ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

A Quantitative Analysis of Bank Lending Relationships

Authors	Kyle Dempsey, and Miguel Faria e Castro
Working Paper Number	2022-033E
Revision Date	March 2025
Citable Link	https://doi.org/10.20955/wp.2022.033
Suggested Citation	Dempsey, K., Faria e Castro, M., 2025; A Quantitative Analysis of Bank Lending Relationships, Federal Reserve Bank of St. Louis Working Paper 2022-033. URL https://doi.org/10.20955/wp.2022.033

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

A Quantitative Analysis of Bank Lending Relationships*

Kyle Dempsey
Ohio State University

Miguel Faria-e-Castro
FRB St. Louis

This version: March 6, 2025

Abstract

We study the aggregate consequences of dynamic lending relationships in a model of heterogeneous banks facing financial frictions. We estimate the model's loan demand system on administrative loan-level data: the market power implied by the estimated strength and persistence of relationships yields a long run reduction in credit of 5.9%. Relationships amplify the negative real effects of credit supply shocks, but mute those of negative credit demand shocks. In a financial crisis which destroys 25% of bank net worth, for example, loan volume drops more than twice as much in our baseline model than in a competitive analog with no relationships, but banks recapitalize faster.

JEL Classification: E44, G21

Keywords: Banking, lending relationships, aggregate dynamics

*This paper was previously circulated under the title "A Quantitative Theory of Relationship Lending." The views expressed here are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System. We thank Asha Bharadwaj for research assistance with the data, and Mike Carter and Jeongwoo Moon for research assistance with computation. We thank seminar participants at the Federal Reserve Board, the Federal Reserve Bank of Chicago, the Federal Reserve Bank of Philadelphia, the Federal Reserve Bank of Richmond, the Bank of Canada, the Bank of England, Queen Mary University of London, London Business School, Bank of Portugal, Nova SBE, as well as conference participants at the 2021 Society for Economic Dynamics, 2021 CEBRA, and the 2021 Econometric Society Winter Meetings for helpful comments. First version: February 2020. Contact: `dempsey.164@osu.edu` or `Miguel.FariaeCastro@stls.frb.org`.

1 Introduction

Banks operate in imperfectly competitive markets in their core business activities: deposit taking and loan making.[1] Market power generates economic profits for banks by enabling them to lend at interest rates above the fair (risk-adjusted) cost of capital and borrow at interest rates below the prevailing risk-free rate.[2] This paper studies the implications of one source of banks' lending market power: long-lasting lending relationships with borrowers. We develop a model in which these relationships introduce dynamic incentives into banks' lending and financing decisions. Specifically, banks internalize the fact that while charging higher interest rates may generate larger profits today, it may also erode relationships and thereby worsen lending prospects tomorrow. The central focus of this paper is to quantify the *aggregate* consequences of banks' lending relationships by showing how they interact with financial frictions at the individual bank and banking industry levels.

We study lending relationships in a dynamic equilibrium model with heterogeneous banks subject to financial constraints, as in Corbae and D'Erasmo (2021) or Bigio and Bianchi (2022). We model lending relationships in a parsimonious way using two key features. First, a borrower may borrow from many banks, but faces costs of adjusting the share of its total lending sourced from each bank. These adjustment costs endow lenders with market power. The adjustment is relative to "relationship capital," which summarizes borrowers' prior loan sourcing decisions in a manner akin to a "deep habit" (Ravn et al., 2006). In this setup, stronger relationships generate higher levels and lower price elasticities of loan demand for individual banks. Likewise, a borrower's total loan demand depends not only on loan interest rates (or a summary statistic of the economy-wide distribution of rates), but on the full joint distribution of rates and relationships across lenders. Second, whether a borrower's relationship with a given bank strengthens or weakens period-to-period depends on the bank's pricing decision. Thus, banks' market power is *dynamic*: banks extract rents commensurate with the inverse elasticity of loan demand (like static monopolists), but these rents are extracted over the life of the relationship.

For clarity, we note from the outset that we do *not* aim to provide a fundamental theory of the origins and evolution of borrower-lender relationships. Rather, we seek to provide a quantitatively tractable and empirically-disciplined framework capable of analyzing how the existence of such relationships shapes banks' behavior, and, in turn, the nature and dynamics of lending markets and the banking industry. In this context, our paper makes four main contributions.

First, to the best of our knowledge, our model is the first that can be used to evaluate how lending relationships shape banking-industry-level and aggregate outcomes in the presence of financial constraints. Importantly, our framework nests a competitive benchmark in which borrowers do not face loan portfolio adjustment costs. In this benchmark, there is no notion of relationships, banks take prices as given, and they choose loan volumes while making the same set of financing decisions

[1] See, for example, Berger and Hannan (1998). Banking industry concentration and market power is well documented both internationally (Fernández de Guevara et al., 2005) and in the U.S., which has experienced a stark secular decline in the number of banks over the last several decades (Prescott and Janicki, 2006).

[2] Such economic profits have long been considered a "feature, not a bug", as they generate franchise value that curbs risk-taking by banks, thereby promoting financial stability (Demsetz et al., 1996).

as in our baseline model. This nesting property allows us to detail quantitatively how the presence of relationships alter bank behavior.

Second, we show that two of the key parameters that govern lending relationships in the model can be directly estimated without solving the full model using loan-level micro-data. This involves combining the law of motion for relationships at the bank level with the individual firm demand for credit from that bank. We then adapt the method of Amiti and Weinstein (2018) to estimate the model-implied bank-level demand curve. This yields estimates of both the *static* interest rate elasticity of loan demand, which pins down the level of loan portfolio adjustment costs, and the *dynamic* impact of banks' market shares on how relationships and loan demand evolve. This leverages the fact that, through the lens of our model, relationship capital is the key non-price shifter of loan demand. The ability to estimate the demand system directly keeps our quantitative analysis tractable, and is therefore a key benefit of our approach to modeling relationships.

Third, we show that financial and relationship capital are complements. Fixing relationship capital, more financially constrained banks tend to charge *higher* interest rates and ration lending to increase profitability, which comes at the expense of relationship capital. Fixing financial capital, banks with weaker relationships tend to charge *lower* interest rates to boost lending and build relationships for the future, which comes at the expense of financial capital. Despite charging lower rates, though, these "relationship constrained" banks lend less in terms of quantities than banks with stronger relationships, as they face weaker demand. Therefore, banks with stronger relationships hold more financial capital to meet the stronger loan demand they face. These forces combine to deliver a strong positive correlation between financial and relationship capital (0.89), and smaller but positive correlations of spreads with both financial and relationship capital (0.35 and 0.59, respectively). Increasing banks' static or dynamic market power strengthens these correlations.

We perform two key analyses which simultaneously demonstrate the model's key mechanisms and validate its quantitative performance along untargeted dimensions. First, we show that our model matches the empirical profile of spreads over the life of a lending relationship. This establishes that banks in our model have realistically strong and persistent market power. In the data, on average, "switching loans" are priced below market in the first year of a new lending relationship, but then above market over the next year and a half. Our baseline model matches this pattern. Moreover, we show that this is a direct function of the estimated strength and persistence of relationships: decreasing the static demand elasticity causes banks to raise rates too quickly, while making relationships more persistent induces them to raise rates too high later in the relationship.

Next, we show that the distribution of banks' capital buffers in our baseline economy lines up well with the empirical distribution. This matters because relationships have an ex ante ambiguous effect on capital buffers. On one hand, banks can expend relationship capital to weather financial shocks by charging high interest rates, dampening any precautionary motives. On the other, the high profits associated with lending at higher rates increase banks' franchise values, strengthening precautionary motives. Our baseline model balances these forces in line with the data, while the alternatives we consider all err on the side of too-small capital buffers. This gap is driven by lower franchise values in a more competitive economy, while it is driven by the ease with which banks can increase loan rates in the face of financial constraints in less competitive economies.

Our fourth main contribution is to show that relationship lending plays an important role in determining how the economy responds to aggregate shocks. In response to a large aggregate financial shock in which all banks lose a fraction of their equity, lending relationships amplify the initial contraction in the credit market and slow the recovery. This is driven by a faster recapitalization in the banking sector, as banks exploit their relationships to rebuild financial capital. This amplification is significant: the drop in lending in response to a 25% negative shock to bank net worth in the relationship economy is over twice as large as the corresponding drop in the competitive benchmark.

We also study how relationship lending impacts the response to shocks to credit demand. In a competitive economy, banks react to a fall in credit demand by trimming their balance sheets, contracting both deposits and net worth. In the relationship economy, on the other hand, banks' additional incentive to maintain relationships leads them to reduce interest rates substantially. While this props up loan volume, it also results in a fall in net worth that is over twice as large as, and more persistent than, the one observed in the competitive economy.

Finally, we study how relationships affect the response to shocks to bank funding costs (e.g. from a monetary tightening). Similar to the financial shock, the presence of relationships helps mute the effects of the shock on loan volumes and the price of credit. The increase in the cost of deposits and the incentive to maintain loan volumes to prevent relationship deterioration induces banks to substitute retained earnings for deposits as a source of financing. Thus shocks to the cost of funding cause a persistent increase in aggregate bank capital in the relationship economy.

Importantly, in addition to comparing our baseline model to a competitive alternative for each aggregate shock, we also compare it to a "fixed relationship" economy in which the overall level and distribution of *static* market power is the same as in the baseline economy, but relationships are permanent. In this case, despite having market power, banks lack the dynamic incentives to maintain their relationships just as in the competitive economy: therefore, the real effects of each of these shocks for this version of the model more closely resembles the competitive case than the baseline relationship case. This highlights the importance of the dynamic nature of lending relationships.

The rest of the paper is structured as follows. The remainder of this section discusses our paper's context in the relevant literatures. Section 2 documents two key empirical observations which motivate our approach to modeling lending relationships. Section 3 presents our model environment. Section 4 describes how we take our model to the data. Sections 5 and 6 present the main results from our quantitative model, with the former focusing on the cross-section and the latter focusing on aggregate dynamics. Section 7 concludes and describes some promising areas for future research.

Related Literature This paper contributes to three distinct literatures in macroeconomics and finance: (i) customer capital in macroeconomic models; (ii) structural models of banking; and (iii) empirical studies of the effects of bank market power.

While the dynamics of customer capital can be related to an older literature on consumption habits, Gourio and Rudanko (2014) provide a seminal formalization of customer capital in a macroeconomic model which we adapt here. Focusing on nonfinancial firms, Gilchrist et al. (2017) argue that the interaction between customer capital dynamics and financial constraints helps explain inflation dynamics in the U.S. during the Great Recession. We model customer capital dynamics with heterogeneous agents similarly, but with a special focus on how customer capital interacts with financial constraints specific to the banking industry. This is critical to understanding recent recessions, since the aggregate capitalization of the banking sector has been argued to be a relevant state variable for macroeconomic performance (Adrian and Boyarchenko, 2012).

The heterogeneous banks in our model take deposits, make loans, and face constraints that depend on their net worth. We therefore contribute to an emerging literature that employs the tools of heterogeneous agent macroeconomic models to study the banking industry. Bigio and Bianchi (2022) incorporate liquidity frictions in the interbank market to study monetary policy implementation. In Corbae and D'Erasmo (2021), the correlation between bank size and market power affects the impacts of capital requirements. Taking the regulatory regime as given, we study how lending relationships affect the overall stability of the banking system. We use our model to study aggregate dynamics similarly to Neri et al. (2010), who introduce a monopolistically competitive banking sector in an otherwise standard monetary DSGE model to study the transmission of standard shocks. Wang et al. (2022) also develop a structural model of banking to study how market power affects the transmission of monetary policy shocks. Boualam (2018) studies how credit relationships are endogenously formed and persist in a setting with search and agency frictions.

Finally, our paper relates to a broader empirical literature that studies the efficiency and stability consequences of banking market power and concentration. Recent work on this topic has been focused on market power in the deposits market. Egan et al. (2017) use detailed branch-level data on deposit quantities and prices to estimate a demand system for deposits at large U.S. banks. They combine these estimates with a dynamic model of bank runs to study the probabilities of counterfactual runs on these large banks. Drechsler et al. (2017) argue that bank market power in the deposit markets gives rise to a new channel of transmission for monetary policy in the U.S. Our notion of customer capital in banking is also related to traditional views on bank franchise value, and the idea that regulatory barriers and market power generate enterprise value beyond the pure accounting value of bank assets and liabilities (Atkeson et al., 2019).

We focus on market power on the loan side of banks' balance sheets. Lending relationships are a form of bank customer capital, which gives individual banks a degree of market power. It has been widely documented that banks value long-lived relationships (Petersen and Rajan, 1994). For example, Berger and Udell (1995) find that banks smooth loan rates when faced with adverse shocks to funding costs, and that this helps them conserve relationships. Banerjee et al. (2021) find similar relationship dynamics using Italian credit registry data around the Lehman episode. There is also an extensive theoretical literature that derives conditions under which the optimal contract between a lender and a borrower shares some of those features under a variety of frictions, such as asymmetric information (Diamond, 1984; Darmouni, 2020) or search frictions and switching costs (Boualam, 2018; Payne, 2018). Our approach is closest to this latter strand: modeling relationships as deep habits can be interpreted as a reduced-form for switching costs. Darmouni (2020), in particular, finds that information frictions are unable to account for the stickiness of borrower-lender relationships in the U.S. syndicated loan market. We take the lending contract and the process for customer capital dynamics as given, and study their macroeconomic implications.

2 Empirical Motivation

We use loan-level micro data for the U.S. to document two facts regarding bank loans: (i) switching between banks is relatively infrequent; and (ii) there exists an interest rate life cycle for new lending relationships, featuring low interest rates in the beginning that rise over the length of the relationship.

2.1 Data

Our main source of data is the Commercial & Industrial loan schedule H.1 of the Federal Reserve's FR Y-14Q dataset (Y-14 for short). This is a quarterly panel of individual loan facilities held in the books of the largest bank holding companies (BHCs) in the US.[3] The Y-14 includes all loan facilities held in the books of covered BHCs with commitments larger than $1 million. It contains detailed information about the characteristics of each loan, such as the identity of the borrower, the type of loan, interest rate, purpose of loan, etc.

We restrict our loan sample along several dimensions. First, we exclude loans to non-US addresses, loans in currencies other than the US dollar, and loans to firms without a US Tax Identification Number (TIN, our main firm identifier). We also exclude loans to the financial and public administration sectors, that is, to any entity classified as a bank or with NAICS code 52 or 92.[4] Due to their different nature and imperfect coverage, we also drop syndicated loans.

2.2 Facts on US bank loan markets

2.2.1 Switching lenders is infrequent.

Figure 1 presents time series plots on the percentage of loans that correspond to "switches," as a percentage of total outstanding loans. Our definition of "switch" is adapted from Ioannidou and Ongena (2010): a loan is considered a switch if it is a new loan and originates from a bank with whom the firm has had no (observable) relationship in the past year. The time series plots show that in terms of both dollar value and loan counts, switches are between 2 and 3.5% of total loans. Thus switching is relatively infrequent.

It is worth noting that due to the characteristics of the Y-14 dataset, we are likely to be *overestimating* the frequency of switching. First, loan observations may enter and/or leave our panel for many reasons other than origination or maturity. A loan may have been originated with a committed

[3] Until 2019, the dataset includes all BHCs with more than $50 billion in assets. From 2019 onwards, only BHCs with more than $100 billion in assets are included.

[4] We also exclude loans made to companies with NAICS codes 5312 (Offices of Real Estate Agents and Brokers) or 551111 (Offices of Bank Holding Companies).

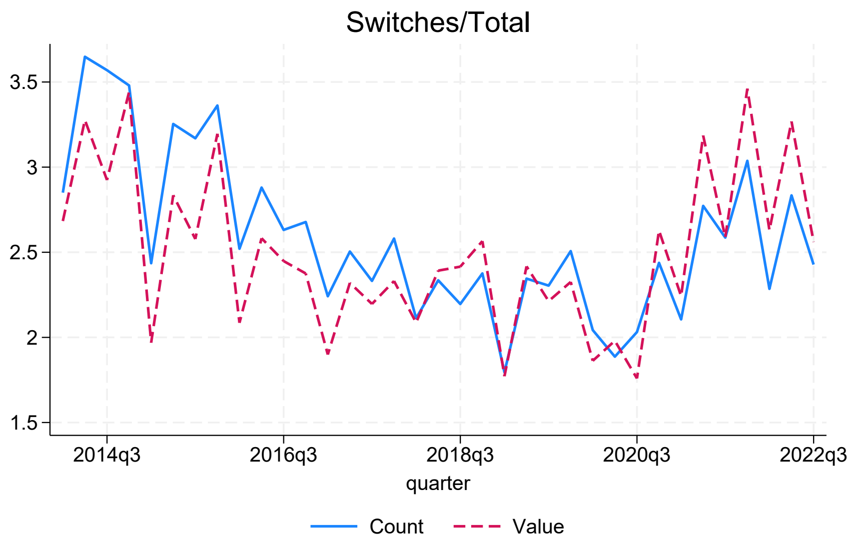

Figure 1: **Switches as a percentage of total outstanding loans**

Notes: See text for details. A loan is classified as a switch if it is (i) a new loan, and (ii) from a bank with which the firm has had no relationship in the past year.

exposure of under \$1 million, with a credit limit increase above \$1 million being renegotiated at a later date. In that case, we only observe the loan after the credit limit has increased. Additionally, banks may not keep originated loans in their portfolios, for example by selling them to other financial institutions. Second, since we only observe credit facilities above \$1 million dollars, we do not observe small firms that borrow lower amounts. It is well documented that large firms tend to have more relationships and switch more often than smaller ones (Petersen and Rajan, 1994). Among studies that use more comprehensive loan-level datasets, Ioannidou and Ongena (2010) find that 3% of all originations are classified as switching loans for Bolivia, while Farinha and Santos (2002) find that on average 4% of all yearly originations involve switching, using data for Portugal. Finally, it is worth pointing out that these facts do not seem to be driven by larger firms that borrow from multiple banks: Figure A.1 in the Appendix repeats the exercise for single-bank firms, and is qualitatively identical.

2.2.2 Interest rates first fall, then rise over the life cycle of a relationship.

We next investigate how interest rates evolve over the life cycle of firm-bank relationships. We follow an approach inspired by Ioannidou and Ongena (2010): we first identify loan originations that correspond to new relationships (again, defined as the absence of an observable relationship between the borrowing firm and lending bank in the previous year), then match those with loan originations from existing relationships that have similar observed characteristics. More specifically, "matched" loans are the same with respect to the following observable characteristics: (i) loan

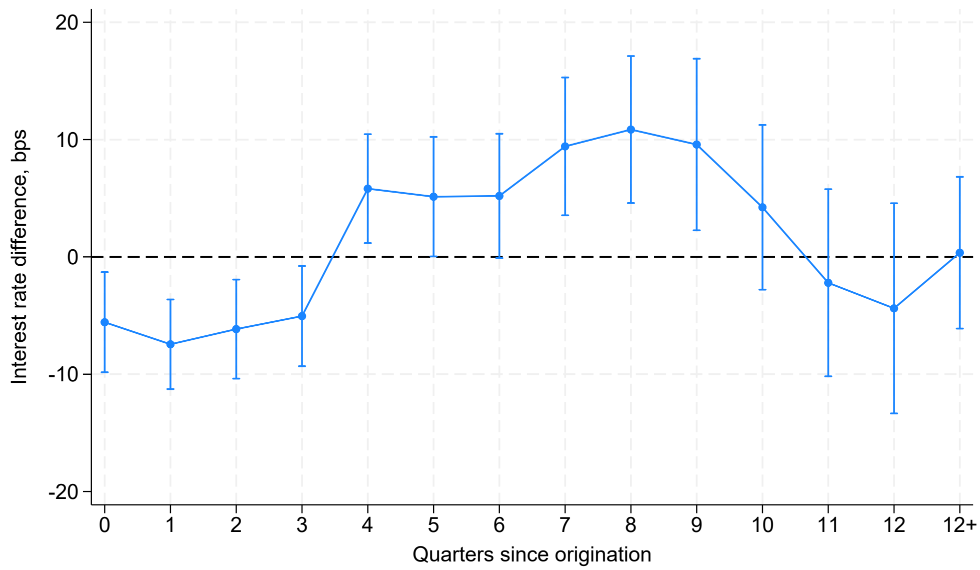

Figure 2: **Average spread between loans for new and existing relationships**

Notes: See text for details. At each time since loan origination, the dot represents the point estimate of γ_i from (1), and the bars represent the associated 90% confidence interval.

origination date; (ii) maturity (in years); (iii) originating bank; (iv) percentile of loan size; (v) loan type (term loan, credit line, or other); (vi) interest rate variability; and (vii) percentile of default probability. Since there are more non-switching loans than switching loans in our data, we match each unique non-switching loan with a similar switching loan, meaning that some switching loans may appear multiple times in the dataset. This procedure generates 20,155 matched loan pairs.

For each pair p and date t, we compute the spread between the switching and non-switching loans, $y_{p,t}$. We then run regressions of the following type:

$$y_{p,t} = \sum_{i=1}^{13} \gamma_i \mathbf{1}[\tau_{p,t} = i] + \epsilon_{p,t} \tag{1}$$

where $\tau_{p,t}$ is time since origination for the matched loan pair p at time t, measured in quarters. Time since origination can alternatively be interpreted as the length of the relationship for the switching loan (by construction, as a switching loan is such that no relationship existed beforehand). We interpret the estimated coefficients $\{\gamma_i\}$ as the average discount or premium that a switching loan obtains relative to a non-switching loan in period i of its life cycle.[5]

Figure 2 summarizes the estimation results for regression (1), depicting the estimated marginal effects along with 90% confidence intervals based on robust standard errors. The figure shows that the average spread between switching and non-switching loans is negative for most of the first year, meaning that switching loans pay lower interest rates on average. After one year, the spread

[5]We consider $i = 1, \ldots, 13$, where each i is a quarter since origination and $i = 13$ stands for all quarters after 12.

becomes positive, meaning that switching loans start paying interest rates that are, on average, higher than those of non-switchers. This positive spread persists for the next 6 quarters, becoming statistically indistinguishable from zero thereafter. This life cycle pattern is consistent with the one detected by Ioannidou and Ongena (2010) for Bolivia: they also find that switchers start out by paying lower interest rates than nonswitchers, but that these interest rates increase over time and eventually exceed those of non-switchers. The authors use these findings to discriminate between alternative theories of firm-bank relationships: these results suggest that relationships are influenced by switching costs that give rise to a hold-up problem. Thus the bank initially attracts the borrower using a "teaser rate," and then exploits the fact that switching is costly to extract surplus. The pattern is also in line with of Banerjee et al. (2021), who use Italian credit registry data to show that interest rates tend to increase with the length of the relationship for different types of loans.

It is worth pointing out that we find spreads of smaller magnitudes than Ioannidou and Ongena (2010). This can be explained by many factors, including lower average interest rates in the US versus Bolivia and the fact that the firms in our sample tend to be larger and safer than those in Bolivia. Both forces lead to compression in interest rates across firms. The $1 million dollar loan cutoff excludes most small enterprises from our sample. As with switching patterns, a very similar pattern emerges when focusing on single-bank firms only, shown in Figure A.2 in the Appendix.

