Macro Shocks and Firm Dynamics with Oligopolistic Financial Intermediaries

Alessandro T. Villa*

June 2, 2025

Abstract

Motivated by a secular increase in the concentration of the U.S. banking industry, I develop a new macroeconomic model with oligopolistic financial intermediaries and heterogeneous firms. Market power allows banks to price discriminate and charge firm-specific markups, exerting stronger market power over productive firms that are more financially constrained. This dampens capital accumulation and amplifies the effects of macroeconomic shocks. During a crisis, banks exploit the higher number of financially constrained firms to extract higher markups, inducing a larger decline in real activity. When a big bank fails, the remaining banks use their increased market power to restrict credit supply, worsening and prolonging the downturn.

JEL Codes: D43, E44, G12, G21, L11.

Keywords: Dynamic Financial Oligopoly, Endogenous Financial Markups, Heterogeneous Firms, Firm Dynamics, Micro-Founded Financial Frictions, Price Discrimination.

*Federal Reserve Bank of Chicago. Mailing address: 230 S. La Salle St., Chicago, IL 60604, USA. Email: alessandro.villa@chi.frb.org. Acknowledgments: I am indebted to my advisors Andrea Lanteri, Lukas Schmid, Giuseppe Lopomo, Francesco Bianchi, and Pietro Peretto for their support and advice. I would also like to thank Joel David, François Gourio, Leonardo Melosi, Pablo D'Erasmo, Nicola Persico, Stijn G. Van Nieuwerburgh, Dean Corbae, Vincenzo Quadrini, Robert Barsky, Virgiliu Midrigan, Roberto Steri, Fabio Schiantarelli, Matthias Kehrig and Cosmin L. Ilut for their helpful comments and encouragements. I benefited from fruitful conversations with numerous seminar participants at Computing in Economics and Finance 2019, BFI Macro Finance Research Program 2020, Econometric Society/Bocconi World Congress 2020, University of Surrey, Duke University, University of Essex, Boston College, Chicago Fed, University of Washington, Federal Reserve Board, Penn State University, and University of Wisconsin-Madison. I declare that I have no relevant or material financial interests that relate to the research described in the paper. All errors are my own.

1 Introduction

The banking industry has become increasingly concentrated over the past two decades, with the asset market share of the five largest U.S. banks rising from 26% in 1996 to 50% in 2018. Moreover, the Lerner index – a metric of market power in the banking industry – increased from 0.2 in 1996 to 0.33 in 2014, pointing to a sizable increase in markups.[1] A large and influential literature has studied the interactions between financial markets, firm, and aggregate dynamics but has typically assumed perfectly competitive financial intermediaries; thus, it does not speak to this trend of increasing banking sector consolidation and markups.[2] Moreover, extensive empirical evidence suggests that banks' market power affects younger and older firms differently.[3]

In this paper, I study the role of imperfect competition in the financial intermediation sector for firm investment and financing dynamics, as well as for the transmission of macroeconomic shocks. First, I develop a dynamic general equilibrium model that incorporates an oligopolistic financial sector (including entry and exit decisions) with heterogeneous firms. The framework formalizes the notion that banks' market power has heterogeneous effects across firms with different characteristics, such as size, age, or productivity. Imperfect competition enables financial intermediaries to charge firm-specific markups that depend on the idiosyncratic characteristics of the firms to which they lend. In particular, banks exert a higher degree of market power on firms that are more financially constrained and have a high marginal productivity of capital. These firms have worse outside options (e.g., a high cost of non-bank finance); hence, they exhibit a higher and less elastic demand for credit. This mechanism creates endogenous financial frictions that slow the growth of firms that operate in more concentrated credit markets. Second, I show that although bank market power persistently dampens capital accumulation, its effects intensify during downturns—when firms' credit needs are greatest—thereby amplifying the transmission of macroeconomic shocks. During a crisis, oligopolistic banks exploit the higher number of financially constrained firms to extract higher markups, inducing a larger misallocation of credits (hence, capital) and a larger decline in real activity. Notably, since my model features non-atomistic banks, I can study market structure changes in the intermediation sector (e.g., the failure of a large intermediary and a new bank entry). When a single big bank fails, surviving banks use their increased market power to restrict credit supply, which amplifies and prolongs the recession. The results suggest that banks' market power should be an important source of concern for policymakers deciding whether to bail out a large intermediary.

[1] See the right panel of Figure 1.

[2] See, for instance, Bernanke and Gertler (1989), Covas and den Haan (2011), Jermann and Quadrini (2012), Khan and Thomas (2013), and Midrigan and Xu (2014).

[3] See, for example, Petersen and Rajan (1995), Rajan and Zingales (1998), Cetorelli and Gambera (2001), Black and Strahan (2002), and Cetorelli and Strahan (2006). Firms reliant on external funding via bank loans, e.g. small and private firms, can become financially constrained when credit conditions tighten (Holmstrom and Tirole, 1997; Diamond and Rajan, 2005; and Chodorow-Reich, 2013). See Subsection 5.2.1 for details.

In summary, the paper makes three contributions. First, to the best of my knowledge, I am the first to develop a macroeconomic model that incorporates oligopolistic banks and heterogeneous firms that formalizes the idea that banks' market power has different effects across firms with different characteristics. The model reveals a mechanism of endogenous financial frictions through which bank competition can play a role in shaping the speed at which firms grow, impacting aggregate productivity and output. More precisely, limited competition enables banks to price discriminate and charge firm-specific markups, exerting a higher degree of market power on firms with a high marginal productivity of capital that are more financially constrained. The resulting dispersion of markups induces credit – and thus capital — misallocation reducing aggregate productivity. Second, I examine how bank market power shapes the transmission of aggregate shocks and find that a marginally more restricted credit supply during periods of elevated demand dampens capital accumulation, thereby amplifying the effects of these shocks. Third, I make a methodological contribution by extending existing heterogeneous firms algorithms to solve for the stationary equilibrium and transitional dynamics in presence of a continuum of non-competitive markets and strategic interactions by exploiting generalized Euler equations.[4]

Succinctly, the model works as follows. Firms make optimal capital structure decisions by balancing equity and debt financing, generating an endogenous dynamic demand for loans. Banks are large (i.e., non-zero mass) players and firms are a continuum of followers in a Stackelberg fashion (i.e., each financial intermediary takes firms' dynamic demand for loans as given and competes to supply funding to each individual firm). Intermediaries make strategic decisions by internalizing the effect of their actions on present and future banks, firms' decisions, and on the aggregate economy. In such an environment, the bank's optimal equilibrium choice of loan supply is determined by solving the aforementioned generalized Euler equations. Each generalized Euler equation is a standard Euler equation, except that it contains a firm-specific elasticity that measures the sensitivity of the future interest rate with respect to the current supply of loans. The model generates firm-specific credit spreads that accrue to banks, which include default premia and markups.

The analysis proceeds in two main steps. First, I develop a stylized two-period model that I use to derive analytical insights on the role of oligopolistic intermediaries for firm dynamics and aggregate outcomes. Second, I build an infinite-horizon version of the model and discuss the impact of the main economic mechanism in greater details. I use the model to show that the time-varying cross-sectional effects of financial markups play a significant role in amplifying the impacts of macroeconomic shocks. I calibrate the model to match several financial and macroeconomic variables using Federal Deposit Insurance Corporation (FDIC) data. Compared to a perfectly competitive setting, the calibrated model incorporating both

[4]Note that the optimal fiscal policy literature uses generalized Euler equations and Markov perfect equilibria in macroeconomics (Klein and Ríos-Rull, 2003; Klein, Krusell, and Ríos-Rull, 2008; and Clymo and Lanteri, 2020).

bank market power and bank default predicts an amplification of aggregate credit spreads of 25 basis points at the peak of the Great Recession, resulting in a 0.5 percentage point decline in output. In the long term, the permanent loss of a major intermediary raises credit spreads by 15 basis points, leading to a sustained 0.2 percentage point decline in output.

In particular, I use the model to investigate the role that banks' market power plays in the transmission of three unexpected aggregate shocks: (i) a decrease to aggregate Total Factor Productivity (TFP) calibrated to the size of the Great Recession and, per se, not sufficiently large to induce a bank failure, (ii) a temporary change to the bank market structure (in the model, the first shock combined with an idiosyncratic shock to the assets of the big banks calibrated to induce one bank to default with a subsequent new bank entry), and (iii) the second shock combined with a permanent increase to the fixed entry cost in the credit market (that induces a permanent change to the bank market structure in the spirit of capturing the long run trend of consolidation in the banking sector). I find that in each of these cases, bank market power plays a significant role in shaping — and in particular, amplifying – the economy's response to the exogenous shock.

A decrease to aggregate TFP, with an associated increase to the aggregate probability of firm default, induces a higher proportion of young, more financially constrained firms with a high marginal productivity of capital. In these conditions, a more concentrated banking sector can control the supply of credit more tightly by extracting higher markups from these firms, leading to higher interest rates. This mechanism allows banks to compensate for the larger losses due to defaults, but it leads to a larger decline in real activity, amplifying the recession. When this shock is also combined with a lower firm entry rate, then imperfect competition in the financial intermediation sector leads to a bigger and delayed amplification effect at the peak that fades away as new firms enter the production sector.[5]

When one large financial intermediary fails, the change in market structure lowers the supply of credit to firms, slowing down the economy. The surviving banks extend more credit to firms in order to capture the market share of the defaulted bank. However, the speed of this adjustment is dampened by the decreased level of competition. As a result of both credit constraints and market power, the aggregate volume of credit drops sharply in the short run. In the long run, I analyze two different scenarios. First, I analyze a scenario where one bank reenters the credit market as the financial and economic conditions ease. Second, I analyze a scenario where no bank reenters (in the model, this is obtained through a permanent increase to the fixed entry cost in the credit market). In the former case, it takes several quarters for the economic conditions to ease sufficiently so that it is profitable for one bank to reenter. This persistently depresses investment, TFP, and output. In the latter case, the resulting increase in banks' market power further amplifies and prolongs the recession and, in the long run, the economy stabilizes at a lower level of total credit, which results in permanent less investment, output, and productivity. The result suggests

[5]For instance, during the Great Recession we observed a decline of the firm entry rate.

that banks' market power may be an important source of concern for policymakers deciding whether to bail out a big bank.

Related Literature This paper contributes to two main strands of literature: (i) firm dynamics under financial frictions, and (ii) macroeconomic models with imperfectly competitive financial intermediaries. A key contribution is to integrate these strands by studying how bank market power affects firm-level investment and financing behavior in a heterogeneous firm environment. While models of firm dynamics often assume competitive credit markets, and macro-finance models with bank competition typically abstract from firm heterogeneity, I bridge this gap by embedding a dynamic financial oligopoly in a heterogeneous firm framework.

Firm Dynamics and Credit Markets. A large literature studies firm dynamics under borrowing constraints or financial frictions (e.g., Gomes, 2001; Cooley and Quadrini, 2001; Hennessy and Whited, 2005, 2007; Jermann and Quadrini, 2012; Khan and Thomas, 2013). However, these models typically assume competitive credit markets. This paper departs from that assumption by introducing imperfectly competitive banks that set cross-sectional financial markups, thereby extending standard macroeconomic frameworks of firm dynamics to include an endogenous, imperfectly competitive credit supply mechanism.

Strategic interaction among banks generates endogenous dispersion in borrowing costs, which in turn affects capital allocation. This mechanism connects to recent work emphasizing dispersion and misallocation (e.g., Lanteri, 2018; David, Schmid, and Zeke, 2022) and aligns with empirical evidence showing that credit market structure can shape firm-level outcomes (e.g., Petersen and Rajan, 1995; Rajan and Zingales, 1998; Black and Strahan, 2002; Cetorelli and Gambera, 2001; Cetorelli and Strahan, 2006).

Macroeconomics with Financial Intermediaries. Several papers analyze the role of financial intermediaries in macroeconomics, either with a focus on banks' imperfect competition (e.g., Corbae and D'Erasmo, 2021) or focusing on the interaction between credit constraints and the financial intermediation sector (e.g., Elenev, Landvoigt, and Van Nieuwerburgh, 2021). While Corbae and D'Erasmo (2021) build a structural model of banking industry dynamics, their focus is on regulatory policy. I instead explore the macroeconomic implications of bank market power through the lens of firm dynamics. Similarly, Elenev, Landvoigt, and Van Nieuwerburgh (2021) examine capital constraints among banks and firms, whereas I highlight a novel amplification channel: the endogenous dispersion of firm-specific loan markups. Another related contribution is Dempsey (2025), who develops a model with heterogeneous firms that can choose between intermediated and direct finance, while heterogeneous banks compete both with one another and with the bond market. His framework features an endogenous concept of banks' comparative advantage in lending relative to non-banks and emphasizes the role of capital requirements in an environment characterized by credit substitutability. While Dempsey (2025) models a continuum of monopolistically competitive banks, my focus is on competition among non-atomistic banks, which gives rise to endogenous cross-sectional dispersion in loan markups and, in turn, shapes firm-level financing and investment decisions.

