NBER WORKING PAPER SERIES

# RELATIONSHIP LENDING AND THE GREAT DEPRESSION: NEW MEASUREMENT AND IMPLICATIONS

Jon Cohen
Kinda Cheryl Hachem
Gary Richardson

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

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2016, Revised September 2017

We thank Joseph Johnson, Padma Sharma, and Daniel Tracht for excellent research assistance. We also thank Ben Bernanke, Gabe Chodorow-Reich, Jonathan Dingel, Joe Haubrich, Erik Hurst, Fabrizio Perri, and Johannes Wieland as well as seminar and conference participants at Chicago Booth, UW-Madison, FRB Cleveland, and NBER Summer Institute 2016 (Monetary Economics) for helpful comments. Kinda Hachem gratefully acknowledges financial support from Chicago Booth. Part of this research was conducted while Gary Richardson was at the Federal Reserve Bank of Richmond. The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Richmond, the Federal Reserve System, or the National Bureau of Economic Research.

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 They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

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

Relationship Lending and the Great Depression: New Measurement and Implications
Jon Cohen, Kinda Cheryl Hachem, and Gary Richardson
NBER Working Paper No. 22891
December 2016, Revised September 2017
JEL No. E44,G01,G21,L14,N22

**ABSTRACT**

We demonstrate empirically that the collapse of long-term lending relationships amplified the severity of the Great Depression and affected the pace of recovery. We overcome the lack of loan-level data for this period by developing a new measure of continuing relationships that can be calculated at any level of aggregation. Our approach 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. After establishing the validity of this approach for a period when microeconomic data on lending relationships do exist, we show, first, that the marginal impact of bank suspensions on economic activity in the early 1930s was significantly higher in areas with more continuing relationships and, second, that the rebuilding of lending relationships destroyed by suspensions helps explain economic performance during the 1937-38 recession.

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

Kinda Cheryl Hachem
University of Chicago
Booth School of Business
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
kinda.hachem@chicagobooth.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, the longest and most severe business cycle in U.S. history, also involved the devastation of a sizeable fraction of the financial sector. Since the pioneering work of Bernanke (1983), a large number of economists and economic historians, with varying degrees of success, have tried to estimate the size and significance of disruptions to the credit channel during the banking panics of the period. 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. This assumption is central to all of these analyses because, in its absence, it is difficult to rationalize why bank suspensions would lead to such a deep and protracted fall in economic activity.

In this paper, we argue that a critical input impaired by bank suspensions in the early 1930s was the soft information that banks had acquired about the quality of their borrowers during multi-period lending relationships. Specifically, we show that disruptions to continuing relationships as a result of these suspensions were a means by which financial shocks were transmitted to the real economy, contributing not only to the severity of the downturn but also to the uneven pace of recovery later in the decade. These results advance the large literature on the Great Depression, provide to the best of our knowledge the first and most direct test of Bernanke's intuition that not all loans were created equal, and are of interest beyond the historical context. Relationship lending continues to be 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 borrowers. However, the soft information on which these relationships depend is intrinsically difficult to transfer from one bank to another. Therefore, 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 only on the basis of easy-to-evaluate collateral. In this respect, then, our findings serve as a graphic illustration as well as a cautionary tale for policymakers of the damage that a financial crisis, through its impact on long-term lending relationships, can inflict on an economy.

Given the dimensions of the downturn and the largely unfettered nature of the banking panics, the Great Depression remains ground zero for studying the non-monetary effects of financial crises. Despite this, efforts to grasp fully the real effects of financial distress on economic activity in the 1930s have been severely constrained by a lack of the type of microeconomic data on bank-firm interactions that informs modern analyses of relationship lending. We overcome this problem by developing a new measure of continuing lending relationships that can be calculated at any level of aggregation. Our measure builds 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 present a model based on Hachem (2011) to provide theoretical foundations. We also validate the idea empirically using transaction-level data from Dealscan, a widely employed database in modern analyses of relationship lending. We find strong evidence that loan rates become less responsive to bank funding costs as relationship length increases. After vetting our measure on several dimensions, we proceed to study the Great Depression.

We use data from the consolidated balance sheets and income statements of national banks in the 1920s to implement our measure for the eve of the Great Depression. These data, reported by the Comptroller of the Currency at a semi-annual frequency for Federal Reserve districts, states, and reserve cities, allow us to infer the weighted average loan rate at each point in time in each of 82 locations. We then calculate the elasticity of the loan rate with respect to the discount rate for each location. Since the twelve Federal Reserve Banks at the time had sufficient latitude to operate largely independent discount windows, there was variation in discount rates across districts (Richardson and Troost (2009)). 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 net it out from the loan rate elasticity. We take cross-sectional differences in net elasticity in the 1920s as indicative of differences in the nature and intensity of relationship lending across locations on the eve of the Great Depression, such that areas with relatively less elastic loan rates were those where continuing relationships prevailed.

Our first set of results based on these data establishes the importance of relationship lending for understanding the real effects of banking distress in the early 1930s. We estimate that bank suspensions alone explain one-eighth of the national economic contraction observed during the Great Depression. Equally important, we are able to show that the marginal impact of bank suspensions on economic activity was much greater in areas with more continuing relationships, other things the same. In other words, the destruction of continuing relationships amplified the severity of the Great Depression. Although Calomiris and Mason (2003), Richardson and Troost (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.[1]

[1] In contrast to the mainstream view, Cole and Ohanian (1999, 2004, 2007) argue that the collapse of the U.S. banking system in the early 1930s had few real economic effects, citing weak (unconditional) correlation between bank failures and economic outcomes, even at the state level. On this point, our results suggest

Our second set of results focuses on the recovery period. When business is dependent on the soft information that banks acquire over time, it is not enough for bank suspensions to cease and new banks to open. These banks also need time to interact with firms, learn about them, and rebuild continuing relationships. With this in mind, we show that cross-sectional differences in the rebuilding of lending relationships destroyed by the bank suspensions of the early 1930s contributed to cross-sectional differences in economic performance during the 1937-38 recession. Specifically, areas that were rebuilding the types of continuing relationships which existed in the 1920s fared better during the 1937-38 recession than otherwise similar areas that failed to rebuild. In contrast, areas that were rebuilding weaker relationships fared worse. The effects of the Great Depression thus appear to have persisted long after the end of the crisis phase and, because of the time required to rebuild destroyed relationships, the banking panics of the early 1930s are also likely to have contributed to what some believe was a surprisingly muted recovery in the wake of such a catastrophic collapse.[2]

The rest of the paper proceeds as follows. Section 2 provides institutional background on the lending activities of Depression-era commercial banks. Section 3 reviews the theory that leads to our new measure of continuing lending relationships and presents evidence from Dealscan. Section 4 describes the empirical implementation of this measure for the 1920s. Sections 5 and 6 then present the key empirical results, with Section 5 using cross-sectional differences in continuing relationships on the eve of the Great Depression to pinpoint the real effects of banking distress in the early 1930s and Section 6 demonstrating how relationship rebuilding in the aftermath of the Depression affected economic performance during the 1937-38 recession. Section 7 concludes.

## 2 Institutional Background

In this section, we review two institutional features of the 1920s and 1930s that motivate our selection of a theoretical model in the next section. First, bank credit was typically extended on a short-term basis and, second, lending relationships were typically uncollateralized.

The annual report of the Comptroller of the Currency provides information on bank balance sheets in the 1920s and 1930s in a format that is consistent across geographic locations. Up until 1928, the Comptroller divided bank loans into four categories: (i) real estate loans, (ii) loans on financial securities, (iii) uncollateralized loans, and (iv) loans collateralized by

that the correlation becomes much more compelling when conditioned on cross-sectional differences in the nature and intensity of lending relationships.

[2]Rajan and Ramcharan (2016) also observe lingering effects of the Depression, although, unlike us, they focus on banking concentration post-WWII, not on economic activity in the late 1930s and the role of relationship lending.

personal security. Based on the Comptroller's description of personal security, the last category consists primarily of loans secured by difficult-to-evaluate collateral such as goods in the process of production and distribution, warehouse receipts, and, in the case of farm loans, future crops.

Loans in all four categories were extended on a short-term basis. Prior to the 1930s, real estate loans had short maturities, commonly 3 to 5 years. Loans on financial securities consisted largely of call loans to brokers and also some call loans directly to firms and individuals, often to facilitate the purchase of stocks, bonds, or other well understood financial securities. All such loans are, by definition, short-term. Loans in the last two categories (i.e., uncollateralized loans and loans collateralized by personal security) were critical sources of working capital for business, financing goods in the process of production, processing, shipment, and sale. Accordingly, they matured in less than a year, typically less than 180 days. To be eligible for use as collateral at a Federal Reserve Bank's discount window, commercial loans had to mature in 90 days or less, unless collateralized by agricultural commodities, in which case they had to mature in 180 days or less.

Of the four categories of bank loans, observers regarded real estate loans and loans on financial securities as transactional (e.g., Currie (1931)), leaving uncollateralized loans and loans collateralized by personal security as the prime venue for relationship building (e.g., Miller (1927), Foulke and Prochnow (1939)). While refinancing was obviously necessary for most real estate loans because of their short maturities, mortgages were often refinanced with different lenders. Uncollateralized loans and loans collateralized by personal security, on the other hand, were usually sequenced by the same lender, an approach that allowed banks to incorporate information acquired during the initial loan period into future credit terms, including, of course, the possibility of discontinuing the lending relationship (e.g., Langston (1921)). On the eve of the Great Depression, almost 50% of national bank loans were uncollateralized, with another 10% collateralized by personal security.

## 3 New Indicator of Continuing Relationships

Within the set of relationship loans, it is important to distinguish between those that are made as part of a continuing relationship and first-time relationship loans. Soft information on borrower quality is accumulated over time through repeated interactions and is hence embodied in continuing relationships, not new ones. Bank failures, such as those that ravaged the U.S. financial system in the early 1930s, destroy this knowledge, disrupt the flow of funds to borrowers and, because it takes time to re-establish these relationships, also impact the pace of recovery. Accordingly, in order to fully capture the effects of financial crises on the real economy, we need to be able to isolate continuing relationships.