Summary Together, these facts suggest that aversion to switching generates a specific form of market power for the lender which evolves over the lending relationship. When attracting a new borrower, the bank is at a disadvantage relative to any incumbent lenders the borrower has. Conditional on forming a relationship, though, the bank gradually develops the advantages of the incumbent, i.e. the ability to charge above-market rates. These facts motivate our parsimonious specification of lending relationships in the next section, which combines costs of adjusting how borrowers source their loans with banks' internalization of these costs in their pricing decisions.

3 A Model of Banks with Lending Relationships

We consider a stationary economy populated by a unit continuum of monopolistically competitive banks $j \in [0,1]$ and a continuum of identical firms $i \in [0,1]$ who borrow from them.[6] Time is discrete and infinite, and there is a single good. The risk-free rate is $\overline{r}$, which defines a risk-free discount price of $\overline{q} = (1+\overline{r})^{-1}$, the wage rate is $\overline{w}$, and the user cost of capital (rental rate) is $\overline{uc} = \overline{r} + \delta$, where δ is the depreciation rate. All these prices are exogenously specified.[7] While the model's focus is on bank behavior, described in Section 3.2, we first present the firms' problem in Section 3.1 since it delivers the demand system banks face and helps introduce notation. Section 3.3 defines the equilibrium. Finally, Section 3.4 discusses the motivation for and implications of the main assumptions in our framework. Proofs of all propositions are contained in Appendix B.

[6] All firms are identical and there are no idiosyncratic shocks to firms, so the model admits a representative firm.

[7] It is straightforward to embed our relationship framework into a general equilibrium model. Since our primary focus is on the banking industry, however, we simplify the model by keeping these prices exogenous.

3.1 Firms: defining bank-specific and aggregate loan demand

Firms operate a decreasing returns production technology using labor n and capital k, producing $y = Ak^{\alpha}n^{\eta}$ units of output for α, η, and $\alpha + \eta \in (0, 1)$ given total factor productivity A. Each period, the firm chooses: (i) how much labor and capital to hire; (ii) how much to borrow; and (iii) the *sourcing* of its borrowing across banks j. The firm is subject to a working capital constraint as in Christiano et al. (2005): total lending must be at least a fraction $\kappa \geq 0$ of its total costs, which include the wage bill and the costs of renting capital. The firm's total loan demand today is L', and the distribution of borrowing across banks is $\mathcal{L}' = \{\ell'_j\}$, where ℓ'_j is the face value of this period's loan from bank j. The discount price of a loan from bank j is q_j, and we denote the set of loan prices across banks by $\mathcal{Q} = \{q_j\}$.

We model lending relationships as follows. For each bank j, we summarize the intensity of a firm's relationship with that bank by s_j; the set of relationships across all banks is $\mathcal{S} = \{s_j\}$. We assume it is costly for a firm to source its loans in a way that deviates from the distribution of relationships. We implement this with a quadratic cost function with scale parameter $\phi \geq 0$, which penalizes deviations in the share of total lending sourced from bank j from the (relative) intensity of the firm's relationship with bank j. For tractability, we assume that borrowers take current relationships as given and do not internalize how current loan sourcing decisions affect future relationships, in the spirit of "external" habits in the literature (e.g. Ravn et al., 2006).

Under this formulation, a firm does not directly care about the "identity" j of any bank from which it borrows; rather, it cares only about the intensity of its relationship with the bank, s, and the loan price the bank offers, q. Therefore, the decision-relevant object which defines borrowing opportunities for the firm is the joint density of prices and relationships across banks, $\mu(q, s)$, which summarizes $\{\mathcal{Q}, \mathcal{S}\}$. The firm's dynamic optimization problem may then be written recursively as:

$$\begin{aligned} W(\mathcal{L};\mu) \quad = \quad & \max_{n,k,L',\{\ell'(q,s)\}} \underbrace{Ak^{\alpha}n^{\eta} - \overline{w}n - \overline{uc}k}_{\text{net operating income}} + \underbrace{L' - \int \ell(q,s)\mathrm{d}\mu(q,s)}_{\text{borrowing net of repayments}} \\ & - \underbrace{\frac{\phi}{2}L' \int \left(\frac{q\ell'(q,s)}{L'} - 1 - (s - S)\right)^2 \mathrm{d}\mu(q,s)}_{\text{sourcing adjustment costs}} + \underbrace{\overline{q}\mathbb{E}\left[W(\mathcal{L}';\mu)\right]}_{\text{continuation value}} \end{aligned} \tag{2}$$

$$\text{subject to [working capital]} \qquad \kappa(\overline{w}n + \overline{uc}k) \leq L' \tag{3}$$

$$\text{[loan sourcing]} \qquad L' \leq \int q\ell'(q,s)\mathrm{d}\mu(q,s) \tag{4}$$

The firm's flow profits in (2) sum net operating income $Ak^{\alpha}n^{\eta} - \overline{w}n - \overline{uc}k$ and net borrowing (new loans less repayments), less adjustment costs. $L'(\mu) \equiv \mathbb{E}_{\mu}[q\ell'(q,s)] = \int q\ell'(q,s)\mathrm{d}\mu(q,s)$, defined in the loan sourcing constraint (4), are total funds borrowed today, and $S(\mu) \equiv \mathbb{E}_{\mu}s = \int s\mathrm{d}\mu(q,s)$ is the average relationship intensity. The firm discounts the future at the risk-free rate and recognizes that, in a stationary equilibrium, the joint distribution of prices and relationship intensities is constant across periods, even if specific banks shift around in the distribution. Constraint (3) is the working capital constraint. Note again that the borrower does not take into account its choice of loan portfolio today on habits tomorrow, hence there is no explicit "law of motion" for s in this problem. This reflects the externality of habits: each individual firm is atomistic and does not internalize the impact of its actions on relationships.

Intuitively, these adjustment costs mean that, all else equal, firms would like to choose their borrowing shares at each bank in line with the relative intensity of their relationship with that bank (i.e., s relative to S), since this implies no adjustment costs. The quadratic functional form is not essential to our results, and it could be seen as a second-order approximation to a more general adjustment cost function that satisfies certain restrictions. We consider this more general case in Appendix B.3. The key advantage of the quadratic specification is that it gives rise to a linear demand system that is amenable to estimation. The following proposition summarizes the loan demand system that arises from the firm's problem:

Proposition 1. ***(Loan demand system)*** *Given a joint distribution of prices and relationship intensities $\mu(q, s)$, bank-specific loan-demand $\ell'(q, s)$ and aggregate loan demand L' satisfy*

$$\frac{q\ell'(q, s; \mu)}{L'(\mu)} = 1 - S + s - \frac{\overline{q}}{\phi}\left[r(q) - R(\mu)\right] \text{ for all } q, s \tag{5}$$

$$L'(\mu) = \kappa(\alpha + \eta)\left[\frac{A\left(\frac{\alpha}{\overline{uc}}\right)^{\alpha}\left(\frac{\eta}{\overline{w}}\right)^{\eta}}{1 + \kappa(\overline{q}\tilde{R}(\mu) - 1)}\right]^{\frac{1}{1-\alpha-\eta}} \tag{6}$$

where $r(q) = q^{-1}$ is the interest rate implied by the bank's loan price, S is the average relationship intensity, $R(\mu) = \mathbb{E}_{\mu}[r(q)]$ is the average interest rate, and $\tilde{R}(\mu)$ is the effective interest rate:

$$\tilde{R}(\mu) = R(\mu) + \mathbb{C}_{\mu}[r(q), s] - \frac{\overline{q}}{2\phi}\mathbb{V}_{\mu}[r(q)] \tag{7}$$

which adjusts the average interest rate for the covariance of interest rates and relationship intensities $\mathbb{C}_{\mu}(r, s)$, and the overall variance of interest rates $\mathbb{V}_{\mu}(r)$.

Equation (5) defines the demand curve faced by a bank with relationship intensity s charging price q as a function of aggregate loan demand, the average interest rate, and the average relationship intensity. The loan demand at a given bank is decreasing in the loan rate spread over the benchmark $r(q) - R(\mu)$, with elasticity governed by the risk-free rate and the adjustment cost. This is a standard price effect: when a given bank's loans are cheap relative to its competition, that bank will capture a higher share of total lending, all else equal. Steeper adjustment costs (higher ϕ) imply a lower elasticity of loan demand with respect to price. In addition, bank-level loan demand increases in the strength of the firm's relationship with that bank, measured by s. Thus, stronger existing lending relationships simultaneously increase the level and lower the price elasticity of loan demand, endowing these banks with more effective market power.

Equation (6) determines aggregate loan demand. Conveniently, the entire joint distribution of loan prices and relationship intensities may be summarized by a single statistic: the effective interest rate $\tilde{R}(\mu)$ from equation (7). This term has three components. First, the average interest rate term $R(\mu)$ conveys that when interest rates are higher on average, aggregate loan demand is lower. Second, loan demand is dampened further when the banks with whom the firm has the strongest relationships charge the highest spreads, as indicated by the covariance term $\mathbb{C}_\mu(r, s)$. Third, holding fixed the previous two terms, greater cross-sectional interest rate variance, $\mathbb{V}_\mu(r)$, burnishes loan demand by creating scope for the firm to gravitate towards cheaper banks.

3.2 Banks: dynamic pricing with relationships and financial constraints

Each bank uses retained earnings (its net worth), newly issued equity $e < 0$, and deposits $d' \geq 0$ (investment in riskless securities if $d' < 0$) to make loans ℓ' at discount price q. Deposits are risk-free (insured) and issued at exogenous price $\overline{q}^d$ for all banks. Banks value dividends $e \geq 0$ and face costs of issuing equity: we follow the dynamic corporate finance literature and model bank preference for dividends via the increasing function $\psi(e)$. The value of positive dividends is simply $\psi(e) = e$ when $e \geq 0$, but equity issuance (i.e. negative dividends) is costly, with $\psi'(e) > 1$ for $e < 0$. Banks' resources can be shifted by a persistent idiosyncratic shock z drawn from an AR(1) process $\Gamma(z, z')$ which proxies excess returns on unmodeled sectors of banks' portfolios. Finally, we assume that banks exit with exogenous iid probability $1 - \pi$ for $\pi \in [0, 1]$ each period. Exiting banks pay out their net worth as a dividend and are replaced with banks with no net worth and no relationships.

Banks face a regulatory capital constraint that specifies that the total value of equity must exceed a fraction χ of total lending. We assume that each bank is monopolistically competitive, setting its loan price while taking as given the bank-specific loan demand function (5), as well as the level of aggregate demand and (the key moments of) the distribution $\mu(q, s)$ described in Proposition 1. Crucially, individual banks internalize the impact of their lending choices today on their relationships tomorrow. We assume that relationships build up over time as a convex combination of the current relationship intensity (coefficient ρ_s) and the share of total loans issued today (coefficient ρ_q).

At the beginning of the period and after the realization of the exit shock, a bank's state is its net worth, n, relationship intensity, s, and idiosyncratic shock, z. A bank's recursive problem is:

$$V(n, s, z; \mu) = \max_{q, e, \ell' \geq 0, d', s', n'} \psi(e) + \overline{q}\mathbb{E}\left[(1 - \pi)\psi(n') + \pi V(n', s', z'; \mu)\right] \tag{8}$$

$$\text{subject to: [budget constraint]} \quad q\ell' + e \leq n + z + \overline{q}^d d' \tag{9}$$

$$\text{[capital requirement]} \quad \chi q\ell' \leq q\ell' - \overline{q}^d d' \tag{10}$$

$$\text{[relationship building]} \quad s' = \rho_q \frac{q\ell'}{L'(\mu)} + \rho_s s \tag{11}$$

$$\text{[market power]} \quad \ell' = \ell(q, s; \mu) \tag{12}$$

$$\text{[net worth accumulation]} \quad n' = \ell' - d' \tag{13}$$

The optimal policies are denoted $g_y(x)$ for $y \in \{q, e, \ell', d', s', n'\}$, where $x = (n, s, z)$.

The bank's objective function (8) reflects its present valuation of dividends net of issuance costs, discounted at factor $\overline{q}$.[8] Banks exit with probability $1-\pi$, at which point they pay out their net worth as a dividend. Constraint (9) is the bank's flow budget constraint: loans and dividends must be financed with either net worth (adjusted for the shock z) or deposits. Constraint (10) is the capital requirement, with equity defined as the difference between assets (loans) and liabilities (deposits). Equation (11) is the law of motion for the intensity of the firm's relationship with the bank, which the bank internalizes. Aggregating this relationship across banks and applying stationarity yields the result that $S = \frac{\pi\rho_q}{1-\pi\rho_s}$, and so the average relationship intensity is independent of policies. Equation (12) imposes the demand curve (5), and equation (13) shows how the bank's net worth evolves as a function of its lending and financing policies. We can establish the following result:

Proposition 2. ***(Optimal lending policies)*** *If $\psi(e)$ is twice continuously differentiable, banks' optimal loan prices satisfy the Euler equation*

$$\frac{\Pi_t + \overline{q}\rho_q \mathbb{E}_t\left[\sum_{i=0}^{\infty}\left(\overline{q}\pi(\rho_q+\rho_s)\right)^i \frac{L_{t+2+i}}{L_{t+1}}\Pi_{t+1+i}\right]}{\frac{\overline{q}}{q_t}\mathbb{E}\left[\psi^e(e_{t+1})\right]} = \epsilon^{-1}(q\ell, q) \tag{14}$$

where Π_t is the bank's net rate of return in period t per unit of loan and $\epsilon^{-1}(q\ell,q)$ is the inverse price elasticity of loan demand, given respectively by

$$\Pi_t = \frac{\overline{q}}{q_t}\mathbb{E}_t\left[\psi^e(e_{t+1})\right] - \psi'(e_t) + \lambda_t(1-\chi) \tag{15}$$

$$\epsilon^{-1}(q\ell, q) = \phi\frac{q}{\overline{q}}\frac{q\ell}{L'} \tag{16}$$

where $\lambda_t = \psi'(e_t) - \frac{\overline{q}}{\overline{q}^d}\pi\mathbb{E}_t\left[\psi^e(e_{t+1})\right] \geq 0$ is the Lagrange multiplier on the capital requirement and $\psi^e(e_{t+1}) \equiv (1-\pi)\psi'(n_{t+1}) + \pi\psi'(e_{t+1})$ is the expected marginal value of internal funds.

Equation (14) has an intuitive interpretation. The left-hand side represents the sum of the bank's discounted marginal net profits associated with increasing its loan price. The choice of loan price today affects not only today's profits (Π_t), but also profits in all future periods (summation term). The weight on this second term increases with the loading on current period lending in the law of motion for relationship intensity ρ_q, since this indicates a stronger dynamic pricing effect. The effective discount rate for future profits is $\overline{q}\pi(\rho_q+\rho_s)$: the first two terms reflect the equilibrium discount factor and the probability of bank survival, while the latter term reflects the *overall* persistence of relationships.[9] The profits in each period (15) reflect the return on loans, less their financing cost, plus the marginal benefit of easing the capital requirement.

This discounted profit stream in (14) must equal the inverse price elasticity of loan demand, $\epsilon^{-1}(q\ell', q)$, which measures the bank's effective market power. As shown in (16), this term is positive due to relationship adjustment costs ($\phi > 0$) and increases with the bank's relative loan share. Consider two extreme cases. First, when the bank's discount factor is zero, (14) implies static

[8]Note that for the individual bank problem, expectations are taken with respect to the idiosyncratic shock z; hence why we make explicit the expectation with respect to μ in the firm problem.

[9]If $\rho_q + \rho_s = 1$, then there is no depreciation in relationships and this term gets its maximal weight.

monopolist pricing: the markup equals the inverse elasticity of demand. The same holds if ρ_q is zero, i.e. if there is no dynamic effect of today's loan price on tomorrow's demand. Second, in the competitive limit ($\phi \to 0$), the price elasticity of loan demand becomes infinite, eliminating the term on the right-hand side of (14). Further, this case admits no notion of relationships, which eliminates the second term on the left-hand side: hence we recover the standard pricing condition $\Pi_t = 0$.

Evolution of bank distribution Given a current distribution of banks over states $m(x)$, the mass of banks next period with a particular x' is

$$\begin{aligned} m'(x';\mu) \quad = \quad & \pi \Bigg\{ \int \mathbf{1}\left[n(x') = g_\ell(x;\mu) - g_d(x;\mu), s(x') = \rho_q \frac{g_q(x;\mu) g_\ell(x;\mu)}{L'(\mu)} + \rho_s s(x) \right] \\ & \times \Gamma(z(x), z(x')) \mathrm{d}m(x;\mu) \Bigg\} + (1-\pi)\mathbf{1}\left[n(x') = 0, s(x') = 0 \right] \overline{\Gamma}(z(x')) \end{aligned} \tag{17}$$

The term in brackets in equation (17) describes state transitions for incumbent banks. For these banks, we require that next period's net worth and relationship intensity be consistent with the policies chosen this period, and that the evolution of the idiosyncratic shocks be consistent with Γ. The second term captures entrant banks, who begin with no net worth, no lending relationships, and idiosyncratic shocks drawn from the ergodic distribution $\overline{\Gamma}(z)$ implied by $\Gamma(z, z')$

3.3 Equilibrium definition

Definition 1. A **stationary recursive equilibrium** consists of: (i) bank-specific and aggregate loan demand functions, $\ell(q, s; \mu)$ for all (q, s) and $L(\mu)$; (ii) bank policy functions $g(n, s, z; \mu)$; (iii) a stationary joint distribution of prices and relationships $\mu(q, s)$; and (iv) a stationary joint distribution of banks over idiosyncratic states $m(n, s, z; \mu)$ which satisfy:

1. **borrower optimality**: bank-specific and aggregate loan demand satisfy (5) and (6);
2. **bank optimality**: banks' optimal policy functions solve the bank problem (8) – (13);
3. **stationarity of bank distribution**: the distribution of banks over idiosyncratic states is a fixed point of the operator defined in (17); and
4. **consistency of distributions**: the joint distribution of prices and relationships is consistent with the bank state distribution and banks' optimal policies:

$$\mu(q, s) = \int \mathbf{1}\left[q = g_q(n, s, z; \mu) \right] m(\mathrm{d}n, s, \mathrm{d}z; \mu) \text{ for all } q, s \tag{18}$$

3.4 Discussion of assumptions

Implementation of lending relationships. Two key elements of the structure of lending relationships in our model bear further comment. First, we assume the representative firm maintains relationships with all banks, and that costs come not only from relationship *formation*, but more generally from relationship *adjustment*.[10] This specification embodies two simple ideas: (i) all else equal, borrowers want to borrow more from banks with whom they have stronger relationships; and (ii) firm-bank relationships strengthen through exposure over time. Our specification of adjustment costs in (2) and the evolution of relationships (11) are exactly consistent with these assumptions. Exposures – and therefore relationships – shift through time for two reasons in our model. First, idiosyncratic shocks render some banks financially constrained, which leads them to charge different prices and lend different amounts than other banks with the same s. Second, the exogenous exit of banks and replacement with new banks yields a natural "life cycle" structure. As banks optimally respond to their financial conditions, they may either build up or expend relationships as a form of "customer capital," as in Gourio and Rudanko (2014). Our specification captures both extensive (it is costly to switch lenders) and intensive (this cost increases in the closeness of the relationship) margins. The former effect can arise, for example, from borrowers having to engage in costly search for lenders, which endows lenders with a type of "local market power" akin to what our model delivers. The intensive margin captures informational asymmetries which are alleviated through exposure and repeat interactions: the more a bank lends to a customer, the more information it acquires relative to other lenders, which can make it harder for the customer to switch.

Second, we assume that the firm does not internalize the formation of lending relationships, while banks do. The former assumption is made purely for tractability, as it is of course reasonable to expect borrowers to respond to developments in their banks' financial conditions by altering exposures to these banks. While possible in principle to allow for the firm to internalize relationship formation, it would require costly iteration between the firm and bank problems in the solution algorithm. By contrast, the demand system in the current framework allows us to solve the model with sole focus on the heterogeneous bank block. The fact that banks do internalize relationship formation shapes their optimal pricing policies, as highlighted in Proposition 2. As will be shown in the quantitative analysis below, this has important implications for how banks respond to financial shocks both at the individual level and in the aggregate.

Specification of relationship adjustment costs. We assume quadratic adjustment costs in loan shares in our baseline model. This specification is attractive for two primary reasons. First, it delivers a simple closed form for bank-specific loan demand (5). Not only does this facilitate computation (see Appendix C), but – more importantly – it also yields a simple structural equation which we can map to the data in order to obtain an empirical estimate of the critical relationship parameters ϕ and ρ_q (see Section 4.2.2). Second, it delivers a single sufficient statistic – the effective interest rate $\tilde{R}(\mu)$ from equation (7) – which summarizes the key economic forces driving *aggregate* outcomes in the model. As we show in Appendix B.3, though, the same central economic forces still hold under a more general specification of adjustment costs.

Macroeconomic models of customer capital typically feature constant elasticity of substitution

[10] Of course, relationship formation is also costly in our model; the specification of adjustment costs in (2) implies that the firm incurs costs for borrowing any positive amount from a bank with no relationships ($s = 0$).

(CES) preferences that feature the level of relationships or customer capital as a preference shifter within the CES aggregator, e.g. Gilchrist et al. (2017). While feasible, this specification raises some issues in our framework. First, what is being aggregated are dollars rather than utility over consumption of goods and services. Second, the CES with customer capital as a preference shifter still features a constant price-elasticity of demand, which does not vary with the intensity of the relationships. We derive the demand system under CES preferences in Appendix B.4.

One way to address the second concern is to aggregate loans across banks using the more general Kimball (1995) aggregator, which allows for a price-elasticity of demand that varies both with price and relationship intensity. This approach is still subject to the conceptual issue of aggregation of dollar loan values. Additionally, as we show in Appendix B.5, the main drawback of this specification is that the resulting bank-specific demand is no longer linear or log-linear and therefore not amenable to direct estimation, which is one of the main advantages of our framework.

Credit risk. In our model, all loans are risk-less. We abstract from borrower credit risk for two main reasons. First, most firms in the sample that we use to calibrate the model have very low default risk (the median 1-year probability of default in our sample is of 0.73%). Second, default risk would complicate the model substantially by making it harder to aggregate outcomes for the borrower across banks. This assumption is not innocuous, as default risk could significantly affect banks' pricing decisions, interacting with their own state-dependent discount factors that arise from equity issuance costs. In particular, it has been shown that ongoing relationships between banks and firms may distort pricing incentives and generate instances of overlending or insurance provision by the bank to the firm (see, for example, Faria-e-Castro et al., 2024). To account for the fact that the model does not feature credit risk, all of our estimation exercises either include explicit controls for default risk, or factors that subsume this risk (such as firm-time fixed effects).

Customer capital in bank liabilities. Our model also abstracts from broader definitions of bank relationships, particularly its accumulation on the liability side of the balance sheet through deposit relationships. For example, Drechsler et al. (2017) argue that imperfect competition in deposit markets is a key factor that modulates the transmission of monetary policy. Polo (2021) expands on this idea and develops a quantitative macroeconomic model where banks accumulate customer capital in deposit markets, showing that this amplifies monetary policy shocks. An interesting extension of our model would feature customer capital accumulation on both sides of the balance sheet, and how the two relate to each other (i.e., whether they are substitutes or complements).