Other papers study bank market power in the context of monetary policy transmission (e.g., Wang, Whited, Wu, and Xiao, 2022; Jamilov and Monacelli, 2025), often relying on markups derived from constant elasticity of substitution (CES) frameworks. In contrast, I use a generalized Euler equation framework that allows for time-varying, endogenous firm-level markups driven by dynamic bank-firm interactions.

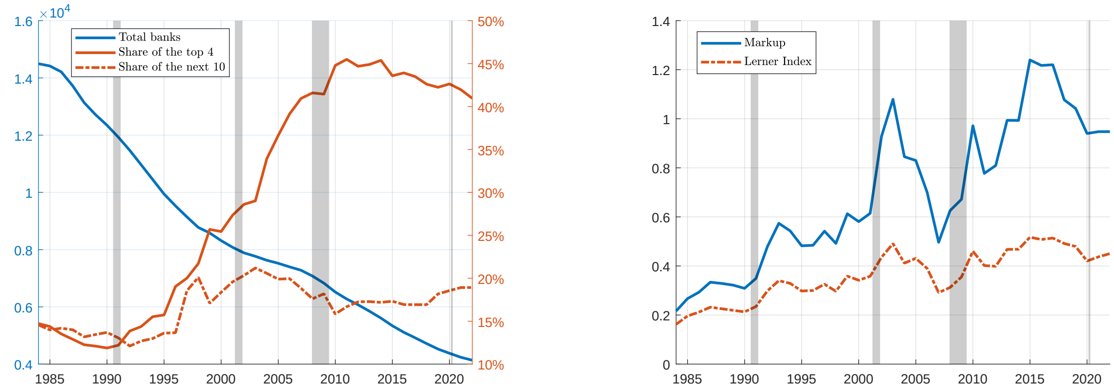
Outline. The rest of the paper is organized as follows. Section 2 reports three salient stylized facts. Section 3 presents a stylized version of the model aimed at delivering basic intuition about the proposed mechanism. Section 4 describes the infinite-horizon version of the model and discusses the main mechanism and its effects in detail. Section 5 explains the calibration and results of the oligopolistic stationary equilibrium. Section 6 illustrates the results of the three aforementioned macroeconomic shocks, such as the failure of one big bank and a new bank entry. Section 7 proposes an extension of the baseline model. Section 8 concludes.

2 Stylized Facts

The purpose of this section is to report the key empirical facts on banks' market structure and market power. Appendix C.4 details the data and the methodology, and supplement the analysis with additional patterns.

(1) Banks Concentration and Markups Increased. I obtain all banks registered in the U.S. from 1984 to 2022 from the Federal Financial Institutions Examination Council's Central Data Repository and combine the bank-level data to construct banks asset-market share and markups in the spirit of De Loecker, Eeckhout, and Unger (2020) and Corbae and D'Erasmo (2021). Figure 1 shows that both the asset-market share of the top 4 banks and the asset-weighted markup have been increasing over time.

Figure 1. Banks Market Share and Markups



Notes: The left panel shows the evolution over time of the total number of registered banks in the U.S., the asset-market share of the top 4 banks, and the asset-market share of the next 10 largest banks. The right panel shows the asset-weighted markup and Lerner Index.

Bank-level markups are calculated as

$$\frac{\textsc{Interest Return on Loans}_t}{\textsc{Cost of Funds}_t + \textsc{Marginal Net Expenses}_t} - 1.$$

Interest rate on loans is calculated as the ratio between interest income on loans and the total loan amount. The cost of funds is calculated as the interest rate on both deposits and Fed funds, relative to the total of deposits and Fed funds. Marginal net expenses is defined as the marginal non-interest expenses net of marginal non-interest income. Marginal non-interest expenses and marginal non-interest income are derived from the trans-log functions (46) and (47) in Appendix C.4.2, along the lines of Demirgüç-Kunt and Martinez Peria (2010). The Lerner index is calculated as

$$\frac{\textsc{Markup}_t}{1 + \textsc{Markup}_t}.$$

Hence, a Lerner index of zero indicates the perfect competition benchmark. The Lerner index has been significantly above zero and increasing over time, pointing to an increase in market power. Appendix C.4.2 explains all these measures in details.

(2) Evolution of Banks Concentration is Positively Correlated with Credit Spread. Given the same database as in point (1), I run the following bank-level regression:

$$\log(R_{b,L,t} - R_{M,t}) = \beta_0 + \beta_1 \times \log L_{b,t} + \beta_2 \times \text{Charge-off}_{b,t} + \beta_3 \times C_{5,t} + \alpha_b + \gamma_t.$$

The independent variable is the log of bank-level credit spread $\log(R_{b,L,t} - R_{M,t})$, calculated as the difference between the loan rate ($R_{b,L,t}$) and the 3-Month T-bill rates (R_M). The dependent variables are (i) the log of the outstanding quantity of loans ($L_{b,t}$), (ii) the charge-off rates, and (iii) the concentration of the biggest 5 banks ($C_{5,t}$). Moreover, α_b indicates bank-level fixed-effects and γ_t indicates time fixed-effects. I present the results in Table C10 in Appendix C.4.2. The coefficient of interest is β_3 which is estimated to be .854 (statistically significant at the 1 percent level), hence implying that a 1 percent increase in $C_{5,t}$ results in a .854 percent increase in the annualized credit spread, ceteris paribus. Table C11 shows that the statistical significance of the coefficients is robust to two-way clustering of standard errors by bank and time.

(3) Controlling for Deal Amount and Proxy for Corporate Default, Smaller Firms Tend to Pay Higher Credit Spreads. In order to investigate interest rates at a firm-bank deal level, I use the Roberts Dealscan-Compustat Linking Database following Chava and Roberts (2008), which connects borrowers and lenders from 1986 to 2012. I run the following firm-bank level regression:

$$\text{All-In Drawn}^d_{f,b,t} = \beta_0 + \beta_1 \times \log K_{f,t} + \beta_2 \times \log L^d_{f,b,t} + \beta_3 \times \text{Altman Z-Score}_{f,t} + \gamma_t,$$

where All-In Drawn$^d_{f,b,t}$ indicates the interest rate spread on a deal d paid by a firm f to a bank b, $K_{f,t}$ indicates the book value of a firm f, $L^d_{f,b,t}$ indicates the deal amount, Altman Z-Score is a proxy for firm's f bankruptcy probability, and γ_t indicates time fixed-effects.[6] I present the results of this baseline specification and several robustness checks in Table C7 in Appendix C.4.1. The coefficient of interest is β_1 which is estimated to be $-.2535$ (statistically significant at the 1 percent level), hence implying that a firm with a size corresponding to the 75$^{\text{th}}$ percentile pays 72 basis points less in annualized credit spreads than a firm with a size corresponding to the 25$^{\text{th}}$ percentile, ceteris paribus. The estimate of β_1 remains negative and significant across specifications. Figure C13 reports the time fixed effect, γ_t, over time. On average, the estimated time fixed effects increase over time, which is consistent with the notion that raising bank concentration has exerted upward pressure on interest rates across all firms. Relatedly, I run the following additional regression which explicitly controls for the interaction between firm size and bank concentration:

$$\begin{aligned}\text{All-In Drawn}^d_{f,b,t} = \beta_0 + \beta_1 \times \log K_{f,t} + \beta_2 \times \log L^d_{f,b,t} + \beta_3 \times \text{Altman Z-Score}_{f,t} + \beta_4 \times \text{C}_{5,t} + \\ \beta_5 \times (\log K_{f,t} \cdot \text{C}_{5,t}).\end{aligned}$$

The new coefficient of interest, β_5, is estimated to be negative, while β_4 is positive and β_1 remains negative. All coefficients are significant at the 1 percent level. This suggests that a higher level of concentration interacts with firm size by increasing interest rates disproportionally more for smaller firms. Table C7 in Appendix C.4.1 reports all estimated coefficients and their significance levels.

[6] Appendix C.4.1 explains the data and all these measures in details.

3 Stylized Model

In this section, I analyze a two-period model designed to provide preliminary intuition for the infinite-horizon model presented in Section 4. An oligopolistic banking sector interacts with a continuum of heterogeneous firms in the presence of idiosyncratic productivity and default shocks. I provide analytical results on how increasing the number of banks B impacts various financial and macroeconomic variables, such as loans, interest rates, expected returns, investment, leverage, and TFP. In the stylized version of the model, there are two dates denoted $t = 0, 1$. Note that, in the stylized model, the number of banks is exogenous but in the infinite-horizon model is endogenous (banks make entry and exit decisions).

Preferences. There is a continuum of identical households with preferences:

$$C_0 + \beta \cdot C_1, \tag{1}$$

where C_t is the household's consumption at time t and $\beta \in (0, 1)$ is the discount factor.

Technology. There is continuum of firms $j \in [0, 1]$. In each period $t = 0, 1$, the output $y_t(j)$ produced by each firm j is given by the production function $y_t(j) = z_t(j) \cdot k_t(j)^\alpha$, where $0 < \alpha < 1$.

Ownership Structure. The households own all banks and firms. Each firm j is characterized by its state vector

$$x_t(j) \equiv \{\{l_{b,t}(j)\}_{b=1}^{B}, r_{l,t}(j), k_t(j), \mathcal{I}_t(j), z_t(j)\},$$

where $l_{b,t}(j)$ denotes the firm's loan by bank b, $r_{l,t}(j)$ (throughout the paper I will also refer to $R_{l,t}(j) = 1 + r_{l,t}(j)$) is the interest rate (charged by all banks), $k_t(j)$ is the firm's capital stock, $z_t(j)$ is the firm's productivity, and $\mathcal{I}_t(j)$ is an indicator function that takes value 1 if the firm has not defaulted. Let $\phi(x_0)$ denote the density function of firms in the economy at $t = 0$.

Markets. There are five markets in the economy: banks' debt, banks' equity, firms' loans, firms' equity, and the market for the representative good.

Banks' equity and debt markets. The household invests in the production sector by supplying equity or debt to banks and by supplying equity to the firms and faces budget constraints:

$$C_0 + \sum_{b=1}^{B} p_b \cdot S_{b,1} + D_1 + \int p_0 \cdot S_1 \, \mathrm{d}\Phi = \sum_{b=1}^{B} (p_b + \pi_{b,0}) \cdot S_{b,0} + \int (p_0 + \tilde{d}_0) \cdot S_0 \, \mathrm{d}\Phi \tag{2}$$

$$C_1 = \sum_{b=1}^{B} \pi_{b,1} \cdot S_{b,1} + R_D \cdot D_{b,1} + \int \mathcal{I} \cdot \tilde{d}_1 \cdot S_1 \, \mathrm{d}\Phi, \tag{3}$$

where p_b, $S_{b,0}$, $S_{b,1}$, $D_{b,1}$, $\pi_{b,0}$, and $\pi_{b,1}$ are, respectively, the bank's share price, the share holdings at $t = 0, 1$, the debt holdings at $t = 1$, and the bank's profit at $t = 0, 1$. Bank b demands equity and debt from the household, in order to finance loans to firms.

Firms' equity market. The household invests in the production sector by supplying equity to the firms, and faces budget constraints (2) and (3), where p_0, S_0, S_1, $\tilde{d}_0$, and $\tilde{d}_1$ are, respectively, the share price, share holdings at $t = 0, 1$, and the dividend of each firm (net of equity issuance cost) at $t = 0, 1$. Firms demand equity from, or distribute dividend to, the household. If a firm decides to issue equity, it incurs a quadratic equity issuance cost at $t = 0$ (with λ_0 being a positive constant):

$$\lambda(d_0) = \begin{cases} \lambda_0 \frac{d_0^2}{2} & \text{if } d_0 \leq 0 \\ 0 & \text{if } d_0 > 0 \end{cases}, \tag{4}$$

where d_0 is a firm dividend at $t = 0$, defined below. The cost of equity issuance could stem from information frictions, such as signaling costs or adverse selection.

