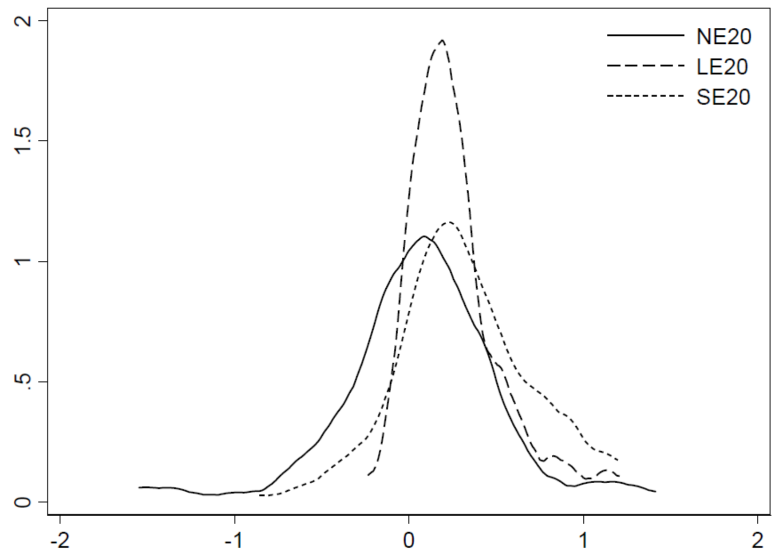
Figure 1: Illustrative Example from NYC Rates



Notes: Data from Board of Governors (1943) and the NBER Macrohistory database. All interest rates are monthly for New York City. Black dots plot the average interest rate on new 90-day brokers' loans, gray dots plot the average interest rate on one-to-six month loans to commercial clients, and the black line plots the discount rate set by the Federal Reserve Bank of New York.

Figure 2: Elasticity Estimates in the 1920s

(a) Kernel Density Estimates



(b) Summary Statistics

	All			Cities			States		
	Mean	Median	Sd	Mean	Median	Sd	Mean	Median	Sd
NE20	0.057	0.062	0.482	-0.161	-0.033	0.509	0.204	0.188	0.406
LE20	0.268	0.219	0.281	0.415	0.318	0.337	0.169	0.160	0.182
SE20	0.326	0.266	0.398	0.254	0.224	0.457	0.374	0.293	0.348

Notes: LE20 is the elasticity of loan returns (β_i^ℓ), SE20 is the elasticity of securities returns (β_i^s), and NE20 is the net elasticity ($NE20_i = \beta_i^s - \beta_i^\ell$). The table reports summary statistics for the full sample ("All"), the 33 reserve cities ("Cities"), and the 49 states, split states, and district remainders ("States"). The figure plots kernel density estimates from an Epanechnikov kernel using the full sample (82 locations).

Online Appendix for

"Relationship Lending and the Great Depression"

Jon Cohen — University of Toronto

Kinda Hachem — UVA Darden and NBER

Gary Richardson — UC Irvine and NBER

December 2019

Appendix A – Proof of Proposition 1

To get the first-stage interest rate R_1^*, define $R_1^* \equiv \overline{R}(\xi)$ so that ξ denotes the firm type that is exactly indifferent between production and speculation when charged R_1^*. In equilibrium, ξ is pinned down by the zero-profit condition for lenders with new firms, namely:

$$\underbrace{\left[q\xi + \int_{\xi}^{1} p(x)\,dx\right]\overline{R}(\xi) - r}_{\text{expected profit from first stage}} + \underbrace{\beta(1-s)\left[\int_{\widetilde{\omega}}^{\widehat{\omega}} \left[p(x)\overline{R}(x) - r\right]dx + \int_{\widehat{\omega}}^{1}\left(p(x)\frac{r}{q} - r\right)dx\right]}_{\text{expected present discounted continuation value}} = 0$$

where $\beta \in (0,1)$ is the lender's discount factor. The lender's continuation value comes from playing the second-stage game with the firm and earning profit on types above $\widetilde{\omega}$.

To get the transactional interest rate R_{TL}^*, define $R_{TL}^* \equiv \overline{R}(\eta)$ so that η denotes the firm type that is exactly indifferent between production and speculation when charged R_{TL}^*. In equilibrium, η is pinned down by the zero-profit condition for transactional lenders, namely:

$$\left[q\eta + \int_{\eta}^{1} p(x)\,dx\right]\overline{R}(\eta) - r = 0$$

There is no second-stage continuation value because transactional loans are one-period contracts where the bank and firm separate with probability 1 at the end of the period.

The elasticities of the relationship lending interest rates, R_1^* as defined above and R_2^* as defined in the main text, are:

$$e_1 = \frac{\left[q\xi + \int_{\xi}^{1} p(x)\,dx\right]\overline{R}'(\xi) + \beta(1-s)\frac{\overline{R}'(\xi)}{\overline{R}(\xi)}\int_{\widetilde{\omega}}^{\widehat{\omega}} p(x)\overline{R}(x)\,dx}{\left[q\xi + \int_{\xi}^{1} p(x)\,dx\right]\overline{R}'(\xi) - \left[p(\xi) - q\right]\overline{R}(\xi)}$$

and:

$$e_2 = \frac{1 - \widehat{\omega} + \widetilde{\omega} + \left[1 - \frac{q}{p(\widetilde{\omega})}\right]\frac{r}{p'(\widetilde{\omega})\overline{R}(\widetilde{\omega}) + p(\widetilde{\omega})\overline{R}'(\widetilde{\omega})}}{1 - \widehat{\omega} + \widetilde{\omega} + \frac{1}{\overline{R}(\widetilde{\omega})}\int_{\widetilde{\omega}}^{\widehat{\omega}} \overline{R}(\omega)\,d\omega}$$

respectively. The elasticity of the transactional interest rate R_{TL}^* is:

$$e_{TL} = \frac{\left[q\eta + \int_{\eta}^{1} p(x)\,dx\right]\overline{R}'(\eta)}{\left[q\eta + \int_{\eta}^{1} p(x)\,dx\right]\overline{R}'(\eta) - \left[p(\eta) - q\right]\overline{R}(\eta)} \equiv f(\eta)$$