By definition, a continuing relationship involves prior interactions between the same lender and borrower pair. This is a widely used measure of relationship strength in modern analyses of relationship lending (e.g., Berger and Udell (1995), Ongena and Smith (2001), Elsas (2005), Chodorow-Reich (2014), Gobbi and Sette (2015)) but it requires transaction-level data that are not available for a representative sample of banks and firms in the 1920s and 1930s. We must, therefore, find another way to detect the presence of continuing lending relationships.[3]

To do this, we propose a procedure that exploits changes in loan pricing over the course of a relationship. Section 3.1 shows theoretically that the returns on continuing relationship loans are less responsive to changes in bank funding costs than are the returns on new relationship loans or the returns on transactional loans. Section 3.2 then shows, using transaction-level data from Dealscan, that the theoretical predictions of the model hold for the modern period. We think of this as a compelling proof of principle. While loan returns were not directly reported by the Comptroller in the 1920s and 1930s, geographically aggregated bank income statements are available and can be used, as explained in Section 4, to estimate the elasticity of loan rates in various locations. The theoretical model and the Dealscan evidence presented here validate our use of cross-sectional differences in these elasticities to infer cross-sectional differences in continuing relationships, subject to the appropriate controls.

## 3.1 Theory

In this section, we sketch a simple model of relationship lending based on Hachem (2011) to fix ideas. There are two stages and three periods. All agents are risk neutral. There is a continuum of firm types, denoted by $\omega$ and distributed uniformly over the unit interval. Types are private information. In each stage, firm $\omega$ can undertake a production project that generates output $\theta_1$ with probability $p(\omega)$, where $p'(\omega) > 0$. For simplicity, we can consider $p(\omega)$ linear. The project fails with probability $1 - p(\omega)$, in which case zero output is generated. 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 $r$. Firms are not endowed with capital, nor can they store capital or output across stages. The credit

[3] In other contexts where data are a constraint, the use of historical narratives has proved fruitful. See Romer and Romer (1989, 2004, 2015). While it may be possible to gain some insight into lending practices through the narrative approach, we lack a source which provides narratives of local conditions in the Depression-era U.S. that is consistent across both time and space.

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 $R_1^*$. The interest rate decision is subject to a risk-shifting problem. In particular, instead of putting capital into production, the firm can invest in a speculative project which produces output $\theta_2 > \theta_1$ with probability $q < p(0)$ and nothing otherwise. Assume $q\theta_2 = p(0)\,\theta_1$ to reduce notation. Banks can detect the presence of output so firms with positive output repay their contracts. However, banks cannot detect the exact value of output so credit contracts cannot be made contingent on realized output.

The loan rate that makes firm $\omega$ indifferent between the two projects is:

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

It is straightforward to show $\overline{R}'(\omega) > 0$. Firm $\omega$ undertakes the production project in the first stage if and only if the interest rate in the first stage does not exceed $\overline{R}(\omega)$. All banks know that this is the best response of a type $\omega$ firm. Firms select banks randomly in the first stage then decide which projects to undertake. 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 $s$. Separated firms become new firms, drawing new types from the uniform distribution and playing the first-stage game again. Firms that are not separated continue to the second stage and have their types discovered by their first-stage bank ("insider"). The insider learns this information by virtue of having interacted with the firm during the first stage. Other banks ("outsiders") can only observe that this firm is not a new firm.[4] Each insider then decides whether to continue the lending relationship and extend another unit of capital to the firm. If the insider wants to keep the firm, the interest rate cannot exceed what outsiders would charge, otherwise the firm will move to one of these outsiders. All banks have equilibrium beliefs so outsiders know that they are being adversely selected.

Interest rate choices in the second stage are still subject to the risk-shifting problem, meaning that firm $\omega$ undertakes the production project in the second stage if and only if the interest rate it is charged in the second stage does not exceed $\overline{R}(\omega)$. Outsiders are perfectly competitive and offer the same zero-profit interest rate to firms that are endogenously separated from their insiders. This rate differs from the zero-profit interest rate $R_1^*$ offered

[4] Hachem (2011) shows that allowing outsiders to also observe whether or not the firm repaid its insider in the first stage does not, in equilibrium, result in outsiders conditioning on credit history. Accordingly, we do not introduce notation for credit history here.

to new (i.e., exogenously separated) firms because banks understand the adverse selection problem. Insiders have an informational advantage over outsiders so the interest rate offered by an insider in the second stage needs not generate zero profits for the insider. Instead, the insider seeks to maximize his profit, subject to the risk-shifting problem and the ability of the firm to move to an outsider.

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

Figure 1 illustrates the equilibrium division of firm types between insiders and outsiders in the second stage, along with the interest rates they are charged.[5] The equilibrium for a given policy rate $r$ can be described with reference to two cutoff types. The first, $\widetilde{\omega}$, is implicitly defined by $p(\widetilde{\omega})\overline{R}(\widetilde{\omega}) \equiv r$ and represents the type on which the insider expects to break even by charging the reservation rate $\overline{R}(\cdot)$. The second, $\widehat{\omega}$, is implicitly defined by $q\overline{R}(\widehat{\omega}) \equiv r$ and represents the lowest type that will choose the production project if charged $\frac{r}{q}$. Types below $\widetilde{\omega}$ move to outsiders and are charged a pooled interest rate $\frac{r}{q}$, prompting them to undertake the speculative project since $\widetilde{\omega} < \widehat{\omega}$. Types above $\widetilde{\omega}$ stay with their insiders and are charged interest rates that lead them to choose the production project. Notice that types between $\widetilde{\omega}$ and $\widehat{\omega}$ are getting policy-invariant interest rates (i.e., $\overline{R}(\cdot)$ does not depend on $r$). If these types were instead charged $\frac{r}{q}$, they would undertake the speculative project. The relationship lender is therefore using his information to mitigate the risk-shifting problem, incentivizing higher repayment rates by not passing through increases in $r$ to some of the borrowers he retains.

For a given $r$, we can integrate over the type space in Figure 1 to get the average interest rate charged in the second stage:

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

To get the first-stage interest rate $R_1^*$, define $R_1^* \equiv \overline{R}(\xi)$ so that $\xi$ denotes the firm type that is exactly indifferent between the production project and the speculative project when charged $R_1^*$. The zero-profit condition for lenders with new firms is then:

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

[5] See Hachem (2011) for a detailed proof.

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 $\widetilde{\omega}$.

The model so far has assumed that all banks have the potential to become relationship lenders in the second stage. As an alternative, we can imagine that fraction $\varphi$ 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 firm and the bank exogenously separate with probability 1 at the end of the period. A transactional loan is therefore similar to a first-stage loan, except that the transactional lender has a continuation value of zero from the firm. Formally, we can write the interest rate on transactional loans as $R^*_{TL} \equiv \overline{R}(\eta)$, where the zero-profit condition for the lender implies:

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

The following proposition 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* $e_2 < \min\{e_1, e_{TL}\}$*.*

***Proof.*** *See Appendix A.* ■

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

$$\mathcal{R} = \varphi\left[sR^*_1 + (1 - s)\, R^*_2\right] + (1 - \varphi)\, R^*_{TL}$$

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 write:

$$\frac{d\mathcal{R}}{dr}\frac{r}{\mathcal{R}} \approx \frac{\varphi s R^*_1}{\mathcal{R}} e_1 + \frac{\varphi(1 - s)\, R^*_2}{\mathcal{R}} e_2 + \frac{(1 - \varphi)\, R^*_{TL}}{\mathcal{R}} e_{TL}$$

Here, $\frac{\varphi s R^*_1}{\mathcal{R}}$ is the fraction of interest income in the economy 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, the expression for $\frac{d\mathcal{R}}{dr}\frac{r}{\mathcal{R}}$ implies that the weighted average interest 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 loans. For our purposes, this means that areas with less elastic loan rates are areas where continuing relationships are more substantial. It is this crucial insight that we exploit in Section 4.

## 3.2 Evidence from Dealscan

We demonstrated in Section 3.1 that the incentive compatible contract offered by a continuing relationship lender reduces the responsiveness of the weighted average loan rate to the policy rate. Interestingly, this insight garners support from an informal survey of country banking practices conducted by Ford (1928) in Northern Texas. While Ford does not discuss the mechanisms that lead to changes in loan rate elasticity over the course of a relationship, he does observe that banks often maintained a constant loan rate for borrowers they wanted to keep. In this section, we use Dealscan, a widely cited source of information on lending relationships for the modern period, to provide more formal evidence on the link between relationship lending and loan rate responsiveness predicted by our model.

Dealscan contains detailed data on contracts in the syndicated loan market since 1985. A syndicated deal typically has one lead arranger (a lead bank) that may also recruit other banks to participate in the financing. For each deal, Dealscan reports the name of the borrower, the names of the lenders, and information on all of the loan contracts that roll up into the deal (i.e., loan amount, interest rate, maturity, whether or not collateral is pledged, purpose of the loan, etc.). We focus on deals that have a single lead arranger and, as a primary purpose, either the provision of "working capital" or for "corporate purposes." We exclude less general types of loans such as those extended for leveraged buyouts, stock buybacks, etc.

For each deal, we calculate the fraction of funding that was extended through fixed as opposed to variable rate loans. Variable rate loans are typically indexed to a prime rate such as LIBOR, which is the average interest rate that banks report they would have to pay to borrow from each other. As a result, the interest expense of the firm changes with the lead arranger's cost of funds over the life of a variable rate loan. The same is not true for a fixed rate loan, which means that we can use the fraction of each deal in fixed rate loans as a measure of how (un)responsive the deal is to the lead arranger's cost of funds.

Table 1 runs deal-level regressions. The dependent variable is the fraction of the deal extended through fixed rate loans, as explained above. We regress this on the number of prior deals between the firm and the lead arranger, which is one of the most commonly used indicators of relationship lending. Prior deals that occurred concurrently are counted as one deal, and the number of prior deals is tabulated before restricting the sample to deals for working capital or corporate purposes. Table 1 includes as controls the weighted average maturity of the deal (i.e., an average over the maturities of all loans in the deal, with weights based on the fraction of the deal in that loan) and a weighted average indicator of whether or not the deal was collateralized. All specifications include dummies for the firm's industry. We also run specifications that control for firm size, using the concordance of Chava and Roberts (2008) to link Dealscan and Compustat. The first panel in Table 1 is estimated using ordinary least squares while the second panel is based on a fractional logit.

We find that the fraction of the deal in fixed rate loans increases with the number of prior interactions between the bank and the firm. This result is robust across all specifications and exhibits the pattern predicted by the theory in Section 3.1. We also find evidence that the positive relationship between the number of prior interactions and the fraction of the deal in fixed rate loans comes specifically from uncollateralized deals.

Table 2 presents two alternative specifications. The first alternative runs loan-level regressions with standard errors clustered at the deal level. The dependent variable and the number of prior deals are still defined at the deal level, but the controls for maturity and collateral are now at the level of the individual loans. The second alternative restricts the sample to deals with only one loan. The dependent variable is then a binary variable equal to one if and only if the loan is fixed rate. Both alternatives deliver messages consistent with Table 1, namely that fixed rate loans are more characteristic of longer relationships than shorter ones, and that this result comes specifically from uncollateralized deals.