Firms' loan market. A finite number B of (identical) banks supply loans to the continuum of firms. Each bank $b = 1 \dots B$ can issue non state-contingent loans $l_{b,1}$ to each firm. Loans are due for repayment in the next period, unless the firm defaults. A firm j takes the interest rate $r_1(j)$ as given and chooses how much to invest and how much to borrow from each bank. Banks take each firm's demand schedule as given and compete à la Cournot, i.e, simultaneously and independently choose their loan portfolios. The Cournot model has the convenient property that the number of firms (banks in my case) is a sufficient statistic to determine market power, rendering it a parsimonious and tractable modeling choice. Alternatively, one could consider a Bertrand model where banks face capacity constraints, in order to rule out perfectly competitive outcomes. As shown by Kreps and Scheinkman (1983), quantity precommitment followed by Bertrand competition would induce Cournot outcomes.

Goods market. The representative household demands goods supplied by all firms.

Shocks. At time 0, firms are heterogeneous with respect to their capital stock k_0 and their idiosyncratic productivity z_0. At time 1, there are two types of idiosyncratic shocks: the firm can default, with exogenous probability $1 - \rho$ and, if it survives, z_1 realizes according to $\log z_1 = \rho_z \log z_0 + \xi_1$ where $\xi_1 \sim \mathcal{N}(0, \sigma_\xi^2)$ and $0 \leq \rho_z < 1$.

Timing. All decisions are taken at $t = 0$. Given the initial distribution of firms with pdf $\phi(x_0)$ (and cdf $\Phi(x_0)$), the timing is as follows: (1) each firm produces output $y_0 = z_0 k_0^\alpha$; (2) each bank finances its supply of loans, $\int l_{b,1}(x_0)\,\mathrm{d}\Phi(x_0)$, by issuing equity and/or debt; (3) each firm takes the interest rate $R_1(x_0)$ as given and chooses how much to invest and the amount of loan to demand from each bank; (4) banks take each firm's demand schedule as given and compete with each others to supply the loans. The outcome is a contract establishing: loan amount $l_{b,1}(x_0)$, interest rate $R_1(x_0)$, and the new level of capital $k_1(x_0)$; and (5) firms distribute dividends $d_0 = z_0 k_0^\alpha + (1-\delta)k_0 - k_1 + \sum_b^B l_{b,1}$ to the household.[7] At $t = 1$, the $1-\rho$ mass of defaulting firms exits the market. For the surviving firms, z_1 is realized and: (1) firms produce output $y_1 = z_1 k_1^\alpha$; (2) firms repay their outstanding debt plus interest $R_1(x_0) \cdot \sum_b^B l_{b,1}(x_0)$; (3) each bank distributes its profit $\int \rho R_1(x_0) l_{b,1}(x_0)\,\mathrm{d}\Phi(x_0)$ to the saver; and (4) firms distribute dividend $d_1 = z_1 k_1^\alpha + (1-\delta)k_1 - R_1 \sum_b^B l_{b,1}$ to the household.

3.1 Agents' Optimization Problems

The representative household maximizes its intertemporal utility (1) subject to the budget constraints (2) and (3), yielding Euler equations that pin down the price of banks' equity $\forall b: \ \beta \pi_{b,1} = p_{b,0}$, and the price of each firm's equity

$$\rho \beta \mathbb{E}_0 \left[\frac{d_1}{p_0}\right] = 1 - \lambda_d(d_0). \tag{5}$$

Firms maximize the net present value of dividends $d_0 + \beta \cdot \mathbb{E}_0\left[\mathcal{I} \cdot d_1\right]$. The firm's optimality condition with respect to the loan requires that the discounted future expected interest rate be one net of the equity issuance cost:

$$\rho \beta R_{l,1} = 1 - \lambda_d(d_0). \tag{6}$$

The firm's first-order condition with respect to capital requires that the future interest rate equals the expected marginal productivity of capital net of depreciation:

$$R_{l,1} = \mathbb{E}_0\left[1 + \alpha z_1 k_1^{\alpha-1} - \delta\right]. \tag{7}$$

Generalized Euler Equation. Banks' strategies map firm characteristics (x_0) onto the current quantity and future interest rate of loans. Given the probability density function $\phi(x_0)$ and cumulative distribution function $\Phi(x_0)$, each bank b chooses $l_{b,1}(x_0)$ to best respond

[7] I assume for simplicity that firms start with $l_{b,0} = 0$.

to other banks' strategies $l_{-b,t}(x_0)$, such that

$$\max_{l_{b,1}(x_0)} \quad \pi = -\int l_{b,1}(x_0)\,\mathrm{d}\Phi(x_0) + \beta \int \rho R_{l,1}(x_0) l_{b,1}(x_0)\,\mathrm{d}\Phi(x_0),$$

subject to equations (5)-(7) for all firms in the distribution.
Each bank's best response is characterized by the following generalized Euler equation (GEE)

$$\forall x_0 : \ \frac{\partial \pi}{\partial l_{b,1}(x_0)} = -1 + \rho\beta \frac{\partial R_{l,1}(x_0)}{\partial l_{b,1}(x_0)} l_{b,1}(x_0) + \rho\beta R_{l,1}(x_0) = 0, \tag{8}$$

where $\frac{\partial R_{l,1}(x_0)}{\partial l_{b,1}(x_0)}$ can be determined by the implicit function theorem on equations (6) and (7). Equation (8) is a generalized Euler equation because it contains the derivatives of each firm's policy functions. Each bank best responds by internalizing the effect of loans on the firms' capital choice $\frac{\partial k_1}{\partial l_{b,1}}$ as well. For ease of notation, I drop the dependency of all optimal choices from x_0. The inverse elasticity $\frac{\partial R_{l,1}}{\partial l_{b,1}} \frac{l_{b,1}}{R_{l,1}}$ implicitly contained in the generalized Euler equation requires the determination of the term $\frac{\partial R_{l,1}}{\partial l_{b,1}}$. From equation (7)

$$\frac{\partial R_{l,1}}{\partial l_{b,1}} \frac{l_{b,1}}{R_{l,1}} = \mathbb{E}_0 \left[\underbrace{\alpha(\alpha-1) z_1 k_1^{\alpha-2}}_{\frac{\partial \mathrm{MPK}}{\partial k_1}} \cdot \frac{\partial k_1}{\partial l_{b,1}} \frac{l_{b,1}}{R_{l,1}} \right], \tag{9}$$

where $\mathrm{MPK} \equiv z_1 \cdot \alpha k_1^{\alpha-1}$. The formula suggests that banks' market power depends on two components. First, banks extract higher markups out of firms that exhibit a lower (i.e., more negative) derivative of the marginal productivity of capital $\frac{\partial \mathrm{MPK}}{\partial k_1}$.[8] These firms are characterized by a less elastic credit demand, since if banks restricted the supply of loans (hence, inducing less investment in physical capital) they would cause: (i) a greater missed production (these firms also exhibit a high MPK) and (ii) a higher interest rate (these firms' levels of capital correspond to a more concave point of the production function). Second, banks think strategically by internalizing the effects their actions have on the firms' investment decisions (i.e., they internalize the impact of their actions on the firms' policy functions). This second effect is captured by the cross-elasticity $\frac{\partial k_1}{\partial l_{b,1}} \frac{l_{b,1}}{R_{l,1}}$. Finally, note that expressions for the two cross-derivatives can be found jointly by taking the total derivatives of equations (6) and (7):

$$\frac{\partial R_{l,1}}{\partial l_{b,1}} = \frac{1 - \rho\beta R_{l,1}}{\rho\beta l_{b,1}} \quad \text{and} \quad \frac{\partial k_1}{\partial l_{b,1}} = \frac{1 - \rho\beta R_{l,1}}{\rho\beta l_{b,1}} \cdot \frac{1}{\mathbb{E}_0 \left[\frac{\partial \mathrm{MPK}}{\partial k_1} \right]}.$$

[8] Note that the term $\alpha - 1$ is always negative; hence, $\frac{\partial \mathrm{MPK}}{\partial k_1}$ and the inverse elasticity $\frac{\partial R_{l,1}}{\partial l_{b,1}} \cdot \frac{l_{b,1}}{R_{l,1}}$ are always negative. Banks exert higher market power when $\frac{\partial R_{l,1}}{\partial l_{b,1}} \cdot \frac{l_{b,1}}{R_{l,1}}$ is smaller.

In equilibrium, for the mass of financially constrained firms ($d_0 < 0$), the degree of imperfect competition (number of banks B) matters. For each firm, the equilibrium is a vector $(k_1^*, R_{l,1}^*, l_{b,1}^*, p_0^*)$ such that equations (5)-(8) hold. For the mass of firms that, in equilibrium, is not financially constrained, the degree of imperfect competition does not matter. For these firms, the solution is given by $(k_1^*, R_{l,1}^*, p_0^*)$ such that equations (5)–(7) hold, and the Modigliani-Miller theorem applies; hence, $l_{b,1}^*$ is undetermined. Note that, in the quantitative model, the household is risk-averse and a tax shield is present. These elements, combined with the non-atomistic behavior of banks, ensure that each firm's optimal capital structure is well defined in equilibrium.[9]

3.1.1 Characterization of the Equilibrium

I now describe intuitively the main mechanism that drives the analytical results presented in this section. First, equation (9) suggests that the higher the MPK of a firm, the higher the marginal value of one unit of loan for that firm that translates in a lower inverse elasticity $\frac{\partial R_{l,1}}{\partial l_{b,1}} \cdot \frac{l_{b,1}}{R_{l,1}}$. Second, the degree of imperfect competition (represented by the number of banks B) is relevant only for the mass of financially constrained firms. Note that being financially constrained is an equilibrium outcome, as it is determined by d_0, which is endogenous, and is itself influenced by the number of banks B. Consequently, banks endogenously exert a greater degree of market power on firms that, in equilibrium, become financially constrained and exhibit a high MPK. Intuitively, these firms have worse outside options (e.g., a high cost of non-bank finance) and one additional unit of investment in physical capital contributes significantly to their future production; hence, they exhibit a less elastic demand for credit. An imperfectly competitive financial sector internalizes that the same financial resources are more valuable for this type of firm and for their future growth paths; therefore, it can charge higher markups.

This mechanism leads financially constrained firms to grow slower in less competitive credit markets. As a result, the dispersion of marginal productivity of capital is higher when there are fewer banks B in the economy. At the same time, a lack of competition in the financial intermediation reduces aggregate productivity, since firms grow toward their efficient level of capital on slower trajectories. This intuitive mechanism is at the base of Proposition I presented in Appendix B. In essence, a higher number of banks (i.e., a higher B) has the following effects: (i) aggregate loans per bank $\int l_{b,1}^* \,\mathrm{d}\Phi$ decreases; (ii) average loan interest rate $\int R_{l,1}^* \,\mathrm{d}\Phi$ decreases; (iii) aggregate physical investment $\int k_1^* - (1-\delta)k_0 \,\mathrm{d}\Phi$ increases; (iv) aggregate expected returns $\int \mathbb{E}_0 \left[d_1^*\right] / p_0^* \,\mathrm{d}\Phi$ decreases; (v) aggregate loans $\int \sum_b^B l_{b,1}^* \,\mathrm{d}\Phi$ increases; (vi) aggregate leverage $\int \sum_b^B l_{b,1}^* / k_1^* \,\mathrm{d}\Phi$ increases; (vii) aggregate TFP $\int k_1^{*\alpha} \,\mathrm{d}\Phi / \left(\int k_1^* \,\mathrm{d}\Phi\right)^\alpha$ increases; (viii) variance of capital $\int k_1^{*2} \,\mathrm{d}\Phi - (\int k_1^* \,\mathrm{d}\Phi)^2$ decreases; (ix) variance of loan interest rates $\int R_{l,1}^{*2} \,\mathrm{d}\Phi - (\int R_{l,1}^* \,\mathrm{d}\Phi)^2$ decreases; and (x) variance of expected

[9] See Appendix B.2 for further details.

returns $\int (\mathbb{E}_0 [d_1^*] / p_0^*)^2 \, \mathrm{d}\Phi - \left(\int \mathbb{E}_0 [d_1^*] / p_0^* \, \mathrm{d}\Phi\right)^2$ decreases.

4 Model

In the two-period model, banks' choices are static. In the infinite-horizon model, each bank faces a dynamic problem that: (i) depends on the same bank's future strategies and other banks' current and future strategies, and (ii) is subject to all firms' dynamic demand for loans; also both the current and future distributions of firms matter. The equilibrium concept used in this section is a Markov perfect equilibrium (e.g., Maskin and Tirole, 2001). Specifically, I characterize the equilibrium using generalized Euler equations in a similar fashion to the optimal fiscal policy literature (see, for instance, Klein and Ríos-Rull, 2003; Klein, Krusell, and Ríos-Rull, 2008; and Clymo and Lanteri, 2020).