The denominators of e_1 and e_{TL} are positive under minimal conditions on $p(\cdot)$; see the online appendix of Hachem (2011) for a formal proof. Accordingly, $e_1 > f(\xi)$ and it will suffice to show $e_2 < f(z)$ for any $z \in (0,1)$ satisfying:

$$\left[qz + \int_z^1 p(x)\,dx\right]\overline{R}(z) \leq r \tag{A.1}$$

The definitions of ξ and η above imply that (A.1) holds with strict equality at $z = \eta$ and strict inequality at $z = \xi$. Use the expressions for e_2 and $f(z)$ to rewrite $e_2 < f(z)$ as:

$$\begin{aligned} &\left[qz + \int_z^1 p(x)\,dx\right]\overline{R}'(z)\left[\frac{[p(\widetilde{\omega}) - q]\,\overline{R}(\widetilde{\omega})}{p'(\widetilde{\omega})\,\overline{R}(\widetilde{\omega}) + p(\widetilde{\omega})\,\overline{R}'(\widetilde{\omega})} - \frac{1}{\overline{R}(\widehat{\omega})}\int_{\widetilde{\omega}}^{\widehat{\omega}} \overline{R}(\omega)\,d\omega\right] \\ < \;\; &[p(z) - q]\,\overline{R}(z)\left[1 - \widehat{\omega} + \widetilde{\omega} + \frac{[p(\widetilde{\omega}) - q]\,\overline{R}(\widetilde{\omega})}{p'(\widetilde{\omega})\,\overline{R}(\widetilde{\omega}) + p(\widetilde{\omega})\,\overline{R}'(\widetilde{\omega})}\right] \end{aligned} \tag{A.2}$$

Linearity of $p(\omega)$ implies $p(\omega) = p(0) + [p(1) - p(0)]\,\omega$ so we can write (A.2) as:

$$\begin{aligned} &\frac{[p(\widetilde{\omega}) - q]\,\overline{R}(\widetilde{\omega})}{p'(\widetilde{\omega})\,\overline{R}(\widetilde{\omega}) + p(\widetilde{\omega})\,\overline{R}'(\widetilde{\omega})} - \frac{1}{\overline{R}(\widehat{\omega})}\int_{\widetilde{\omega}}^{\widehat{\omega}} \overline{R}(\omega)\,d\omega \\ < \;\; &\frac{[p(z) - q]\,z}{p(0) - q}\,\frac{p(z) - q}{qz + \int_z^1 p(x)\,dx}\left[1 - \widehat{\omega} + \widetilde{\omega} + \frac{[p(\widetilde{\omega}) - q]\,\overline{R}(\widetilde{\omega})}{p'(\widetilde{\omega})\,\overline{R}(\widetilde{\omega}) + p(\widetilde{\omega})\,\overline{R}'(\widetilde{\omega})}\right] \end{aligned}$$

where we have also used $p(0)\,\theta_1 = q\theta_2$. From (A.1):

$$\frac{p(z) - q}{qz + \int_z^1 p(x)\,dx} \geq \frac{[p(1) - p(0)]\,\theta_1 z}{r}$$

so a sufficient condition for (A.2) is:

$$\begin{aligned} &\frac{[p(\widetilde{\omega}) - q]\,\overline{R}(\widetilde{\omega})}{p'(\widetilde{\omega})\,\overline{R}(\widetilde{\omega}) + p(\widetilde{\omega})\,\overline{R}'(\widetilde{\omega})} - \frac{1}{\overline{R}(\widehat{\omega})}\int_{\widetilde{\omega}}^{\widehat{\omega}} \overline{R}(\omega)\,d\omega \\ < \;\; &\frac{[p(z) - q]\,z^2}{p(0) - q}\,\frac{[p(1) - p(0)]\,\theta_1}{r}\left[1 - \widehat{\omega} + \widetilde{\omega} + \frac{[p(\widetilde{\omega}) - q]\,\overline{R}(\widetilde{\omega})}{p'(\widetilde{\omega})\,\overline{R}(\widetilde{\omega}) + p(\widetilde{\omega})\,\overline{R}'(\widetilde{\omega})}\right] \end{aligned}$$

Going through the algebra, we can rewrite the sufficient condition as:

$$\begin{aligned} &\frac{p(\widetilde{\omega})\,\widetilde{\omega}^2}{[p(1) - p(0)]\,\widetilde{\omega} + p(\widetilde{\omega})\,\frac{p(0) - q}{p(\widetilde{\omega}) - q}} - \frac{q(\widehat{\omega} - \widetilde{\omega})\,\widetilde{\omega}}{p(\widetilde{\omega}) - q} \\ < \;\; &\frac{[p(z) - q]\,z^2}{p(0) - q}\left[1 - \widehat{\omega} + \widetilde{\omega} + \frac{[p(\widetilde{\omega}) - q]\,\overline{R}(\widetilde{\omega})}{p'(\widetilde{\omega})\,\overline{R}(\widetilde{\omega}) + p(\widetilde{\omega})\,\overline{R}'(\widetilde{\omega})}\right] + \frac{\int_{\widetilde{\omega}}^{\widehat{\omega}}\left[\overline{R}(\omega) - \overline{R}(\widetilde{\omega})\right]d\omega}{[p(1) - p(0)]\,\frac{\theta_1}{q}} \end{aligned}$$

The right-hand side is positive so it will be enough for the left-hand side to be negative.