4 Mapping the Model to the Data

We parameterize our model in three steps. First, we assign values externally (i.e. outside the solution of the model) to standard parameters in the macroeconomics and banking literature. Second, we directly estimate two of our model's unique relationship lending parameters – the adjustment cost ϕ and the persistence ρ_q parameters – from the micro-data using a semi-structural approach. Third, we jointly estimate the remaining parameters so that the model's stationary equilibrium matches a series of relevant banking industry moments. We describe each of these steps in turn. The full parameterization of the model is summarized in Table 2.

4.1 Externally set parameters

We set nine parameters externally. The risk-free quarterly discount price $\overline{q}$ implies an annualized risk-free rate of $\overline{r}_{\text{ann}} = 2\%$, in line with recent macroeconomic data. We set the interest rate on deposits $\overline{q}^d$ to be consistent with this risk-free rate and an annualized liquidity premium of 17 bps (van Binsbergen et al., 2022). The capital requirement is $\chi = 8\%$, in line with current capital requirements for large bank holding companies in the US. Since all exit is exogenous in the model, we set the bank exit rate equal to the historical average quarterly bank exit frequency, $1-\pi = 0.72\%$. We set total returns to scale for the firm to be consistent with a profit share of 5%, a capital share of 0.4, and a labor share of 0.6. The user cost of capital is set to be consistent with an annual interest rate of 2% and depreciation rate of 7%. Finally, we normalize the wage rate $\overline{w}$ to imply a marginal factor cost of one and the steady state level of aggregate TFP $\overline{A} = 1$.

4.2 Directly estimated parameters

Our model features three parameters that are not standard in models of banking and financial frictions: the cost of adjusting relationships, ϕ, and the parameters governing the persistence of relationships at the bank-level, ρ_q and ρ_s. We directly estimate the first two on micro data exploiting the relevant model demand equations, and internally calibrate ρ_s in the next step. In particular, we use loan-level data from the Federal Reserve's FR Y-14Q data to estimate the equation for bank-specific loan demand in (5) with an instrumental variables approach and obtain estimates for ϕ and ρ_q. We now describe the data and the estimation procedure in detail.

4.2.1 Data and sample selection

We use the Y-14 loan-level dataset described in Section 2 to construct a "relationship panel" at the firm-BHC-quarter level, where the quantity of credit $\ell_{f,b,t}$ is the total value of loans outstanding of firm f owed to BHC b at quarter t, and the interest rate $r_{f,b,t}$ is the average rate on those loans, weighted by utilized loan value.[11] After all sample restrictions, our final panel runs from 2013Q1 to 2022Q2 and includes 3.365 million observations, for 242,623 distinct firms and 41 distinct BHCs.

4.2.2 Estimating parameters using bank-level loan demand

To estimate ϕ and ρ_q, we exploit the fact that these parameters appear in both the bank-specific demand curve (5) and the law of motion for relationship intensity (11). Given data on loan quantities

[11] We also impose some additional restrictions that are aimed at eliminating observations that are likely errors: we drop all loans with interest rates less or equal than zero or above 50%, as well as loans for which the size of the commitment is non-positive, or the utilized amount is larger than the commitment.

and interest rates, as well as controls for unobservable bank and firm characteristics that could influence these relations, we can substitute for $s_{f,b,t}$ in (11) using (5) to estimate these two parameters using linear regression in a single step:

$$\frac{\ell_{f,b,t}}{L_{f,t}} = 1 - S_{f,t} - \frac{\bar{q}}{\phi}(r_{f,b,t} - r_{f,t}) + \rho_q \frac{\ell_{f,b,t-1}}{L_{f,t-1}} + \rho_s s_{f,b,t-1} \tag{19}$$

where $L_{f,t} \equiv \sum_b \ell_{f,b,t}$ is total borrowing by a particular firm across all banks, and $r_{f,t} \equiv \sum_b \frac{\ell_{f,b,t}}{L_{f,t}} r_{f,b,t}$ is the balance-weighted average interest rate paid by a particular firm across all banks. The goal is to regress the loan share of each bank within a given firm on: (i) the spread between the interest rate charged by the bank to that firm and the average rate paid by the firm, which should identify ϕ; and (ii) the lagged loan share of each bank within the firm, which should identify ρ_q.

There are three main challenges to directly estimating (19): (i) unobserved heterogeneity, (ii) the presence of terms such as $S_{f,t}, s_{f,b,t-1}$, which are not directly measurable in the data, and (iii) endogeneity. We now explain how we address each of these problems in order.

First, there could be unobserved heterogeneity among lenders and borrowers that could affect the demand of a firm for credit from a particular bank. To address this, we include firm-time fixed effects $\alpha_{f,t}$, which help control for fluctuations in overall credit demand by the firm that are unrelated to the characteristics of its relationship with different banks. We also include bank fixed effects α_b to control for time-invariant bank-level characteristics such as different business models or degrees or risk tolerance. Additionally, we include several time-varying bank-level controls $X_{b,t}$ that help control for changes in bank characteristics over time.[12]

Second, equation (19) contains extra terms that are not directly observable in the data: $S_{f,t}$ and $s_{f,b,t-1}$. $S_{f,t}$ is a time-varying firm characteristic and should therefore be subsumed by the firm-time fixed effects. There is no simple way, however, to account for $s_{b,f,t-1}$. We control for this variable by postulating that there should be a positive association between relationship intensity and the length of the relationship between a bank and a firm, which is observable in our dataset and that we include as a control as $\tau_{f,b,t}$.[13]

Third, equation (19) is a demand curve, and thus OLS estimates suffer from classical simultaneity bias. We address this issue by constructing an instrument for bank-specific credit supply shocks following Amiti and Weinstein (2018). Specifically, we first estimate the following regression

$$r_{f,b,t} - r_{f,t} = \gamma_{f,t} + \gamma_{b,t} + v_{f,b,t} \tag{20}$$

where $\gamma_{f,t}$ and $\gamma_{b,t}$ are firm-time and bank-time fixed effects. The idea is that the firm-time fixed effect controls for any factor that is related to the demand for credit, while the bank-time fixed

[12]We include bank size, measured as the log of total assets, bank leverage (liabilities over assets), a measure of asset liquidity (liquid assets over assets), the loan ratio (total loans to assets) and the deposit ratio (total deposits to liabilities). All these variables are measured from the FR Y-9C.

[13]We have experimented with different measures of relationship length: "continuous relationship length", which resets with each observable break, or "maximum relationship length", which counts the time since the earliest start of a relationship in the sample. Both approaches yield similar results.

	(1)	(2)	(3)	(4)
spread, $r_{f,b,t} - r_{f,t}$	-12.898***	-19.391***	-7.850***	-9.906**
	(1.574)	(2.664)	(0.785)	(3.991)
lagged loan share, $\ell_{f,b,t-1}/L_{f,t-1}$	0.622***	0.568***	0.563***	0.528***
	(0.008)	(0.009)	(0.009)	(0.010)
Firm identifier	TIN	TIN	ISL cell	ISL cell
Observations	74,121	60,332	259,972	229,764
Firm-Quarter FE	✓	✓	✓	✓
Bank FE	✓	✓	✓	✓
Bank & Relationship controls	✓	✓	✓	✓
Model	OLS	IV	OLS	IV

Standard errors in parentheses, clustered at the BHC level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1: **Estimating ϕ and ρ_q using firm- and cell-level data**

effect absorbs all variation that is related to the supply of credit, and is by construction orthogonal to demand. We therefore use $\hat{\gamma}_{b,t}$ as a valid instrument for the credit spread $r_{f,b,t} - r_{f,t}$. The final form of the equation (19) that we estimate is:

$$\frac{\ell_{f,b,t}}{L_{f,t}} = \beta(r_{f,b,t} - r_{f,t}) + \rho_q \frac{\ell_{f,b,t-1}}{L_{f,t-1}} + \alpha_{f,t} + \alpha_b + \gamma X_{b,t} + \zeta \tau_{f,b,t} + u_{f,b,t} \tag{21}$$

Estimation results are reported in Table 1. The relatively low number of observations (compared to the size of the full sample) is due to two factors: first, our identification strategy relies on multi-bank firms, that is, firms that borrow from multiple banks, while the vast majority of firms in our data borrow from one bank only. We elaborate on and address this issue below. Second, we estimate (21) only on loans originated in the last 4 quarters. The model features one-period debt, and so the firm can effectively adjust its demand for debt across banks every period. In reality, firms borrow at many different maturities, and it is not clear that a firm will find it advantageous or even feasible to constantly prepay debt that was contracted in the past. Older loans are likely to be priced at rates that no longer reflect aggregate credit market conditions, and so including them could bias our estimates in the direction of estimating a lower credit demand elasticity.

The first column reports the simple OLS results, while column (2) reports the estimation results using the credit supply shock instrument. Both specifications include the full set of fixed effects. In the instrumental variables specification the point estimate $\hat{\beta} = -19.391$ implies $\hat{\phi} = 0.051$. The implied $\hat{\rho}_q$ can be read directly from the second line and is estimated to be equal to 0.568.

One issue with our estimation method that also applies more broadly to the identification approach of Amiti and Weinstein (2018) is that it relies on firms that borrow from multiple banks in order to isolate demand from supply effects. It is well documented across space and time that the vast majority of firms borrow from a single lender; in our sample over 80% of firms maintain a relationship with a single lender. This means that our reliance on multi-lender firms for identification precludes the use of the majority of our data. Degryse et al. (2019) address this issue by defining borrowers at the industry-size-location level, instead of at the firm level. The identification assumption is that the demand for credit should be relatively stable among firms of the same industry (I), size (S), and location (L). We apply their methodology and define "ISL cells" where industry is the 3-digit NAICS, location is the CBSA of the borrower's address, and size is the borrower's decile in terms of total assets. This generates a total of 82,685 unique cells in our sample.

The results for this alternative estimation procedure are reported in columns (3) and (4) of Table 1. This sample – now almost four times larger – yields a smaller coefficient (in absolute value): the IV estimate for the slope parameter is $\hat{\beta} = -9.906$, which yields a larger estimate for the elasticity, $\hat{\phi} = 0.100$. The estimated $\hat{\rho}_q = 0.528$ is very similar. To balance the comparative advantages of these approaches, in our quantitative model we use the average of the estimates for the TIN and ISL versions of the IV regressions in columns (2) and (4).

4.3 Jointly estimated

4.3.1 Parameters and targets

The remaining parameters are jointly estimated so that the model matches a series of targets from the data, given the externally set and directly estimated parameters. These parameters and target moments are summarized in Table 2.C. The working capital parameter κ determines the level of overall loan demand given output. Therefore, this parameter is informed by the level of business debt relative to total output, which averages 71.5% for the U.S. economy. This parameter also determines the intensity of loan demand, which shapes how much market power banks have.

Given π and ρ_q, the relationship persistence parameter ρ_s determines the "baseline" loan demand $1 - S = \frac{1-\pi(\rho_s+\rho_q)}{1-\pi\rho_s}$ banks face (i.e. the intercept of the demand curve at a net interest rate of 0 for a bank with no relationships). Intuitively, the lower is ρ_s, the higher is $1 - S$, which means that banks face a higher baseline level of loan demand at a given price: that is, they have greater effective market power. Therefore, to discipline this parameter, we target banks' average net interest margin. Because our model has no default, and risk premia are an important component of loan spreads and therefore of net interest margins, we target a "no-default" net interest margin which filters out default risk premia from loan spreads. To construct this measure, we use Y-14 data to regress interest rates on new loans (originated in the last 4 quarters) on the originating bank's reported estimate for the 1-year probability of default. We then sum the constant and the residual, and call this the "zero-default interest rate." We compute the average zero-default interest rate for each bank in our sample, weighted by loan size, and subtract the average interest expense on deposits computed from the Call Reports. This measure averages 1.8% in our sample period, and our estimated ρ_s is consistent with $1 - S = 0.056$.

We assume that the shocks to bank net worth follow an AR(1) process with mean $\overline{z} = 0$, persistence ρ_z, and standard deviation of innovations σ_z.[14] We model convex costs of issuing equity

[14] That is, we assume $z' = \rho_z z + (1-\rho_z)\overline{z} + \varepsilon_z$, where $\varepsilon_z \sim \mathcal{N}(0, \sigma_z)$. This shock process is discretized over a grid

	Description	Value	Target	Data	Model
Panel A: Externally Assigned Parameters					
$\overline{r}_{\text{ann}}$	Annualized risk-free rate	2%	Quarterly discount price $\overline{q} = (1+\overline{r}_{\text{ann}})^{-\frac{1}{4}}$		
ν_{ann}	Deposit liquidity premium	0.17%	Quarterly deposit price $\overline{q}^d = (1+\overline{r}_{\text{ann}} - \nu_{\text{ann}})^{-\frac{1}{4}}$		
χ	Capital requirement	8%	Current US bank regulation		
π	Bank survival rate	0.9928	Quarterly bank exit rate of 0.72%		
α	Capital share	0.38	Profit share of 5%, capital share of 0.4		
η	Labor share	0.57	Profit share of 5%, labor share of 0.6		
$\overline{uc}$	Ann. user cost of capital	9%	2% interest plus 7% depreciation rate		
$\overline{w}$	Wage rate	3.78	Normalize factor costs to 1		
$\overline{A}$	Aggregate TFP	1	Normalization		
Panel B: Directly Estimated Parameters					
ϕ	Lending share adj. costs	0.068	Average of estimates, Section 4.2.2		
ρ_q	Mkt. share impact on rels.	0.548	"		
Panel C: Internally Calibrated Parameters					
κ	Working capital constraint	0.755	Business debt to GDP ratio	71.5%	71.5%
ρ_s	Persistence of relationships	0.427	Average net interest margin	1.8%	1.3%
$\overline{\psi}$	Marginal equity issuance cost	0.750	Gross equity issuance / NW	1.1%	1.9%
ρ_z	Persistence of net worth shocks	0.450	Net dividend payouts / NW	5.8%	1.1%
σ_z	Std. dev. of net worth shocks	0.006	Average bank leverage	87.7%	87.4%

Table 2: **Summary of calibration**

Notes: Firm leverage and business debt to GDP are sourced from the Flow of Funds. The leverage moment corresponds to corporate firms. Gross equity issuance and net dividend payout rates are computed following Baron (2020). The net interest margin is computed using Y-14Q interest rate on new loans (originated in the last four quarters), residualized from firm 1-year probability default, and deposit expense data from the Call Reports. All moments are averaged between 2009Q1 and 2020Q3.

by using a piece-wise linear cost function in which $\psi'(e) = 1$ for $e \geq 0$ and $\psi'(e) = 1 + \overline{\psi} > 1$ for $e < 0$. This specification is commonly used in the dynamic corporate finance literature (Hennessy and Whited, 2005). Further, it allows us to prove that banks' optimal financing policies have the property that if the capital requirement is slack, then the bank does not issue equity, which is a useful result computationally.

The parameters ρ_z, σ_z, and $\overline{\psi}$, then, are closely related to the financing choices banks make, and so we discipline them with moments of the data describing these choices. Given the costs of issuing equity and the relative cheapness of deposits, banks generally prefer to finance using deposits, and so our model closely matches the high average leverage of 87.7% we observe in the banking sector. It is worth noting, however, that the capital requirement is not always binding in our model, and so we obtain a distribution of capital buffers as in the data (Corbae and D'Erasmo, 2021). Beyond deposits, banks can respond to financial shocks by retaining earnings, cutting dividends, or issuing

of size $N_z = 11$ using the Adda-Cooper method.

new equity. We ensure realistic behavior along this dimension by targeting the average gross equity issuance and net dividend payout rates of 1.1% and 5.8% (Baron, 2020).

4.3.2 Solving the model

Internally calibrating the five parameters above requires iteratively solving the model over the parameter space. Since computing our model requires several non-standard steps, we describe our algorithm at a high level before proceeding. Appendix C contains the detailed algorithm.

The main complication in solving for a stationary equilibrium is that equilibrium is described not by a small vector of aggregate prices, but by the entire joint distribution of prices and relationships. However, given guesses of the distribution of banks over idiosyncratic states, $m(x)$, and bank pricing policy functions, $g_q(x)$, we can use the consistency condition (18) to infer the implied joint distribution of prices and relationships $\mu(q, s)$. Given this distribution, we can compute the demand-relevant summary statistics $R(\mu)$ and $\tilde{R}(\mu)$, which are the necessary inputs to bank-specific and aggregate loan demand according to equations (5) and (6). Finally, we can use these implied demand curves to solve for updates of banks' optimal policies, which in turn deliver an implied update to the initial guess of the distribution of banks. This procedure can be repeated until convergence on both policy functions and the distribution in order to obtain a stationary equilibrium.

5 Model Mechanics and the Role of Relationships

In this section, we use the steady state of our baseline economy and several variants to explain the key mechanisms in our model of relationship lending. This provides the underpinnings for understanding how relationship lending alters aggregate dynamics, which is the focus of the next section.

Throughout our quantitative analysis, we focus on two main versions of the model: (i) the **baseline**, whose calibration was described in the previous section; and (ii) a **competitive** version of the model where banks take market interest rates as given and choose how much to lend. Details of this second economy are presented in Appendix B.6. In this model, the lack of adjustment costs in the borrower's problem removes any meaningful notion of relationships. This implies a single equilibrium lending rate is taken as given by all banks. Banks then choose ℓ' directly, and a bank's state is fully described by (n, z). The competitive version of the model features banks with no market power since they face an infinite price elasticity of loan demand at each date.

We also report results for three other variants: (iii) a **low elasticity** version, where adjustment costs ϕ are higher than in the baseline, giving banks have more effective market power; (iv) a **low punishment** version in which $\rho_q \to 0$ so that banks face the same static loan demand elasticity but do not sever relationships as easily as in the baseline model; and (v) a **fixed relationship** version in which the distribution of s is taken from the baseline economy, but we assume that s is a permanent type, removing the dynamic incentives to maintain relationships.[15]

[15] The low elasticity model has ϕ 60% above the baseline. The low punishment model has $\hat{\rho}_q$ 80% lower than the baseline. When ρ_q is changed to $\hat{\rho}_q$, ρ_s is also changed to $\hat{\rho}_s$ so that S has the same value as in the baseline; that

		baseline (i)	comp. (ii)	low elas. (iii)	low pun. (iv)	fixed rel. (v)
Panel A: pricing and lending						
effective interest rate (pp, ann.)	$\tilde{R}(\mu)$	3.65	1.99	4.71	4.39	3.66
= average interest rate	$R(\mu)$	3.55	2.03	4.54	3.75	3.64
+ covariance term	$\mathbb{C}_\mu(r,s)$	0.10	-	0.18	0.68	0.04
+ variance term	$\mathbb{V}_\mu(r)$	-0.01	-	-0.02	-0.05	-0.02
loan volume	$L'(\mu)$	0.68	0.72	0.65	0.66	0.67
loan-weighted average interest rate		3.64	1.99	4.70	4.36	3.63
Panel B: banking industry moments						
average net worth		0.090	0.094	0.079	0.078	0.104
coefficient of variation, net worth		0.33	0.79	0.28	0.41	0.29
coefficient of variation, relationships		0.24	-	0.22	0.48	0.24
correlation, net worth and relationships		0.89	-	0.89	0.76	-0.05
correlation, net worth and spread		0.35	-	0.46	0.48	-0.59
correlation, relationships and spread		0.59	-	0.70	0.90	0.66
Panel C: surplus metrics (pp change relative to baseline)						
bank value		-	-71.4	36.7	57.8	-5.2
firm value		-	5.8	-3.6	-2.6	0.0
total value		-	2.8	-2.0	-0.2	-0.2

Table 3: **Cross-sectional and aggregate results across model variants**

Notes: In Panel A, all pricing moments are expressed in annualized net percentage points. In Panel B, all net worth objects are computed using total beginning-of-period net worth, $n + z$. In Panel C, we report the percentage point difference in the indicated welfare metric relative to the baseline model.

5.1 How do lending relationships shape industry-level and aggregate outcomes?

5.1.1 Loan rates decrease as competition increases.

Table 3.A presents statistics on interest rates and loan quantities, confirming basic insights about competition in our model environment. The effective interest rate $\tilde{R}$ varies sharply with the degree of competition, dropping 45.5% in the competitive model relative to the baseline. This lower effective interest rate raises loan volumes by 5.9%. Raising banks' static market power in the low elasticity economy increases the effective interest rate by 29.0%. When banks' static market power remains unchanged but becomes less sensitive to pricing decisions in the low punishment economy, effective interest rates increase by 20.3%. Finally, we recalibrate ϕ in the fixed relationship economy such that the equilibrium effective interest rate equals that from the baseline model: this facilitates

is, $\hat{\rho}_s = \frac{S-\pi\hat{\rho}_q}{\pi S}$, and reducing ρ_q implies increasing ρ_s. In the fixed relationship model, we recalibrate ϕ to match the average net interest margin in the baseline model. This is necessary because the removal of the dynamic incentives effectively increases banks' market power, and so ϕ must be reduced for comparability. See Appendix B.7 for details.

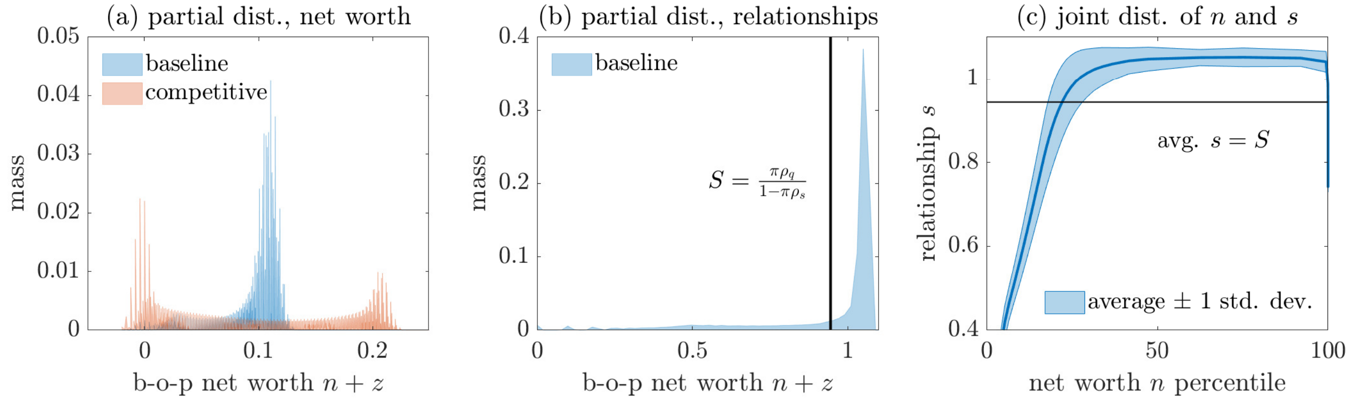

Figure 3: **Distributions of the endogenous state variables across models**

Notes: Panels (a) and (b) plot the partial distributions for total beginning-of-period net worth $(n + z)$ and relationships (s), respectively. Average net worth for each model is presented in Table 3. The average relationship intensity, represented by the black line in panel (b), is common across all our relationship models by construction. Panel (c) plots the average relationship intensity with a one standard deviation band conditional on each level of net worth in the baseline model.

apples-to-apples comparisons of aggregate dynamics in Section 6 below.