# 4 Empirical Implementation for the 1920s

The Dealscan results provide validation of the idea that the interest rates on loans in continuing relationships are less responsive to changes in bank funding costs than are the interest rates on other loans. In the rest of the paper, we exploit this price-based feature of continuing relationships to study the Great Depression, a period where, up to this point, the lack of loan-level data has hidden from view the impact of bank failures on relationship lending.

## 4.1 Methodology

Recall from the theoretical discussion in Section 3 that the elasticity of the average loan rate with respect to bank funding costs will be lower in an area where continuing relationships are more substantial. Here, we explain how we translate this prediction into a price-based indicator of continuing relationships using the data available for the 1920s and 1930s. Data sources will be described in more detail in Section 4.2. For now, it is sufficient to note that we have available for this period the balance sheets and income statements of U.S. commercial banks aggregated by geographic region.

It is necessary, first, to determine bank funding costs. From the foundation of the Fed through the 1930s, the discount window of each Federal Reserve Bank was operated in a manner that made the discount rate the marginal cost of funding a commercial loan in each district. The Fed expected member banks to discount commercial loans to accommodate seasonal and cyclical peaks in demands for credit, one of the principle motivations for the creation of the Federal Reserve System. The Fed also needed banks to discount loans because they were its primary source of earned income. From the early 1920s through the early 1930s, the Fed was not allowed to hold government bonds, corporate securities, derivatives, or equities so, to cover its operating costs, the Fed encouraged member banks to use the discount window. Overall, then, banks that wanted to expand their commercial loan book knew they could always raise the necessary funds by discounting those loans at their district Fed, making the discount rate the relevant cost of funds.[6]

The next step is to measure the responsiveness of loan rates to these costs. Let $\beta_i^\ell$ denote the elasticity of loan returns with respect to the discount rate during the pre-Depression era in location $i$. For each $i$, we estimate $\beta_i^\ell$ by running the regression:

$$\log\left(\mathit{ReturnOnLoans}_{i,t}\right) = \alpha_i^\ell + \beta_i^\ell \log\left(\mathit{DiscountRate}_{i,t}\right) + \varepsilon_{i,t}^\ell$$

where $t$ denotes time, specifically six-month periods from 1923 to 1929 inclusive as this is the frequency for which commercial bank data are available in the 1920s. The dependent variable, $\mathit{ReturnOnLoans}_{i,t}$, 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 $t$.

The intercepts $\alpha_i^\ell$ in the above regressions control for any heterogeneity across locations that was constant over time. To also control against the possibility that there was time-varying heterogeneity unrelated to relationship lending in the 1920s, we will consider the estimate of $\beta_i^\ell$ relative to the elasticity of other returns in location $i$. The identifying assumption is that time-varying heterogeneity unrelated to relationship lending affects the return elasticity of all interest-earning assets in a given location in roughly the same way. To this end, we estimate the elasticity of securities returns with respect to the discount rate in each location $i$ during the pre-Depression era by running the regression:

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

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

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 $\beta_i^s$ is the same as for $\beta_i^\ell$.

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

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

If $NE20_i > NE20_j$ for two locations $i$ and $j$, it means that loan rates respond less in location $i$ than in location $j$, relative to interest rates on other financial products, when the discount rate changes. 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. When running regressions later on, we will also present specifications where $\beta_i^\ell$ and $\beta_i^s$ are included as separate regressors rather than restricting their coefficients to sum to zero by including only $NE20_i$.

## 4.2 Data Sources

Historical data on commercial bank assets come from the annual reports of the Comptroller of the Currency. These reports tabulate the balance sheets of commercial banks aggregated by Federal Reserve district, state, and major municipalities (principally financial centers then termed reserve cities) for June and December of each year. The Comptroller's balance sheet data are reported separately for banks with national charters and banks with state charters. In contrast, income statement data, described in more detail below, are only tabulated for nationally chartered banks. Since the income statements are necessary to impute loan rates, we limit our sample to national banks.

Of particular importance from the balance sheet data are the earning assets of national banks. These assets include loans, government bonds, and other financial securities. Until 1928, national bank loans were classified according to the four categories discussed in Section 2. Starting in 1929, the Comptroller reduced the number of categories to three, combining uncollateralized loans and loans collateralized by personal security into one category called "all other loans" and refining somewhat the cutoffs between the three remaining categories in line with the Fed's 1929 direct action campaign. Despite these refinements, the "all other loans" category remains a valid source of information on the sum of unsecured loans and loans on personal collateral. For example, we find that for the U.S. as a whole, the fraction of "all other loans" in June 1929 is roughly of the same magnitude as the fraction of loans secured by either no collateral or personal collateral in June 1928. Moreover, the cross-sectional correlation between the fraction of "all other loans" in June 1929 and the fraction of loans secured by either no collateral or personal collateral in June 1928 is 0.75, large enough to suggest that the two definitions are reasonably similar and thus interchangeable.

The Comptroller's annual reports also published earnings and expense tables for nationally chartered banks aggregated at the district, state, and city level. One table starts in 1919 and contains data for the months of January through June for each year. Another table starts in 1925 and contains data for the months of July through December for each year. Although, to the best of our knowledge, the literature has made little use of these income statements, they are, in fact, a valuable input into the construction of our price-based indicator.

Starting in 1926, the Comptroller reported earnings on loans separately from earnings on government bonds and other securities. Prior to 1926, when the Comptroller did not separate earnings by asset class, we can construct estimates. Specifically, we multiply the stock of securities held on bank balance sheets by market yields to estimate earnings on publicly-traded securities. Market yields appear in *Banking and Monetary Statistics, 1914 to 1941*. This tome recapitulates information previously published in annual and monthly reports of the Federal Reserve Board and Fed District Banks as well as trade publications such as the *Commercial and Financial Chronicle* and the *Wall Street Journal*. We also estimate interest earned from balances at other banks and interest earned on Fed securities. We then subtract these estimates from the total interest income contained in the Comptroller's report to get an estimate of loan income. This procedure delivers reasonable predictions when applied to 1926 to 1929, a period where we can compare against actuals.[7]

Dividing income by the stock of assets at the end of the period, we obtain the average loan returns and average securities returns that we need to run the elasticity regressions explained in Section 4.1. Federal Reserve discount rates are then obtained from *Banking and Monetary Statistics, 1914 to 1941*. As the discount rate in each Federal Reserve district often changed over the course of the six-month-periods that make up the time units in our panels, we calculate an average time-weighted rate in effect for the whole period. This weighted rate is a summation over the product of the interest rate and the number of days in which it was in effect, divided by the total number of days in the period.

The geographic disaggregation in the Comptroller's report allows us to construct complete time series on assets and income for 82 locations: 33 reserve cities, 31 states (net of reserve cities) each fully contained within a single Federal Reserve district, 12 parts of states (net of reserve cities) each fully contained within a single district, and 6 district residuals. Since discount rates varied across districts and, since we need to use the appropriate discount rate when running our elasticity regressions, we focus on locations that are fully contained in

[7]For loans, the correlation between predicted returns and actual returns is 0.95. For securities, it is 0.83.

a single district. The issue arises because some states are split across two districts. For a district that has only one split state, we isolate the part of this state contained in the district by subtracting from district-level data the relevant city-level data as well as state-level data for states fully contained in the district. We can then subtract the isolated part for this state from the state's total in order to ascertain the part of the state contained in another district. As long as this other district does not have more than two split states, we can repeat the process to back out any additional splits in the other district. For some districts, where there are simply too many split states to be fully identified by this iterative procedure, we define district remainders.

### 4.3 Net Elasticity Estimates

In this section, we present our net elasticity estimates and report their correlation with observables. Our goal is to show that these estimates actually do capture variations in the nature and intensity of relationship lending and do not, instead, represent spurious patterns in the data or other potentially confounding factors. Once demonstrated, we proceed to our analysis of the Great Depression.

Figure 2 plots the distributions of $\beta_i^{\ell}$, $\beta_i^{s}$, and $NE20_i$, as estimated from the data. Overall, the distribution of the net elasticities is roughly normal with mean slightly greater than zero. Since we are interested primarily in the relative ranking of locations (e.g., whether $NE20_i$ is above or below $NE20_j$, as explained above), we do not infer much from the mean of the distribution. If anything, there may be a downward bias in our calculations because of the approximation used to separate interest income in the Comptroller's report into income from loans and income from securities pre-1926. As discussed in Section 4.2, the approximation uses market yields to estimate the amount of interest that banks earned on the various components of their securities portfolios. It is an approximation because market yields are averages over different issues (e.g., the market yield on municipal bonds is an average over multiple municipalities). Accordingly, these yields are likely smoother than the yields in any individual location, understating fluctuations in securities income and overstating, by subtraction, fluctuations in loan income.[8] As we have no reason to believe that the precision of our approximations varies systematically across locations, particularly in the early 1920s when institutional investing was less advanced, the ordinal ranking of net elasticities remains informative. The main takeaway from Figure 2, that there is significant dispersion in net elasticity across locations, is exactly what we require for our analysis.

[8] If we were to dispense with the approximation and calculate net elasticities using only actuals from 1926 to 1937 (stopping in 1929 would be too short), we would find that all net elasticities are positive. However, because we want to separate the 1920s from the 1930s, we cannot use only actuals from 1926 to 1937.

In Figure 3, we plot the net elasticity, $NE20_i$, against the average fraction of national bank loans made on little to no collateral in location $i$ during the 1920s. Recall from Section 2 that observers in the 1920s and 1930s viewed uncollateralized loans as the ones most closely associated with relationship building. We would therefore expect an area with more continuing relationships, as measured by its net elasticity, to also have more uncollateralized lending. This positive correlation shows up in Figure 3.

Table 3 reports additional cross-sectional correlations, specifically between $NE20_i$ and a variety of economic indicators from the 1920 population census. We also report correlations with banking variables in 1920. We can see from Table 3 that net elasticity tended to be higher in areas that were more rural and/or had more small banks, characteristics that arise frequently in historical anecdotes about relationship lenders (e.g., Ford (1928)). This also makes sense theoretically. 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 can be mapped into Section 3.1 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.