What we want to show is therefore:

$$\frac{p\left(\widetilde{\omega}\right)\widetilde{\omega}}{\left[p\left(1\right)-p\left(0\right)\right]\widetilde{\omega}+p\left(\widetilde{\omega}\right)\frac{p(0)-q}{p(\widetilde{\omega})-q}}<\frac{q\left(\widehat{\omega}-\widetilde{\omega}\right)}{p\left(\widetilde{\omega}\right)-q} \tag{A.3}$$

Use $p\left(\widetilde{\omega}\right)\overline{R}\left(\widetilde{\omega}\right)=q\overline{R}\left(\widehat{\omega}\right)$ to isolate:

$$\widehat{\omega}=\frac{\left[p\left(0\right)-q\right]p\left(\widetilde{\omega}\right)\widetilde{\omega}}{q\left[p\left(\widetilde{\omega}\right)-q\right]-\left[p\left(1\right)-p\left(0\right)\right]p\left(\widetilde{\omega}\right)\widetilde{\omega}}$$

We can then rewrite (A.3) as:

$$\left[\left(p\left(\widetilde{\omega}\right)\right)^{2}+q\left[p\left(\widetilde{\omega}\right)-q\right]\right]\left[p\left(1\right)-p\left(0\right)\right]\widetilde{\omega}>p\left(\widetilde{\omega}\right)q\left[p\left(\widetilde{\omega}\right)-p\left(0\right)\right]$$

This simplifies to $p\left(\widetilde{\omega}\right)>q$, which is true. ■

References

Hachem, K. 2011. "Relationship Lending and the Transmission of Monetary Policy." *Journal of Monetary Economics*, 58(6-8), pp. 590-600.

Appendix B – Summary Statistics

Table B.1 reports summary statistics for the variables used in Eq. (2) and (3), namely the return on loans, the return on securities, and the discount rate. We also report summary statistics for the interest paid on (demand and time) deposits.

There are 82 locations in our final sample: 33 reserve cities, 31 states (net of the reserve cities), 12 parts of states (that aggregate up to 6 states net of the reserve cities), and 6 district remainders. The earnings and expense tables used to construct the time series dimension of the panel are available annually to semi-annually. As explained in Section 4.2, the Comptroller of the Currency began publishing a July-to-December table in 1925, where July-to-December means that the table aggregates over the last six months of the year. Prior to that, the Comptroller published annual earnings and expense tables with June 30 as the year-end. The table ending June 30 naturally became a January-to-June table after the publication of the first July-to-December table.

Starting with the July-to-December table for 1926, the Comptroller also began separating earnings by asset class (i.e., loan income reported separately from securities income). For prior observations, we use market yields and data on the composition of the securities portfolios of national banks in each location to impute an income breakdown between loans and securities. Detailed data on securities portfolios is reported by the Comptroller for states and reserve cities but not for districts. Constructing the split states and district remainders requires district-level data (see again Section 4.2) so, for these locations, we do not impute income breakdowns prior to the second half of 1926.

Tables B.2 to B.5 report summary statistics for the variables used in Section 5 and Appendix C. We present means, medians, and standard deviations first for the full sample (82 locations) then separately for the 33 reserve cities and 49 states, split states, and district remainders.

Table B.2 reports summary statistics for the dependent variable in Eq. (4). Retail33/29 is the ratio of retail sales in 1933 to retail sales in 1929. This is the dependent variable used in Table 4, as well as in Table C.2 and column (1) of Table C.3 in Appendix C. Retail35/29 is the ratio of retail sales in 1935 to retail sales in 1929. This is the dependent variable used in column (2) of Table C.3.

Table B.3 reports summary statistics for the control variables used in various estimations of Eq. (4), specifically the 1920 census variables defined in Table 3 and banking sector HHI as defined in Section 4.3.

Table B.4 reports summary statistics for the banking distress indicators defined in Section 5.1. SuspNum32 is the number of national banks suspended from 1930 to 1932 as a fraction of the number of national banks in 1929. SuspVal32 is deposits in national banks suspended from 1930 to 1932 as a fraction of deposits in national banks in 1929.

Table B.5 reports summary statistics for the variables used to instrument banking distress in Appendix C. The variable names are as defined in the notes to Table C.1.

Table B.6 reports summary statistics for alternative measures of economic activity based on the Census of Manufactures. Table C.5 in Appendix C runs two-stage versions of Eq. (4) with these measures as the dependent variable instead of the change in retail sales. MfgPrd33/29 is the value of products manufactured in 1933 as a fraction of the value of products manufactured in 1929. MfgVa33/29 is value added by manufacture in 1933 as a fraction of value added by manufacture in 1929. MfgPrd35/29 and MfgVa35/29 are constructed similarly but using the relevant manufacturing outcomes from 1935 instead of 1933.

Table B.7 presents cross-sectional correlations between our net elasticity measure (NE20) and the variables in Tables B.2 to B.6. Cross-sectional correlations between NE20 and the census controls in Table B.3 are reported in Table 3 so we do not repeat them here.

Table B.1:
Returns Panel

	a) $I_t^{dec} = 0$			b) $I_t^{dec} = 1$			p-value (H_0: a=b)	
	Obs	Mean	Sd	Obs	Mean	Sd	H_1: a<b	H_1: a>b
Discount Rate								
All	502	0.0423	0.0036	392	0.0426	0.0054	0.1978	0.8022
Cities	231	0.0423	0.0034	165	0.0423	0.0052	0.4904	0.5096
States	271	0.0423	0.0038	227	0.0428	0.0056	0.1479	0.8521
Paid on Deposits								
All	502	0.0214	0.0047	392	0.0210	0.0043	0.9074	0.0926
Cities	231	0.0192	0.0037	165	0.0189	0.0033	0.7739	0.2261
States	271	0.0233	0.0046	227	0.0225	0.0043	0.9729	0.0271
Return on Loans								
All	502	0.0622	0.0094	392	0.0647	0.0101	0.0001	0.9999
Cities	231	0.0586	0.0082	165	0.0598	0.0067	0.0601	0.9399
States	271	0.0653	0.0093	227	0.0683	0.0106	0.0005	0.9995
Securities Return								
All	502	0.0436	0.0074	392	0.0433	0.0076	0.6956	0.3044
Cities	231	0.0440	0.0074	165	0.0448	0.0074	0.1541	0.8459
States	271	0.0432	0.0074	227	0.0423	0.0076	0.9212	0.0788

Notes: $I_t^{dec} = 0$ restricts the sample to periods ending in June; $I_t^{dec} = 1$ restricts the sample to periods ending in December. "All" indicates that summary statistics are calculated using all 82 locations. "Cities" uses only the 33 reserve cities while "States" uses only the 49 states, split states, and district remainders. The last two columns test for seasonality in interest rates. For each rate, the null hypothesis is that its mean is equal in the first and second halves of the year. The alternative hypothesis in column "H_1: a<b" is that the interest rate is lower in the first half of the year. The alternative hypothesis in column "H_1: a>b" is that the interest rate is lower in the second half of the year. A p-value below 0.1 indicates that the null hypothesis is rejected in favor of the alternative at the 10% significance level.