The next three rows decompose these differences using the three components of $\tilde{R}$ from (7). The bulk of the difference stems from higher average rates: banks exercise their market power. The positive covariance between relationships and rates behaves similarly, but with a smaller magnitude: banks with stronger relationships can charge higher rates and have the same loan volume (naturally, this effect is absent in the perfectly competitive economy, which has no notion of relationships). The covariance term is notably large (68 bps, or 15.5% of the total effective interest rate) for the low punishment model: in this case, not only are banks with stronger relationships inclined to charge higher rates, but the greater persistence in relationships induces a more unequal bank distribution. Finally, there is a small attenuation effect arising from interest rate dispersion: greater variance in loan rates provides more scope for borrowers to substitute into cheaper borrowing. This effect, however, is quantitatively small across model specifications.

5.1.2 Relationships compress the distribution of bank net worth.

Table 3.B focuses on moments related to the distribution of financial capital (net worth) and relationships in the banking industry. Net worth is slightly lower on average and less dispersed in the baseline economy than in the competitive economy. Figure 3 plots the partial distributions of net worth and relationships, and the joint distribution of net worth and relationships in the baseline model (the latter two objects are not defined for the competitive economy). The compression of the net worth distribution in Figure 3(a) combines two main forces. First, bank market power generates higher franchise values and lower optimal loan issuance in the baseline model. Banks ration quantities to keep markups high, as is standard in models of imperfect competition. Thus banks tend to cluster at lower levels of net worth than unconstrained banks in the competitive case, where there

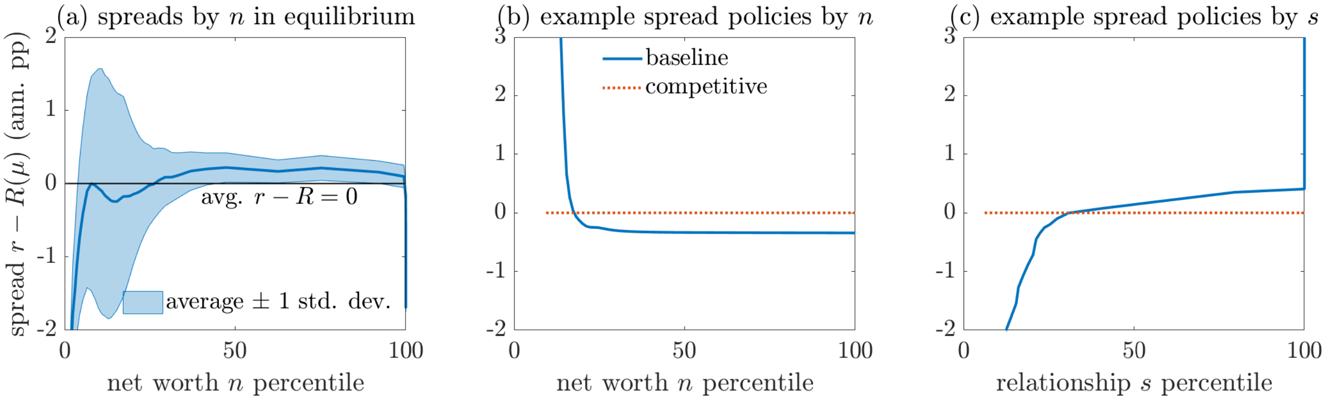

Figure 4: **Steady state loan pricing policies**

Notes: Panel (a) plots the first and second moments of the pricing policies (expressed as annualized percentage point spreads over the average interest rate) conditional on a given level of net worth over the distribution of banks in the baseline model. Variation in pricing at each level of n comes from dispersion in s and z. Panels (b) and (c) plot sample pricing policy functions over the distribution of net worth and relationships: spreads are zero by definition in the competitive case. Panels (b) and (c) each fix $z = 0$, the median level, and s or n at the 25th percentile from.

is a large concentration of banks to the right of the dense part of the distribution in the baseline model. Counteracting this first effect, though, is the fact that profitability is lower – and therefore financial constraints are effectively tighter – in the competitive economy.[16] Correspondingly, there are far more banks with very low net worth in the competitive economy.

The low elasticity economy features lower average, more compressed net worth than the baseline. This stems from the fact that more profitable lending means banks are better able to smooth dividends in the face of idiosyncratic uncertainty. Increasing *dynamic* market power instead, as in the low punishment economy, also yields lower average net worth but with more dispersion. On the one hand, more persistent relationships give banks an incentive to accumulate more net worth to lend large quantities at high spreads. On the other, the dominance of banks with strong relationships in this model makes it hard for smaller banks to compete and grow, and so there is a long left tail of smaller banks in this case. Quantitatively, the second effect dominates.

5.1.3 Financial and relationship capital are complements.

The last rows in Table 3.B present correlations between the key state variables and interest rate spreads. In the simple relationship models (baseline, low elasticity, and low punishment), there is a strong positive correlation between net worth and relationships. Figure 3(c) depicts the range of relationships across the distribution of net worth in the baseline economy, which visually confirms this correlation: banks with less financial capital have weaker relationships. While relationship strength increases across the entire net worth distribution, this rise is especially sharp over the bottom quartile of the distribution. Notably, this effect is absent in the fixed relationship model,

[16]Additionally, since each unit of lending is less profitable, the "life cycle" of net worth tends to have a much flatter profile in the more competitive economies (shown in Panel (c) of Figure 6(c) and discussed below).

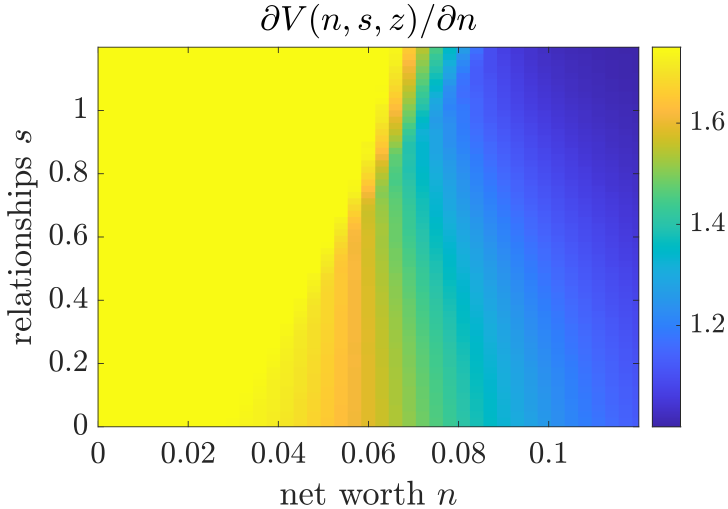

Figure 5: **Complementarity of financial and relationship capital**

Notes: This figure plots a numerical approximation to the slope of the bank value function with respect to net worth with z fixed at the median level in the baseline model. The color at each point represents the marginal value of net worth at that point, with shading governed by the color bar to the right of the figure.

since banks "endowed" with strong relationships need no financial capital buffers to preserve them.

In our baseline model, the correlations of spreads with both net worth and relationships are positive. This result combines two effects. First, less capitalized banks – who are constrained by the capital requirement and the prospect of having to issue equity – charge high spreads and lend small amounts, as shown in Figure 4(b). Second, banks with weak relationships tend to price competitively – even below market, with $r - R < 0$ – to build relationships for the future, as shown in Figure 4(c). With stronger relationships, banks can sustain lending above market interest rates. In isolation, these two forces would lead the correlation between spreads and net worth (relationships) to be negative (positive). These forces are tempered, however, by the strong positive correlation between net worth and relationships described above. That is, the banks who are financially constrained and would like to charge high spreads tend to be the very same banks who have weak relationships and therefore would like to charge low spreads. These two effects roughly offset each other for all but the smallest banks (for whom the second effect dominates), as highlighted in Figure 4(a), which plots the joint distribution of net worth and spreads in the baseline model.

How do changes in the competitive landscape alter these effects? With more market power (static or dynamic), constrained banks are able to charge higher interest rates while sustaining similar levels of lending. This naturally strengthens the correlations between spreads and relationships, as well as between spreads and net worth. Here again, though, the fixed relationship model is quite different: since relationships are stable, the banks with the strongest relationships charge the highest spreads and have the weakest capital buffering incentives, and so the correlation between spreads and net worth is sharply negative.

Our model, then, sheds new light on financial constraints in banking. Measuring banks' net worth provides information on banks' pricing and lending decisions, but the degree to which these policy functions are actually elastic with respect to net worth, and the levels of net worth at which this elasticity manifests, can vary considerably with relationships and the competitive landscape.

Ultimately, we find that financial and relationship capital are complements. These two types of capital are complements if more net worth delivers more value to a bank with stronger relationships, i.e. if the cross-partial of the value function, $\frac{\partial^2 V}{\partial n \partial s}$ is positive. Figure 5 plots the partial of the value function with respect to net worth over a range of bank states to confirm this point: everywhere in the state space, this object is weakly increasing as relationships strengthen. Quantitatively, the marginal valuation of net worth is bounded between 1 and $1 + \overline{\psi}$ given our specification of equity issuance costs. The value $1 + \overline{\psi}$ obtains towards the top left in Figure 5, where banks are financially constrained but have a valuable loan franchise due to strong relationships: here, banks are in fact quite willing to issue equity. Moving down and to the right in the figure towards more financial capital and weaker relationships, banks are less likely to find it optimal to issue new equity, and so the marginal valuation of net worth moves towards 1.

5.1.4 Relationships redistribute surplus from borrowers to lenders.

A natural question for comparing our model economies is: how does the nature of lending relationships affect the allocation of economic surplus between banks and firms? Table 3.C addresses these questions by comparing bank, firm, and total value for each model variant relative to the baseline economy. While our model is not well-suited to traditional welfare analysis, we can measure both bank and firm values directly using value functions since both objective functions are denominated in dollars. Total bank value is measured as $\overline{V} \equiv \int V(x) \mathrm{d}m(x)$, total firm value is the solution to the firm problem, denoted $\overline{W}$, and total value is simply the sum $\overline{V} + \overline{W}$.[17]

Making the banking sector more competitive reallocates surplus from banks to firms: total bank value drops by more than 70% in the competitive economy relative to the baseline, while firm value increases nearly 6%. Similarly, making the banking sector less competitive – either statically or dynamically – works in the opposite direction. Interpreting the changes in the fixed relationship model is less intuitive: while banks face no dynamic discipline on their pricing decisions, the demand elasticity is much higher to facilitate comparison with the baseline model. While these effects cancel in the aggregate, they do involve a slight loss of surplus for banks and a slight gain for firms.

5.1.5 Relationships shape banks' life cycles.

Figure 6 investigates the life cycle of a bank beginning from entry across all versions of our model. In all cases, new banks start out with no net worth (panel (c)), which does not allow for much lending (panel (a)). In the less competitive economies, banks optimally price below market (panel (b)) in order to build relationships (panel (d)). Given the persistence of relationships in the low punishment

[17] Note that by computing total value as the sum of bank and firm value, we are implicitly weighting the two sectors by their profit shares in the economy. Since firm profits are large relative to bank profits, then, the total value metrics closely follow the firm value metrics.

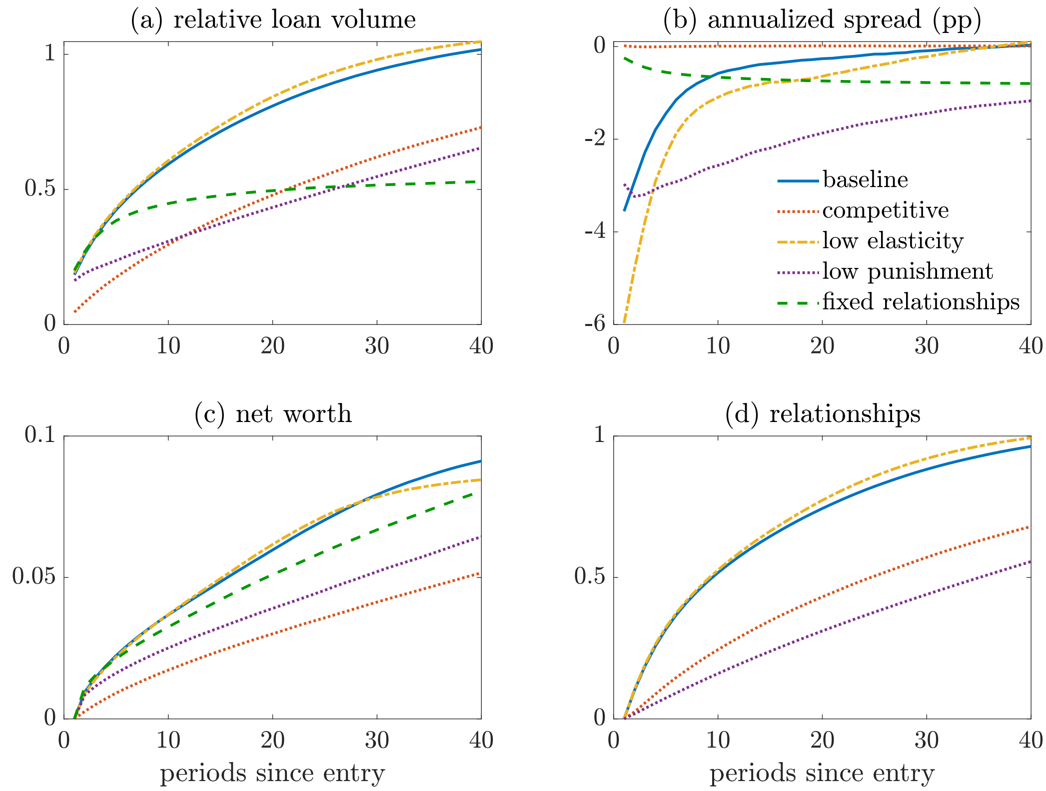

Figure 6: **Average bank life cycle across model variants**

Notes: For each version of the model, we simulate paths of $T = 40$ periods for $N = 10,000$ entrant banks with initial states $n = 0$, $s = 0$, and z drawn from the ergodic distribution $\overline{\Gamma}(z)$. Each plot presents averages of the indicated metrics across the N banks for each date $t = 1, ..., T$.

case, the required period of pricing below market is much longer. By contrast, in the competitive model, banks simply price at the market rate and increase loan volume gradually alongside net worth. In the fixed relationship model, a bank's only goal is to reach the desired level of net worth: in the short run, they do this by building up loan volumes. Notably, the results in Figure 6 suggest that there is a "sweet spot" with regard to the accumulation of net worth: in the competitive economy, low profit margins across all banks make it hard for banks to accumulate financial capital, while in the less competitive low punishment economy the intensity with which small, weak relationship banks must compete against more established banks renders profits similarly low. As banks accumulate financial and relationship capital, they lend more while increasing spreads.

5.2 Empirical validation

This section validates our model along two dimensions. First, we demonstrate that our model matches how spreads evolve over a lending relationship in the data by appropriately balancing two competing forces. Second, we document that our baseline model generates incentives to build up capital buffers that are more in line with the data than the alternatives we consider.

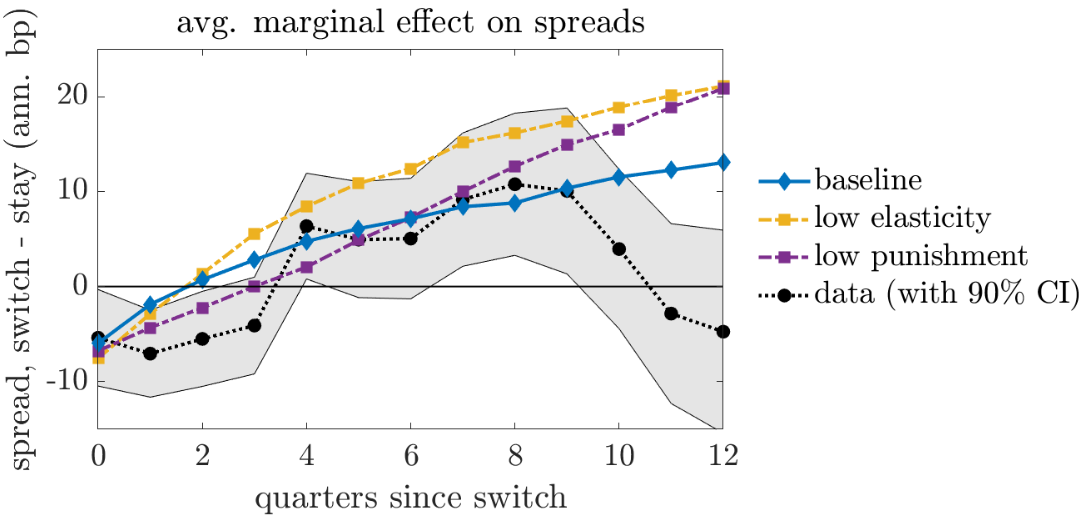

Figure 7: **Differences in spreads from switching relative to non-switching loans**

Notes: This figure plots the differences in spreads between switching and non-switching loans for several model variants and in the data. For detailed descriptions of the construction of this figure, see Section 2.2.2 for the data and Section 5.2 for the model.

5.2.1 Evolution of spreads over a relationship

A key novelty in our framework is the link between persistent lending relationships and banks' pricing decisions. To have confidence in our model's predictions about how relationships affect aggregates, then, we must first provide evidence that our model delivers pricing implications given changes in relationships that are in line with the data.

Recall that Figure 2 shows the average difference in interest rates for loans from new banks ("switches") compared to loans from banks with existing relationships. The key insights from this figure are that: (i) switching loans have are relatively cheap immediately upon and in the first year following the switch; and (ii) this pattern reverses in the second year after the switch.

Does our model deliver similar patterns? We address this question by simulating a panel of banks drawn randomly from the stationary distribution and study how they behave after a share δ_s of their relationship capital s is destroyed. Holding the rest of the banks' states fixed, this reduction in relationship capital accounts for differences in lending practices solely attributable to differences in relationships, as in our empirical analysis. We choose δ_s for each model variant to match the average initial drop in spreads immediately upon switching of 5.4 bps.[18]

Figure 7 plots the difference between switching and non-switching loan spreads for our baseline model and the two variants (low elasticity and low punishment) in which relationships evolve. The empirical relationship is reproduced as the black line and shaded region. Our baseline model matches

[18]It is perhaps natural to think of a true "switch" as the case $\delta_s = 100\%$. This implementation, however, is far too extreme with a representative borrower: a bank with $s = 0$ faces extremely small loan demand under equation (5), and so prices extremely aggressively to build it up. Empirically, both incumbent and switching lenders have *portfolios* of borrowers, and so the drop in relationship intensity we implement must be more marginal.

the dynamics of spreads in the data quite closely along several dimensions. In both model and data, banks price below market upon switching, but then steadily charge slightly above-market spreads thereafter. Our model correctly captures – both qualitatively and in terms of magnitudes – the gradual increase in spreads after the "honeymoon phase" following the switch; over the entire three-year span depicted in Figure 7, the relative increase in the spread hovers between 5 and 15 bps as in data. While the magnitudes vary across different institutional contexts, this is also qualitatively consistent with the findings of Ioannidou and Ongena (2010) and Banerjee et al. (2021).

When we repeat this exercise for the low elasticity and low punishment models, the comparison with the data is less favorable. While both match the initial drop in spreads by construction, they fail to match the data in the subsequent periods in different ways. In the low elasticity model, banks leverage high switching costs to take immediate advantage of their newly captive borrower, charging much higher spreads than we observe in the data. Spreads remain persistently high over the life of the relationship. In the low punishment case, the spreads remain low for longer, but the profile does not level off as in the data. This is because banks need more time to accumulate sufficient relationship capital to increase rates without losing borrowers. Once these relationships are built, though, charging higher spreads does not erode the relationship as much as in the baseline.

5.2.2 Relationships and capital buffers

As discussed throughout Section 5.1, relationships interact with financial constraints to shape lending and financing decisions, as well as the distribution of banks in the industry. Do these interactions deliver financing policies in line with what we see in the data? A key measure of banks' financing and lending policies – and one with important implications for financial stability – is how capitalized banks are. This is typically measured using the capital buffer, which is the ratio of a bank's equity value to its asset value. Figure 8 plots the distribution of capital buffers in the data, our baseline model, and the competitive model. The insights here are similar to those from Table 3 and Figure 3. Our baseline model has a strong precautionary motive: the average capital buffer is 12.7% compared with 11.7% and 12.3% in the competitive economy and the data, respectively. Furthermore, the greatest density of banks occurs at capital buffers in the range of 11 to 15% as in the data.

By comparison, these clusters occur elsewhere in each alternative model (see Appendix Figure D.3). In the competitive case with smaller franchise values, the biggest clusters are at the capital requirement and far above the data. In the low elasticity economy, high returns to lending effectively provide banks insurance against shocks, and so banks cluster between 10% and 12%. In the low punishment economy, banks with strong relationships retain even lower capital buffers (clustering between 8% and 10%), but there is a fat tail resulting from the fact that many banks have built up net worth but do not yet have strong lending relationships. In the fixed relationship model, capital buffering motives blend the baseline model and the competitive model: since the distribution of relationships is the same, the shape of the distribution matches the baseline, but since many banks with weak relationships have no hope to grow them and insure against shocks using relationship capital, they pile up at even higher capital buffers.

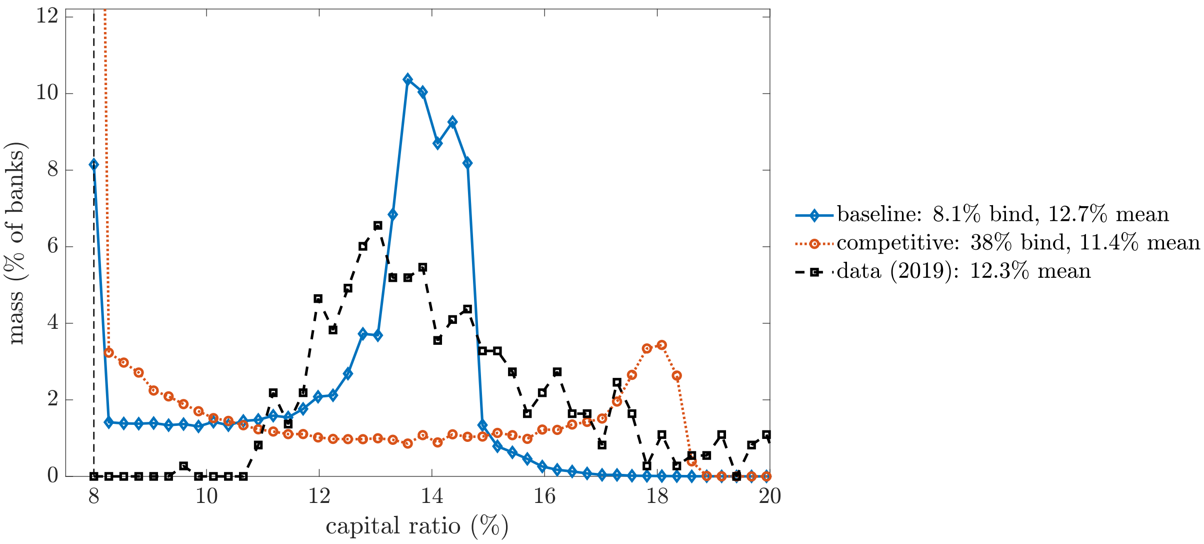

Figure 8: **Capital buffers and relationships**

Notes: This figure represents the distribution of capital buffers across economies. Each is depicted as a probability mass function with 50 bins between the minimum, $\chi = 8\%$, and 20% (the share of banks with buffers above 20% in negligible in each model). The legend includes the share of banks for whom the capital requirement actually binds. The capital buffer is defined as the ratio of book equity to book assets, $\frac{g_q(x)g_{\ell'}(x)-\bar{q}^d g_{d'}(x)}{g_q(x)g_{\ell'}(x)}$. The data refers to the total capital ratio reported in form FR Y-9C (BHCA7205), averaged across the four quarters of 2019.