In this section, I build a dynamic framework to study firms' financing-investment decisions when banks are big (i.e., non-atomistic), strategically interact with each others, and face idiosyncratic firms' default risk.[10] Households derive utility from a non-durable consumption good, own the shares of the banks, and supply deposits. Banks issue debt and use both their internal resources and debt to purchase firms' loans. Firms make investment decisions, taking into account the fact that debt provides a tax shield and issuing new equity is increasingly costly. The key feature of the framework is the simultaneous presence of strategic interactions among financial institutions, general equilibrium, macroeconomic shocks, and heterogeneous firms. Note that each firm stipulates an idiosyncratic contract with the banks: in equilibrium, banks have different degrees of market power on each single firm in function of its idiosyncratic characteristics.

I now describe the model and define the stationary oligopolistic equilibrium, in which all aggregate quantities and prices are constant over time. I overcome the computational challenge by proposing GEE-based algorithms to solve for the oligopolistic stationary equilibrium and related transitional dynamics in the presence of strategic interactions, general equilibrium, and heterogeneous firms. The algorithms are detailed in Appendix A.

4.1 Environment

Time is discrete $t = 0, 1, \ldots$ and the horizon is infinite. There is an endogenous number B of identical big banks in equilibrium and an endogenous number $\mathcal{B}$ of identical small banks in equilibrium. Differently from big banks, small banks operate under capacity constraints. All banks are owned by a continuum of identical and infinitely lived households, equivalent to one representative household. The household also owns a continuum of firms $j \in [0, 1]$.

[10] Specifically, I interpret the financial intermediation sector as a succession of decision makers – one at each date t – without commitment to future realized quantity of loans supplied.

Preferences. The household ranks stream of consumption C_t according to the following lifetime utility function:

$$\sum_{t=0}^{\infty} \beta^t \cdot u(C_t), \tag{10}$$

where $\beta \in (0,1)$ is the household's discount factor, and $u_c > 0$, $u_{cc} < 0$.

Technology. In each period t, the output $y_t(j)$ produced by each firm $j \in [0,1]$ is given by the production function $y_t(j) = Z_t \cdot z_t(j) \cdot k_t(j)^{\alpha}$, $0 < \alpha < 1$, where Z_t is an aggregate TFP shock (which is used in Section 6 in the form of an unexpected shock) and $z_t(j)$ indicates the idiosyncratic firm's j productivity with characteristics specified in the paragraph below.

Ownership Structure. The household owns all the banks and the entire mass of firms $j \in [0,1]$. Each firm j is characterized by its state vector

$$x_t(j) \equiv \{\{l_{b,t}(j)\}_{b=1}^{B_t}, l^s_{b,t}(j)\}_{b=1}^{\mathcal{B}_t}, r_{l,t}(j), k_t(j), \mathcal{I}_t(j), z_t(j)\},$$

where $l_{b,t}(j)$ denotes the firm's loan by bank $b = 1, ..., B$, $l^s_{b,t}(j)$ denotes the firm's loan by a small bank $b = 1, ..., \mathcal{B}$, $r_{l,t}(j)$ is the interest rate (charged by all banks), $k_t(j)$ is the firm's capital stock, $\mathcal{I}_t(j)$ is a dummy variable that takes value of one if the firm has not defaulted, and $z_t(j)$ is the firm's idiosyncratic productivity. Let $\phi_t(x_t)$ denote the density function of firms in each given period after firms default and reenter, with associated cumulative distribution $\Phi_t(x_t)$. The household can save through the banks or through firm's equity, in this latter case it incurs the equity issuance cost (4). Throughout the paper, I will omit the explicit dependency on j of the firm specific variables contained in $x_t(j)$ for notational convenience.

Markets. There are six markets in the economy: banks' debt, banks' equity, firms' loans, firms' equity, interbank market, and the market for the representative good.

Banks' equity and debt markets. The household invests in the production sector by supplying equity or debt to banks and faces the budget constraint:[11]

$$
\begin{aligned}
&C_t + \sum_{b=1}^{B_{t+1}} \left(p_{b,t} S_{b,t+1} + D_{b,t+1}\right) + \sum_{b=1}^{\mathcal{B}_{t+1}} \left(p^s_{b,t} S^s_{b,t+1} + D^s_{b,t+1}\right) = \int \mathcal{I}_t \tilde{d}_t + (1 - \mathcal{I}_t)\tilde{d}_0 \, \mathrm{d}j \\
&\sum_{b=1}^{\overline{B}_t} \left((p_{b,t} + \pi_{b,t}) S_{b,t} + R_{D,t} D_{b,t}\right) + \sum_{b=1}^{\overline{\mathcal{B}}_t} \left((p^s_{b,t} + \pi^s_{b,t}) S^s_{b,t} + R^s_{D,t} D^s_{b,t}\right),
\end{aligned}
\tag{11}
$$

where $p_{b,t}$, $S_{b,t}$, $S_{b,t+1}$, $D_{b,t}$, $D_{b,t+1}$, $R_{D,t}$, and $\pi_{b,t}$ are, respectively, the big bank's share price, the share holdings at t and $t+1$, the big bank's debt holdings at t and $t+1$, the interest rate on the bank's debt, and the big bank's profit at t. The corresponding variables denoted with suffix s refer to the small banks instead. Each bank b demands equity and debt from the representative household, in order to finance loans to firms.

Firms' equity market. The household owns all the firms and, for the firm that have not defaulted, it receives each firm's dividend $\tilde{d}_t$ (net of equity issuance cost) at t. Firms demand equity from, or distribute dividends to, the household. If a firm decides to issue equity, it incurs the equity issuance cost $\lambda(d_t)$ given by equation (4), where d_t is a firm dividend at time t, defined below. The corresponding mass of firms that have defaulted is replaced by a mass of newly entering firms with characteristics x_0, which produce a dividend $\tilde{d}_0$. This dividend can potentially be negative in the case of an equity issuance.

Firms' loan market. The B_t big banks and the $\mathcal{B}_t$ small banks supply loans to a continuum of firms. Each bank $b = [1 \ldots B_t]$ and $b = [1 \ldots \mathcal{B}_t]$ can issue non-state-contingent loans $l_{b,t+1}$ and $l^s_{b,t+1}$ to each firm. Loans are due for repayment in the next period, unless the firm defaults. A firm j takes the interest rate $r_{l,t+1}(j)$ as given and chooses how much to invest and how much to borrow from each bank. Banks take each firm's demand schedule as given and compete dynamically à la Cournot. Differently from big banks, small banks are subject to capacity constraints. The process determines the total amount of loans banks supply to each firm which, together with the firm's demand schedule, pins down the firm-specific interest rate $r_{l,t+1}(j)$. At time t, each bank and firm commit to such an interest rate.

Interbank market. A big bank b can lend $M_{b,t}$ to other big banks that will be repaid in the following period at rate $r_{M,t+1}$. Since all big banks are identical, in equilibrium $\forall b: \ M_{b,t} = 0$. Similarly, a small bank b can lend $M^s_{b,t}$ to other small banks that will be repaid in the following period at rate $r_{M,t+1}$. Since all small banks are identical, in equilibrium $\forall b: \ M^s_{b,t} = 0$. For simplicity, I assume no cross big-small banks lending is allowed.

Goods market. The representative household demands goods supplied by all firms.

[11]Note that $\overline{B}_t \equiv \max\{B_t, B_{t+1}\}$ and $\overline{\mathcal{B}}_t \equiv \max\{\mathcal{B}_t, \mathcal{B}_{t+1}\}$. If a bank decides to exit, then $p_{b,t} = 0$, but it still needs to be taken into account because it can liquidate its assets in the form of a dividend. If a bank decides to enter, it is owned by the household and requires an initial equity injection—i.e., a negative dividend—to operate.

Shocks. At time t, $\mathcal{I}_t$ takes the value of one with probability ρ_t, and zero with probability $1-\rho_t$. A new mass of firms re-enters the economy so that the total mass remains constant over time. I relax this assumption in Subsection 6.2. Reentering firms are assigned characteristics x_0, and I set their initial capital k_0 as described in Section 5. I assume that, in the event of default, banks recover $\nu k+k_0$ from the defaulting firm's capital. The portion k_0 is transferred to the household (the owner of both firms and banks) and is recycled to finance newly entering firms. This assumption ensures that new firm entry is not fueled by exogenous resource injections, but rather by reallocating existing capital within the economy. For simplicity, I set the initial amount of the loan owed to the big bank as $l_{b,0} = 0$ for all $b = [1, \ldots, \mathcal{B}]$, and the initial amount owed to the small bank as $l_b^s = 0$ for all $b = [1, \ldots, \mathcal{B}]$. Firms have heterogeneous productivity $z_t(j)$, with $z_0(j)$ drawn at the birth of the firm from a Gaussian probability distribution function with mean 1 and variance σ_z^2.[12] Finally, in Section 7, I also propose and solve an extension of the model with firms endogenous default decisions, which yields consistent results with the baseline model.

Government. The government imposes proportional taxes τ on all firms' production. Firms can deduct loan interest and depreciated capital from their taxes. The government runs a balanced budget constraint. That is, the government uses the aggregate revenue from taxes

$$T_t = \tau \left[\int \mathcal{I}_t \left(Z_t z_t k_t^{\alpha} - \sum_{b=1}^{B_t} r_{l,t} l_{b,t} - \sum_{b=1}^{\mathcal{B}_t} r_{l,t} l_{b,t}^s - \delta k_t \right) \mathrm{d}j + \int (1-\mathcal{I}_t) \left(Z_t z_0 k_0^{\alpha} - \delta k_0 \right) \mathrm{d}j \right],$$

to finance an exogenous government expenditure that exactly balances T_t at each point in time.

Timing. The aggregate state of the economy at time t is

$$X_t \equiv \{ \{D_{b,t}\}_{b=1}^{B_t}, \{D_{b,t}^s\}_{b=1}^{\mathcal{B}_t}, r_{D,t}, \{M_{b,t}\}_{b=1}^{B_t}, \{M_{b,t}^s\}_{b=1}^{\mathcal{B}_t}, r_{M,t}, B_t, \mathcal{B}_t, Z_t, \phi_t(x_t) \}.$$

Given X_t, the timing is as follows: (1) a mass $1-\rho_t$ of firms defaults, (2) each surviving firm produces output $y_t = Z_t z_t k_t^{\alpha}$ and repay its debt $\sum_b^{B_t} R_{l,t} l_{b,t} + \sum_b^{\mathcal{B}_t} R_{l,t} l_{b,t}^s$ to all incumbent banks; (3) a mass $1-\rho_t$ of firms enter the production sector with characteristics x_0; (4) each incumbent bank decides whether to exit; (5) potential banks entrants decide whether to enter, this requires an initial equity injection from the household; (6) each bank finances its supply of loans, $\int l_{b,t+1}(x_t)\,\mathrm{d}j$, by issuing equity and/or debt; (7) each firm takes the interest rate $r_{l,t+1}(x_t)$ as given and chooses how much to invest and the amount of loan to demand from each bank; (8) banks take each firm's demand schedule as given and compete

[12] Similarly, in spirit, to Melitz (2003) and Asplund and Nocke (2006).

with each others to supply the loans. The outcome is a contract establishing: loan amount $l_{b,t}(x_t)$, interest rate $r_{l,t+1}(x_t)$, and new level of capital $k_{t+1}(x_t)$; (9) firms distribute dividends $d_t = (1-\tau)\left[Z_t z_t k_t^{\alpha} - \sum_b^{B_t} r_{l,t} l_{b,t} - \sum_b^{\mathcal{B}_t} r_{l,t} l_{b,t}^s\right] + \tau\delta k_t - \tilde{i}_t$ to the household, where $\tilde{i}_t = i_t + \sum_{b=1}^{B_t} l_{b,t} - \sum_{b=1}^{B_{t+1}} l_{b,t+1} + \sum_{b=1}^{\mathcal{B}_t} l_{b,t}^s - \sum_{b=1}^{\mathcal{B}_{t+1}} l_{b,t+1}^s$ and i denotes investment in physical capital; (10) bank b distributes profit to the household. To simplify notation, I omit explicit time subscripts from variables throughout the remainder of the model exposition. It is understood that (x, X) refers to (x_t, X_t), and (x', X') refers to (x_{t+1}, X_{t+1}).