Table B.2:
Retail Sales Ratios

	All			Cities			States		
	Mean	Med.	Sd	Mean	Med.	Sd	Mean	Med.	Sd
Retail33/29	0.515	0.517	0.058	0.523	0.517	0.053	0.509	0.506	0.061
Retail35/29	0.685	0.676	0.075	0.670	0.658	0.070	0.695	0.680	0.078

Notes: Retail33/29 is the ratio of retail sales in 1933 to retail sales in 1929. Retail35/29 uses retail sales in 1935 instead of 1933. Data from the Census of Business.

Table B.3:
Census Controls and HHI

	All			Cities			States		
	Mean	Med.	Sd	Mean	Med.	Sd	Mean	Med.	Sd
LogArea	8.670	9.047	2.447	6.155	6.413	1.105	10.363	10.845	1.419
LogPop	13.309	13.449	1.254	12.437	12.228	1.009	13.897	13.873	1.049
Urban	0.567	0.554	0.286	0.842	0.831	0.117	0.381	0.311	0.204
Nwnp	0.563	0.568	0.154	0.543	0.544	0.159	0.576	0.594	0.151
Age1844	0.434	0.427	0.040	0.473	0.476	0.026	0.408	0.402	0.024
School1620	0.277	0.269	0.062	0.268	0.261	0.064	0.283	0.272	0.061
LogMfgEstPc	-5.986	-5.972	0.375	-5.841	-5.849	0.301	-6.084	-6.059	0.390
MfgWork	0.080	0.065	0.054	0.087	0.083	0.040	0.075	0.047	0.062
LogMfgVa	11.075	11.112	0.575	11.263	11.294	0.411	10.949	10.872	0.637
LogFarmsPc	-3.690	-3.312	1.632	-4.959	-4.320	1.754	-2.835	-2.618	0.769
Acres	0.560	0.604	0.269	0.579	0.684	0.305	0.548	0.583	0.245
LogAvgCrop	7.653	7.717	0.476	7.622	7.773	0.529	7.673	7.677	0.442
LogAvgValue	9.452	9.429	0.652	9.741	9.665	0.538	9.258	9.168	0.654
HomeOwnClr	0.268	0.268	0.083	0.214	0.219	0.058	0.305	0.324	0.077
HHI	0.217	0.203	0.107	0.179	0.167	0.093	0.243	0.250	0.110

Notes: HHI measures local banking market concentration in 1929. All other variables are from the 1920 population census and are as defined in Table 3.

Table B.4:
Banking Distress Indicators

	All			Cities			States		
	Mean	Med.	Sd	Mean	Med.	Sd	Mean	Med.	Sd
SuspNum32	0.096	0.077	0.104	0.057	0	0.115	0.123	0.100	0.087
SuspVal32	0.060	0.030	0.091	0.027	0	0.082	0.081	0.047	0.091

Notes: SuspNum32 and SuspVal32 measure respectively the number and deposits of national banks suspended from 1930 to 1932 relative to the number and deposits of all national banks in 1929.

Table B.5:
Instruments for Banking Distress

	All			Cities			States		
	Mean	Med.	Sd	Mean	Med.	Sd	Mean	Med.	Sd
LogAsset29	11.909	11.859	1.143	11.761	11.669	1.181	12.008	11.968	1.117
LogNum29	3.523	3.882	1.477	2.223	2.079	0.740	4.398	4.489	1.171
RealEst29	0.038	0.038	0.014	0.033	0.033	0.016	0.041	0.041	0.011
Capital29	0.123	0.121	0.029	0.111	0.102	0.027	0.130	0.126	0.028
DueTo29	0.110	0.084	0.082	0.189	0.174	0.069	0.058	0.053	0.034
DueFrom29	0.120	0.124	0.047	0.157	0.162	0.037	0.095	0.100	0.036
DemDep29	0.395	0.396	0.088	0.413	0.439	0.079	0.383	0.377	0.093
Cash29	0.064	0.063	0.010	0.067	0.067	0.012	0.062	0.063	0.008

Notes: All variables as defined in Table C.1 (see Appendix C) and based on data for national banks in 1929.

Table B.6:
Manufacturing Ratios

	All			Cities			States		
	Mean	Med.	Sd	Mean	Med.	Sd	Mean	Med.	Sd
MfgPrd33/29	0.444	0.435	0.085	0.435	0.433	0.056	0.450	0.442	0.100
MfgPrd35/29	0.626	0.611	0.150	0.657	0.623	0.163	0.605	0.604	0.139
MfgVa33/29	0.478	0.471	0.084	0.490	0.484	0.073	0.470	0.456	0.091
MfgVa35/29	0.592	0.588	0.123	0.635	0.598	0.131	0.562	0.570	0.109

Notes: MfgPrd33/29 (MfgPrd35/29) is the value of products manufactured in 1933 (1935) as a fraction of the value of products manufactured in 1929. MfgVa33/29 (MfgVa35/29) is value added by manufacture in 1933 (1935) as a fraction of value added by manufacture in 1929. Data from the Census of Manufactures.