As derived in Section 3.1, our price-based approach to measuring continuing relationships rests on the idea that differences in loan rates are driven by differences in lending practices, not by differences in funding practices. As a final confirmation, then, we would like to show that the patterns we find in loan rates are not driven by patterns in deposit rates. This is relevant to consider as Neumark and Sharpe (1992) have argued that deposit rates are slower to rise in concentrated markets. One may therefore wonder whether loan rates are stickier in some areas because market concentration makes deposit rates in those areas stickier, not because there are more continuing relationships. We already have some evidence that refutes this: Table 3 shows that locations with more concentrated banking markets, as measured by a lower number of banks per capita, had lower net elasticities and hence loan rates that were less, not more, sticky. However, we can also test the relevance of deposit rates more formally by running Granger causality tests for each location in our sample. Figure 4 plots the results. The median p-value for the null hypothesis that deposit rates did not cause loan rates in the 1920s is 0.11. In other words, for more than half of the locations in our sample, we accept the hypothesis that loan rates were not Granger-caused by deposit rates. For locations where we reject this hypothesis, we also tend to reject the opposite hypothesis that deposit rates were not Granger-caused by loan rates. It can also be seen from Figure 4 that there is a mildly positive cross-sectional correlation between the p-values of the Granger tests and our net elasticity estimates. This suggests that loan rates were less likely to be Granger-caused by deposit rates in precisely those locations where our price-based indicator finds the most evidence of continuing lending relationships.

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

We now use our new, net elasticity measure to show that cross-sectional differences in continuing lending relationships during the 1920s played a fundamental role in determining the real effects of banking distress in the early 1930s.

## 5.1 Empirical Specification

We measure economic outcomes using retail sales. Unlike construction contracts and business failures, which are leading and lagging indicators of economic activity respectively, retail sales are generally thought of as a contemporaneous indicator. In addition, few other measures of economic activity are readily available at the same frequency and level of disaggregation. We obtain retail sales for each location in our sample by aggregating the appropriate counties in the Census of Business. This census is available for the years 1929, 1933, 1935, and 1939. A subsample of the firms from the 1935 census was also surveyed in 1937. The 1937 survey reports what the 1935 results would have been had they been based on the same subsample, making it possible to scale up the 1937 results and thus approximate total retail sales for 1937. Although we focus primarily on the 1933 survey in this section, other years are considered later in the paper.

Our first step is to define an indicator of banking distress that provides an accurate representation of its nature in the early 1930s. An example will elucidate what needs to be done. Suppose bank suspensions in location $i$ amount 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 describe accurately 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 will use $SuspNum32_i$ as our main indicator of banking distress in each location $i$, with $SuspVal32_i$ as a control.

Define $Sales_{i,t}$ as retail sales in location $i$ in year $t$. Using the 82 locations discussed at the end of Section 4.2, we estimate:

$$\begin{aligned}\frac{Sales_{i,1933}}{Sales_{i,1929}} \quad = \quad & {}_0 + {}_1 NE20_i + {}_2 SuspNum32_i + {}_3 NE20_i * SuspNum32_i \\ & + {}_4 SuspVal32_i + {}_5 NE20_i * SuspVal32_i + \Box X_i + \upsilon_i\end{aligned}$$

where $X_i$ is a vector of controls. For reference, the cross-sectional correlation between $NE20_i$ and $SuspNum32_i$ is 0.17 while that between $NE20_i$ and $SuspVal32_i$ is 0.21.

If banking panics had real effects because they destroyed soft yet valuable information embodied in continuing lending relationships, we should find ${}_3 < 0$. That is, the negative effect of bank suspensions on economic activity should have been more pronounced in locations with more continuing relationships, as measured by higher values of $NE20_i$, than in comparable locations with fewer such relationships. Intuitively, firms in continuing relationships lose a critical source of working capital when their banks are suddenly suspended. If the soft information garnered through relationships is either unimportant or easy to redeploy, then the credit challenges faced by these firms should be no greater than those confronted by comparable firms in otherwise comparable locations and, thus, we should find relatively little difference in their performance. If, on the other hand, 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 their operations.

## 5.2 Regression Results

The regression results are reported in Table 4. In columns (1) to (4), the control vector $X_i$ includes district fixed effects and an indicator variable for reserve cities. In column (5), we add demographic and economic controls from the 1920 population census as well as controls for banking market structure.

The first column of Table 4 reports baseline results (i.e., estimation of the regression in Section 5.1 by ordinary least squares with simple controls). The estimates of ${}_2$ and ${}_3$ are negative and statistically significant, whereas the estimates of ${}_1$, ${}_4$, and ${}_5$ are not statistically different from zero. The magnitudes of ${}_2$ and ${}_3$ suggest that suspending 10% of national banks in the early 1930s would have led to a 3.77% decline in retail sales between 1929 and 1933 in locations where the net elasticity in the 1920s was one standard deviation above the mean. This is almost double a 2.16% decline in locations where net elasticity was at the mean. The impact of banking distress on retail sales during the Great Depression was therefore more severe in areas with more continuing relationships.

In the second column, we separate the net elasticity $NE20_i$ into its components: the elasticity of loan returns $\beta_i^\ell$ and the elasticity of securities returns $\beta_i^s$. The second column can therefore be interpreted as the regression in the first column without the restriction that the coefficients on $\beta_i^\ell$ and $\beta_i^s$ sum to zero. With $\beta_i^s$ as a separate regressor, $\beta_i^\ell$ is the relevant price-based indicator of continuing relationships. As explained in Section 3.1, lower loan rate elasticity indicates more continuing relationships. Therefore, to confirm the baseline result that banking distress had a more severe effect on retail sales in areas with more continuing relationships, we should find a positive and statistically significant coefficient on the interaction between $\beta_i^\ell$ and $SuspNum32_i$. This is exactly what we do find, as can be seen in the second column of Table 4.

In the third column, we rerun the baseline results using purged indicators of banking distress. Specifically, we attempt to remove from $SuspNum32_i$ and $SuspVal32_i$ bank suspensions that could plausibly be attributed to bad loans made in the 1920s. From the perspective of destroying continuing relationships, it does not matter why exactly banks were suspended as long as the suspensions were not driven by the relationships themselves. The low correlation found in Section 5.1 between bank suspensions and our measure of continuing relationships already suggests that banking distress was not more common in relationship lending areas but, as an additional exercise, we can run the baseline regression expunging suspensions of any delinquent loans to more narrowly proxy the unpredictable component of the banking distress (e.g., panics) that hit in the 1930s. To this end, we define an indicator of pre-Depression loan losses for each location $i$:

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

Data on loan losses come from the Comptroller's report and aggregate across all types of lending. The cross-sectional correlation between $NE20_i$ and $avgLLpre_i$ is 0.29, which is also low. We run Tobit regressions of $SuspNum32_i$ and $SuspVal32_i$ on $avgLLpre_i$, including as controls $NE20_i$, district fixed effects, and the indicator for reserve cities.[9] Using the estimated coefficients on $avgLLpre_i$ from the Tobit regressions, we obtain predictions for the banking distress indicators when all controls are set to zero. We then subtract these predictions from the actuals, and re-censor so that all negative values are recorded as zero, to generate the purged indicators of banking distress. Results using these indicators are shown in the third column of Table 4 and confirm that the negative effect of banking distress on retail sales was more severe in areas with more continuing relationships.

In the fourth column, we instrument the original (unpurged) indicators of banking distress. One may worry that lower demand for retail goods drives both a drop in retail sales and an increase in bank suspensions as firm profitability falls and defaults rise. Retail sales involve tradeable goods so district fixed effects should already soak up many of the demand-side determinants. To this point, the Federal Reserve Act of 1913 explicitly drew district boundaries around large regional markets to accord with "the convenience and customary course of business" in 1914, including trade, transportation, and communication links. In contrast, the cost of production, which is strongly influenced by the cost and availability of credit, was determined locally during the 1920s and 1930s. This is confirmed by a number of surveys conducted at the time, including the Bureau of Foreign and Domestic Commerce's *Consumer Debt Study* and the *Survey of Reports of Credit and Capital Difficulties* compiled by the Bureau of the Census. Nevertheless, it will still be useful to verify that our results are robust to instrumenting the indicators of banking distress. Calomiris and Mason (2003) have argued that the following variables, all measured in 1929, are acceptable instruments: logged banking assets, real estate owned by banks as a fraction of non-cash banking assets, and bank capital as a fraction of banking assets. Accordingly, we will check the robustness of our baseline results to the use of these instruments, adding cash-to-deposit and loan-to-deposit ratios in 1929 to improve the first-stage estimation.[10] The fourth column of Table 4 shows that, once banking distress is instrumented, the only significant predictor of the change in retail sales from 1929 to 1933 is the interaction between $SuspNum32_i$ and $NE20_i$ and, as before, the coefficient is negative. In an alternative specification, we removed the loan-to-deposit ratio as an instrument and replaced it with the ratio of demand deposits to

[9] The coefficient on $avgLLpre_i$ is positive and statistically significant in both Tobit regressions, suggesting that some of the bank suspensions in the early 1930s were predictable based on loan losses in the late 1920s. The coefficient on $NE20_i$ is not statistically significant in either regression.

[10] The identifying assumption of Calomiris and Mason (2003) is that shocks in the 1930s were not just a continuation of shocks in the 1920s. The two ratios we add are informative about the ability of banks in different areas to withstand deposit withdrawals (positive for the former, negative for the latter) and should therefore have predictive power for the amount of banking distress, at least in the early stages of the crisis.

banking assets in 1929. The results were very similar.

In the fifth column, we add demographic and economic variables from the 1920 population census to the vector of controls in the baseline regression (see Table 3 for a list of these variables). We also control for banking concentration – specifically, the log of banks per capita in 1920 or $LogBankpc_i$ as defined in Table 3 – as well as its interactions with the banking distress indicators.[11] There are two noteworthy results. First, the coefficient on the interaction between $SuspNum32_i$ and $NE20_i$ remains negative and statistically significant, confirming that bank suspensions have worse effects in areas with more continuing relationships. Second, the coefficients on $SuspVal32_i$ and its interactions are now also statistically significant. The coefficient on $SuspVal32_i$ is negative, meaning that a larger banking shock leads to a bigger decline in retail sales. In contrast, the coefficient on $SuspVal32_i * NE20_i$ is positive, as is the coefficient on $SuspVal32_i * LogBankpc_i$. Increasing $SuspVal32_i$ without also increasing $SuspNum32_i$ is akin to considering suspensions of larger banks. In the early 1930s, one respect in which the suspension of large banks was less damaging to the economy than the suspension of small banks was the rate at which depositors could access their funds. Large banks during this period were liquidated quite rapidly, both in absolute terms and relative to smaller banks, meaning that depositors in large banks were able to access their funds much more quickly than depositors in small banks. Receivership data for 1931 and 1932 show 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. Depositors in banks that failed with $2-6 million in deposits received only 3.0 cents per month while depositors in banks that failed with less than $2 million in deposits received 2.0 cents per month. In practice, suspended 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. Given that depositors of large suspended banks still had at least partial access to their funds, the suspension of a large bank created space in the business landscape for surviving banks to increase their deposits and thus their lending activities. The coefficients in the fifth column of Table 4 suggest that this reallocation was most valuable in less concentrated banking markets (as captured by more banks per capita) and in areas where business was more dependent on continuing relationships (as measured by higher net elasticity in the 1920s).