4.2 Household

I now describe the household's problem in recursive form. Let $V_H(X)$ be the value function of the household with banks' debt holdings $\mathbf{D} = [D_1 \dots D_B]$, and banks' equity holdings $\mathbf{S} = [S_1 \dots S_B]$. This function satisfies the following functional equation:

$$V_H(X) = \max_{\mathbf{S}', \mathbf{D}'} \quad u(C) + \beta \cdot V_H(X') \tag{12}$$

subject to the budget constraint (11). The left-hand side of the budget equation (11) reports the household's expenditures: household aggregate consumption, banks b's equity and debt purchases. The right-hand side reports the household's resources: each bank b's equity holdings and debt, and firms' dividends.

The household takes the future banks' debt market rate r'_D as given, together with future banks' profits, and purchases banks' debt and equity according to:

$$\forall b: \quad 1 = \mathcal{M}'(X, X') \cdot \frac{p'_b + \pi'_b}{p_b} \tag{13}$$

$$\forall b: \quad 1 = \mathcal{M}'(X, X') \cdot R'_D, \tag{14}$$

where $\mathcal{M}' \equiv \beta \frac{u_c(C')}{u_c(C)}$, and π_b is the profit of bank b distributed as a dividend to the household.

4.3 Firms

I now characterize firm j's problem in recursive form. For convenience, I omit the index notation j, i.e., all idiosyncratic time t state variables $x(j) \equiv \{\{l_b(j)\}_{b=1}^{B}, l_b^s(j)\}_{b=1}^{\mathcal{B}}, r_l(j), k(j), z(j)\}$ and corresponding choices are noted without index j. Let $V_F(x, X)$ be the value function of the firm j with loan holdings $[l_1 \dots l_B]$ and $[l_1^s \dots l_B^s]$ from each bank and capital k. This function satisfies the following functional equation:

$$V_F(x, X) \;=\; \max_{\{l'_b\}_{b=1}^{B'}, \{l_b^{s\prime}\}_{b=1}^{\mathcal{B}'}, k'} \; d - \lambda(d) + \mathbb{E}\left[\mathcal{I}' \cdot \mathcal{M}'(X, X') \cdot V_F(x', X') \mid (x, X)\right],$$

subject to

$$
\begin{aligned}
k' &= k(1-\delta) + i \\
i &= \tilde{i} - \sum_{b=1}^{B} l_b + \sum_{b=1}^{B'} l'_b - \sum_{b=1}^{\mathcal{B}} l^s_b + \sum_{b=1}^{\mathcal{B}'} l^{s\prime}_b \\
d &= (1-\tau)\left[Zzk^{\alpha} - \sum_{b=1}^{B} r_l l_b - \sum_{b=1}^{\mathcal{B}} r_l l^s_b \right] + \tau\delta k - \tilde{i} \\
\lambda(d) &= \mathbb{I}[d \leq 0] \cdot \lambda_0 \frac{d^2}{2} + \mathbb{I}[d > 0] \cdot 0,
\end{aligned}
$$

where $\mathcal{M}'$ is the discount factor of the household, as described in the previous subsection, and $\mathbb{I}$ is the indicator function with takes value of one when the argument is true and zero otherwise. Each firm takes the future loans' market rate r'_l as given and finances itself through internal financing (production and equity issuance) and external financing (loans from banks). The first-order condition with respect to k' is

$$
1 - \lambda_d(d) = \mathbb{E}\left[\mathcal{I}' \cdot \mathcal{M}'(X, X') \cdot \left(1 + (1-\tau)\left(Z'z'\alpha k'^{\alpha-1} - \delta \right) \right) \cdot (1 - \lambda_d(d')) \mid (x, X) \right], \tag{15}
$$

which weight in the marginal cost of investing in one unit of physical capital today against the marginal benefit of the expected discounted future marginal production net of depreciation, both adjusted by the marginal cost of equity issuance in case dividends are negative. The first-order condition with respect to $l^{\chi\prime}_b$, with $\chi = \{., s\}$, is

$$
1 - \lambda_d(d) = \mathbb{E}\left[\mathcal{I}' \cdot \mathcal{M}'(X, X') \cdot (1 + (1-\tau) r'_l) \cdot (1 - \lambda_d(d')) \mid (x, X) \right], \tag{16}
$$

which weight in the marginal benefit of borrowing one unit of loan today against the expected discounted marginal cost of the future interest repayment net of the tax shield, both adjusted by the marginal cost of equity issuance in case dividends are negative.

4.4 Banks

First, I present the decision problem of the incumbent banks. Then, I present the decision problem of the new potential entrants. Banks can be either big or small, i.e. $\chi = \{., s\}$. Consistently with the notation adopted above, big banks do not have a suffix whereas small banks have a suffix s. Note that banks are subject to aggregate constraints that contain the entire distribution of firms. In equilibrium, and given the exit shocks, there is a mapping between a firm's age age(j) and the vector $\{\{l_b(j)\}_{b=1}^{B}, l^s_b(j)\}_{b=1}^{\mathcal{B}}, r_l(j), k(j), \mathcal{I}(j)\}$. Hence, for computational purposes, it is convenient to represent the state space simply as $x(j) \equiv$

$\{\text{age}(j), z(j)\}$. Given the exit shocks, the distribution of firms is simply given by $\phi(\text{age}, z)$, which is not directly affected by the actions of the banks.

4.4.1 Incumbents

An incumbent big bank $b = [1...B]$, or an incumbent small bank $b = [1...\mathcal{B}]$, chooses the new level of debt to demand from the household and the new level of loans to offer to each firm. Formally, the strategy space is defined as:

$$\mathcal{S}_b^{\chi\prime}(x, X) \equiv \{D_b^{\chi\prime}(X), l_b^{\chi\prime}(x, X)\}.$$

The new amount of debt issued ($\Delta D_b^{\chi\prime} = D_b^{\chi\prime} - D_b^{\chi}$) and internal financing F^{χ} is chosen to provide enough coverage for the change in interbank lending and aggregate loans:

$$F^{\chi} + \Delta D_b^{\chi\prime} = \Delta M_b^{\chi\prime} - \int \mathcal{I} \cdot l_b^{\chi} + (1 - \mathcal{I}) \cdot \mathcal{R}^{\chi} - \big(\mathcal{I} \cdot l_b^{\chi\prime}(x, X) + (1 - \mathcal{I}) \cdot l_b^{\chi\prime}(x_0, X)\big)\, \mathrm{d}j, \tag{17}$$

where I assume that, in the event of a firm's default, each bank recovers a fraction of the outstanding debt proportional to the firm's capital and to its share of total debt holdings. Specifically, $\mathcal{R}^{\chi} = \nu k \cdot l_b^{\chi} / (\sum_{b=1}^{B} l_b + \sum_{b=1}^{\mathcal{B}} l_b^s)$, so that the total recovery across all banks equals νk. In addition, each bank also recovers $k_0 \cdot l_b^{\chi} / (\sum_{b=1}^{B} l_b + \sum_{b=1}^{\mathcal{B}} l_b^s)$ from the defaulting firm's capital, implying that the total amount k_0 is recovered collectively by all banks. This portion k_0 is then transferred to the household (the owner of both firms and banks) and recycled to finance newly entering firms. Since banks have no control over this amount, it does not appear in their dividend payouts; rather, it is assumed to be automatically transferred to shareholders and used to support new firm entry within the same period. I now describe the bank's problem in recursive form. Let $V_b^{\chi}(X)$ be the value function of a bank b. This function satisfies the following functional equation:

$$\tilde{V}_b^{\chi}(X) = \max_{\{D_b', r_D'\},\ M_b',\ \{l_b^{\chi\prime}(x,X), r_l'(x,X)\}} \quad \pi_b^{\chi} + \mathcal{M}'(X, X') \cdot V_b^{\chi}(X') \tag{18}$$

where each incumbent bank makes a exit decision

$$V_b^{\chi}(X) = \max\{0, \tilde{V}_b^{\chi}(X)\} \tag{19}$$

subject to: (i) equation (17), (ii) the household's interest rate-quantity schedule jointly defined by equations (13) and (14), (iii) each firm's interest rate-quantity schedule jointly defined by equations (15) and (16). Bank b's profit π_b is given by:

$$\pi_b^\chi = \int \mathcal{I} \cdot r_l \cdot l_b^\chi \,\mathrm{d}j + r_M M_b^\chi - r_D D_b^\chi - F^\chi. \tag{20}$$

Future market rates $r'_D(X)$ and $r'_l(x, X)$ adjust consistently with the interest rate-quantity schedules. Each bank b issues bank debt according to a generalized Euler equation:

$$1 = \mathcal{M}'(X, X') \cdot R'_D(X, X') \cdot \left(1 + \eta_D^{\chi\prime}(X, X')\right), \tag{21}$$

where $\eta_D^{\chi\prime}$ is the inverse elasticity $\frac{\partial R'_D}{\partial D_b^{\chi\prime}} \cdot \frac{D_b^{\chi\prime}}{R'_D}$ between debt and its rate. In principle, equation (21) is a best response function that captures the trade-off that a bank faces issuing new debt. Every new unit of debt increases the current investing capacity but needs to be repaid tomorrow at the contracted interest rate. Moreover, since $\eta_D^{\chi\prime}$ is non-negative, when a bank issues new debt it also increases the market rate of deposits, incurring an additional future marginal cost. In equilibrium, $\eta_D^{\chi\prime}$ is zero, as implied by equations (21) and (14). Without aggregate risk, households are indifferent to the financing structure of the banks. Small banks are identical to big banks except that their loans supply choice is subject to a capacity constraint. Specifically, pick a small bank $b \in [1...\mathcal{B}]$ and a big bank $\tilde{b} \in [1...B]$, the capacity constraint reads:

$$\forall (b, \tilde{b}) : \; l_b^{s\prime} \leq \kappa \cdot l'_{\tilde{b}}, \tag{22}$$

which, intuitively, captures the idea that small banks face steep scaling costs, represented by a fraction $0 < \kappa < 1$ relative to the size of the large banks. Note that, in equilibrium, $\forall (b, \tilde{b}) : \; l_b^{s\prime} = \kappa \cdot l'_{\tilde{b}}$, since small banks are identical to big banks except the capacity constraint. Hence, if given the opportunity, small banks would develop the same size of big banks. In equilibrium, all firms borrow from all banks. Although a firm may borrow different amounts from large and small banks, the borrowing proportions are determined solely by the capacity constraint of small banks. On the bank side, there is a generalized Euler equation (23) for each firm type in the distribution with characteristics x, which closes the model and, when taken together with the firm's optimization problem, determines the firm-specific big-bank loan level, interest rate, and new level of physical capital. Given this, the allocation between big and small banks is determined by the small bank capacity constraint (22). An interesting avenue for future research is to allow firms to choose from which banks to borrow based not only on capacity constraints, but also on additional considerations—such as the probability of default of each bank.[13]

The generalized Euler equation that arises from the big bank loan first-order condition

[13]For instance, if certain firms borrow disproportionately more from smaller banks (which may be more likely to fail), this could create interesting spillover effects.

is

$$1 = \mathbb{E}\left[\mathcal{M}'(X,X')\cdot\left(\mathcal{I}'\cdot R_l'(x,X,x',X')\cdot(1+\eta_l'(x,X,x',X'))+(1-\mathcal{I}')\cdot\frac{\partial\mathcal{R}'}{\partial {l_b}'}\right)\ \mid (x,X)\right], \tag{23}$$

and, similarly, for the small banks' the first-order condition is

$$1+\mathcal{Z} = \mathbb{E}\left[\mathcal{M}'(X,X')\cdot\left(\mathcal{I}'\cdot R_l'(x,X,x',X')\cdot\left(1+\eta_l^{s\prime}(x,X,x',X')\right)+(1-\mathcal{I}')\cdot\frac{\partial\mathcal{R}^{s\prime}}{\partial l_b^{s\prime}}\right)\ \mid (x,X)\right], \tag{24}$$

where $\mathcal{Z}$ is the Lagrange multiplier on (22) and $\eta_l^{\chi\prime}(x,X,x',X') \equiv \frac{\partial R_l'}{\partial l_b^{\chi\prime}}\cdot\frac{l_b^{\chi\prime}(x,X)}{R_l'(x,X,x',X')}$ is the firm-specific inverse elasticity between loans and their rates. Equations (23) and (24) are functional equations that depend on the idiosyncratic characteristics of each firm. In equilibrium, equation (24) is unnecessary since equation (22) is binding, as discussed above. Nevertheless, the presence of small banks is important in shaping the inverse elasticity since it enters directly in $\eta_l^{\chi\prime}(x,X,x',X')$ and $\mathcal{M}'(X,X')$. See Subsection 4.4.2 for details on how to calculate $\eta_l^{\chi\prime}(x,X,x',X')$.