Table B.7:
Cross-Sectional Correlations

	Correlation with NE20		
	All	Cities	States
HHI	0.164	0.160	-0.007
SuspNum32	0.151	-0.081	0.171
SuspVal32	0.206	0.056	0.151
LogAsset29	-0.074	-0.188	-0.067
LogNum29	0.309	-0.082	0.136
RealEst29	0.140	-0.099	0.204
Capital29	-0.115	-0.324	-0.229
DueTo29	-0.180	0.331	-0.011
DueFrom29	-0.116	0.321	0.059
DemDep29	-0.121	-0.042	-0.080
Cash29	-0.059	0.134	-0.091
Retail33/29	-0.271	-0.297	-0.218
Retail35/29	-0.202	-0.288	-0.294
MfgPrd33/29	-0.061	-0.225	-0.054
MfgPrd35/29	-0.107	-0.084	-0.010
MfgVa33/29	-0.155	-0.051	-0.175
MfgVa35/29	-0.274	-0.115	-0.256

Notes: NE20 is the measure of continuing relationships on the eve of the Great Depression. All other variables are as defined in Tables B.2 to B.6. Cross-sectional correlations between NE20 and the census controls in Table B.3 appear in Table 3 and are hence omitted here.

Appendix C – Two-Stage Estimation Results

This appendix provides details on the two-stage estimation results discussed in Section 5.3 of the main text.

First-Stage Regression

Calomiris and Mason (2003) argue that the following variables, all measured in 1929, are valid instruments for loan supply during the Great Depression: average bank size, the fraction of non-cash banking assets in real estate, and the ratio of bank capital to total banking assets. They explain that real estate and capital would have differed across locations in 1929 mainly because of differences in exposure to earlier agricultural crises while average bank size would have differed mainly because of differences in pre-existing regulations that restricted where banks could open and branch. Their identifying assumption is that the shocks of the Great Depression were not just a continuation of the shocks that produced cross-sectional differences in the instruments in 1929.

We follow the same approach, adding interbank linkages in 1929 to the list of instruments. Recent work by Mitchener and Richardson (2019) shows that banks that normally received large amounts of interbank deposits cut lending to firms in the early 1930s when those deposits were withdrawn. Interbank linkages thus amplified the contraction in bank lending through loan supply, providing an additional instrument. A longer list of instruments is useful for us given that we need to instrument both the number and value of national bank suspensions in a sample that includes a mix of states and reserve cities.

We measure interbank linkages using data on deposits received from other banks (due-tos) and deposits held at other banks (due-froms). We add demand deposits and cash holdings of national banks in 1929 as controls to ensure that dues proxy for the channels in Mitchener and Richardson (2019), not for liquid assets and liabilities more generally. We also interact all instruments with the dummy variable for reserve cities to allow for the possibility that the instruments predict loan supply differently in cities versus states.

Table C.1 reports the first-stage regression results in which the instruments are used to predict banking distress. We use a Tobit regression since $SuspNum32_i$ and $SuspVal32_i$ as defined in Section 5.1 are censored from below at zero. In effect, our two-stage procedure uses the latent variables behind $SuspNum32_i$ and $SuspVal32_i$ as the indicators of banking distress in Eq. (4). The instrumented indicators are then linear predictions from the Tobit estimation.[1]

[1] We therefore refrain from comparing the magnitudes of the estimated second-stage coefficients to the OLS estimates in Table 4.

Average bank size helps predict banking outcomes, consistent with the first-stage results of Calomiris and Mason (2003). Holding constant the total value of banking sector assets in Table C.1, distress is increasing in the total number of banks, and, holding constant the total number of banks, distress is decreasing in the total value of banking sector assets, especially outside reserve cities. The capital ratio of banks also enters Table C.1 with the expected negative sign and is a statistically significant predictor of the suspension rate. Due-tos and due-froms are statistically significant as well. More due-tos predict more banking distress in reserve cities but less distress elsewhere; more due-froms predict less banking distress in reserve cities but more distress elsewhere. These are the signs one would expect to find based on the pyramid structure of interbank deposits described in Mitchener and Richardson (2019). Lastly, Table C.1 shows that a higher fraction of banking assets financed by demand deposits (i.e., highly runnable retail liabilities) predicts more banking distress unless cash holdings are also a larger share of assets. These signs are again as expected, along with being statistically significant.[2]

Second-Stage Regressions

Second-stage results are presented in Tables C.2 and C.3. In all specifications, we use suspensions-based measures of distress, as we believe they provide a more broad-based notion of distress than outright failures, and we include district fixed effects, the reserve city dummy, and the variables from the 1920 census as controls.

Column (1) of Table C.2 adds no further controls, as in the first column of Table 4. Column (2) controls for concentration of the local banking market, as in the second column of Table 4. The rest of Table C.2 provides further evidence of the economic content of net elasticity over and above other local characteristics by including interactions between the census variables and banking distress in the control vector X_i. The size of our sample does not permit including all interactions at once so we interact distress (specifically the instrumented indicators of banking distress predicted by the first-stage regressions in Table C.1) with one census variable at a time.

As discussed in Section 4.3, net elasticity tended to be higher in areas that were more rural, as measured by a smaller urban population and/or more farming activity. More rural areas also tended to have smaller banks, as indicated by a very high correlation between the urbanization rate and the log of average bank size. Columns (3) and (4) in Table C.2 run

[2] The F-statistic for the first-stage estimation in Table C.1 is 6.93. This is below the magic number of 10 targeted by some papers, but, in our opinion, not that far below and actually quite sensible given our application and the fact that we have only 82 data points in the cross-section. Thus, while we do not want to overstate the two-stage estimates, we do trust them to be informative.

the two-stage version of Eq. (4) when X_i interacts instrumented banking distress with either Urban or LogFarmsPc, both as defined in Table 3. In column (5), instrumented distress is instead interacted with the log of average bank size in 1920. Columns (6) to (17) repeat the exercise for each of the remaining census variables in Table 3. In all regressions, X_i still includes district fixed effects, the reserve city dummy, and the other controls in Table 3.

In all columns in Table C.2, the coefficient on the interaction between the instrumented $SuspNum32_i$ variable and $NE20_i$ is negative and statistically significant and the magnitude of this coefficient is stable across columns. Our results are therefore robust to controlling for interactions between banking distress and local characteristics from the 1920 census.