[11] We exclude the bank size variable in Table 3 ($LogBankSize_i$) from the vector of controls because its correlation with $LogBankpc_i$ is -0.83. Including the remaining variables ($DepRatio_i$ and $DDratio_i$) as controls does not change the results.

## 5.3 Discussion

The main takeaway from Table 4 is that banking panics had a more pronounced negative effect on retail sales in locations with more continuing relationships. Across all specifications, the coefficient on the interaction between $SuspNum32_i$ and $NE20_i$ was negative and statistically significant. We also found some evidence that efforts by receivers and clearinghouses to ensure a rapid release of funds to depositors of large suspended banks had a more pronounced positive effect on retail sales in locations with more continuing relationships. While this was observed most clearly in the last column of Table 4, with a positive and statistically significant coefficient on the interaction between $SuspVal32_i$ and $NE20_i$, the positive coefficient is also visible in other columns, albeit without the same statistical significance.

In Table 5, we repeat the analysis using the retail sales ratio in 1935, rather than the retail sales ratio in 1933, as the dependent variable. The coefficient on the interaction between $SuspNum32_i$ and $NE20_i$ is still negative but no longer statistically significant, except in the last column which is the regression with census controls. The interaction between $SuspVal32_i$ and $NE20_i$ is also still positive and statistically significant in this column. In Table 6, we use the retail sales ratio in 1937 as the dependent variable and find that all statistical significance disappears. In other words, bank suspensions from the early 1930s no longer had a direct effect on retail sales by 1937. We emphasize that this is only a statement about direct effects. We will turn to indirect effects in Section 6.

The estimated coefficients can also 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 (i.e., $SuspNum32_i = SuspVal32_i = 0$ for all locations $i$) then compare this to the fitted values when all regressors are as observed in the data. Depending on which column in Table 4 is used to generate the comparison, we find that total retail trade in the U.S. would have been 1% to 4% higher in 1933 had there been no banking distress. If we set only $SuspNum32_i = 0$ 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 trade would have been 3% to 9% higher in 1933 had there been no banking distress. In both counterfactuals, the upper bounds come from the instrumental variables regression (fourth column) while the lower bounds come from the regression with census controls (fifth column). That being said, the regression with census controls delivers more sizeable estimates for 1935 than it does for 1933. In particular, it predicts that total retail trade would have been 3% higher in 1935 had there been no banking distress and 5% higher had there also been no mitigating effect.

Are the aggregate effects just derived economically significant? This is a question about both the magnitude of the effects we found and the importance of retail trade for the broader economy. Geographically disaggregated measures of total economic activity (e.g., GNP) are not available for this period, but 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 for the period 1920-1929 of GNP growth on retail sales growth, with the latter instrumented by its one period lag. We start in 1920 because that is when the Federal Reserve began publishing its monthly index of retail sales. We seasonally adjust this index using Census software then take annual averages to match the frequency of the GNP estimates available from Romer (1989) and Balke and Gordon (1989). We end in 1929 to ensure that our results are not distorted by the Great Depression, the 1937-38 recession, or WWII. We also eliminate the post-WWII period since the ratio of retail sales to GNP declines markedly after the war. The multipliers from various specifications are reported in the first row of Table 7. Overall, we find that a 4% drop in retail sales is consistent with a 6-7% decline in nominal GNP and a 3-4% fall in real GNP, which is about one-eighth of the economic contraction experienced during the Great Depression. This would seem to leave little doubt about the economic significance of the results.

# 6 Relationship Rebuilding and the 1937-38 Recession

In this section, we show that cross-sectional differences in the rebuilding of lending relationships destroyed during the Great Depression are important for understanding cross-sectional differences in economic performance later in the decade. Specifically, areas that appear to have been rebuilding relationships in the mold of those destroyed by the banking panics of the early 1930s fared better than others during the 1937-38 recession. This finding is important for two reasons. First, it sheds light on the heterogeneous nature of recovery from the Depression. Second, it strengthens the interpretation of our previous results by serving as evidence that what we measured to be continuing relationships on the eve of the Great Depression do indeed have value in a period where they are not forcibly destroyed.

## 6.1 Empirical Specification

The spirit of the analysis is similar to Section 5 in that we want to see how continuing relationships affect performance in a crisis. The crisis we considered in Section 5 was the Great Depression. The crisis we are considering here is the 1937-38 recession. However, there are two differences between these events that will affect both the regressions we run and the intuition behind the results: the magnitude of the crisis and the precision with which we can detect continuing relationships on its eve.

Consider first the magnitude. The Great Depression involved bank runs by depositors and swaths of bank suspensions. In contrast, the 1937-38 recession involved an increase in bank funding costs that did not spiral out of control, possibly benefitting from the introduction of deposit insurance in 1934. Friedman and Schwartz (1963) attribute higher bank funding costs in the late 1930s to the Fed's decision to begin doubling reserve requirements in August 1936. Hanes (2006) and Calomiris et al (2011) instead assign a more important role to the Treasury's decision to sterilize gold inflows starting in December 1936. The bottom line is the same in both cases: contractionary monetary policy was introduced in the U.S. in 1936. The model in Section 3.1 predicts that continuing relationships will mitigate policy tightening by extending policy-invariant loan rates to some borrowers, thus enabling them to continue operations uninterrupted.[12] Of course, the ability of continuing relationships to absorb a cost-of-funds shock is conditional on bank survival. If the banks are instead destroyed, as they were during the runs of the early 1930s, the continuing relationships are also destroyed and, as we saw in Section 5, locations dependent on those relationships suffer greatly. Whether continuing relationships absorb or amplify potential crises is, therefore, sensitive to the nature of the shock.

A second difference between the Great Depression and the 1937-38 recession is the precision with which we can detect continuing relationships on the eve of the crisis. We were able to construct our price-based indicator of continuing relationships on the eve of the Great Depression using data from the 1920s. This is much more difficult to do for the 1937-38 recession for the obvious reason that early years of the decade were scarred by the Depression, giving us only three years (1934-36) of recovery to work with, too few to produce reliable estimates of the elasticity of asset returns using the semi-annual data published by the Comptroller.[13] If we were to estimate elasticities using data from 1930 to 1936, we would find that the cross-sectional correlation between net elasticity in the 1920s and the change in net elasticity from the 1920s to the 1930s is -0.90, which only reinforces the point that the Great Depression destroyed continuing relationships in areas where such relationships existed. We therefore need a proxy to measure the extent to which continuing relationships had been (re)built by 1936.

We select as our proxy the interaction between two variables prior to the 1937-38 recession. The first variable is $NE20_i$, our price-based indicator of continuing relationships in the

[12] By making loan rates less responsive to changes in bank funding costs, continuing relationships also make firm project choice, and hence production, less responsive to those changes. See Hachem (2011) for a formal treatment of this issue.

[13] Note also that the first real test for the majority of post-Depression lenders may have been the 1937-38 recession itself.

1920s. The idea is that locations where continuing relationships had previously flourished are likely to be the ones where they will again thrive, on the assumption that the preferences of banks and the needs of firms were, for the most part, stable over time. We will return to this point when discussing control variables in Section 6.2. The second variable, $CollLite30_i$, is the average fraction of bank loans in location $i$ that were either uncollateralized or collateralized by personal security between 1934 and 1936. We know from Section 2 that observers in the 1920s and 1930s viewed uncollateralized loans as the most natural loans for relationship building. It then stands to reason that locations with high values of $NE20_i$ and $CollLite30_i$ are the most likely candidates for relationship rebuilding: they had a history of continuing relationships and they were still making the types of loans that were an input into such relationships.

To study whether continuing relationships affected the severity of the 1937-38 recession, we run the regression:

$$\frac{Sales_{i,1939}}{Sales_{i,1937}} = \lambda_0 + \lambda_1 NE20_i + \lambda_2 CollLite30_i + \lambda_3 NE20_i * CollLite30_i + \Pi Z_i + \zeta_i$$

where $Z_i$ is a vector of controls described in more detail below. If continuing relationships absorbed the contractionary monetary policy shocks that others have argued occurred in the U.S. in 1936, we should find $\lambda_3 > 0$. That is, the late 1930s recession should have been less pronounced in those locations where we would expect to find the greatest concentration of rebuilt lending relationships.

## 6.2 Results and Discussion

The regression results are reported in Table 8.[14] In column (1), the control vector $Z_i$ includes district fixed effects and an indicator variable for reserve cities. In column (2), we add as a control the excess reserve ratio in location $i$ on the eve of the 1937-38 recession and interact it with the main regressors. This ratio, $ExcessRR_i$, is calculated as the ratio of excess reserves to demand deposits, both measured in June 1936. In column (3), we instrument the excess reserve ratio and its interactions to address potential endogeneity concerns.

There are two reasons why banks in some areas may have been holding high excess reserve ratios in June 1936. The first is backward-looking and is the one that motivates inclusion of $ExcessRR_i$ in the vector of controls: the banking panics of the early 1930s may have

[14] Recall from Section 4.2 that uncollateralized lending for the 1930s is based on the "all other loans" category in the Comptroller's report. The Comptroller separated "all other loans" from "loans on securities" for cities and states but not districts so we cannot use any of the split states or district residuals in the analysis here. As a result, the number of locations is less than 82.

prompted bankers in those areas to hold precautionary reserve cushions. The problem is that heterogeneous changes in liquidity management from the 1920s to the 1930s confound the ability of continuing relationships in the pre-Depression era to predict the rise of similar relationships once the Depression had subsided. Controlling for excess reserve ratios in June 1936 helps control for the effect of the Great Depression on banker preferences, making $NE20_i * CollLite30_i$ a better predictor of continuing relationships on the eve of the 1937-38 recession.[15]

The second possible reason for high excess reserve ratios in June 1936 is forward-looking: bankers in those areas may have opted for large reserve cushions in anticipation of poor economic conditions ahead. This introduces an endogeneity concern so we use instrumental variables to isolate only the backward-looking component of excess reserve ratios in June 1936. Changes in liquidity management from the 1920s to the 1930s are more likely to have occurred in locations where banks were burned by liquidity problems in the early 1930s, motivating the following two instruments for $ExcessRR_i$. The first instrument is the ratio of cash to demand deposits in location $i$ in 1929. Areas where banks held less cash relative to demand deposits on the eve of the Great Depression would have been more vulnerable once depositors began to panic and withdraw funds. The second instrument is the purged indicator of banking distress in the early 1930s based on the number of suspended banks in location $i$ (see Section 5.2). This captures the extent to which banks with otherwise good assets succumbed to the panics.