Takeaways Section 5.2 has shown that the model specification with parameters (ϕ, ρ_q) directly estimated from the micro data as described in Section 4.2 is better at reproducing untargeted features of the data than models with alternative values for each of these parameters. This gives us confidence that if we were to use a more traditional macro approach to parameter calibration, choosing values to target moments such as the evolution of spreads over relationships and the distribution of capital buffers, we would have likely arrived at similar values for these parameters. It is worth noting that this approach would likely have led our estimates of the persistence of lending relationships to increase (i.e. for ρ_q to decrease and ρ_s to increase). This is because the low punishment model shares key features in common with the data in both experiments: slower rate increases over the first year of the relationship in Figure 7, and a longer right tail of capital buffers in Figure D.3.

6 Aggregate Dynamics

We now analyze how lending relationships shape how the economy responds to aggregate shocks. We consider three types of negative aggregate shocks: (i) a "financial crisis", where all banks experience a proportional decline in their net worth; (ii) a shock to the deposit funding cost of banks, $\bar{q}^d$, and (iii) a negative shock to loan demand. We assume that these are one-time, unanticipated shocks. To keep the discussion focused, we consider only three of the five model variants from the previous section: baseline, competitive, and fixed relationships. Results for the other two model variants are available in Appendix D.4. These results are qualitatively similar to the baseline.

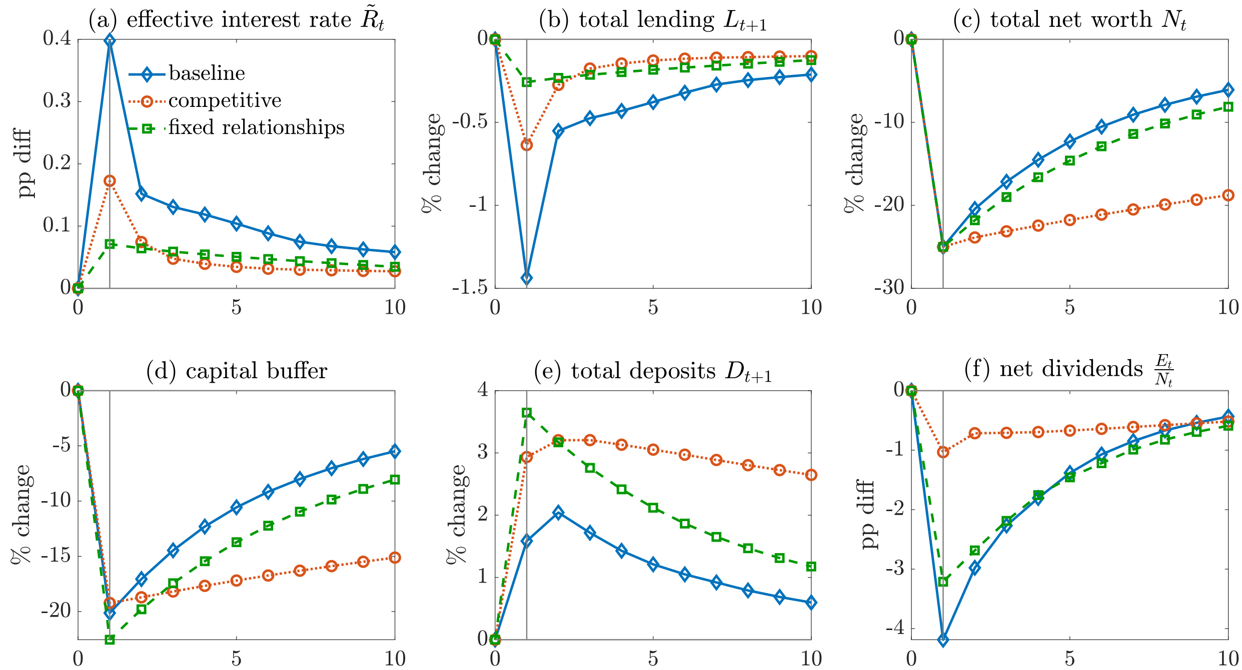

Figure 9: **Aggregate shock to bank net worth**

Notes: This figure plots percent deviations (panels (b), (c), (d), and (e)) or percentage point differences (panels (a) and (f)) from steady state. The aggregate shock reduces each bank's net worth by 25% at date 1.

6.1 Net worth or "financial crisis" shock

Figure 9 plots the response of aggregate variables to a negative aggregate shock in which the net worth of every bank unexpectedly declines by 25%. This shock is consistent with unexpected losses arising from other business lines within the bank, such as mortgage lending and/or MBS holdings during the 2007-08 financial crisis. Panels (a) through (f) present the effects of the shock on the effective interest rate, total lending, total net worth, average capital buffer, total deposits, and net dividend payouts, respectively. While the responses are qualitatively similar across the economies, there are important quantitative differences.

Panels (a) and (b) show that the increase in the effective interest rate and corresponding decrease in lending are larger on impact and more persistent in the baseline economy than the competitive one. On impact, for example, the drop in lending is 24 bps or 150% larger. Panel (c) shows that these dynamics allow banks to recapitalize much more quickly in the baseline (the initial drop in net worth is the same by construction). These results come from banks with stronger relationships exploiting the complementarity between relationship and financial capital, which they are unable to do in the competitive case. Faced with tighter financial constraints, they raise rates on "captive" borrowers. Since all banks respond this way, though, and since relationships evolve based on *relative* loan volume, in equilibrium there is no net erosion of relationships associated with this strategy. As a result, this advantage persists over the life of the recovery.

On impact, banks in both economies replace their net worth with deposit financing and cut dividends, with the former effect leading to smaller capital buffers (note that the capital buffer depends on the behavior of both lending and net worth). In the economy with relationships, though, the faster recovery of bank net worth promotes quicker rebuilding of capital buffers. This implies that while lending relationships amplify the real effects of the financial shock, they also promote financial stability and resilience within the banking sector in the subsequent recovery.

The response is materially different in the fixed relationship economy. Here, as in the baseline, banks exploit market power to recapitalize faster, and so the financing patterns in these two cases are similar. On the real side, though, the drop in net worth translates into only a small (3 bp) increase in the effective interest rate. This is because banks in this economy on average retain much larger capital buffers than in the baseline model (see Figure D.3), and so there is less pass-through from the financial shock into lending. These results suggest that even in economies with the same "average" level of market power, the dynamics associated with that market power are crucial in shaping the real responses to aggregate shocks.

6.2 Funding cost shock

Figure 10 considers a negative shock to $\overline{q}^d$, such that banks' cost of funding increases from 2% to 4% (annualized). The shock is persistent with $\rho = 0.5$. This is a stylized way of modeling a monetary policy shock and studying its transmission through the bank lending channel in the context of our model.[19] This exercise is similar to the one performed by Wang et al. (2022), among others.

The baseline and competitive economies respond similarly on the real side. Banks largely pass the increased cost of funding through to borrowers, yielding a persistent increase (drop) in loan rates (volumes). Thus both models generate standard transmission of monetary policy to quantities of credit, which affects real activity via the working capital constraint.

While almost all (94.5%) of the increase in funding costs is passed through in the competitive case, this pass-through is muted in the relationship model (63.2%).[20] The former result arises because net interest margins are razor thin in the competitive model: just 0.2% in steady state, 90% lower than in the baseline. The latter result arises because, in the relationship model, well-capitalized banks – i.e. those with greater capital buffers and less direct dependence on deposit financing – seize on this shock and their relatively advantageous financial position as an opportunity to build relationships. As a result, these banks raise interest rates by less. This dampens the upward pressure on the effective interest rate $\tilde{R}$ through both the level and covariance components. Notably, this latter effect is shut down in the fixed relationship model since there is no incentive to build relationships. Therefore, in this case the real response in the fixed relationship model more closely resembles the competitive model than the baseline.[21]

[19] Of course, the monetary policy (Federal Funds) rate and banks' marginal cost of funds (i.e. the deposit rate) are not exactly the same thing. However, as documented, for example, in Drechsler et al. (2017) and Wang et al. (2022), there is strong co-movement between the two objects. To facilitate discussion, however, in this section we use the terms funding cost shock and monetary policy shock synonymously.

[20] We measure "pass-through" as the increase in $\tilde{R}_t$ relative to the increase in the funding rate $1/\overline{q}^d$.

[21] It is worth noting that models of static market power may be calibrated to deliver lower levels of pass-through as in the data (see, for example, Wang et al. (2022)). Since this is not the primary focus of our analysis in this paper,

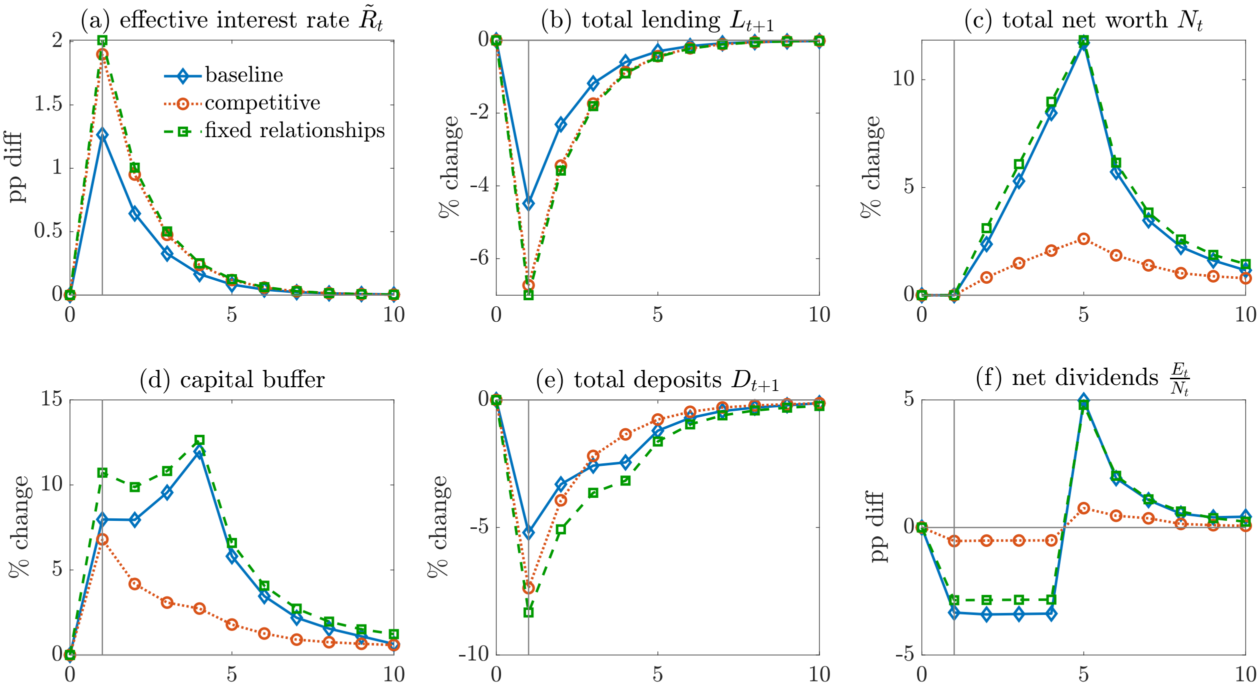

Figure 10: **Aggregate shock to cost of funding**

Notes: This figure plots percent deviations (panels (b), (c), (d), and (e)) or percentage point differences (panels (a) and (f)) from steady state. The aggregate shock raises the cost of funding $\overline{r}^d_{\text{ann}}$ from 2% to 4% at date 1, returning to 2% with persistence 0.5.

Raising the relative cost of deposit financing promotes the build-up of capital buffers, and so bank net worth rises in all three economies. Banks substitute away from deposits and into retained earnings to fund their loans as long as the funding cost remains elevated. As part of this substitution, banks cut dividends and/or issue equity (depending on their initial net worth position). Given the substantial gap in profit margins, though, this effect is much more pronounced in the baseline and fixed relationship models: for example, the increase in net worth (capital buffers) peaks at 11.7% (12.0%) above steady state in the baseline model, as opposed to 2.6% (6.8%) in the competitive model, and the relative drop in total dividends on impact is -3.3 pp as opposed to -0.5 pp.

Ultimately, the pass-through of a funding cost shock to the loan market plays out similarly in the three economies, but is relatively muted in the baseline relative to both the competitive and fixed relationship models. The nature of competition generates stark differences in how capital buffers react to a tightening of monetary policy, which has natural implications for financial stability.

6.3 Loan demand shock

Figure 11 plots the effects of a persistent contraction in loan demand, modeled as a 1% drop in TFP A with a persistence of 0.5. The headline result is that the presence of lending relationships sharply

though, our fixed relationship model is calibrated to be comparable to our baseline model (same average spreads), not to generate this feature of the data.

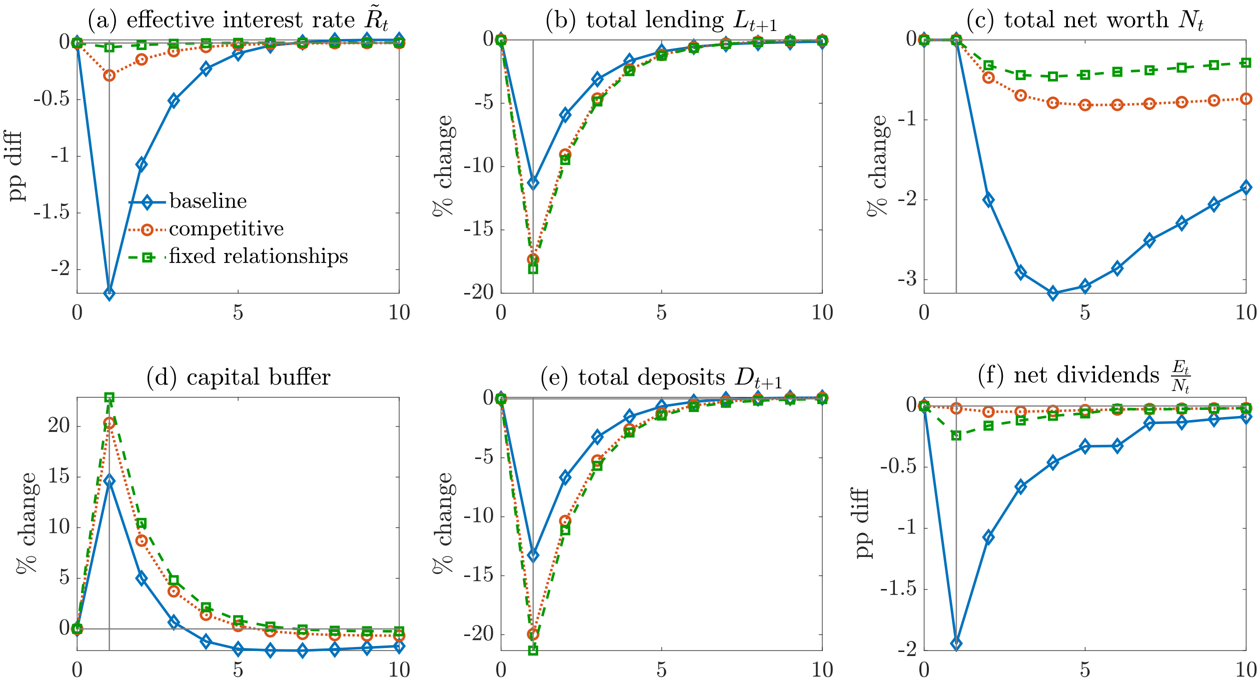

Figure 11: **Aggregate shock to loan demand**

Notes: This figure plots percent deviations (panels (b), (c), (d), and (e)) or percentage point differences (panels (a) and (f)) from steady state. The aggregate shock lowers TFP by 1% at date 1, recovering with persistence 0.5.

dampens the contraction in loan volume (and therefore output) associated with the drop in TFP in the baseline model relative to both the competitive and fixed relationship models. Whereas the equilibrium loan rate drops only 0.29 pp on impact in the competitive model and 0.04 pp in the fixed relationship model, it drops 2.21 pp in the baseline. This translates into an initial drop in total lending that is 35.0% smaller in the baseline economy than the competitive one.

The bolstering of lending we observe in the relationship model arises from banks' dynamic incentives: despite lending being less profitable today given the drop in aggregate demand, banks would still like to improve their relative position by lending more to build and maintain relationships for tomorrow. As in the case of the funding cost shock above, since the primary difference between the baseline and competitive economies arises from banks' dynamic incentives, the fixed relationship economy – with no such incentives, despite market power as in the baseline – more closely resembles the competitive economy.

Turning to financing, this sharp drop in profitability in the relationship model causes a much larger and more persistent contraction in banks' net worth than in the competitive model. This drop in net worth, however, is smaller in magnitude than the drop in lending, and so capital buffers increase on net. This drop in net worth occurs despite cuts in deposits and dividends: taken together, these results imply that the drop in net worth is driven by lending unprofitably in the early aftermath of the initial shock. By contrast, banks in the competitive model respond to the negative shock to loan demand by simply shrinking: they cut their deposit base and *increase* dividend payouts. Clearly

shock	**net worth**			**funding cost**			**TFP**		
model	base	comp	fixed	base	comp	fixed	base	comp	fixed
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
bank value	-5.7	-23.1	-7.4	-0.5	-0.2	-0.2	-1.9	-1.0	-0.3
firm value	-0.0	-0.3	-0.0	-0.0	-0.1	-0.1	-0.1	-0.1	-0.2
total value	-0.3	-0.5	-0.0	-0.0	-0.1	-0.0	-0.2	-0.2	-0.2

Table 4: **Analysis of surplus across aggregate shocks**

Notes: Each moment is the percentage point difference in the indicated value from the date of impact of the shock relative to the steady state for the indicated model economy.

this difference in response stems from the fact that the competitive banks have no reason to lend at undesirable terms absent dynamic relationship considerations. While relationships temper the impact of the shock on the real side, this happens with no appreciable cost to financial stability (as measured by capital buffers), despite the amplified contraction in bank net worth.

6.4 Summary and analysis of surplus

A key theme emerges from the analysis of the three shocks across the three models considered above: the *dynamic incentives* associated with the form of market power in our baseline model of lending relationships make it behave materially differently than either the competitive model or the fixed relationships model with only static market power. We can support this insight, documented via real effects and financing patterns in Figures 9 through 11, by measuring economic surplus across agents in the model relative to steady state for each shock. This also allows us to compare the total costs of each shock across economies. Concretely, we modify the analysis from Section 5.1.4 by computing our surplus metrics as of the impact date of the shock.[22] The results of this analysis are presented in Table 4.

Since all these shocks are negative by construction, both banks and firms lose relative to steady state in all cases. For banks, unsurprisingly the net worth shock is particularly acute, resulting in losses in value of more than 5% relative to steady state in each case. The effect is especially strong (23.1%) in the competitive model, which makes sense because banks in this model cannot increase spreads to absorb the loss of financial capital. Since banks recapitalize faster recapitalization and steady state firm values are lower in the less competitive models, there are only small reductions in firm value relative to the steady state. This furthers the idea that bank market power – captured in franchise value – has benefits for financial stability.

While the effect sizes for the funding cost and TFP shocks are smaller, the results are intuitive.

[22]Specifically, we replace the bank and firm steady state values $\overline{V}$ and $\overline{W}$ with V_1 and W_1, where, for example,

$$W_t = Ak_t^\alpha n_t^\eta - \overline{w}n_t - \overline{uc}k_t + L_{t+1} - \int \ell_t(q,s)\mathrm{d}\mu_{t-1}(q,s) - \frac{\phi}{2}L_{t+1}\int\left(\frac{q\ell_{t+1}(q,s)}{L_{t+1}} - 1 - (s-S)\right)^2 \mathrm{d}\mu_t(q,s) + \overline{q}\overline{W}_{t+1}$$

As discussed in Section 6.2, both the competitive and fixed relationship models exhibit high rates of pass-through of the increase in funding costs to borrowers. Therefore, in these cases, the reduction in bank value is small relative to the baseline, while the reduction in firm value is large relative to the baseline. The effects are largely reversed in the case of the TFP shock: in this case, banks with dynamic relationships as in the baseline have incentive to dampen the reduction in credit demand to build relationships, which is costly in the short term. Therefore the loss in value to banks is appreciably larger in the baseline model than in either the competitive or fixed relationship model.

7 Conclusion and Directions for Future Research

This paper presents a quantitative framework with which to evaluate the aggregate consequences of lending relationships. Our model environment combines standard features from the literature on heterogeneous banks subject to financial constraints with two novel elements: (i) loan sourcing adjustment costs for borrowers; and (ii) internalization of relationship formation by banks. These elements yield a tractable but rich model of relationships which is amenable to direct estimation of key relationship parameters and efficient computation, despite the richness of heterogeneity and financing choices within the banking sector.

Quantitatively, we present four primary results. First, our baseline model matches the profile of interest rate spreads over the life of a bank-borrower relationship that we observe in the data. Our model gets both the static and dynamic components of the market power which arises from relationships right, as model variants which vary either element struggle to match this empirical profile. Second, we show that financial and relationship capital are complements at the bank level and therefore correlated in the equilibrium distribution of banks in the model. Third, we show that the equilibrium relationship between the degree of lending relationships and the strength of banks' capital buffers is non-monotone: our baseline model features stronger precautionary motives than both more and less competitive alternatives. Fourth, we show that relationships shape both the real and financial impacts of aggregate shocks to both credit supply and credit demand.

Our analysis suggests several promising directions for future research. The first is related to two important long-terms in the U.S. banking industry. Over the past several decades, the commercial banking industry has consolidated, and alternative non-bank financial intermediaries (shadow banks, "fintech") have grown. On the one hand, industry consolidation directly leads to reduced competition (which could translate in a lower "punishment" parameter ρ_q, or a lower elasticity ϕ), but this effect could be counteracted by the increasingly diverse geographical footprint of major U.S. bank holding companies (Oberfield et al., 2024). Thus it is possible that the U.S. banking industry has become a more competitive oligopoly, which could contribute to a rise in ρ_q or ϕ. On the other hand, the emergence of new digital-only lenders, such as the case of many fintech companies, could reduce the importance of relationships in banking, as lending terms and decisions are increasingly determined by centralized algorithms. Our framework is well-suited to measuring the aggregate impacts of these forces along this transition.

Second, the fact that different countries' banking sectors are structured in ways very different from the U.S. might imply different interactions between relationship capital and financial frictions than those described in the present paper, which was intended to model the U.S. banking sector. For example, it is well-documented that the Canadian banking sector is considerably more consolidated than the U.S. financial sector. Third, our results on aggregate dynamics suggest a rich interplay between relationships and bank capitalization, which could have important implications for regulation and financial stability. We leave these and other avenues for future research.