Equations (23) and (24) are the best response functions that capture the trade-off that a bank faces issuing a new unit of loan to a specific firm. Every new unit of loan decreases the current bank's dividend but produces a marginal income tomorrow at the contracted interest rate. Moreover, since $\eta_l^{\chi\prime}(x,X,x',X')$ is non-positive, when banks issue new loans they are also decreasing the future market rate of loans, incurring a marginal loss in the future. Note that banks best respond internalizing the effects that their actions have on aggregate quantities and all firms' choices (e.g., if a bank changes the quantity of loan offered to a firm, that firm might decide to re-optimize and adopt a different capital structure as a function of the credit market conditions). All these equations, together with an Euler equation that regulates the banks' behavior on the interbank market

$$1 = \mathcal{M}'(X,X')\cdot R_M'(X,X'), \tag{25}$$

captures the decision making behavior of each bank. The outcome of the game played by the banks at time t is a contract that pins down the firm-specific intermediation margin $R_l'(x,X,x',X') - R_D'(X,X')$. In principle, this margin can be decomposed into: (i) firm-specific loan's intermediation margin $(R_l'(x,X,x',X') - R_M'(X,X'))$ and (ii) debt's intermediation margin $(R_M'(X,X') - R_D'(X,X'))$. Note that the debt intermediation margin is zero, since $\eta_D^{\chi\prime} = 0$, hence $R_D'(X,X') = R_M'(X,X')$.

4.4.2 Calculation of the Inverse Elasticity η_l^χ

This subsection contains the calculation of the inverse elasticity η_l^χ in the generalized Euler equations (23) and (24). First, combine the firms' first-order conditions (15) and (16) to define the two functions:

$$f\left(x, X, x', X'\right) \equiv Z' z' \alpha k^{'\alpha-1} - \delta - R_l' + 1 = 0,$$
$$g\left(x, X, x', X'\right) \equiv \rho \cdot \mathcal{M}' \cdot (1-\lambda_d^{\chi'})\left((1-\tau)(R_l'-1)+1\right) - 1 + \lambda_d^\chi = 0.$$

Second, compute the total derivatives of these two functions with respect to k' and $l_b^{\chi'}$. Hence, solve the resulting linear system to get the following expression:

$$\frac{\partial R_l'}{\partial l_b^{\chi'}} = \frac{f_{k'} \cdot g_{l_b^{\chi'}} - f_{l_b^{\chi'}} \cdot g_{k'}}{f_{R_L'} \cdot g_{k'} - f_{k'} \cdot g_{R_L'}}.$$

Note that $f_{l_b^{\chi'}} = 0$, $f_{R_l'} = -1$, and $f_{k'} = Z' z' \alpha(\alpha-1) k^{'\alpha-2}$. Therefore, the elasticity $\eta_l^{\chi'}$ is given by

$$\eta_l^{\chi'} = \frac{\partial R_l'}{\partial l_b^{\chi'}} \frac{l_b^{\chi'}}{R_l'} = -\frac{f_{k'} \cdot g_{l_b^{\chi'}}}{g_{k'} + f_{k'} \cdot g_{R_l'}} \frac{l_b^{\chi'}}{R_l'}, \tag{26}$$

where the partial derivatives of the g function are

$$g_{k'} = \rho \frac{\partial \mathcal{M}'}{\partial k'}(1-\lambda_d(d'))\left((1-\tau)r_l'+1\right) - \rho \mathcal{M}' \frac{\partial \lambda_d(d')}{\partial k'}\left((1-\tau)r_l'+1\right) + \frac{\partial \lambda_d(d)}{\partial k'}, \tag{27}$$

$$g_{l_b^{\chi'}} = \rho \frac{\partial \mathcal{M}'}{\partial l_b^{\chi'}}(1-\lambda_d(d'))\left((1-\tau)r_l'+1\right) - \rho \mathcal{M}' \frac{\partial \lambda_d(d')}{\partial l_b^{\chi'}}\left((1-\tau)r_l'+1\right) + \frac{\partial \lambda_d(d)}{\partial l_b^{\chi'}}, \tag{28}$$

$$g_{R_l'} = \rho \frac{\partial \mathcal{M}'}{\partial R_l'}(1-\lambda_d(d'))\left((1-\tau)r_l'+1\right) - \rho \mathcal{M}' \frac{\partial \lambda_d(d')}{\partial R_l'}\left((1-\tau)r_l'+1\right) + \rho \mathcal{M}'(1-\tau) + \frac{\partial \lambda_d(d)}{\partial R_l'}. \tag{29}$$

Equations (27)-(29) contain all other banks strategies $\mathbf{l}_{-b}^\chi = [l_1^\chi, ..., l_B^\chi] \backslash \{l_b^\chi\}$ in the discount factor $\mathcal{M}'$ and its derivatives, in the marginal equity issuance costs $\lambda_d(d)$ and $\lambda_d(d')$, and also the derivatives of the future firms' policy functions inside the derivatives of $\mathcal{M}'$ and inside the derivatives of $\lambda_d(d')$. Since all bank strategies enter into $\mathcal{M}'$ via aggregate consumption, $\mathcal{M}'$ serves as a pivotal channel for strategic interactions. In the stationary equilibrium, $\mathcal{M}' = \beta$; however, for any $\gamma > 0$, its derivatives remain nonzero (see Appendix B.2 for details). This underscores the essential role of utility curvature: without it, aggregate consumption would not enter the stochastic discount factor and these strategic channels would vanish.

4.4.3 Entry Conditions

The new numbers of banks, B' and $\mathcal{B}'$, are the biggest integers such that

$$\mathcal{M}(X, X') \cdot V_b^{\chi}(X') \geq F_E^{\chi} \quad \text{for } \chi \in \{., s\}, \tag{30}$$

where F_E^{χ} denotes the fixed entry cost for big ($\chi = .$) and small ($\chi = s$) banks, respectively. The updated numbers of banks satisfy $B' = B - B^{\text{Default}} + B^{\text{Entry}}$ and $\mathcal{B}' = \mathcal{B} - \mathcal{B}^{\text{Default}} + \mathcal{B}^{\text{Entry}}$, and are components of the vector X'. Here, $(B^{\text{Default}}, \mathcal{B}^{\text{Default}})$ are the numbers of big and small banks that default, as determined by conditions (19), and $(B^{\text{Entry}}, \mathcal{B}^{\text{Entry}})$ are the numbers of new entrants. Entry of a new bank requires an initial capital injection from the household.

4.5 Oligopolistic Equilibrium

The aggregate resource constraint of the economy is:

$$C + \int I + \Lambda \, \mathrm{d}j + T = \int Y \, \mathrm{d}j + \int R \, \mathrm{d}j, \tag{31}$$

where $I \equiv \mathcal{I} \cdot i(x, X) + (1 - \mathcal{I}) \cdot i(x_0, X)$, $\Lambda \equiv \mathcal{I} \cdot \lambda(x, X) + (1 - \mathcal{I}) \cdot \lambda(x_0, X)$, $Y \equiv \mathcal{I} \cdot Zzk^{\alpha} + (1 - \mathcal{I}) \cdot Zz_0 k_0^{\alpha}$, and $R \equiv (1 - \mathcal{I}) \cdot \left(\sum_{b=1}^{B} \mathcal{R} + \sum_{b=1}^{\mathcal{B}} \mathcal{R}^s\right)$. The total production of the economy on the right-hand side of equation (31) can: (i) be consumed by the household, (ii) be used for aggregate investment in physical capital (in case some dividends are negative, some resources are spent to pay the equity issuance cost λ), and (iii) be paid in taxes. Note also that the right-hand side includes the resources recovered from the mass of defaulting firms, captured by the term R.

A formal definition of the notion of *Recursive Stationary Oligopolistic Equilibrium* is presented in Definition 4.1, and its extension to the dynamic case is discussed in Section 6.

Definition 4.1. A **Recursive Stationary Oligopolistic Equilibrium** is a Markov perfect equilibrium where i) the banks' debt holdings ($\{D_b\}_{b=1}^{B}$,$\{D_b^s\}_{b=1}^{\mathcal{B}}$) and the relative market rate R_D; ii) the banks' share holdings ($\{S_b\}_{b=1}^{B}$, $\{S_b^s\}_{b=1}^{\mathcal{B}}$) and the relative market prices ($\{p_b\}_{b=1}^{B}$, $\{p_b^s\}_{b=1}^{\mathcal{B}}$); iii) the interbank debt holdings ($\{M_b\}_{b=1}^{B}$, $\{M_b^s\}_{b=1}^{\mathcal{B}}$) and the relative market rate R_M; iv) the household's consumption C; v) the distribution $\phi(x)$; vi) the policy functions: $k'(x)$, $l'(x)$, $l^{s\prime}(x)$, and $R_l'(x)$; vii) the number of banks (B,$\mathcal{B}$) are such that i) the household's problem is solved–i.e, equations 13) hold; ii) each firm's problem is solved–i.e, equations (15) and (16); iii) each incumbent bank is best responding to all other banks–i.e, equations (21), (23), (24), and (25) hold; iv) it is not profitable for new banks to enter the intermediation market as per entry conditions (30) and incumbent banks have no incentive to default; v) and all markets clear: 1) the good market clears –i.e., equation (31) holds;

2) each bank's equity market clears –i.e., $\forall b : S_b^\chi = 1$; 3) the interbank market clears –i.e., $\forall b : M_b^\chi = 0$.

5 Calibration

The choices of parameter values are described below and summarized in Table C1.

Household Preferences. A period in the model coincides with a quarter, consistent with the frequency in the data. I set $\beta = 0.995$ to match a quarterly stationary bank's debt rate (or interbank debt rate) of 0.50%, which is consistent with the average of the 3-month T-bill rates calculated between 1997 and 2017. I use a constant relative risk aversion utility function with degree of relative risk aversion $\gamma = 1$. Hence, $u(C) = \log(C)$.

Firms. The quarterly depreciation rate δ is set to a standard value of 3%. In the baseline, I set the curvature of the reduced-form profit function α to 0.34.[14] I then conduct a sensitivity analysis by increasing θ to 0.75, which yields $\alpha = 0.51$, holding all other parameters constant. In this case, the banking sector profitability rises to 26.4%, compared to 16.5% under the baseline value of $\alpha = 0.34$. Banks profitability increases further to 30.5% when I raise θ to approximately 0.777, which implies $\alpha = 0.55$.[15] As shown in Figure C7 in Appendix C.1, when $\alpha = 0.51$, banks maintain positive markups across a broader segment of the firm distribution—particularly among lower-productivity firms. As α increases, the marginal return on capital is lower for very small firms but diminishes at a slower rate as capital rises. This implies that banks internalize a shift in marginal value: one dollar lent to a tiny firm yields less marginal return when α is higher, but becomes more valuable as the firm accumulates capital. This analysis suggests that the baseline value $\alpha = 0.34$ is conservative in terms of implied bank markups. See Appendix B.5 for further details.

Firms' income tax is set to 19.7%, which was the effective corporate income tax in the U.S. in 2021. This taxation rate is also broadly consistent with the ratio between taxes on corporate income over corporate profit (both time series are retrievable from FRED) between 1997 to 2017. I use the net charge-off rates of the Commercial and Industrial Loans (C&I) to pin down the risk of default, so that the stationary equilibrium value of ρ is $\bar{\rho} = 1 - 0.21\%$ (quarterly) consistent with the average of the time series between 1997

[14] This value can be obtained using the mapping implied by a standard Cobb–Douglas production function with decreasing returns to scale. Specifically, consider the functional form $Zz^v(k^a n^{1-a})^\theta$, where k is capital, n denotes labor input (not explicitly modeled in this paper), $a = 0.35$ is the capital share, and $\theta = 0.6$ governs returns to scale. Defining $v = 1 - (1 - a)\theta$, the reduced form curvature of profits with respect to capital is $\alpha = \frac{a\theta}{v} = \frac{0.35\times0.6}{1-0.65\times0.6} \approx 0.34$.

[15] These values of α are broadly consistent with estimates reported in the literature—for example, Cooper and Haltiwanger (2006) and Hennessy and Whited (2007) document values of 0.592 and 0.627, respectively.

and 2017.[16] The equity issuance cost λ_0 is set to 0.8 to match an annualized frequency of equity issuance of 0.042. The frequency of equity issuance is computed from a sample of non-financial, unregulated firms from Compustat. Several studies (e.g., Gomes, 2001 and Hennessy and Whited, 2007) incorporate an equity issuance cost that includes both fixed and proportional components. In Appendix C.2, I explore an alternative functional form for the cost of equity issuance, which incorporates a fixed cost through a logit function to ensure differentiability. In each period, I also assume that a new mass of firms replaces the mass of firms that defaulted $1-\bar{\rho}$, starting from age zero with capital k_0, which is set jointly with other parameters as explained below. In Subsection 6.2, I consider an extension where firms entry change over time and not necessarily match the firm exit mass.