The third column of Table 8 presents the results. We find that $\lambda_3$ (the coefficient on $NE20_i * CollLite30_i$) is positive and statistically significant while the coefficient on the interaction between $NE20_i * CollLite30_i$ and $ExcessRR_i$ is negative and statistically significant. The magnitudes of the estimated coefficients imply that economic performance during the late 1930s recession was increasing in $NE20_i * CollLite30_i$ for any location $i$ with an excess reserve ratio below 8.4%. For comparison, the mean of $ExcessRR_i$ across locations was 7.0% with a standard deviation of 3.9%.

We would argue that the Great Depression did not give rise to a preference for liquidity among banks in locations where our two instruments predict low excess reserve ratios in June 1936. Accordingly, any relationships rebuilt in these locations are the ones most likely to have been cut from the same cloth as the continuing relationships that existed in the 1920s. The positive coefficient on $NE20_i * CollLite30_i$ for these locations substantiates this claim in that the 1937-38 recession was less pronounced in areas where more continuing relationships

[15] The seminal work of Diamond and Dybvig (1983) expounds the fundamental fragility of banks vis-à-vis liquidity so it seems reasonable to focus specifically on the preference for liquidity when considering the effect of the Depression on banker preferences.

had been built, consistent with the shock-absorbing nature of these relationships as shown in Section 3.1. The negative coefficient on $NE20_i * CollLite30_i$ for locations where our instruments predict high excess reserve ratios then provides evidence that the recession was more pronounced when attempts at relationship rebuilding had been made but banks were less dedicated to maintaining them than they had been in the 1920s. Bankers chastened by the Great Depression into holding large excess reserves may also have been inclined to preemptively discontinue even good lending relationships in order to expand their liquidity buffers at the first sign of trouble. Such behavior would cause economic activity to contract for the very same reason that it did during the Depression: the soft information about borrowers cannot be easily redeployed.

We also find that the coefficient on $NE20_i$ in the third column of Table 8 is negative and statistically significant while the coefficient on the interaction between $NE20_i$ and $ExcessRR_i$ is positive and statistically significant. At the average value of $ExcessRR_i$, the magnitudes of the estimated coefficients imply that economic performance during the late 1930s recession was decreasing in $NE20_i$ for any location $i$ with $CollLite30_i$ below 0.56. For comparison, the mean of $CollLite30_i$ across locations was 0.55 with a standard deviation of 0.12. Section 5 established that bank suspensions in the early 1930s had a more contractionary effect on economic activity in locations with more continuing relationships. It would therefore appear that these locations stayed weak into the late 1930s unless they continued making the types of loans that observers argued built relationship capital (i.e., uncollateralized loans). The interaction between $NE20_i$ and $CollLite30_i$ is therefore important for capturing relationship rebuilding and, as discussed above, relationship rebuilding in the aftermath of the Great Depression helped mitigate the 1937-38 recession when the relationships reproduced those that had existed in the 1920s.

# 7 Conclusion

We proposed in this paper a novel measure of continuing lending relationships that can be calculated using data from geographically aggregated financial statements, resolving the data limitations of the 1920s and 1930s and pinpointing 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 involve loan rates that are less responsive to changes in bank funding costs. We discussed theoretical foundations for this idea and validated it empirically using loan-level data from Dealscan, a commonly cited database in modern analyses of relationship lending. We then implemented our measure to study whether the Great Depression was amplified by the destruction of continuing relationships and whether the rebuilding of these relationships affected the subsequent recovery.

We showed that the marginal impact of bank suspensions on economic activity in the early 1930s was more negative in areas that had more continuing relationships on the eve of the Great Depression. We also showed that relationship lending played an important role during the recovery period. In particular, areas that rebuilt the types of continuing relationships that they embraced in the 1920s fared better during the 1937-38 recession than otherwise similar areas that did not rebuild. In contrast, areas that rebuilt relationships but, based on their reserve holdings, appear to have been less dedicated to these relationships, fared worse than otherwise similar areas that failed to rebuild.

In keeping with Bernanke (1983) and others, we have focused on the Great Depression. Given the dimensions of the downturn and the scope of the banking panics, this episode remains ground zero for studying financial crises. However, the value of our methodology transcends resolution of data limitations for the 1920s and 1930s. Policymakers working in real time, often in crisis situations, usually have only aggregate data to provide them with information about what is happening at more disaggregated levels. Our method permits them to use these aggregates to extract this more detailed information. Our findings also transcend the historical context of the Great Depression. In many parts of the world where relationship loans constitute a major part of bank assets, bank failures are likely to have serious and potentially long-lasting deleterious effects on the real economy. Fortunately, policymakers, fortified with this knowledge, have the means to mitigate the damage if they act swiftly and strongly enough to shore up the banks and salvage the soft yet crucial information embodied in their continuing lending relationships.

# References


Balke, N. and R. Gordon. 1989. "The Estimation of Prewar Gross National Product: Methodology and New Evidence." *Journal of Political Economy*, 97(1), pp. 38-92.

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

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

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

Calomiris, C., J. Mason, and D. Wheelock. 2011. "Did Doubling Reserve Requirements Cause the Recession of 1937-1938? A Microeconomic Approach." NBER Working Paper No. 16688.

Carlson, M. and J. Rose. 2015. "Credit Availability and the Collapse of the Banking Sector in the 1930s." *Journal of Money, Credit, and Banking*, 47(7), pp. 1239-1271.

Chava, S. and M. Roberts. 2008. "How Does Financing Impact Investment? The Role of Debt Covenants." *Journal of Finance*, 63(5), pp. 2085-2121.

Chodorow-Reich, G. 2014. "The Employment Effects of Credit Market Disruptions: Firm-Level Evidence from the 2008-9 Financial Crisis." *Quarterly Journal of Economics*, 129(1), pp. 1-59.

Cole, H. and L. Ohanian. 1999. "The Great Depression in the United States from a Neoclassical Perspective." *Federal Reserve Bank of Minneapolis Quarterly Review*, 23(1), pp. 2-24.

Cole, H. and L. Ohanian. 2004. "New Deal Policies and the Persistence of the Great Depression: A General Equilibrium Analysis." *Journal of Political Economy*, 112(4), pp. 779-816.

Cole, H. and L. Ohanian. 2007. "A Second Look at the U.S. Great Depression from a Neoclassical Perspective." In Kehoe, T. and E. Prescott (eds.), *Great Depressions of the Twentieth Century*, Federal Reserve Bank of Minneapolis, pp. 21-57.

Currie, L. 1931. "The Decline of the Commercial Loan." *Quarterly Journal of Economics*, 45(4), pp. 698-709.

Diamond, D. and P. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy*, 91(3), pp. 401-419.

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

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

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

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

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

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

Hanes, C. 2006. "The Liquidity Trap and U.S. Interest Rates in the 1930s." *Journal of Money, Credit, and Banking*, 38(1), pp. 163-194.

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

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

Neumark, D. and S. 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, S. and D. Smith. 2001. "The Duration of Bank Relationships." *Journal of Financial Economics*, 61(3), pp. 449-475.

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

Richardson, G. and W. Troost. 2009. "Monetary Intervention Mitigated Banking Panics During the Great Depression: Quasi-Experimental Evidence from the Federal Reserve District Border in Mississippi, 1929 to 1933." *Journal of Political Economy*, 117(6), pp. 1031-1073.

Romer, C. 1989. "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869-1908." *Journal of Political Economy*, 97(1), pp. 1-37.

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

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

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

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

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

Table 1:
Fraction of Deal Amount in Fixed Rate Loans

| | Ordinary Least Squares | | | Fractional Logit | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Prior Deals | 0.00357*** | 0.00311*** | 0.00799*** | 0.0549*** | 0.0154* | 0.0865*** |
| | (9.39e-05) | (0.00016) | (0.00066) | (0.00558) | (0.00790) | (0.0170) |
| Maturity | 0.00128*** | 0.00125*** | 0.00086*** | 0.0101*** | 0.00885*** | 0.0138*** |
| | (7.05e-05) | (7.11e-05) | (8.66e-05) | (0.00057) | (0.00057) | (0.00127) |
| Secured | -0.0906*** | -0.0888*** | 0.0140*** | -1.383*** | -1.311*** | -0.0420 |
| | (0.00245) | (0.00257) | (0.00325) | (0.0506) | (0.0542) | (0.0792) |
| Prior Deals × Maturity | | 1.26e-05*** | 3.82e-06 | | 0.00146*** | -0.00044** |
| | | (3.49e-06) | (1.05e-05) | | (0.00028) | (0.00021) |
| Prior Deals × Secured | | -0.00534*** | -0.0172*** | | -0.147** | -0.267*** |
| | | (0.00182) | (0.00144) | | (0.0705) | (0.0650) |
| Log (Firm Assets) | | | 0.0344*** | | | 0.366*** |
| | | | (0.0010) | | | (0.0145) |
| Observations | 53,730 | 53,730 | 24,959 | 53,730 | 53,730 | 24,959 |
| Industry Dummies | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| R-squared | 0.271 | 0.272 | 0.400 | | | |