Column (1) in Table C.3 extends the specification in the second column of Table C.2 by interacting $NE20_i$ and the instrumented banking distress indicators with the reserve city dummy to explore whether the results differ across the types of locations that comprise our sample. The coefficient on the interaction between the instrumented $SuspNum32_i$ variable and $NE20_i$ remains negative. The unconditional estimate is -1.460 and statistically significant. The estimate for reserve cities is -0.426 (-1.460+1.034) and we cannot reject the null hypothesis that the difference relative to the unconditional estimate is simply driven by noise. Column (2) in Table C.3 repeats the analysis using the change in retail sales from 1929 to 1935 as the dependent variable. We find similar results as in the first column.

Aggregate Implications

The estimated coefficients can be used to make statements about the aggregate implications of banking distress in the early 1930s. Specifically, we can predict what retail sales would have been had there been no bank suspensions (e.g., $SuspNum32_i = SuspVal32_i = 0$ for all locations i) then compare this to the fitted values when all regressors are as observed in the data. We find that retail sales in the U.S. would have been almost 4% higher in 1933 had there been no banking distress. If we set only the $SuspNum32_i$ variable to zero to try to isolate the impact of the initial distress without the mitigating effect from the rapid liquidation of large suspended banks, we find that total retail sales would have been on the order of 10% higher.

Although geographically disaggregated measures of total economic activity (e.g., GNP) are not available for the period we study, we can perform a back-of-the-envelope calculation to translate the total retail sales decline into an aggregate GNP effect. The "multiplier" we use is the coefficient from a simple regression of GNP growth on retail sales growth and a constant. We use annual data aggregated at the national level to run the regression, instrumenting retail sales growth by its one period lag. The coefficient on instrumented retail sales growth constitutes a rough multiplier that we can use for our back-of-the-envelope calculation. The sample period for the regression is 1920 to 1929. 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 second row of Table C.4. Using the GNP estimates from Romer (1989), a 1% increase in retail sales is consistent with a 1.7% increase in nominal GNP and a 0.7% increase in real GNP. Using the GNP estimates from Balke and Gordon (1989), a 1% increase in retail sales is consistent with a 1.6% increase in nominal GNP and a 0.9% increase in real GNP. Overall, then, a 4% drop in retail sales is consistent with a 6-7% drop in nominal GNP and a 3-4% drop in real GNP, which is about one-eighth of the economic contraction experienced during the Great Depression.

Manufacturing Outcomes

We now consider alternative measures of economic activity as the dependent variable in Eq. (4), specifically changes in manufacturing outcomes from the Census of Manufactures instead of changes in retail sales from the Census of Business.

The locations for which we can calculate $NE20_i$ include reserve cities, states, split states, and district remainders, so, to run regressions with the change in manufacturing outcomes as the dependent variable, we need county-level manufacturing data that can be aggregated into the appropriate units i. We consider manufacturing outcomes here since other measures of economic activity, such as unemployment rates, are not available at the county level until 1937. Construction contracts and business failures are also unavailable by county.

The Census of Manufactures is the go-to source for county-level manufacturing data, but it involves substantially more top-coding than the retail sales data in the Census of Business. Top-coding occurs when two or more counties in a state are aggregated into a state residual. Top-coding is common for counties with few manufacturers or one dominant manufacturer that would otherwise be revealed. It is also a function of the precision of the survey. The number of top-coded counties is notably high in the 1933 Census of Manufactures. The issue also exists in 1929 and 1935 but to a lesser extent. Our measures of manufacturing activity will therefore be noisier than our measure of retail sales, particularly in 1933. We present the following results with this caveat in mind.

We run two versions of Eq. (4) using a two-stage procedure as above. In the first version, the dependent variable is the value of products manufactured in 1933 as a fraction of the value of products manufactured in 1929. In the second version, the dependent variable is value added by manufacture in 1933 as a fraction of value added by manufacture in 1929. Value added captures the value of products net of the cost of materials, fuel, purchased energy, and contract work used in production. Both versions are also run using data from 1935 instead of 1933. The banking distress indicators are instrumented as described earlier and the second-stage regression controls for district fixed effects, the reserve city dummy, the variables from the 1920 census, and HHI of the local banking market.

The results are reported in Table C.5. They echo the results obtained using retail sales. In all columns of Table C.5, the coefficient on the interaction between the instrumented $SuspNum32_i$ variable and $NE20_i$ is negative, indicating that the marginal impact of banking distress on manufacturing activity was more severe in areas with more continuing relationships. This interaction is statistically significant for 1935 but not 1933, which is not surprising given the noise introduced into the manufacturing data by the top-coding issues discussed above.

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.

Calomiris, C. and J. Mason. 2003. "Consequences of Bank Distress During the Great Depression." *American Economic Review*, 93(3), pp. 937-947.

Mitchener, K. and G. Richardson. 2019. "Network Contagion and Interbank Amplification during the Great Depression." *Journal of Political Economy*, 127(2) pp. 465-507.

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.

Table C.1:
First Stage Tobit Regressions for Two-Stage Procedure

VARIABLES	(1) SuspNum32	(2) SuspVal32
LogAsset29	-0.089***	-0.094**
	(0.031)	(0.036)
LogNum29	0.072**	0.079**
	(0.031)	(0.036)
RealEst29	2.219	2.035
	(1.527)	(2.336)
Capital29	-1.234*	-1.049
	(0.694)	(0.791)
DueTo29	-1.474***	-1.028*
	(0.470)	(0.608)
DueFrom29	2.580***	1.944*
	(0.752)	(1.057)
DemDep29	1.383*	1.078
	(0.706)	(0.908)
Cash29	12.300***	1.179
	(4.519)	(5.760)
DemDep29 x Cash29	-29.364***	-16.509
	(10.034)	(12.688)
City	-0.807	-1.102
	(1.044)	(1.162)
LogAsset29 x City	0.100*	-0.002
	(0.055)	(0.058)
LogNum29 x City	0.049	0.096
	(0.055)	(0.062)
RealEst29 x City	-3.581	-4.760
	(2.409)	(3.096)
Capital29 x City	-0.335	-0.115
	(1.194)	(1.026)
DueTo29 x City	2.093***	2.479***
	(0.563)	(0.747)
DueFrom29 x City	-3.161***	-3.974***
	(1.147)	(1.448)
DemDep29 x City	-0.429	3.967**
	(1.773)	(1.907)
Cash29 x City	-1.533	18.748
	(10.282)	(12.674)
DemDep29 x Cash29 x City	3.783	-57.817*
	(25.217)	(30.498)
Observations	82	82
District FE	YES	YES