References

Adrian, Tobias and Nina Boyarchenko, "Intermediary Leverage Cycles and Financial Stability," Staff Reports 567, Federal Reserve Bank of New York 2012.

Amiti, Mary and David E. Weinstein, "How Much Do Idiosyncratic Bank Shocks Affect Investment? Evidence from Matched Bank-Firm Loan Data," *Journal of Political Economy*, 2018, *126* (2), 525–587.

Atkeson, Andrew G., Adrien d'Avernas, Andrea L. Eisfeldt, and Pierre-Olivier Weill, "Government Guarantees and the Valuation of American Banks," *NBER Macroeconomics Annual*, 2019, *33* (1), 81–145.

Banerjee, Ryan N., Leonardo Gambacorta, and Enrico Sette, "The real effects of relationship lending," *Journal of Financial Intermediation*, 2021, *48* (C).

Baron, Matthew, "Countercyclical Bank Equity Issuance," *The Review of Financial Studies*, 2020, *33*, 4186–4230.

Berger, Allen N. and Gregory F. Udell, "Relationship Lending and Lines of Credit in Small Firm Finance," *The Journal of Business*, 1995, *68* (3), 351–381.

_ **and Timothy H. Hannan**, "The Efficiency Cost of Market Power in the Banking Industry: A Test of the Quiet Life and Related Hypotheses," *The Review of Economics and Statistics*, 1998, *80* (3), 454–465.

Bigio, Saki and Javier Bianchi, "Banks, Liquidity Management, and Monetary Policy," *Econometrica*, January 2022, *90*, 391–454.

Boualam, Yasser, "Credit Markets and Relationship Capital," Working Paper, Kenan-Flagler Business School 2018.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 2005, *113* (1), 1–45.

Corbae, Dean and Pablo D'Erasmo, "Capital Requirements in a Quantitative Model of Banking Industry Dynamics," *Econometrica*, November 2021, *89* (6), 2975–3023.

Darmouni, Olivier, "Informational Frictions and the Credit Crunch," *The Journal of Finance*, 2020, *75* (4), 2055–2094.

de Guevara, Juan Fernández, Joaquín Maudos, and Francisco Pérez, "Market Power in European Banking Sectors," *Journal of Financial Services Research*, Apr 2005, *27* (2), 109–137.

Degryse, Hans, Olivier De Jonghe, Sanja Jakovljević, Klaas Mulier, and Glenn Schepens, "Identifying credit supply shocks with bank-firm data: Methods and applications," *Journal of Financial Intermediation*, 2019, *40* (C).

Demsetz, Rebecca, Philip E. Strahan, and Marc R. Saidenberg, "Banks with Something to Lose: the Disciplinary Role of Franchise Value," *Economic Policy Review*, October 1996, pp. 1–14.

Diamond, Douglas W., "Financial Intermediation and Delegated Monitoring," *The Review of Economic Studies*, 1984, *51* (3), 393–414.

Dotsey, Michael and Robert G. King, "Implications of state-dependent pricing for dynamic macroeconomic models," *Journal of Monetary Economics*, January 2005, *52* (1), 213–242.

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

Egan, Mark, Ali Hortaçsu, and Gregor Matvos, "Deposit Competition and Financial Fragility: Evidence from the US Banking Sector," *American Economic Review*, January 2017, *107* (1), 169–216.

Faria-e-Castro, Miguel, Pascal Paul, and Juan M. Sanchez, "Evergreening," *Journal of Financial Economics*, 2024, *153* (C).

Farinha, Luísa A. and Joao Santos, "Switching from Single to Multiple Bank Lending Relationships: Determinants and Implications," *Journal of Financial Intermediation*, 2002, *11* (2), 124–151.

Gilchrist, Simon, Raphael Schoenle, Jae Sim, and Egon Zakrajsek, "Inflation Dynamics during the Financial Crisis," *American Economic Review*, March 2017, *107* (3), 785–823.

Gourio, François and Leena Rudanko, "Customer Capital," *Review of Economic Studies*, 2014, *81* (3), 1102–1136.

Hennessy, Christopher A. and Toni M. Whited, "Debt Dynamics," *Journal of Finance*, 2005, *60*, 1129–1165.

Ioannidou, Vasso and Steven Ongena, "Time for a Change: Loan Conditions and Bank Behavior when Firms Switch Banks," *The Journal of Finance*, 2010, *65* (5), 1847–1877.

Kimball, Miles S., "The Quantitative Analytics of the Basic Neomonetarist Model," *Journal of Money, Credit and Banking*, 1995, *27* (4), 1241–1277.

Neri, Andrea Geraliand Stefano, Luca Sessa, and Federico M. Signoretti, "Credit and Banking in a DSGE Model of the Euro Area," *Journal of Money, Credit and Banking*, 2010, *42*, 107–141.

Oberfield, Ezra, Esteban Rossi-Hansberg, Nicholas Trachter, and Derek T. Wenning, "Banks in Space," NBER Working Papers 32256, National Bureau of Economic Research, Inc March 2024.

Payne, Jonathan, "The Disruption of Long Term Bank Credit," Working Papers 2018-2, Princeton University. Economics Department. December 2018.

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

Polo, Alberto, "Imperfect pass-through to deposit rates and monetary policy transmission," Bank of England working papers 933, Bank of England July 2021.

Prescott, Edward Simpson and Hubert P. Janicki, "Changes in the size distribution of U.S. banks: 1960-2005," *Economic Quarterly*, 2006, (fall), 291–316.

Ravn, Morten, Stephanie Schmitt-Grohé, and Martín Uribe, "Deep Habits," *Review of Economic Studies*, 2006, *73* (1), 195–218.

van Binsbergen, Jules H., William F. Diamond, and Marco Grotteria, "Risk-free interest rates," *Journal of Financial Economics*, 2022, *143* (1), 1–29.

Wang, Yifei, Toni M. Whited, Yufeng Wu, and Kairong Xiao, "Bank Market Power and Monetary Policy Transmission: Evidence from a Structural Estimation," *The Journal of Finance*, 2022, *77* (4), 2093–2141.

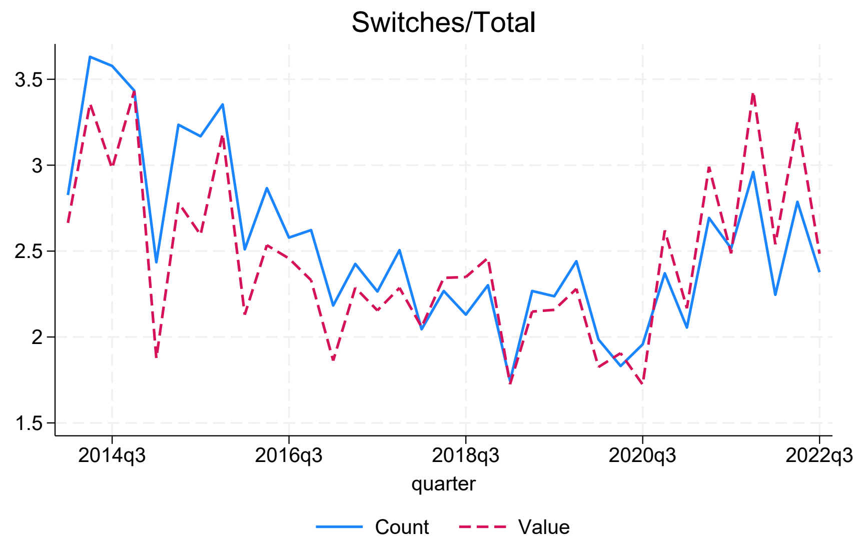

Figure A.1: **Switches as a percentage of total outstanding loans, single-bank firms only**

Appendix for "A Quantitative Theory of Relationship Lending"

A Data Appendix

A.1 Facts on Loan Markets: Single-Bank Firms

Figures A.1 and A.2 replicate figures 1 and 2 for single-bank firms. We define a single-bank firm as a firm that maintains at most one observable relationship with a bank in a quarter.

B Model Appendix

B.1 Proof of Proposition 1: Loan Demand System

First note that cost-minimization implies an optimal capital-labor ratio that allows us to express optimal labor as a function of the choice of capital

$$n = \frac{\overline{uc}}{\overline{w}} \frac{\eta}{\alpha} k \tag{B.1}$$

This implies that total costs can be written as $\overline{uc}k + \overline{w}n = \overline{uc}k\frac{\alpha+\eta}{\alpha}$. Begin by placing multipliers $\lambda \geq 0$ on constraint (3) and $\zeta \geq 0$ on constraint (4) and taking first order conditions in the borrower's

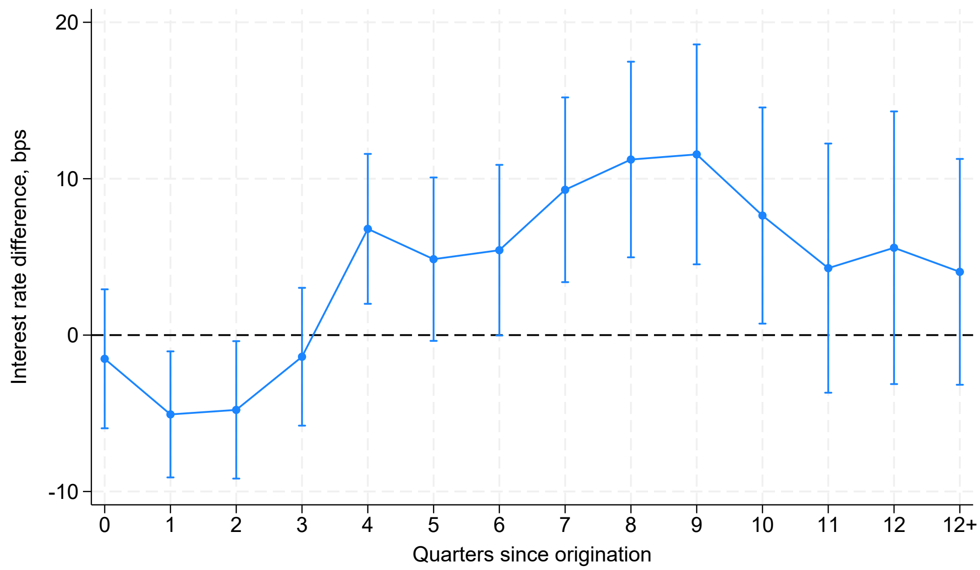

Figure A.2: **Average spread between loans for new and existing relationships, single-bank firms only**

Notes: See text for details. At each time since loan origination, the dot represents the point estimate of γ_i from (1), and the bars represent the associated 90% confidence interval.

problem (2) – (4):

$$[k] \; A\left(\frac{\alpha}{\overline{uc}}\right)^{1-\eta}\left(\frac{\eta}{\overline{w}}\right)^{\eta} k^{\alpha+\eta-1} = 1+\lambda\kappa$$

$$[L'] \; 1-\frac{\phi}{2}\int\left(\frac{q\ell'(q,s)}{L'}-s+S-1\right)^2 \mathrm{d}\mu(q,s) - \phi L'\int\left(-\frac{q\ell'(q,s)}{(L')^2}\right)\left(\frac{q\ell'(q,s)}{L'}-s+S-1\right)\mathrm{d}\mu(q,s)$$
$$+\lambda-\zeta=0$$

$$[\ell'(q,s)] \; -\phi L'\frac{q}{L'}\left(\frac{q\ell'(q,s)}{L'}-s+S-1\right)\mathrm{d}\mu(q,s)+\overline{q}\mathbb{E}\left[V_{\mathcal{L}'}(\ell';\mu)\right]+\zeta q\mathrm{d}\mu(q,s)=0$$

With the envelope condition $W_\ell(\mathcal{L};\mu) = -\mathrm{d}\mu(q,s)$, we obtain the optimality conditions:

- for capital demand:

$$k = \left[\frac{A\left(\frac{\alpha}{\overline{uc}}\right)^{1-\eta}\left(\frac{\eta}{\overline{w}}\right)^{\eta}}{1+\lambda\kappa}\right]^{\frac{1}{1-\alpha-\eta}} \tag{B.2}$$

  Applying the binding working capital constraint (3) again, using (B.1) and (B.2) gives aggregate loan demand:

$$L' = \kappa(\alpha+\eta)\left[\frac{A\left(\frac{\alpha}{\overline{uc}}\right)^{\alpha}\left(\frac{\eta}{\overline{w}}\right)^{\eta}}{1+\lambda\kappa}\right]^{\frac{1}{1-\alpha-\eta}} \tag{B.3}$$

  where all we need to do is solve for λ.

- for bank-specific loan demand:[23]

$$\zeta = \frac{\overline{q}}{q} + \phi\left(\frac{q\ell'(q,s)}{L'} - s + S - 1\right) \text{ for all } (q,s) \tag{B.4}$$

  Recognizing that equation (B.4) holds for all (q,s), we can integrate the right-hand side over the distribution μ to obtain:

$$\zeta = \overline{q}\int \frac{\mathrm{d}\mu(q,s)}{q} + \phi\int \frac{q\ell'(q,s)}{L'}\mathrm{d}\mu(q,s) - \phi\int s\mathrm{d}\mu(q,s) + \phi(S-1) = \overline{q}R \tag{B.5}$$

  Plugging (B.5) back into (B.4) and rearranging terms gives us our bank-specific loan demand equation (5).

- for total loan demand:

$$\begin{aligned}
1+\lambda-\zeta &= \frac{\phi}{2}\int\left(\frac{q\ell'(q,s)}{L'} - s + S - 1\right)^2 \mathrm{d}\mu(q,s) - \frac{\phi}{L'}\int q\ell'(q,s)\left(\frac{q\ell'(q,s)}{L'} - s + S - 1\right)\mathrm{d}\mu(q,s) \\
&= \phi\int\left(\frac{q\ell'(q,s)}{L'} - s + S - 1\right)\left(\frac{1}{2}\left(\frac{q\ell'(q,s)}{L'} - s + S - 1\right) - \frac{q\ell'(q,s)}{L'}\right)\mathrm{d}\mu(q,s) \\
&= -\frac{\phi}{2}\left[\int\left(\frac{q\ell'(q,s)}{L'}\right)^2 \mathrm{d}\mu(q,s) - \int (s-S+1)^2 \mathrm{d}\mu(q,s)\right] \\
&= -\frac{\phi}{2}\left[\mathbb{V}_\mu\left(\frac{q\ell'}{L'}\right) - \mathbb{V}_\mu(s)\right]
\end{aligned} \tag{B.6}$$

We can use (5) to simplify (B.6):

$$\begin{aligned}
\mathbb{V}_\mu\left(\frac{q\ell'}{L'}\right) - \mathbb{V}_\mu(s) &= \mathbb{V}_\mu\left(1 + S - s - \frac{\overline{q}}{\phi}(r-R)\right) - \mathbb{V}_\mu(s) \\
&= \mathbb{V}_\mu(s) + \mathbb{V}_\mu\left(\frac{\overline{q}}{\phi}(r-R)\right) + 2\mathbb{C}_\mu\left(s, -\frac{\overline{q}}{\phi}(r-R)\right) - \mathbb{V}_\mu(s) \\
&= \left(\frac{\overline{q}}{\phi}\right)^2 \mathbb{V}_\mu(r) - 2\frac{\overline{q}}{\phi}\mathbb{C}_\mu(s,r)
\end{aligned} \tag{B.7}$$

This delivers the aggregate demand equation (6) since we now have

$$\begin{aligned}
1 + \lambda - \overline{q}R - \phi(1-S) &= -\frac{\phi}{2}\left[\left(\frac{\overline{q}}{\phi}\right)^2 \mathbb{V}_\mu(r) - 2\frac{\overline{q}}{\phi}\mathbb{C}_\mu(s,r)\right] \\
\implies \lambda &= \overline{q}\left[R + \mathbb{C}_\mu(s,r) - \frac{1}{2}\frac{\overline{q}}{\phi}\mathbb{V}_\mu(r)\right] - 1
\end{aligned}$$

where the term in brackets is equal to $\tilde{R}$ from equation (7) in the main text.

[23] A useful result here is if $X = \mathbb{E}[x]$, then $\mathbb{E}[x/X] = \mathbb{E}[x]/X = X/X = 1$, and similarly

$$\mathbb{V}\left(\frac{x}{X}\right) = \mathbb{E}\left[\left(\frac{x}{X} - \mathbb{E}\left(\frac{x}{X}\right)\right)^2\right] = \mathbb{E}\left[\left(\frac{x}{X} - 1\right)^2\right] = \mathbb{E}\left[\left(\frac{x}{X}\right)^2\right] - 2\mathbb{E}\left(\frac{x}{X}\right) + 1 = \mathbb{E}\left[\left(\frac{x}{X}\right)^2\right] - 1$$

B.2 Proof of Proposition 2: Bank Financing and Lending

Since $\psi'(e) > 0$, the budget constraint (9) must bind, and so we can eliminate e from the set of control variables. Mechanically, conditions (11), (12), and (13) must bind with $\ell(q,s)$ given by (5), and so we may further eliminate s', ℓ', and n'. This leaves us with a two-control-variable problem (dropping explicit dependence on μ to ease notation):

$$V(n,s,z) = \max_{q,d'} \psi\left(\overline{q}^d d' + z + n - q\ell(q,s)\right) + \overline{q}\mathbb{E}\left[(1-\pi)\psi(n'(q,d',s)) + \pi V\left(n'(q,d',s), s'(q,s); z'\right)\right]$$

subject to $\quad [\lambda] \quad \overline{q}^d d' \leq (1-\chi)q\ell(q,s)$

where it is understood that $n'(q,d',s) = \ell(q,s) - d'$ and $s'(q,s) = \rho_q \frac{q\ell'(q,s)}{L'} + \rho_s s$ (we keep these general for now). Taking first order conditions, we obtain:

$$\begin{aligned} \frac{\partial q\ell}{\partial q}\psi'(e) &= \overline{q}(1-\pi)\mathbb{E}\left[\psi'(n')\frac{\partial n'}{\partial q}\right] + \overline{q}\pi\mathbb{E}\left[V_n\left(n',s',z'\right)\frac{\partial n'}{\partial q} + V_s\left(n',s',z'\right)\frac{\partial s'}{\partial q}\right] \qquad \text{(B.8)} \\ &+ \lambda(1-\chi)\frac{\partial q\ell}{\partial q} \\ \overline{q}^d\psi'(e) &= -\overline{q}(1-\pi)\mathbb{E}\left[\psi'(n')\frac{\partial n'}{\partial d'}\right] - \overline{q}\pi\mathbb{E}\left[V_n\left(n',s',z'\right)\frac{\partial n'}{\partial d'}\right] + \lambda\overline{q}^d \qquad \text{(B.9)} \end{aligned}$$

The relevant envelope conditions are:

$$\begin{aligned} V_n(n,s,z) &= \psi'(e) \qquad \text{(B.10)} \\ V_s(n,s,z) &= q\frac{\partial \ell}{\partial s}\left[\lambda(1-\chi) - \psi'(e)\right] + \overline{q}(1-\pi)\mathbb{E}\left[\psi'(n')\frac{\partial n'}{\partial s}\right] \qquad \text{(B.11)} \\ &+ \overline{q}\pi\mathbb{E}\left[V_n(n',s',z')\frac{\partial n'}{\partial s} + V_s(n',s',z')\frac{\partial s'}{\partial s}\right] \end{aligned}$$

In addition, the ancillary derivatives for accumulating state variables are

$$\frac{\partial n'}{\partial q} = \frac{\partial \ell}{\partial q} \text{ and } \frac{\partial n'}{\partial s} = \frac{\partial \ell}{\partial s} \text{ and } \frac{\partial n'}{\partial d'} = -1 \qquad \text{(B.12)}$$

$$\frac{\partial s'}{\partial q} = \frac{\rho_q}{L'}\frac{\partial q\ell}{\partial q} \text{ and } \frac{\partial s'}{\partial s} = \frac{\rho_q}{L'}q\frac{\partial \ell}{\partial s} + \rho_s \qquad \text{(B.13)}$$

Finally, it is useful to define the expected marginal value of funds tomorrow as

$$\psi^e(e') \equiv \pi\psi'(n') + (1-\pi)\psi'(e')$$

Turning first to financing results, combining equations (B.12) and (B.10) with (B.9) yields

$$\lambda = \psi'(e) - \frac{\overline{q}}{\overline{q}^d}\pi\mathbb{E}\left[\psi^e(e')\right] \qquad \text{(B.14)}$$

If the capital requirement is slack, then the bank equates the marginal value of internal funds today to the expected marginal value of funds tomorrow. If the capital requirement is binding, then $\lambda > 0$ and the marginal value of funds today may exceed the expected value in the following period.

Before considering the pricing policy, it is useful to simplify the envelope condition for relationship intensity (B.11). Using (B.12) and (B.10) and switching to sequential notation, we can first write

$$V_{s,t} = q_t \frac{\partial \ell_{t+1}}{\partial s_t} \left[\underbrace{\lambda_t (1-\chi) - \psi'(e_t) + \frac{\overline{q}}{q_t} \mathbb{E}_t \psi^e(e_{t+1})}_{\equiv \Pi_t} \right] + \overline{q}\pi \frac{\partial s_{t+1}}{\partial s_t} \mathbb{E}_t \left(V_{s,t+1} \right)$$

where the term in brackets represents the static flow profits associated with an additional unit of lending defined in (15). From equation (5) we know that $\frac{\partial \ell'}{\partial s} = \frac{L'}{q}$ which implies that $\frac{\partial s'}{\partial s} = \rho_q + \rho_s$, so this can be simplified further:

$$V_{s,t} = L_{t+1}\Pi_t + \overline{q}\pi(\rho_q + \rho_s)\mathbb{E}_t \left(V_{s,t+1} \right) \tag{B.15}$$

Iterating on equation (B.15) yields

$$\begin{aligned} V_{s,t} &= L_{t+1}\Pi_t + \overline{q}\pi(\rho_q + \rho_s)\mathbb{E}_t \left[L_{t+2}\Pi_{t+1} + \overline{q}\pi(\rho_q + \rho_s)\mathbb{E}_{t+1} \left(V_{s,t+2} \right) \right] \\ &= \ldots \\ &= \sum_{i=0}^{\infty} \left(\overline{q}\pi(\rho_q + \rho_s) \right)^i L_{t+i+1}\Pi_{t+i} \end{aligned} \tag{B.16}$$

Next, combine equations (B.8), (B.12), (B.13), (B.10), and the simplifications above to obtain a modified version of the pricing optimality condition

$$\frac{\partial q\ell}{\partial q}\psi'(e) = \overline{q}\frac{\partial \ell}{\partial q}\mathbb{E}[\psi^e(e')] + \overline{q}\pi\frac{\rho_q}{L'}\frac{\partial q\ell}{\partial q}\mathbb{E}\left[V_s'\right] + \lambda(1-\chi)\frac{\partial q\ell}{\partial q}$$

We can simplify by dividing through by $\frac{\partial q\ell}{\partial q}$, which involves recognizing that since $\frac{\partial q\ell}{\partial q} = q\frac{\partial \ell}{\partial q} + \ell$,

$$\frac{\frac{\partial \ell}{\partial q}}{\frac{\partial q\ell}{\partial q}} = \frac{\frac{\frac{\partial q\ell}{\partial q} - \ell}{q}}{\frac{\partial q\ell}{\partial q}} = \frac{1}{q}\left(1 - \epsilon^{-1}(q\ell, q)\right)$$

where $\epsilon(q\ell, q)$ denotes the elasticity of total loan demand, $q\ell$, with respect to loan price, q, so that ϵ^{-1} is the inverse elasticity. Then, combining this expression and the simplified envelope condition (B.16), we obtain the expression from (14):

$$\Pi_t + \overline{q}\pi\rho_q\mathbb{E}_t\left[\sum_{i=0}^{\infty}\left(\overline{q}\pi(\rho_q + \rho_s)\right)^i \frac{L_{t+2+i}}{L_{t+1}}\Pi_{t+1+i}\right] = \epsilon^{-1}(q_t\ell_{t+1}, q_t)\frac{\overline{q}}{q_t}\mathbb{E}_t\left[\psi^e(e_{t+1})\right] \tag{B.17}$$

The last part of the proof is to give the form of the inverse elasticity term in equation (16). To derive this, simply compute the derivative of $q\ell$ with respect to q in equation (5):

$$\frac{\partial q\ell}{\partial q} = -L'\frac{\overline{q}}{\phi}\left(-\frac{1}{q}\right)^2 = \frac{\overline{q}}{\phi}\frac{L'}{q^2} \implies \epsilon(q\ell, q) \equiv \frac{\frac{\partial q\ell}{\partial q}}{\frac{q\ell}{q}} = \frac{1}{q}\frac{\overline{q}}{\phi}\frac{L'}{q\ell}$$

B.3 General adjustment cost function

Assume the quadratic adjustment cost function in (2) is replaced by:

$$L' \int \phi\left(\frac{q\ell'(q,s)}{L'}, s\right) \mathrm{d}\mu(q,s)$$

where $\phi(\cdot)$ is a generic penalty function that allows for a more general relationship between relative relationship intensity and loan share. Note that this specification still embeds that total adjustment costs scale with the total size of the loan portfolio.