Notes: Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1

Table 2:
Alternative Specifications with Dealscan Data

| | Loan Level, S.E. Clustered at Deal Level | | | | Deals with One Loan | | |
|---|---|---|---|---|---|---|---|
| | OLS | | Fractional Logit | | Binary Dep Variable, Logit | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Prior Deals | 0.00364*** | 0.00786*** | 0.0195** | 0.0636*** | 0.0597*** | 0.0302*** | 0.0589*** |
| | (0.00023) | (0.00051) | (0.00757) | (0.0129) | (0.00536) | (0.00879) | (0.0129) |
| Maturity | 0.00065*** | 0.00072*** | 0.00483*** | 0.0117*** | 0.0123*** | 0.0110*** | 0.0151*** |
| | (7.07e-05) | (0.00010) | (0.00047) | (0.00158) | (0.00069) | (0.00075) | (0.00125) |
| Secured | -0.0901*** | 0.0157*** | -1.295*** | -0.0447 | -1.405*** | -1.307*** | 0.139 |
| | (0.00308) | (0.00383) | (0.0620) | (0.0891) | (0.0552) | (0.0606) | (0.0952) |
| Prior Deals × Maturity | 2.87e-05*** | -8.36e-06 | 0.00149*** | -0.00020 | | 0.00105*** | -0.00021 |
| | (5.82e-06) | (6.24e-06) | (0.00024) | (0.00018) | | (0.00028) | (0.00018) |
| Prior Deals × Secured | -0.00745*** | -0.0275*** | -0.235*** | -0.478*** | | -0.225*** | -0.757*** |
| | (0.00151) | (0.00182) | (0.0660) | (0.121) | | (0.0763) | (0.142) |
| Log (Firm Assets) | | 0.0362*** | | 0.373*** | | | 0.427*** |
| | | (0.00114) | | (0.0157) | | | (0.0178) |
| Observations | 62,178 | 23,601 | 62,178 | 23,601 | 41,119 | 41,119 | 16,659 |
| Industry Dummies | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| R-squared | 0.267 | 0.408 | | | | | |

Notes: Robust standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1

Table 3:
Cross-Sectional Correlations

| | | Correlation with NE20 |
|---|---|---|
| LogArea | log(area in square miles) | 0.460 |
| Urban | % population urban in 1920 | -0.543 |
| Nwnp | % population with native white parents in 1920 | 0.325 |
| Age1844 | % population aged 18 to 44 in 1920 | -0.435 |
| School1620 | % of aged 16 to 20 in school in 1920 | 0.329 |
| LogMfgEst | log(number of mfg establishments in 1920) | -0.084 |
| LogMfgSize | log(workers per mfg establishment in 1920) | -0.297 |
| MfgWork | mfg workers as % of population in 1920 | -0.390 |
| LogMfgVa | log(value added per mfg establishment in 1920) | -0.314 |
| LogFarms | log(number of farms in 1920) | 0.491 |
| Acres | farm land as % of area in 1920 | 0.291 |
| LogAvgAcre | log(farm acres / number of farms in 1920) | 0.344 |
| LogAvgCrop | log(crop value / number of farms in 1920) | 0.186 |
| LogAvgValue | log(value of farm land, equip, etc / no. of farms in 1920) | -0.124 |
| OwnerOp | % of farms owner-operated in 1920 | -0.080 |
| HomeOwnClr | home ownership rate in 1920 | 0.457 |
| LogBankpc | log(number of national banks / population in millions in 1920) | 0.479 |
| LogBankSize | log(banking assets / number of banks in 1920) | -0.511 |
| DepRatio | demand and time deposits / banking assets in 1920 | 0.276 |
| DDratio | demand deposits / demand and time deposits in 1920 | -0.302 |

Notes: All variables are from the 1920 population census, except the last four which are from the Comptroller's annual report in 1920.

Table 4:
Retail Sales Ratio, 1933 / 1929

| | Baseline | Comp. Elast. | Purged Susp. | IV Susp. | Extra Controls |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| NE20 | 0.00777 | | 0.0131 | 0.0252 | 0.0164 |
| | (0.0174) | | (0.0175) | (0.0304) | (0.0195) |
| SuspNum32 | -0.201** | -0.257*** | -0.201*** | -0.326 | 0.0717 |
| | (0.0668) | (0.0654) | (0.0610) | (0.225) | (0.391) |
| SuspNum32 × NE20 | -0.327* | | -0.832* | -0.554* | -0.624** |
| | (0.162) | | (0.458) | (0.292) | (0.280) |
| SuspVal32 | 0.141 | 0.197 | 0.0881 | 0.393 | -0.563** |
| | (0.0971) | (0.139) | (0.0990) | (0.338) | (0.186) |
| SuspVal32 × NE20 | 0.330 | | 1.032 | 0.133 | 0.714** |
| | (0.269) | | (0.658) | (0.657) | (0.254) |
| $\beta^{\ell}$ | | -0.0210 | | | |
| | | (0.0414) | | | |
| SuspNum32 × $\beta^{\ell}$ | | 0.433* | | | |
| | | (0.229) | | | |
| SuspVal32 × $\beta^{\ell}$ | | -0.493 | | | |
| | | (0.957) | | | |
| $\beta^{s}$ | | 0.00025 | | | |
| | | (0.0172) | | | |
| SuspNum32 × $\beta^{s}$ | | -0.301 | | | |
| | | (0.194) | | | |
| SuspVal32 × $\beta^{s}$ | | 0.318 | | | |
| | | (0.241) | | | |
| LogBankpc | | | | | 0.0108 |
| | | | | | (0.0352) |
| SuspNum32 × LogBankpc | | | | | -0.0424 |
| | | | | | (0.114) |
| SuspVal32 × LogBankpc | | | | | 0.153** |
| | | | | | (0.0505) |
| Observations | 82 | 82 | 82 | 82 | 81 |
| District and RC Dummies | ✓ | ✓ | ✓ | ✓ | ✓ |
| Census Controls | × | × | × | × | ✓ |
| R-squared | 0.496 | 0.500 | 0.498 | | 0.781 |
| J-Statistic (p-value) | | | | 0.830 | |
| AR Wald Test (p-value) | | | | 0.000 | |

Notes: All columns include district fixed effects and a dummy variable for reserve cities. Column (1) is the baseline regression in Section 5.1. Column (2) splits NE20 into the component elasticities defined in Section 4.1. Column (3) purges SuspNum32 and SuspVal32 of bank suspensions that could plausibly be attributed to bad loans made in the 1920s. Column (4) instruments SuspNum32 and SuspVal32 following Calomiris and Mason (2003). Column (5) adds controls from Table 3 to the baseline regression, specifically all the variables from the 1920 population census and LogBankpc. Standard errors clustered at the district level are in parentheses. ***p<0.01, **p<0.05, *p<0.1

Table 5:
Retail Sales Ratio, 1935 / 1929

| | Baseline | Comp. Elast. | Purged Susp. | IV Susp. | Extra Controls |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| NE20 | -0.0198 | | -0.00819 | -0.0104 | -0.0189 |
| | (0.0129) | | (0.0138) | (0.0409) | (0.0245) |
| SuspNum32 | -0.291** | -0.383* | -0.366** | -0.307 | 0.593 |
| | (0.117) | (0.176) | (0.130) | (0.357) | (0.430) |
| SuspNum32 × NE20 | -0.227 | | -0.913 | -0.267 | -0.531** |
| | (0.296) | | (0.816) | (0.583) | (0.211) |
| SuspVal32 | 0.250 | 0.206 | 0.131 | 0.447 | -1.312* |
| | (0.216) | (0.233) | (0.208) | (0.530) | (0.614) |
| SuspVal32 × NE20 | 0.192 | | 1.378 | -0.140 | 0.706** |
| | (0.396) | | (1.180) | (0.705) | (0.314) |
| $\beta^{\ell}$ | | 0.0223 | | | |
| | | (0.0584) | | | |
| SuspNum32 × $\beta^{\ell}$ | | 0.177 | | | |
| | | (0.299) | | | |
| SuspVal32 × $\beta^{\ell}$ | | 0.489 | | | |
| | | (1.065) | | | |
| $\beta^{s}$ | | -0.0181 | | | |
| | | (0.0182) | | | |
| SuspNum32 × $\beta^{s}$ | | -0.116 | | | |
| | | (0.292) | | | |
| SuspVal32 × $\beta^{s}$ | | 0.0754 | | | |
| | | (0.344) | | | |
| LogBankpc | | | | | 0.0111 |
| | | | | | (0.0399) |
| SuspNum32 × LogBankpc | | | | | -0.225* |
| | | | | | (0.117) |
| SuspVal32 × LogBankpc | | | | | 0.359** |
| | | | | | (0.153) |
| Observations | 82 | 82 | 82 | 82 | 81 |
| District and RC Dummies | ✓ | ✓ | ✓ | ✓ | ✓ |
| Census Controls | × | × | × | × | ✓ |
| R-squared | 0.474 | 0.485 | 0.489 | | 0.755 |
| J-Statistic (p-value) | | | | 0.737 | |
| AR Wald Test (p-value) | | | | 0.017 | |

Notes: All columns include district fixed effects and a dummy variable for reserve cities. Column (1) is the baseline regression in Section 5.1. Column (2) splits NE20 into the component elasticities defined in Section 4.1. Column (3) purges SuspNum32 and SuspVal32 of bank suspensions that could plausibly be attributed to bad loans made in the 1920s. Column (4) instruments SuspNum32 and SuspVal32 following Calomiris and Mason (2003). Column (5) adds controls from Table 3 to the baseline regression, specifically all the variables from the 1920 population census and LogBankpc. Standard errors clustered at the district level are in parentheses. ***p<0.01, **p<0.05, *p<0.1

Table 6:
Retail Sales Ratio, 1937 / 1929

| | Baseline | Comp. Elast. | Purged Susp. | IV Susp. | Extra Controls |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| NE20 | -0.431 | | -0.384 | -4.134 | 0.280 |
| | (0.613) | | (0.470) | (4.980) | (1.124) |
| SuspNum32 | -0.817 | -0.973 | -0.586 | 22.61 | -1.496 |
| | (0.833) | (1.397) | (0.777) | (28.81) | (18.35) |
| SuspNum32 × NE20 | 7.359 | | 11.23 | 55.30 | 17.12 |
| | (9.418) | | (11.13) | (64.47) | (19.20) |
| SuspVal32 | 0.582 | 7.020 | 0.0781 | -19.47 | -36.55 |
| | (1.074) | (6.194) | (1.457) | (40.48) | (36.54) |
| SuspVal32 × NE20 | -7.963 | | -7.669 | 8.671 | -32.74 |
| | (8.744) | | (9.094) | (77.52) | (28.51) |
| $\beta^{\ell}$ | | 3.184 | | | |
| | | (2.773) | | | |
| SuspNum32 × $\beta^{\ell}$ | | -13.40 | | | |
| | | (13.94) | | | |
| SuspVal32 × $\beta^{\ell}$ | | -12.92 | | | |
| | | (13.69) | | | |
| $\beta^{s}$ | | 0.431 | | | |
| | | (0.361) | | | |
| SuspNum32 × $\beta^{s}$ | | 12.59 | | | |
| | | (13.95) | | | |
| SuspVal32 × $\beta^{s}$ | | -16.92 | | | |
| | | (15.71) | | | |
| LogBankpc | | | | | -2.425 |
| | | | | | (3.273) |
| SuspNum32 × LogBankpc | | | | | -0.184 |
| | | | | | (5.273) |
| SuspVal32 × LogBankpc | | | | | 9.874 |
| | | | | | (9.852) |
| Observations | 82 | 82 | 82 | 82 | 81 |
| District and RC Dummies | ✓ | ✓ | ✓ | ✓ | ✓ |
| Census Controls | × | × | × | × | ✓ |
| R-squared | 0.060 | 0.076 | 0.059 | | 0.258 |
| J-Statistic (p-value) | | | | 0.964 | |
| AR Wald Test (p-value) | | | | 0.918 | |