Notes: All regressors from national bank data in 1929. LogAsset29 is logged banking assets; LogNum29 is logged number of banks; RealEst29 is fraction of non-cash assets in real estate; Capital29 is capital-to-asset ratio; DueTo29 is deposits received from other banks as a share of assets; DueFrom29 is deposits held at other banks as a share of assets; DemDep29 is ratio of demand deposits to assets; Cash29 is cash-to-asset ratio. City is a dummy for reserve cities. Robust standard errors in brackets. ***p<0.01, **p<0.05, *p<0.1

Table C.2:
Two-Stage Estimates, Retail Sales Ratio, 1933/29

VARIABLES	(1) Baseline	(2) Control: HHI	(3) Control: Urban	(4) Control: LogFarmsPc	(5) Control: LogBkSize
NE20	0.055	0.051	0.047	0.058	0.055
	(0.036)	(0.040)	(0.040)	(0.042)	(0.041)
SuspNum32	-0.027	-0.065	-0.248	0.095	-0.200
	(0.123)	(0.251)	(0.496)	(0.486)	(1.113)
SuspNum32 x NE20	-0.909**	-0.838**	-0.899**	-1.001**	-1.008**
	(0.393)	(0.419)	(0.413)	(0.459)	(0.455)
SuspVal32	-0.115	-0.304	0.389	-0.083	0.650
	(0.114)	(0.246)	(0.646)	(0.517)	(1.400)
SuspVal32 x NE20	0.802**	0.709*	0.747*	0.877*	0.882**
	(0.395)	(0.419)	(0.423)	(0.455)	(0.447)
Control		0.076	0.039	0.065	-0.001
		(0.127)	(0.135)	(0.041)	(0.025)
SuspNum32 x Control		-0.256	0.213	0.030	0.013
		(1.192)	(0.585)	(0.108)	(0.118)
SuspVal32 x Control		1.464	-0.585	0.007	-0.080
		(1.323)	(0.753)	(0.110)	(0.147)
Observations	82	82	82	82	82
Adjusted R-squared	0.518	0.549	0.523	0.508	0.508
District FE	YES	YES	YES	YES	YES
Census Controls	YES	YES	YES	YES	YES

Notes: The dependent variable is retail sales in 1933 as a fraction of retail sales in 1929. NE20 is the measure of continuing relationships on the eve of the Great Depression. The indicators of banking distress, SuspNum32 and SuspVal32, are defined using national bank suspensions and instrumented as per Table C.1. The heading of each column specifies the control variable interacted with instrumented distress: in column (2), HHI measures local banking market concentration in 1929; in column (5), LogBkSize is the log of average bank size in 1920; variables in all other headings are as defined in Table 3. All columns include the remaining variables in Table 3 as controls. Bootstrapped standard errors clustered at the district level are in brackets. ***p<0.01, **p<0.05, *p<0.1

Table C.2 (Continued):
Two-Stage Estimates, Retail Sales Ratio, 1933/29

VARIABLES	(6) Control: LogArea	(7) Control: LogPop	(8) Control: Nwnp	(9) Control: Age1844	(10) Control: School1620	(11) Control: LogMfgEstPc
NE20	0.054	0.065*	0.066	0.055	0.060	0.060
	(0.038)	(0.039)	(0.042)	(0.041)	(0.039)	(0.042)
SuspNum32	0.217	1.112	-0.217	-0.823	-0.225	-1.051
	(0.472)	(1.545)	(0.428)	(1.610)	(0.677)	(2.890)
SuspNum32 x NE20	-0.985**	-1.037**	-1.047**	-0.932**	-0.965**	-1.014**
	(0.417)	(0.441)	(0.462)	(0.445)	(0.422)	(0.455)
SuspVal32	-0.579	-1.103	-0.146	0.936	-0.125	0.377
	(0.517)	(1.705)	(0.451)	(2.002)	(0.676)	(3.267)
SuspVal32 x NE20	0.868**	0.943**	0.915**	0.814*	0.815*	0.906**
	(0.415)	(0.447)	(0.454)	(0.450)	(0.422)	(0.457)
Control	-0.041	0.047**	-0.235**	0.894	0.079	-0.047
	(0.028)	(0.024)	(0.105)	(0.899)	(0.276)	(0.046)
SuspNum32 x Control	-0.039	-0.086	0.362	1.654	0.689	-0.176
	(0.064)	(0.116)	(0.757)	(3.451)	(2.263)	(0.498)
SuspVal32 x Control	0.074	0.074	0.000	-2.200	0.009	0.087
	(0.076)	(0.131)	(0.795)	(4.203)	(2.230)	(0.564)
Observations	82	82	82	82	82	82
Adjusted R-squared	0.527	0.510	0.515	0.506	0.509	0.503
District FE	YES	YES	YES	YES	YES	YES
Census Controls	YES	YES	YES	YES	YES	YES

Notes: See previous.