Extending the same analysis from Appendix B.1 shows that this specification gives rise to the modified demand system:

$$-\overline{q}\,(r-R) = \phi_1\left(\frac{q\ell'(q,s)}{L'}, s\right) - \underbrace{\int \phi_1\left(\frac{q\ell'(q,s)}{L'}, s\right) \mathrm{d}\mu(q,s)}_{\equiv \Phi_1} \tag{B.18}$$

$$L' = \kappa(\alpha+\eta)\left[\frac{A\left(\frac{\alpha}{\overline{uc}}\right)^{\alpha}\left(\frac{\eta}{\overline{w}}\right)^{\eta}}{1+\tilde{\Lambda}(\mu)\kappa}\right]^{\frac{1}{1-\alpha-\eta}} \tag{B.19}$$

$$\text{where } \tilde{\Lambda}(\mu) = \overline{q}R + \underbrace{\int \phi\left(\frac{q\ell'(q,s)}{L'}, s\right) \mathrm{d}\mu(q,s)}_{\equiv \Phi} - \int \left(\frac{q\ell'(q,s)}{L'} - 1\right) \phi_1\left(\frac{q\ell'(q,s)}{L'}, s\right) \mathrm{d}\mu(q,s) - 1$$

Equation (B.18) is the analog of (5) in the main text; likewise, equation (B.19) is the analog of (6) in the main text. The former equation still takes the form of specifying loan demand as a function of a pricing penalty term and the marginal cost of relationship adjustment. Likewise, the latter specifies aggregate demand as a function of average interest rates, a term describing aggregate adjustment costs (akin to the covariance term in (7)), and marginal adjustment costs. In particular, assuming that ϕ_1 is invertible, we can write the demand function as

$$\frac{q\ell'(q,s)}{L'} = (\phi_1^{-1})\left(\Phi_1 - \bar{q}(r-R), s\right)$$

This demand function satisfies the same properties as the one that arises from quadratic adjustment costs as long as ϕ_1^{-1} is increasing in both of its arguments. That is, demand rises with more relationship intensity and/or with lower interest rate spreads.

Note the change of notation from $\tilde{R}(\mu)$ in the main text. This is because Λ is actually the multiplier on the working capital constraint, which measures the excess borrowing costs. The analog of $\tilde{R}$ in this context would be such that it solves $\tilde{\Lambda} = \overline{q}\tilde{R} - 1$, or

$$\tilde{R} = R + \bar{q}^{-1}\left[\Phi - \Phi_1 + \int \phi_1\left(\frac{q\ell'(q,s)}{L'}, s\right)\frac{q\ell'(q,s)}{L'} \mathrm{d}\mu(q,s)\right]$$

B.4 CES loan demand

This subsection describes the model with CES loan demand. The firm's problem can be written as

$$\begin{aligned} W(\mathcal{L};\mu) \quad &= \quad \max_{n,k,L',\{\ell'(q,s)\}} Ak^{\alpha}n^{\eta} - \overline{w}n - \overline{uc}k + L' - \int \ell(q,s)\mathrm{d}\mu(q,s) + \overline{q}\mathbb{E}\left[W(\mathcal{L}';\mu)\right] \\ \text{subject to} \quad & \quad \kappa(\overline{w}n + \overline{uc}k) \leq L' \\ & \quad L' \leq \left[\int (s^{\theta}q\ell'(q,s))^{\frac{\varepsilon-1}{\varepsilon}}\mathrm{d}\mu(q,s)\right]^{\frac{\varepsilon}{\varepsilon-1}} \end{aligned}$$

Note that we include the relationship term s directly in the CES for loan demand, in the spirit of how Gilchrist et al. (2017) interpret customer capital in product markets. θ is a parameter that affects how the relationship intensity influences the contribution of borrowing from a particular firm to total borrowing. This is interpreted as a preference shifter in the customer capital literature.

Define $\tilde{R}^{CES}$ as a habit-weighted geometric mean of interest rates:

$$\tilde{R}^{CES} \equiv \frac{1}{\left[\int (s^{\theta}q)^{\varepsilon-1}\mathrm{d}\mu(q,s)\right]^{\frac{1}{\varepsilon-1}}} \tag{B.20}$$

Then, we can show that the two-tier demand system becomes

$$\begin{aligned} \frac{q\ell'(q,s)}{L'} &= s^{\theta(\varepsilon-1)}\left(\frac{1/q}{\tilde{R}^{CES}}\right)^{-\varepsilon} \\ L' &= \kappa(\alpha+\eta)\left[\frac{A\left(\frac{\alpha}{\overline{uc}}\right)^{\alpha}\left(\frac{\eta}{\overline{w}}\right)^{\eta}}{1+\kappa(\bar{q}\tilde{R}^{CES}-1)}\right]^{\frac{1}{1-\alpha-\eta}} \end{aligned}$$

As it is well know, the price-elasticity of demand with respect to $R = 1/q$ is equal to $-\varepsilon$, and therefore does not vary either with price or the intensity of relationships.

We can then take logs of the individual demand function to write an estimable version:

$$\log\left(\frac{q\ell'(q,s)}{L'}\right) = -\varepsilon r + \varepsilon\log\tilde{R}^{CES} + \theta(\varepsilon-1)\log s$$

where we use the fact that $\log(1/q) \simeq -r$. The above condition can be estimated using the techniques described in the main body of the text. In this case, while $\tilde{R}^{CES}$ is no longer an average spread, it is subsumed in firm-time FE.

B.5 Kimball loan demand

This subsection describes a specification for loan demand following Kimball (1995). The firm's problem can be written as

$$\begin{aligned} W(\mathcal{L};\mu) \quad &= \quad \max_{n,k,L',\{\ell'(q,s)\}} Ak^{\alpha}n^{\eta} - \overline{w}n - \overline{uc}k + L' - \int \ell(q,s)\mathrm{d}\mu(q,s) + \overline{q}\mathbb{E}\left[W(\mathcal{L}';\mu)\right] \\ \text{subject to} \quad & \quad \kappa(\overline{w}n - \overline{uc}k) \leq L' \\ & \quad 1 = \int G\left(s^{\theta}\frac{q\ell'(q,s)}{L'}\right)\mathrm{d}\mu(q,s) \end{aligned}$$

where G is a general aggregator. We follow Dotsey and King (2005) in assuming that this aggregator takes the form

$$G(x) = \frac{\omega}{1+\omega\nu}\left[(1+\nu)x - \nu\right]^{\frac{1+\omega\nu}{\omega(1+\nu)}} + 1 - \frac{\omega}{1+\omega\nu}$$

Note that this aggregator becomes a standard CES when $\nu = 0$, with $\omega = \frac{\varepsilon-1}{\varepsilon}$. The relevant effective price in this case is defined as

$$\tilde{R}^K = \left[\int \left(\frac{1}{s^\theta}\frac{1}{q}\right)^{\frac{1+\omega\nu}{1-\omega}} \mathrm{d}\mu(q,s)\right]^{\frac{1-\omega}{1+\omega\nu}} \tag{B.21}$$

This allows us to write the bank-specific demand function as

$$\frac{q\ell'(q,s)}{L'} = \frac{1}{1+\nu}\frac{1}{s^\theta}\left[\left(\frac{1}{s^\theta}\frac{R}{\tilde{R}^K}\right)^{\frac{\omega(1+\nu)}{1-\omega}} + \nu\right] \tag{B.22}$$

where $R \equiv 1/q$. The firm's aggregate credit demand is still given by an expression of the form

$$L' = \kappa(\alpha+\eta)\left[\frac{A\left(\frac{\alpha}{\overline{uc}}\right)^{\alpha}\left(\frac{\eta}{\overline{w}}\right)^{\eta}}{1+\lambda\kappa}\right]^{\frac{1}{1-\alpha-\eta}}$$

where λ is now given by a more involved expression:

$$\lambda = \tilde{R}^K q \int \left\{\left[(1+\nu)\frac{1}{s^\theta}\frac{q\ell'(q,s)}{L'} - \nu\right]^{\frac{1-\omega}{\omega(1+\nu)}} \frac{1}{s^\theta}\frac{q\ell'(q,s)}{L'}\right\} \mathrm{d}\mu(q,s) - 1$$

The Kimball demand function has advantages and disadvantages over the CES specification. The main advantage is that unlike in the CES case, the price-elasticity of demand under Kimball (1995) is no longer constant and varies with both price and relationship intensity,

$$\epsilon(q\ell, R) = \frac{\omega(1+\nu)}{1-\omega}\frac{\left(\frac{1}{s^\theta}\frac{R}{\tilde{R}^K}\right)^{\frac{\omega(1+\nu)}{1-\omega}}}{\left(\frac{1}{s^\theta}\frac{R}{\tilde{R}^K}\right)^{\frac{\omega(1+\nu)}{1-\omega}} + \nu}$$

One disadvantage, however, is that the bank-specific demand function (B.22) no longer has a functional form that is amenable to direct estimation with linear methods. In particular, the right-hand side depends on the bank-specific interest rate, the relationship intensity term, and on the aggregate time-varying object $\tilde{R}^K$, and these terms cannot be disentangled with either linear or log-linear transformations of this expression.

B.6 Model with perfect competition

The perfectly competitive version of our model corresponds to the case in which there are no adjustment costs; that is, the case when $\phi = 0$ in the borrower's objective function (2). In this case, the state variable s is completely redundant. Furthermore, there is no reason for the borrower to diversify its loan portfolio, and in fact bank-specific demand is not well-defined and so in equilibrium all banks must charge the same loan price, $Q = R^{-1}$.

Correspondingly, the problem of the borrower is simply to choose labor, capital, and total loan demand per the following problem:

$$W(\mathcal{L};R) \quad = \quad \max_{n,k,L'} Ak^{\alpha}n^{\eta} - \overline{w}n - \overline{uc}k + \frac{L'}{R} - L + \overline{q}\mathbb{E}\left[W(\mathcal{L}';R)\right] \tag{B.23}$$

$$\text{subject to} \qquad \kappa(\overline{w}n + \overline{uc}k) \leq \frac{L'}{R} \tag{B.24}$$

The objective function (B.23) is modified relative to the original objective (2) to reflect that there are no loan sourcing considerations in this model and there is only a single equilibrium loan price. As a result, the loan sourcing constraint (4) is obviated in this version of the model. Finally, observe that the working capital constraint (B.24) is the same as the original constraint (3), with the modification that discount prices are accounted for directly on L' rather than on the individual ℓ'. The solution to this problem yields the same aggregate demand curve as in equation (6), with the modification that the effective interest rate $\tilde{R}(\mu)$ is replaced by the single equilibrium interest rate R:

$$L'(R) = \kappa(\alpha+\eta)\left[\frac{A\left(\frac{\alpha}{\overline{uc}}\right)^{\alpha}\left(\frac{\eta}{\overline{w}}\right)^{\eta}}{1+\kappa(\overline{q}R-1)}\right]^{\frac{1}{1-\alpha-\eta}} \tag{B.25}$$

The problem of the banks is similarly stripped down:

$$V\left(n,z;R\right) \quad = \quad \max_{e,\ell'\geq 0,d',n'} \psi(e) + \overline{q}\pi\mathbb{E}\left[V\left(n',z';R\right)\right] \tag{B.26}$$

$$\text{subject to} \qquad q\ell' + e \leq n + z + \overline{q}^{d}d' \tag{B.27}$$

$$\chi q\ell' \leq q\ell' - \overline{q}^{d}d' \tag{B.28}$$

$$n' = \ell' - d' \tag{B.29}$$

The only change in the objective function in (B.26) relative to the baseline (8) is the elimination of the state variable s from the value function and the removal of the loan price q from the set of control variables. Constraints (B.27), (B.28), and (B.29) are identical to their counterparts from the baseline model, (9), (10), and (13), respectively. Since banks do not face bank-specific demand curves and the state variable s has no meaning in this version of the model, constraints (12) and (11) become irrelevant in this case.

A stationary recursive competitive equilibrium for this version of the model is defined in the standard way. The main differences relative to the equilibrium definition from the main text are that now borrower optimality specifies only aggregate demand, and the distributional consistency condition is replaced by the simple market clearing condition that aggregate demand equals aggregate supply, integrated across the entire equilibrium distribution of banks.

$$L'(R) = \int \ell'(n,z)\mathrm{d}m(n,z) \tag{B.30}$$

B.7 Fixed relationship model

In the "fixed relationship" version of our model, we assume that each bank draws a permanent s from the equilibrium distribution of s from the baseline economy, $\tilde{m}(s)$. We assume that each bank faces the same bank-level demand curve conditional on s as in the baseline model (i.e. equation (5)), but that banks' lending policies and therefore the equilibrium values of R and $\tilde{R}$ (and therefore L') adjust. The key change in this model relative to the baseline model, therefore, is that banks in this version of the model have no dynamic incentive to preserve their market power, since the demand shifter s is fixed. Concretely, banks in this model solve:

$$\begin{aligned} V(n,s,z;\mu) \quad &= \quad \max_{q,e,\ell'\geq 0,d',n'} \psi(e)+\overline{q}\mathbb{E}\left[(1-\pi)\psi(n')+\pi V(n',s,z';\mu)\right] \\ \text{subject to:} \quad & \quad q\ell'+e\leq n+z+\overline{q}^d d' \\ & \quad \chi q\ell' \leq q\ell'-\overline{q}^d d' \\ & \quad \ell'=\ell(q,s;\mu) \\ & \quad n'=\ell'-d' \end{aligned} \tag{B.31}$$

The only change in this problem relative to the baseline in Section 3.2 is that there is no evolution from s to s' based on the law of motion (11). This of course implies that we can not determine the average s, and therefore the demand shifter $1-S$, in the same way as the baseline model. In order to standardize between the baseline model and this version, then, we simply set S in the fixed relationship model equal to the value from the baseline.

C Computational Algorithms

C.1 Bank problem

The solution to the bank problem takes the current distribution of prices and relationships, μ, as given, and so the notation in this section suppresses that argument.

1. **Compute continuation value.** Given a current guess V_0 of V, compute the expected continuation value (denoted $\overline{W}$) for all (n,s,z):

$$\overline{W}(n,s,z)=(1-\pi)\psi(n)+\pi\sum_{z'}\Gamma(z,z')V_0\left(n,s,z'\right) \tag{C.1}$$

2. **Solve for optimal policies.** Fix (n,s,z). Solve for optimal policies and the updated value function by considering two candidate policies: (i) one with a binding capital requirement binds and unrestricted dividends; and (ii) one with a slack capital requirement and non-negative dividends. This approach leverages the result that if the capital requirement is slack, the bank will not issue equity, which follows from $\overline{q}^d>\overline{q}$ and the kink in the equity issuance cost function at 0.

   In practice, this approach has three advantages. First, finding optimal policies while assuming the capital requirement binds is a straightforward univariate search, since the d' associated with each q is determined by the capital requirement, and the dividend is determined by the flow budget constraint. Second, as will be shown below, the set of states and prices for which a slack capital requirement is feasible is much smaller than those for which a binding capital requirement is feasible, and so we can eliminate some more costly bivariate search. Third, given the kink in the equity issuance cost function at zero, this approach is more stable numerically.

   (a) *Binding capital requirement:* Solve for optimal q policies using golden section search. At each candidate q, the implied loan demand $\ell(q,s)$ and next period customer capital $s'(q;s)$ via (5) and (11), respectively, the associated deposits $d'(q;s)$ are determined by the binding capital requirement (10), the associated dividend $e(q;n,s,z)$ is determined by the flow budget constraint (9), and the implied next period net worth $n'(q,d';s)$ is given by the law of motion (13). The search range over q is bounded below by the lowest price such that loan demand is non-negative, which is computed from (5):

$$q_{\min}(s) = \left[R(\mu) + \frac{\phi}{\overline{q}}(1 - S + s)\right]^{-1} \tag{C.2}$$

Set an arbitrary upper bound $q_{\max}$ to define the search area; after solving be sure to check that this does not bind. For each candidate q, given the implied d', the value of the action is given by(q, d') according to

$$v(q, d'(q;s); n, s, z) = \psi\left(\overline{q}^d d'(q;s) + n + z - q\ell(q,s)\right) + \overline{q}\overline{W}\left(n'(q, d'(q;s); s), s'(q;s), z\right) \tag{C.3}$$

This step requires interpolation on n' and s'. Denote the value associated with the binding capital requirement by

$$v_b(n, s, z) = \max_{q \in [q_{\min}(s), q_{\max}]} v(q, d'; n, s, z) \tag{C.4}$$

(b) *Slack capital requirement:* Implement *nested* golden section with the q choice as the outer loop and deposits d' in the inner loop. For each candidate q, we consider deposits d' in the range $[d_{\min}(q; n, s, z), d_{\max}(q; s)]$, where $d_{\max}(q; s)$ comes from the binding capital requirement (10) and $d_{\min}$ comes from the restriction that dividends must be non-negative and the flow budget (9):

$$d_{\min}(q; n, s, z) = \frac{q\ell(q;s) - n - z}{\overline{q}^d} \tag{C.5}$$

In order for this interval to be well-defined, we must have $d_{\min} < d_{\max}$, which occurs when $\chi q\ell'(q,s) \leq n + z$, or $q \leq \tilde{q}(n, s, z)$ where:

$$\tilde{q}(s, n, z) = \left[R + \frac{\phi}{\overline{q}}\left(1 - S + s - \frac{n+z}{\chi L'}\right)\right]^{-1} \tag{C.6}$$

Similarly, it must be the case that for the lowest feasible non-negative dividend ($e = 0$), $\overline{q}^d d' \geq 0$. Therefore we must have $q\ell'(q, s) \geq n + z$, or $q \geq \hat{q}$ where:

$$\hat{q}(n, s, z) = \left[R + \frac{\phi}{\overline{q}}\left(1 - S + s - \frac{n+z}{L'}\right)\right]^{-1} \tag{C.7}$$

Therefore, the optimal policy for a slack capital requirement can be determined by solving:

$$w_s(q; n, s, z) = \max_{d' \in [d_{\min}(n,s,z), d_{\max}(s)]} v(q, d'; n, s, z) \tag{C.8}$$

$$v_s(n, s, z) = \max_{q \in [\hat{q}(n,s,z), \tilde{q}(n,s,z)]} w_s(q; n, s, z) \tag{C.9}$$

The implied dividend, loan demand, future net worth, and future relationships are determined exactly as in the case of the binding capital requirement. Note that a slack capital requirement is only feasible if the price interval $[\hat{q}, \tilde{q}]$ is non-empty, which only occurs when $n + z \geq 0$; otherwise, the capital requirement must bind and the bank issues equity.

(c) If both policies are feasible, set the update V_1 of V to

$$V_1(n, s, z) = \max\{v_b(n, s, z), v_s(n, s, z)\} \tag{C.10}$$

If only the binding capital requirement policies are feasible, set $V_1(n, s, z) = v_b(n, s, z)$. Assign the policy functions to be those from the relevant sub-problem as well.

3. **Evaluate convergence of the value function and decision rules.** Let $\Delta_V = \max_x |V_1(x) - V_0(x)|$, $\Delta_q = \max_x |q_1(x) - q_0(x)|$, and $\Delta_{d'} = \max_x |d'_1(x) - d'_0(x)|$. If $\max_{i \in \{V,q,d'\}} \{\Delta_i\} < \varepsilon$, a pre-specified tolerance parameter, then the problem is solved; otherwise, set $V_0 = V_1$, $q_0 = q_1$, and $d'_0 = d'_1$ and return to step 1.

C.2 Steady state

1. Begin with a guess of bank pricing and deposit policies $q_0(x)$ and $d'_0(x)$, the bank value function $V_0(x)$, and the distribution of banks over states $m_0(x)$.

2. Use the consistency condition (18) to obtain the $\mu(q, s)$ implied by $(q_0(x), m_0(x))$. Use (7) to compute the implied $\tilde{R}_0$ and R_0. Given $\tilde{R}_0$, compute L'_0.

3. Solve for banks' updated policies and value function given R, L' and $V_0(x)$ using the algorithm described above. Denote these objects by $q_1(x)$, $d'_1(x)$ and $V_1(x)$.

4. Solve for the distribution of banks over idiosyncratic states implied by the policies above, $m_1(x)$; that is, iterate to convergence on equation (17).

5. Compute the aggregates implied by $(m_1(x), q_1(x))$; denote these objects by $\tilde{R}_1$ and R_1.

6. Assess convergence of the bank policies, value, and distribution, as well as the equilibrium aggregates R, $\tilde{R}$, and L'. That is, compute $\delta_V = \max_x |V_1(x) - V_0(x)|$, $\delta_q = \max_x |q_1(x) - q_0(x)|$, $\delta_{d'} = \max_x |d'_1(x) - d'_0(x)|$, $\delta_m = \max_x |m_1(x) - m_0(x)|$, $\delta_R = |R_1 - R_0|$, and $\delta_{\tilde{R}} = |\tilde{R}_1 - \tilde{R}_0|$. If $\max_{i \in \{V,q,d',m,R,\tilde{R}\}} \delta_i < \bar{\varepsilon}$, a pre-specified tolerance, then the model is solved. Otherwise, we set $V_0 = V_1$, $q_0 = q_1$, $d'_0 = d'_1$, $m_0 = m_1$, $R_0 = \psi R_0 + (1 - \psi) R_1$, and $\tilde{R}_0 = \psi \tilde{R}_0 + (1 - \psi) \tilde{R}_1$, where $\psi \in (0, 1)$ is a relaxation parameter, and return to step 1.