Notes: All columns include district fixed effects and a dummy variable for reserve cities. Column (1) is the baseline regression in Section 5.1. Column (2) splits NE20 into the component elasticities defined in Section 4.1. Column (3) purges SuspNum32 and SuspVal32 of bank suspensions that could plausibly be attributed to bad loans made in the 1920s. Column (4) instruments SuspNum32 and SuspVal32 following Calomiris and Mason (2003). Column (5) adds controls from Table 3 to the baseline regression, specifically all the variables from the 1920 population census and LogBankpc. Standard errors clustered at the district level are in parentheses. ***p<0.01, **p<0.05, *p<0.1

Table 7:
GNP Multiplier for Retail Sales

| | Nominal GNP Growth | | Real GNP Growth | |
|---|---|---|---|---|
| | Romer | Balke-Gordon | Romer | Balke-Gordon |
| RetailGrowth | 1.690*** | 1.565*** | 0.696*** | 0.896*** |
| | (0.242) | (0.299) | (0.163) | (0.242) |
| Constant | -0.0102 | -0.00644 | 0.0278*** | 0.0246*** |
| | (0.00998) | (0.0115) | (0.00512) | (0.0056) |

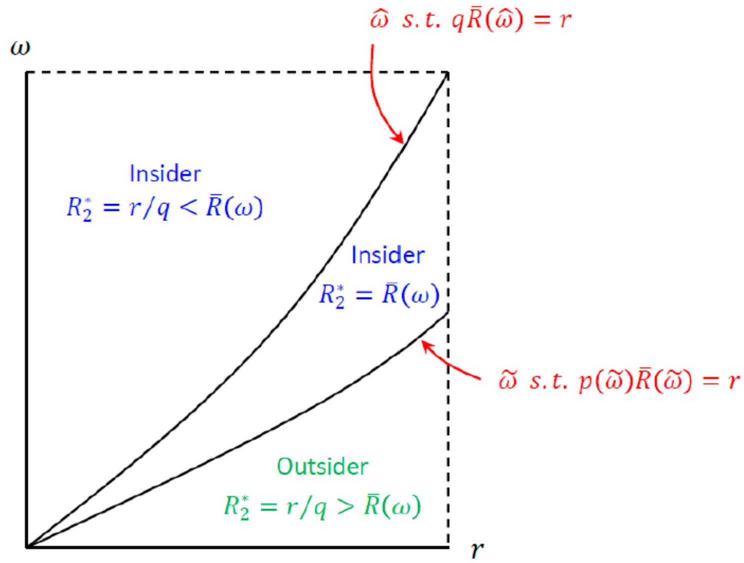
Notes: Data are annual for the period 1920-1929. RetailGrowth is instrumented using its one period lag. Robust standard errors are in parentheses. ***p<0.01, **p<0.05, *p<0.1

Table 8:
Retail Sales Ratio, 1939 / 1937

| | OLS | OLS | IV |
|---|---|---|---|
| | (1) | (2) | (3) |
| NE20 | 0.612 | -0.777 | -5.256** |
| | (0.387) | (1.363) | (2.361) |
| CollLite30 | -0.114 | -0.562 | -1.695 |
| | (0.436) | (0.499) | (1.582) |
| NE20 × CollLite30 | -0.890 | 1.421 | 9.155** |
| | (0.595) | (2.321) | (4.024) |
| ExcessRR | | -3.170 | -14.12 |
| | | (5.587) | (17.14) |
| NE20 × ExcessRR | | 15.60 | 62.88** |
| | | (11.21) | (26.77) |
| CollLite30 × ExcessRR | | 6.274 | 24.77 |
| | | (8.082) | (26.88) |
| NE20 × CollLite30 × ExcessRR | | -26.00 | -109.0*** |
| | | (18.90) | (42.32) |
| Observations | 64 | 63 | 63 |
| District and RC Dummies | ✓ | ✓ | ✓ |
| R-squared | 0.619 | 0.653 | |
| J-Statistic (p-value) | | | 0.328 |
| AR Wald Test (p-value) | | | 0.000 |

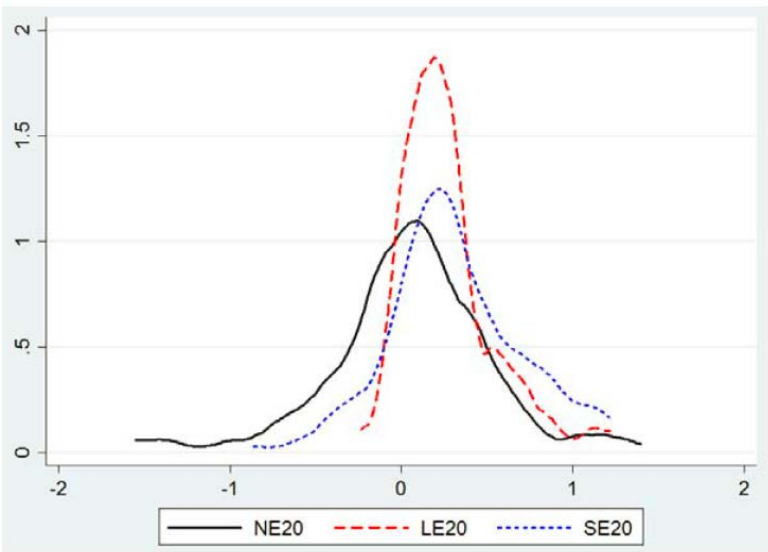
Notes: Column (3) instruments ExcessRR to isolate the component that can be explained by differences in liquidity management as a result of the Great Depression. Standard errors clustered at the district level are in parentheses. ***$p<0.01$, **$p<0.05$, *$p<0.1$

Figure 1:
Equilibrium Loan Rates in the Second Stage



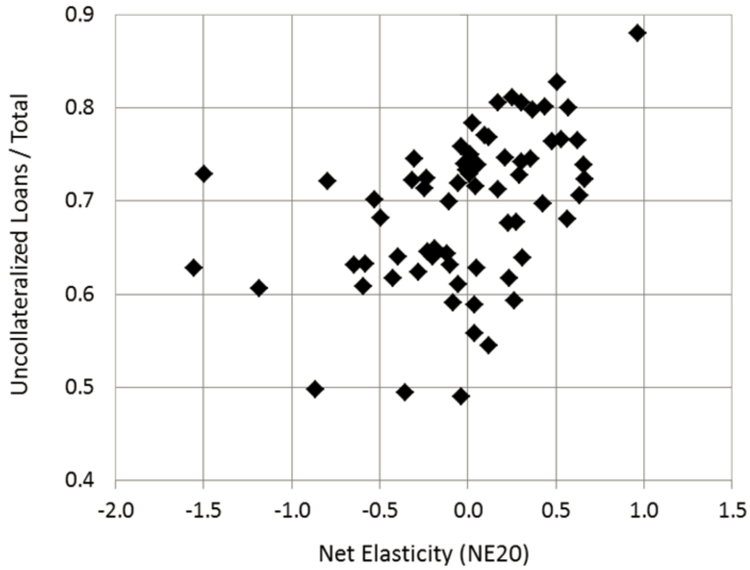

Notes: The policy rate is $r$ and firm types are denoted by $\omega$. The interest rate that makes $\omega$ indifferent towards risk-shifting is $\overline{R}(\omega)$. The information sets of insiders and outsiders are as defined in Section 3.1.

Figure 2:
Distribution of Elasticity Estimates in the 1920s



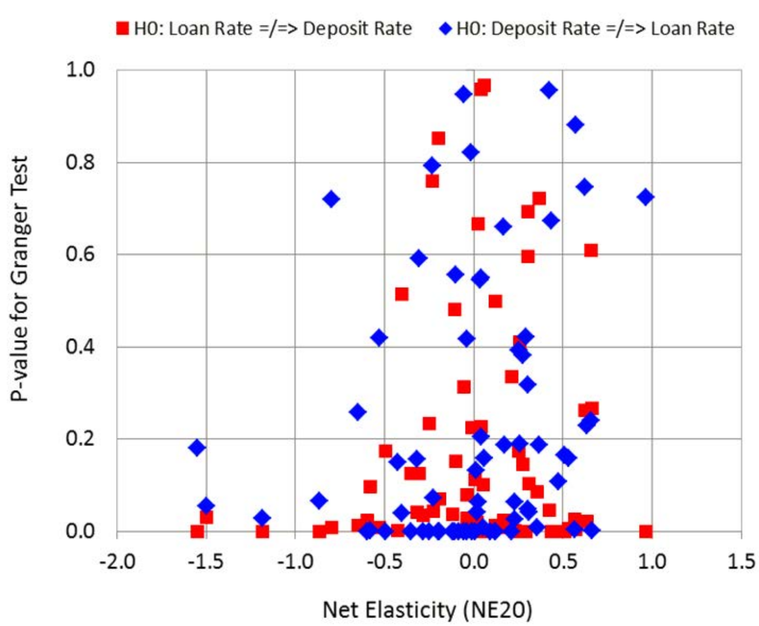

Notes: LE20 is the elasticity of loan returns ($\beta_i^\ell$), SE20 is the elasticity of securities returns ($\beta_i^s$), and NE20 is the net elasticity ($NE20_i = \beta_i^s \square \beta_i^\ell$).

Figure 3:
Net Elasticities by Location in the 1920s




Notes: Each point represents a location. Uncollateralized loans include loans collateralized by personal security.

Figure 4:
Granger Causality Test Results




Notes: Each point represents a location. Tests are based on semi-annual data from 1925 to 1929 inclusive.

# Appendix A

This appendix proves Proposition 1 in the main text.

The elasticities of the relationship lending interest rates are:

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

and:

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

while the elasticity of the transactional interest rate is:

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

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 $\xi$ and $\eta$ in Section 3.1 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):

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

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

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

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

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

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:

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

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

$$\widehat{\omega} = \frac{[p(0) \square q]\,p(\widetilde{\omega})\,\widetilde{\omega}}{q\,[p(\widetilde{\omega}) \square q] \square [p(1) \square p(0)]\,p(\widetilde{\omega})\,\widetilde{\omega}}$$

We can then rewrite (A.3) as:

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

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