Table C.2 (Continued):
Two-Stage Estimates, Retail Sales Ratio, 1933/29

VARIABLES	(12) Control: MfgWork	(13) Control: LogMfgVa	(14) Control: Acres	(15) Control: LogAvgCrop	(16) Control: LogAvgVal	(17) Control: H.Own
NE20	0.058	0.044	0.056	0.054	0.058	0.055
	(0.043)	(0.039)	(0.043)	(0.037)	(0.038)	(0.040)
SuspNum32	0.066	0.186	-0.052	1.226	1.014	-0.038
	(0.328)	(2.878)	(0.334)	(2.448)	(2.666)	(0.438)
SuspNum32 x NE20	-1.011**	-0.854**	-0.914*	-0.918**	-0.930**	-0.936**
	(0.456)	(0.423)	(0.523)	(0.414)	(0.422)	(0.461)
SuspVal32	-0.111	1.328	-0.092	-2.257	-1.802	-0.182
	(0.317)	(2.780)	(0.325)	(2.321)	(2.948)	(0.460)
SuspVal32 x NE20	0.911**	0.777*	0.805	0.811**	0.828*	0.823*
	(0.457)	(0.423)	(0.499)	(0.404)	(0.426)	(0.462)
Control	0.444	-0.020	-0.169**	0.025	0.011	0.419**
	(0.592)	(0.034)	(0.080)	(0.032)	(0.031)	(0.211)
SuspNum32 x Control	-1.051	-0.020	0.050	-0.158	-0.104	-0.007
	(3.284)	(0.255)	(0.554)	(0.316)	(0.273)	(1.805)
SuspVal32 x Control	-0.194	-0.128	-0.047	0.273	0.170	0.304
	(3.084)	(0.246)	(0.469)	(0.299)	(0.299)	(2.088)
Observations	82	82	82	82	82	82
Adjusted R-squared	0.514	0.528	0.498	0.520	0.507	0.499
District FE	YES	YES	YES	YES	YES	YES
Census Controls	YES	YES	YES	YES	YES	YES

Notes: See previous. LogAvgVal abbreviates LogAvgValue; H.Own abbreviates HomeOwnClr.

Table C.3:
Two-Stage Estimates, Retail Sales Ratio, Extended

VARIABLES	(1) 1933/29	(2) 1935/29
NE20	0.102	0.121
	(0.064)	(0.093)
SuspNum32	-0.336	0.016
	(0.511)	(0.824)
SuspNum32 x NE20	-1.460**	-1.939*
	(0.737)	(1.145)
SuspVal32	0.183	-0.061
	(0.629)	(0.931)
SuspVal32 x NE20	1.424*	1.931
	(0.814)	(1.199)
HHI	0.089	0.275
	(0.130)	(0.221)
SuspNum32 x HHI	-0.013	-0.059
	(1.410)	(2.458)
SuspVal32 x HHI	1.023	0.782
	(1.619)	(2.550)
NE20 x City	-0.123	-0.136
	(0.085)	(0.133)
SuspNum32 x City	0.234	-0.066
	(0.382)	(0.615)
SuspNum32 x NE20 x City	1.034	1.314
	(1.087)	(1.750)
SuspVal32 x City	-0.442	-0.165
	(0.473)	(0.711)
SuspVal32 x NE20 x City	-1.328	-1.523
	(1.096)	(1.752)
Observations	82	82
Adjusted R-squared	0.637	0.487
District FE	YES	YES
Census Controls	YES	YES

Notes: In column (1), the dependent variable is retail sales in 1933 as a fraction of retail sales in 1929. Column (2) uses retail sales in 1935 instead of 1933. NE20 is the measure of continuing relationships on the eve of the Great Depression. The indicators of banking distress, SuspNum32 and SuspVal32, are defined using national bank suspensions and instrumented as per Table C.1. HHI measures local banking market concentration in 1929. City is a dummy for reserve cities. Bootstrapped standard errors clustered at the district level are in brackets. ***p<0.01, **p<0.05, *p<0.1

Table C.4:
GNP Multipliers for Retail Sales

	(1) Nominal Romer	(2) Nominal Balke-Gordon	(3) Real Romer	(4) Real Balke-Gordon
Constant	-0.010	-0.006	0.028***	0.025***
	(0.010)	(0.012)	(0.005)	(0.006)
%ΔRetail	1.690***	1.565***	0.696***	0.896***
	(0.242)	(0.299)	(0.163)	(0.242)

Notes: The dependent variable is nominal GNP growth in columns (1)-(2) and real GNP growth in columns (3)-(4). GNP estimates from Romer (1989) are used in columns (1) and (3). GNP estimates from Balke and Gordon (1989) are used in columns (2) and (4). %ΔRetail is retail sales growth instrumented using its one period lag. Robust standard errors are in brackets. ***p<0.01, **p<0.05, *p<0.1

Table C.5:
Two-Stage Estimates, Manufacturing Outcomes

VARIABLES	(1) MfgPrd 1933/29	(2) MfgPrd 1935/29	(3) MfgVa 1933/29	(4) MfgVa 1935/29
NE20	0.043	0.141	-0.016	0.069
	(0.084)	(0.129)	(0.073)	(0.100)
SuspNum32	0.157	0.026	0.110	-0.279
	(0.478)	(0.894)	(0.474)	(0.647)
SuspNum32 x NE20	-1.176	-2.808**	-0.596	-2.090*
	(0.853)	(1.410)	(0.833)	(1.097)
SuspVal32	-0.583	-0.643	-0.476	-0.241
	(0.492)	(0.828)	(0.488)	(0.612)
SuspVal32 x NE20	0.939	1.936	0.267	1.395
	(0.878)	(1.416)	(0.831)	(1.106)
HHI	0.208	0.120	-0.013	-0.233
	(0.222)	(0.356)	(0.225)	(0.296)
SuspNum32 x HHI	-2.296	-2.381	-1.607	0.331
	(2.389)	(3.732)	(2.282)	(2.820)
SuspVal32 x HHI	3.404	4.164	2.825	0.833
	(2.708)	(4.066)	(2.714)	(3.041)
Observations	82	82	82	82
Adjusted R-squared	0.025	0.135	0.121	0.260
District FE	YES	YES	YES	YES
Census Controls	YES	YES	YES	YES

Notes: In column (1), the dependent variable is the value of products manufactured in 1933 as a fraction of the value of products manufactured in 1929. Column (2) uses the value of products manufactured in 1935 instead of 1933. In column (3), the dependent variable is value added by manufacture in 1933 as a fraction of value added by manufacture in 1929. Column (4) uses value added in 1935 instead of 1933. NE20 is the measure of continuing relationships on the eve of the Great Depression. The indicators of banking distress, SuspNum32 and SuspVal32, are defined using national bank suspensions and instrumented as per Table C.1. HHI measures local banking market concentration in 1929. Bootstrapped standard errors clustered at the district level are in brackets. ***$p<0.01$, **$p<0.05$, *$p<0.1$