Note that the parameter k_0, which governs the initial capital stock of reentering firms, plays an important role in both the recovery mechanism and the determination of firm leverage. In the event of default, banks recover $\nu k + k_0$ from the defaulting firm's capital. I assume that k_0 is transferred to the household (the shareholder of both firms and banks) and recycled to finance newly entering firms. This assumption ensures that entry is not fueled by exogenous resource injections, but rather by reallocating existing capital within the economy. Consequently, k_0 directly influences the recovery rate, alongside the parameter ν, which governs the share of physical capital recovered by lenders. Beyond its role in recovery, k_0 also affects firms' leverage decisions: a smaller k_0 implies lower initial production capacity, requiring firms to borrow more if they aim to grow quickly. Banks internalize this demand for external financing and adjust contract terms accordingly.

Banks. I set the capacity constraints parameter κ to 0.1 consistent with the ratio between the tenth biggest bank and the first (see Figure C14 in Appendix C.4.2), in order to mimic the dominant-fringe market structure observed in the data. While this calibration anchors the model to an empirical reference point, Appendix B.4 shows that the stationary equilibrium depends only on the product $\mathcal{B}\kappa$, and similar aggregate outcomes can be obtained under alternative configurations that accommodate a longer tail of small banks.

The fixed entry cost F_E for big banks is chosen to match the aggregate profitability of the banking sector. For simplicity, I assume $F_E^s = \kappa \cdot F_E$. Given the small bank capacity constraint $l_b^{s'} = \kappa \cdot l_b$, the small bank value function is proportional: $V_b^s(X') = \kappa \cdot V_b(X')$. Under these assumptions, the entry conditions for big and small banks collapse to the same condition. Specifically, the small bank entry condition $\mathcal{M}(X, X') \cdot V_b^s(X') \geq F_E^s$ becomes $\mathcal{M}(X, X') \cdot V_b(X') \geq F_E^s/\kappa$, which is identical to the big bank entry condition. In stationary equilibrium, $X = X'$, with $B = B'$ and $\mathcal{B} = \mathcal{B}'$. To anchor the composition of entrants, I first determine the number of big banks B' such that adding one more big bank would render

[16] I also verify numerically that the firm's equity value remains strictly positive for all firms in the distribution, hence introducing endogenous firm default in this baseline model would be immaterial under the current calibration. See Section 7 for an extension that incorporates endogenous firm default.

entry unprofitable, and then choose the number of small banks $\mathcal{B}'$ to finely adjust the degree of market competition and match aggregate profitability.

The aggregate profitability of the banking sector is calculated as *net operating income* over *total interest income*, which can be obtained from the *Quarterly Income and Expense of FDIC-Insured Commercial Banks and Savings Institutions*.[17] In particular, I use the average of a quarterly time series from 1984 and 2018. To summarize, I target: (i) the profitability of the banking sector through the fixed entry cost F_E, (ii) the frequency of equity issuance through the parameter λ_0, (iii) the market leverage through the parameter k_0, (iv) the recovery rate through the parameter ν, and (v) the standard deviation of investment rate through the parameter σ_z. The model targets a cross-sectional standard deviation of investment rates of 0.337, in line with other studies; see Khan and Thomas (2013), Clementi and Palazzo (2016), and Lanteri (2018).[18] All remaining moments are untargeted and reported for validation. Table I summarizes all targeted and untargeted moments.

Table I. Stationary Equilibrium and Annualized Moments

Targeted	**Description**	**Moment**	**Model**	**Data**
Yes	Profit/Revenue	$\pi_b / \int r_l(x,X) l_b d\Phi$	16.5%	16.2%
Yes	Freq. of Equity Iss.	$4 \int (d(x,X) < 0) d\Phi$	4.3%	4.2%
Yes	Market Leverage	$\int (B l_b(x,X) + \mathcal{B} l_b^s(x,X)) / V_F(x,X) \, d\Phi$	38%	34%
Yes	Std. Investment Rate	$2\sqrt{\int (i/k)^2 d\Phi - \left(\int (i/k) d\Phi\right)^2}$	0.333	0.337
Yes	Recovery Rate	$\int \mathcal{I}(\nu k + k_0) / (\mathcal{I} B l_b(x,X) + \mathcal{I} \mathcal{B} l_b^s(x,X)) d\Phi$	0.46	0.51
No	Capital to GDP	$K/(4Y)$	2.2	2.2
No	Investment to K	$4I/K$	14%	16%
No	Debt Adjust. to K	$\int 4(B \Delta l_b'(x,X) + \mathcal{B} \Delta l_b^{s\prime}(x,X)) \, d\Phi / K$	0.92%	0.62%
No	Num. Big Banks	B	3	-
No	Num. Small Banks	$\mathcal{B}$	7	-
No	Std. MPK	$2\sqrt{\int (Zz\alpha k^{\alpha-1})^2 d\Phi - \left(\int Zz\alpha k^{\alpha-1} d\Phi\right)^2}$	0.16	0.68

Notes: This table reports the targeted and untargeted aggregated annualized moments. Capital and investment in the data are computed from current-cost net stock of private fixed assets, retrievable from the U.S. Bureau of Economic Analysis. Loans in the data are computed from the C&I Loans retrievable from FRED. Leverage in the data is computed as leverage of non-financial corporate business (debt as a percentage of the market value of corporate equities) also retrievable from FRED. All moments refers to the average of the yearly time series from 1997 to 2017.

The calibrated model features an economy populated by 3 big and 7 small banks. More generally, I view the number of banks in the model as a proxy for the degree of strategic interaction in the market rather than to match the institutional count. See Appendix B.4 for further discussion. Note also that the banking sector in the U.S. is characterized by a dominant-fringe market structure. As an example, in 2018, the biggest four banks in terms

[17] https://www.fdic.gov/analysis/quarterly-banking-profile/index.html.

[18] This is consistent with micro-level evidence from Cooper and Haltiwanger (2006).

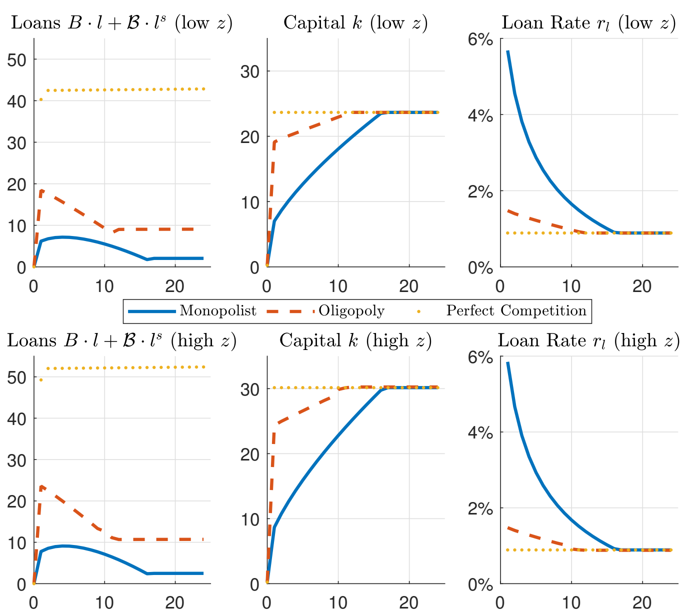
of assets were: (i) JP Morgan Chase & Co with a market share of 14.45%, (ii) Bank of America Corp. with a market share of 11.73%, (iii) Wells Fargo & Company with a market share of 11.17%, and (iv) Citigroup Inc. with a market share of 9.32%. The fifth bank was U.S. Bancorp with a significantly smaller market share of 3.02%, less than a third that of Citigroup Inc.

The value for the standard deviation of MPK, 0.68, is sourced from David, Schmid, and Zeke (2022). My model generates, through bank market power alone, a value of 0.16 (∼23.5% of the total). I view the fact that the channel of bank market power can explain a relevant portion of the standard deviation of MPK but not all of it as natural, as there could be several other frictions at play as also described by David, Schmid, and Zeke (2022).

5.1 Firm Dynamics in the Stationary Equilibrium

I now describe the key properties of the stationary equilibrium of the calibrated model, with a greater focus on the role of strategic interactions. Figure 2 reports the relevant stationary equilibrium policy functions across firms of different ages and productivity levels.

Figure 2. Stationary Equilibrium Policy Functions



Notes: This figure reports the stationary equilibrium policies for loan quantity (left panels), physical capital (middle panels), and loan interest rate (right panels), across firms of different ages and productivity levels. The x-axes indicate firm age.

Firms can reach their capital objectives by (i) investing internal resources, (ii) issuing equity, or (iii) demanding external financing resources on the loan market. A more concentrated banking sector reduces the credit availability in the economy. Firms with a high marginal productivity of capital and worse outside options, likely smaller or highly leveraged, exhibit a higher and less elastic credit demand. Therefore, this type of firm is more exposed to the negative effects of the lack of competition in the banking sector. This is the same intuition captured by equation (9) in the stylized model of Section 3. Since markups are endogenous in the cross-section of firms, banks endogenously exert a higher degree of market power on firms with a high marginal productivity of capital and lower internal resources; hence, fewer outside options but a higher marginal value from growth. These firms need banks' credits and would otherwise incur an equity issuance cost to finance their growth, in case their current production alone would not be sufficient to sustain the desired level of physical investment. An imperfectly competitive financial sector internalizes that the same financial resources are more valuable for firms with a higher marginal productivity of capital and fewer outside options (e.g., a higher marginal equity issuance cost) and, therefore, can charge higher markups. This creates a mechanism of endogenous financial friction, as captured by the central panels of Figure 2.

Through this mechanism, limited competition in the financial sector not only induces credit misallocation—slowing firm growth—but also lowers aggregate productivity. This outcome is illustrated in Figure C1. The mechanism aligns with the empirical observation that firms more dependent on bank credit—such as small and private firms (Holmstrom and Tirole, 1997, Diamond and Rajan, 2005, Chodorow-Reich, 2013, Saunders, Spina, Steffen, and Streitz, 2024)—are disproportionately affected by limited competition, while firms with access to alternative funding sources, such as public bond markets, are less exposed to these frictions (Chava and Purnanandam, 2011).

As noted by Dempsey (2025), "while commercial banks account for a sizable subset of total lending in the U.S. economy, other lenders also supply significant amounts of credit." While the model can, in principle, be interpreted as referring to financial intermediaries more broadly, the quantitative calibration focuses specifically on bank lending. Incorporating non-bank financial institutions into the model could help capture substitution across credit channels and offer a more comprehensive picture of aggregate credit dynamics. Importantly, the mechanism of endogenous financial frictions described in this section would remain operative in such an environment. In connection with Figure 2, substitution toward non-bank finance appears more likely in the middle market segment, where relatively larger firms have better access to alternative sources of funding.

Figure C2 in Appendix C.1 reports the inverse elasticities $\eta_l'(x, X, x', X')$ from the generalized Euler equation (23), evaluated in the stationary equilibrium for firms of different ages. These elasticities, which are endogenous and depend on current and future firm-level characteristics, are directly linked to the equilibrium trajectories of markups. A lower inverse elasticity translates into a higher financial markup. Furthermore, Figure C2 in Appendix C.1 shows that the higher the concentration of the banking sector, the lower the inverse elasticities and the longer it takes for a firm to reach its efficient level of capital. Under perfect competition, the inverse elasticity is constant in the cross-section of firms. Long-lived firms still experience a non-zero elasticity because of the interaction of bank market power with the curvature of the household utility function and the presence of a tax shield, so the level of loan is always well-defined.[19]

5.2 Validating the Mechanism

Qualitatively, this mechanism is consistent with stylized fact 3 ("controlling for deal amount and a proxy for corporate default, smaller firms tend to pay higher credit spreads"), with parameters of interest reported in Table C7 in Appendix C.4.1. This is qualitatively consistent with the model since, as suggested by the right panels of Figure 2, interest rates (hence, credit spreads) are bigger for the smaller firms, holding default risk fixed.

Quantitatively, I compare the stationary equilibrium results from Section 5 with the empirical regression presented in Table C7. It is worth noting that this quantitative analysis should be interpreted with caution, as the Compustat-DealScan dataset includes public firms, which are inherently large, whereas the model suggests that smaller firms are the most affected by market power. Nevertheless, the variation in firm size within Compustat-DealScan can still be utilized to conduct the analysis.