C.3 Perfect foresight transitions / impulse responses

The maintained assumption throughout these steps is that both the initial and terminal steady states are known, that the initial distribution of banks over idiosyncratic states may be computed directly given the initial steady state, and that bank policies may be solved backwards given the value function implied by the terminal steady state.

1. Update the initial distribution of banks over idiosyncratic states, $m_0(x)$, to reflect the incidence of the shock being simulated.

2. Guess a sequence of aggregate prices $\{\tilde{R}^0_t, R^0_t\}_{t=1}^T$. A natural initial guess is that these prices are equal to their steady state values at all dates t.

3. Using the terminal value function $V_{T+1}(x)$ and the path of aggregate prices computed in the step above, solve backwards to obtain the sequence of bank policy functions $G = \{q_t(x), d_{t+1}(x)\}_{t=1}^T$.

4. Given the sequence of policy functions G, compute the implied sequence of distributions of banks over idiosyncratic states, $M = \{m_t(x)\}_{t=1}^{T}$.

5. Use the consistency condition to compute the implied sequence of aggregate prices $\{\tilde{R}_t^1, R_t^1\}$ consistent with this sequence of distributions.

6. Assess the convergence of aggregate prices: that is, compute $\Delta_{\tilde{R}} = \max_t \left|\tilde{R}_t^1 - \tilde{R}_t^0\right|$ and $\Delta_R = \max_t \left|R_t^1 - R_t^0\right|$ If $\max\{\Delta_{\tilde{R}}, \Delta_R\} < \bar{\varepsilon}$, the transition path has been solved. Otherwise, update the guesses of the aggregate price via relaxation as in the steady state algorithm above and return to step 2.

D Additional Quantitative Results

D.1 Spread and lending policies across model economies

Figure D.1 is the analog of Figure 4 from the main text, except we show lending policies directly (expressed in units of relative loan volume, $\frac{q\ell'}{L'}$), rather than spreads. Panel (a) shows that the joint correlation between loan shares and net worth follows the pattern for relationships and net worth, a direct by-product of the accumulation process (11). Panel (b) shows that lending policies are elastic with respect to net worth in only the bottom quartile of the net worth distribution in the relationship models, but for the bottom 50% in the competitive case. For higher levels of net worth, these policy functions are essentially flat. Panel (c) shows that lending is very sensitive to customer capital in the bottom quartile of relationships in the baseline and low elasticity models, while it is flat for the competitive model and very sensitive to relationships over the whole range of s for the low punishment model. This last result reflects the extremely strong incentives to build up relationships in a world of highly persistent relationships.

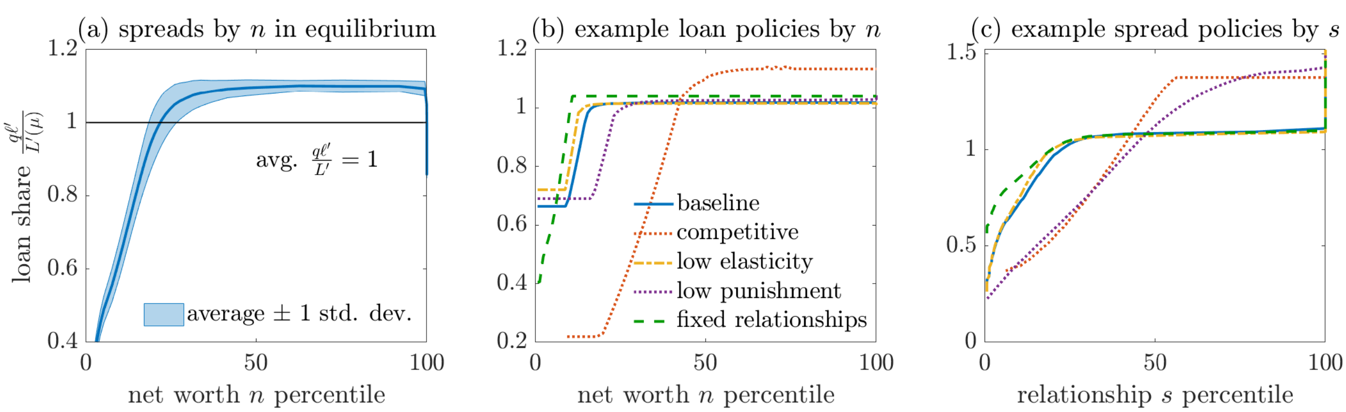

Figure D.1: **Steady state lending policies**

Notes: Panel (a) plots the first and second moments of the lending policies (expressed as loan volume relative to average) conditional on net worth over the distribution of banks in the baseline model. Variation at each level of n comes from dispersion in s and z. Panels (b) and (c) plot sample lending policy functions over net worth and relationships, respectively, for the indicated model variants. Panels (b) and (c) each fix $z = 0$, and panel (b) / (c) fixes s / n at the 25th percentile from the baseline economy.

Figure D.2 provides more detail on Figure 4 from the main text by comparing the baseline model to the low elasticity and low punishment models, rather than only the competitive model. The main result here, echoing the results from above, is that the low punishment world features much greater sensitivity of policies to relationships.

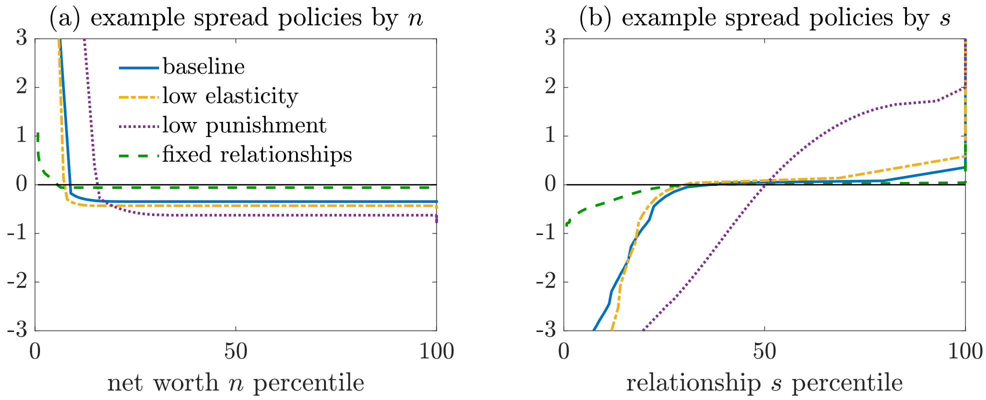

Figure D.2: **Steady state loan pricing policies**

Notes: Panels (a) and (b) plot sample pricing policy functions over the equilibrium distribution of net worth and relationships, respectively, for the indicated model variants. Panels (a) and (b) each fix $z = 0$. Panel (a) fixes s, and panel (b) fixes n, at the 25th percentile from the baseline economy.

D.2 Capital buffers in alternative models

Figure D.3 is the analog of Figure 8 from the main text but with information on the low elasticity, low punishment, and fixed relationship models. The low elasticity and low punishment models both feature lower average capital buffers owing to banks' ability to use their market power to extract surplus in response to a financial shock. In the fixed relationship model, the shape of the distribution matches the baseline by construction, but is shifted right due to the recalibrated level of static market power.

D.3 How do relationships affect individual banks' responses to financial shocks?

One of key questions we investigate in this paper is: how does the presence of relationships affect the way banks respond to financial shocks? To this end, we investigate individual banks' responses to a negative shock to net worth across the stationary distribution. The average responses for each variant of the model are presented in Figure D.4. Panel (a) shows that the recovery of net worth is much faster in the baseline model than the competitive model. Panel (b) explains why: in the baseline, the bank expends relationship capital to rebuild net worth. This is done by raising spreads (panel (d)) and cutting loan volume (panel (c)). In the competitive case, instead, there can be no positive spreads: banks must cut back on lending, and the reaccumulation of capital is slower with lower interest rates. This experiment showcases how relationship capital can help serve as a buffer for banks to weather shocks to their financial capital.

Comparing the baseline to the low elasticity and low punishment versions of the model further highlights these dynamics. In each of the less competitive versions of the model with relationships, the bank has more market power and therefore more ability to lend less at higher rates. This leads the banks to recapitalize even more quickly in each of these cases. In the low elasticity case, this is accomplished through extremely high spreads on impact, whereas in the low punishment case, this is accomplished through lower but more persistent increases.

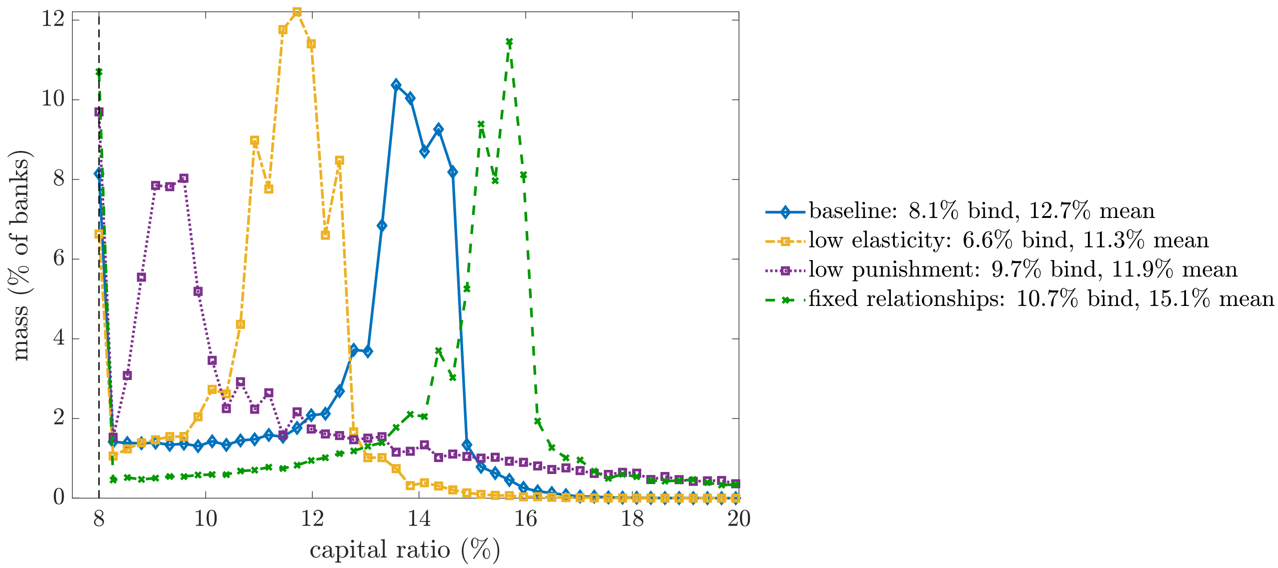

Figure D.3: **Capital buffers and relationships**

Notes: This figure represents the distribution of capital buffers across economies. Each is depicted as a probability mass function with 50 bins between the minimum, $\chi = 8\%$, and 20% (the share of banks with buffers above 20% in negligible in each model). The legend includes the share of banks for whom the capital requirement actually binds. The capital buffer is defined as the ratio of book equity to book assets, $\frac{g_q(x)g_{\ell'}(x) - \overline{q}^d g_{d'}(x)}{g_q(x)g_{\ell'}(x)}$. The data refers to the total capital ratio reported in form FR Y-9C (BHCA7205), averaged across the four quarters of 2019.

D.4 Response to aggregate shocks across model economies

Figures D.5 –D.7 plot the responses of selected variables to shocks to the aggregate shocks we consider in the main text in Figures 9 – 11 in the baseline, low elasticity, and low punishment economies. The patterns documented here broadly mirror the baseline economy, with any differences explainable by the increased market power in these alternatives relative to the baseline.

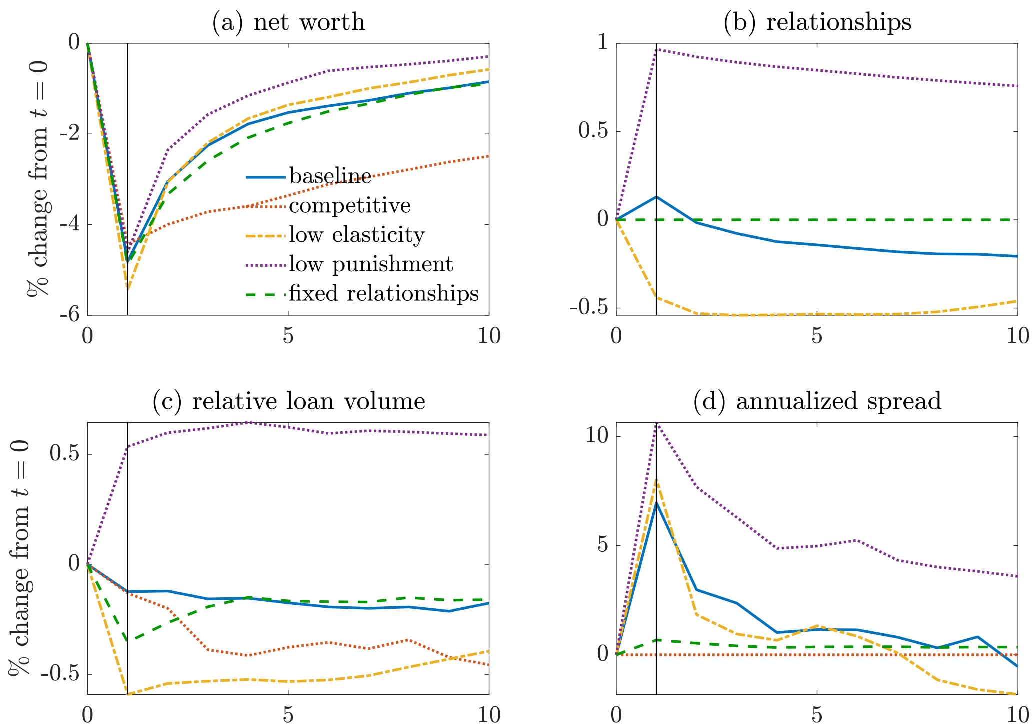

Figure D.4: **Average recoveries from an idiosyncratic shock to net worth**

Notes: In each panel, each line is constructed as follows. First, we simulate a panel of banks for T periods from the stationary distribution of the model. Then, we simulate the same banks (initial conditions and idiosyncratic shocks) with a one-time exogenous 5% drop in net worth at date 1. Last, for each indicated variable, we compute the difference between these two paths for each bank, average across banks, and then plot the result. This differencing procedure controls for the natural evolution of banks resulting from the life cycle features of the model.

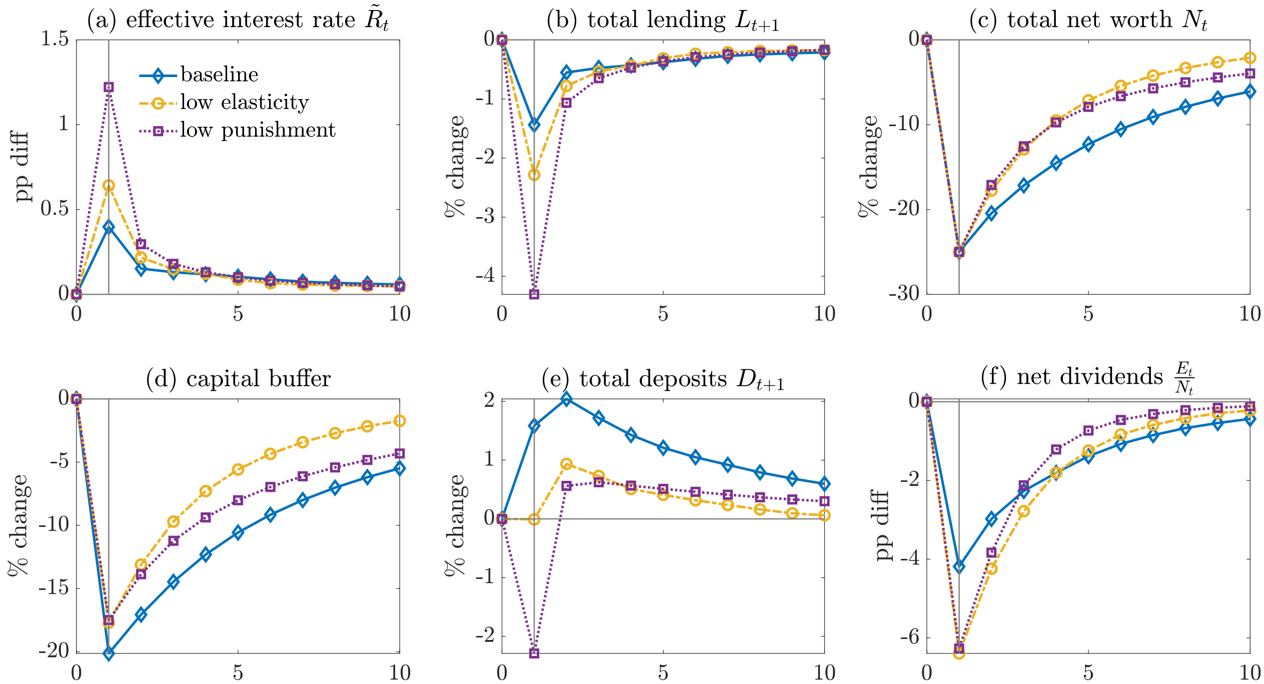

Figure D.5: **Aggregate shock to bank net worth: alternative relationships**

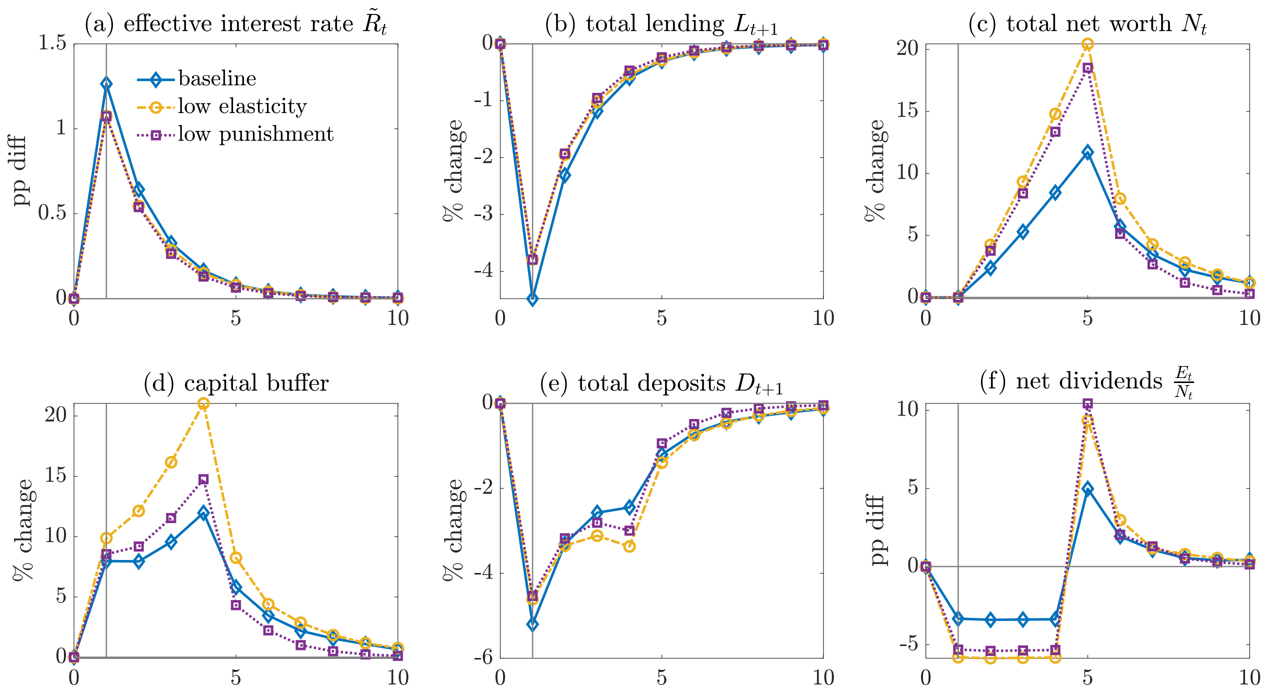

Figure D.6: **Aggregate shock to cost of funding: alternative relationships**

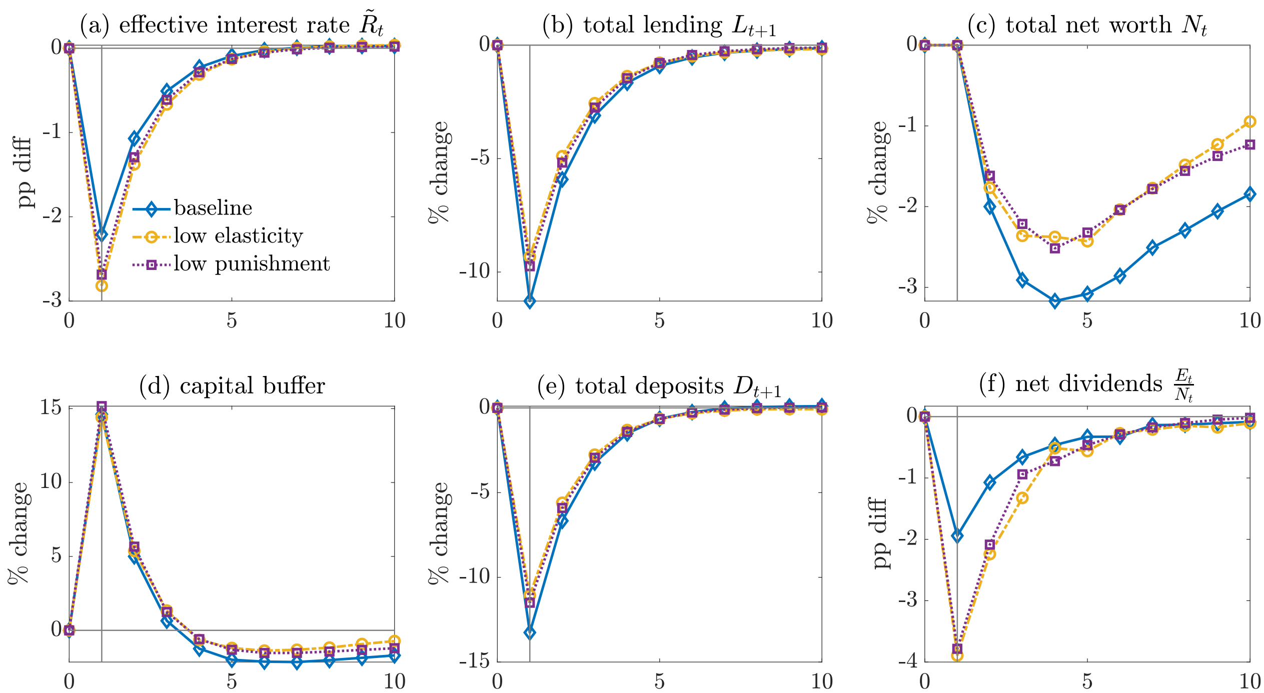

Figure D.7: **Aggregate shock to loan demand: alternative relationships**