As presented in Table C1, firms begin with a calibrated initial capital (k_0) of 0.237. In the calibrated oligopoly depicted in Figure 2, k_0 corresponds to a quarterly credit spread (relative to perfect competition) of approximately 0.587%.[20] Over time, firms grow to their efficient capital level and ultimately pay the competitive rate, implying that credit spreads relative to the perfect competition scenario converge to zero. In the calibrated oligopoly (illustrated by the dashed-red line), this occurs at a capital level of approximately 27.6, aggregating firms of varying productivity. Therefore, if default risk is held constant as in the model, regression (1) in Table C7 predicts the subsequent decline in the (annualized) credit spread:

$$\underbrace{-0.2535 \cdot [\log(26.7) - \log(0.237)]}_{\simeq -1.2\%} - 0.0083 \cdot [\log(9.8743) - \log(B \cdot l_0 + \mathcal{B} \cdot l_0^s)] - 0.0373 \cdot \underbrace{0}_{\text{Const. Altman}} .$$

The second term is negative since, consistently with the model, firms start with zero loans.[21] Hence, the second term contributes to a reduction in interest rates. Thus, focusing exclusively on the interest rate reduction due to changes in firm size represents a conservative estimate. For comparison with the model, I divide -1.2% by 4, resulting in a quarterly rate reduction of -0.3%. Within this range, the model predicts a quarterly credit spread

[19] See Appendix B.2 for further details.

[20] Under perfect competition, interest rates reflect the risk-free rate plus compensation for default risk, which is uniform across firms in the baseline model. Thus, credit spreads represent the portion not attributable solely to default risk and, within the model's framework, are interpreted as markups.

[21] Although an exact zero is not feasible in logarithmic terms, the reader can think of small initial values.

reduction of -0.587%, demonstrating a comparable magnitude. The fact that the model yields a larger reduction aligns with the observation that the Compustat-DealScan dataset primarily consists of public firms, which are inherently larger and therefore less impacted by market power, consistent with the model's predictions.

5.2.1 Relationship to the Empirical Literature

An extensive body of empirical research has explored the impact of bank competition on firms, with seminal works by Rajan and Zingales (1998), Black and Strahan (2002); Cetorelli and Gambera (2001); and Cetorelli and Strahan (2006). While evidence varies, the consensus aligns with my model, indicating that banks' market power reduces the total credit available, though the effect differs across firms. Young firms face higher credit demand and are more affected by limited competition (see, Freixas and Rochet, 1997). The less competitive the conditions in the credit market, the lower the incentive for lenders to finance newcomers as documented by Petersen and Rajan (1995) and Cetorelli and Strahan (2006). Although my model does not focus on capturing complex features of relationship banking, the desire for inter-temporal smoothing of banks' profits and households' consumption – embedded in the dynamic contract of equations (15), (16), and (23) – captures the fact that creditors, when contracting markups, take into account the expected stream of future dividends of the firms, as well as their own future profits. Hence, in my model, banks balance markups intertemporally in the spirit of relationship lending. This important economic force is consistent with Petersen and Rajan (1995) and Cetorelli and Strahan (2006).

6 Aggregate Shocks

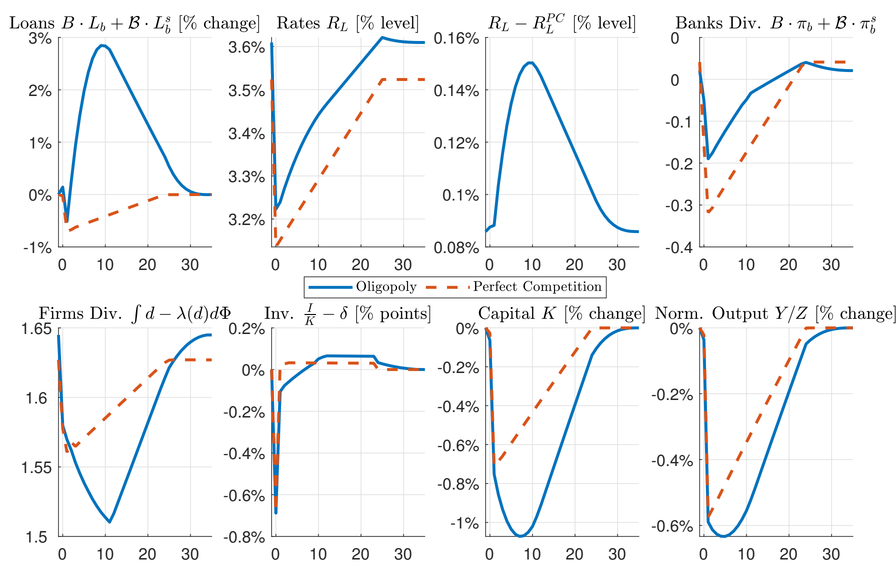
This section analyzes the role that banks' market power plays in the transmission of macroeconomic shocks. The section includes three unexpected aggregate shocks: (i) a decrease to aggregate TFP combined with an increase to the aggregate firms' default probability, (ii) a temporary change to the bank market structure (in the model, the first shock combined with an idiosyncratic shock to the assets of the big banks calibrated to push one bank to default), and (iii) a permanent change to the bank market structure (in the model, the second shock combined with a permanent increase to the fixed cost F_E to enter the banks' market, calibrated to push one bank to default without a subsequent new bank reentry on the equilibrium path).

I compute the transitional dynamics of the model initialized at the stationary equilibrium as defined in Section 5. Then, I hit the economy with the unexpected aggregate shocks and, depending on the type of shock, the economy converges to the old (or a new) stationary equilibrium in the long run. Several papers assume agents did not foresee the aggregate shocks of the Great Recession (e.g., Guerrieri and Lorenzoni, 2017). Along the transitional dynamics, after the shock, I assume all agents can perfectly foresee the paths of all aggregate variables. In order to compute the equilibrium dynamics, I find sequences of: (i) aggregate consumption $\{C_t\}_{t=0}^{T}$, and (ii) firms' distributions $\{\phi_t(x_t)\}_{t=0}^{T}$; such that households maximize utilities, all markets clear in each period and the firms' distributions evolve according to: (i) the firms' policy functions, (ii) the banks' generalized Euler equations (23) and (iii) the idiosyncratic default shocks.

6.1 A TFP Shock

This section investigates the effects of banks' market power when the economy is hit at time $t = 0$ by an unexpected negative aggregate TFP shock of 2.5% (in correspondence of which the quarterly firms' survival probability ρ decreases by 0.7%) as shown by Figure C4 in Appendix C.1. The shock is calibrated to match the size of the Great Recession.

Figure 3. TFP Shock, Financing, and Real Activity



Notes: This figure reports the transitional dynamics of: (i) the aggregate loans $B \int l_{b,t}(x_t, X_t)\, \mathrm{d}\Phi_t$, (ii) the annualized aggregate interest rate $\int r_{l,t}(x_t, X_t)\, \mathrm{d}\Phi_t$, (iii) the annualized aggregate interest rate spread calculated as the difference between the annualized aggregate interest rate $\int r_{l,t}(x_t, X_t)\, \mathrm{d}\Phi_t$ under the calibrated oligopoly and the annualized aggregate interest rate $\int r_{l,t}^{\mathrm{PC}}(x_t, X_t)\, \mathrm{d}\Phi_t^{\mathrm{PC}}$ under perfect competition, (iv) the aggregate banks dividend $\sum_{b=1}^{B_t} \pi_{b,t}$, (v) the annualized aggregate firms dividend $\int d_t(x_t, X_t) - \lambda(d_t(x_t, X_t))\, \mathrm{d}\Phi_t$, (vi) the aggregate physical investment $(\int i_t(x_t, X_t)\, \mathrm{d}\Phi_t)/(\int k_t\, \mathrm{d}\Phi_t)$ net of depreciation (quarterly), (vii) the aggregate capital $\int k_t\, \mathrm{d}\Phi_t$ and (viii) the aggregate output (normalized to remove the exogenous component Z_t to isolate the effects on the endogenous component) $\int k_t^{\alpha}(x_t, X_t)\, \mathrm{d}\Phi_t$, following the TFP shock reported in Figure C4 in Appendix C.1. X-axes report time t.

Note that the banks entry condition (30) and the banks exit condition (19) must hold during the transition. Given the calibration of Table I, this shock does not induce any new bank to enter or an incumbent bank to exit. Figure 3 reports the dynamic responses calculated (i) with the calibrated oligopolistic banking sector of Section 5 (solid line) and (ii) with the corresponding perfectly competitive banking sector (dashed line). As discussed in Subsection 5.1, a more concentrated banking sector can extract higher rents out of the financially constrained firms with a high marginal productivity of capital. In the stationary equilibrium, this mechanism endogenously creates slower growth trajectories as a function of the banks' market structure. In the dynamics, this mechanism of endogenous financial frictions interacts with the higher density of financially constrained firms, generating a higher credit demand that drives up markups as shown in Figure 3.[22] Therefore, during the transitional dynamics, interest rates rise more under oligopolistic competition than under perfect competition. Moreover, total loans rise under oligopolistic competition. Banks exploit their market power not only to extract higher markups, but also to prevent their profits from falling as much as in the perfectly competitive case. When the shock hits, banks incur large losses and require an equity injection from households (i.e., a negative bank dividend). Nevertheless, the shock is calibrated to ensure that no bank defaults in equilibrium, as $\forall b = \{1, \ldots, B\} : \ \tilde{V}_{b,0} > 0$ and $\forall b = \{1, \ldots, \mathcal{B}\} : \ \tilde{V}^s_{b,0} > 0$, where $\tilde{V}_{b,0}$ and $\tilde{V}^s_{b,0}$ denote the time-0 value functions of large and small banks, respectively.

In summary, in such conditions, a concentrated banking sector exploits its market power to extract higher interest rates. This mechanism induces a larger decline in real activity in terms of aggregate investment, capital, and output. Finally, note that the decline in real activity captured by Figure 3 further constrains firms by restraining households' capacity to support firms with equity issuance (the aggregate firms' dividend declines). This mechanism creates a vicious cycle that further reduces the interest rate-quantity loan elasticities $\eta_l'(x, X, x', X')$ and boosts banks' interest rates. Throughout the paper, I assume that the mass of defaulting firms is replaced, at each date t, by an equal mass of new entrant firms. In Subsection 6.2, I examine the effects of an entry rate temporarily below the exit rate, as occurred during the Great Recession.

6.2 The Effects of the Firms' Entry Rate

During the Great Recession, the firms' default rate increases but not all firms immediately re-enter the market. In this section, I let the mass of exiting firms evolves according to the evolution of the default rate in the shock reported in Figure C4 in Appendix C.1. Differently from before, I keep the entering mass of firms initially lower than the exit mass of firms and higher after 10 periods. The entry rate, along the shock, is calibrated so that the mass

[22] This is consistent with the idea that firms more dependent on external funding via bank loans, such as small and private firms, can become more financially constrained when credit conditions tighten (Holmstrom and Tirole, 1997; Diamond and Rajan, 2005; Chodorow-Reich, 2013; Saunders, Spina, Steffen, and Streitz, 2024).

of firms in the production sector drops in the short-run by about 1 percent and returns to mass 1 in the long-run. In Figure C3 in Appendix C.1, I refer to this type of shock with the label *Variable Entry Mass*. In contrast, the label *Entry Mass = Exit Mass* refers to the shock already analyzed in the previous Subsection 6.1, whose effects on real activity are reported in Figure 3. Figure C3 suggests that, when the firms' default rate increases but not all firms immediately re-enter the market, then imperfect competition in the financial intermediation sector leads to a bigger and delayed amplification effect at the peak that fades away as new firms enter the production sector. The intuition is linked to the economic mechanism discussed in Subsection 6.1. Banks can extract higher interest rates for longer as firms slowly re-enter the production sector (as shown by the left panel of Figure C3). A concentrated banking sector extracts higher rents out of the financially constrained firms that slowly re-enter the market and need credits. Output eventually converges back to its pre-shock level as the mass of producing firms reverts back to 1.

6.3 Changes in Bank Market Structure

One salient ingredient of my framework is the presence of non-atomistic financial intermediaries. Thanks to this feature, I can study market structure changes, such as a bank failure, and a bank reentry. The economy is hit at time $t = 0$ by the same shock of Subsection 6.1, which is reported in Figure C4 in Appendix C.1. Furthermore, I introduce an idiosyncratic shock ζ_t^χ to the assets of the banks such that equation (20) becomes:

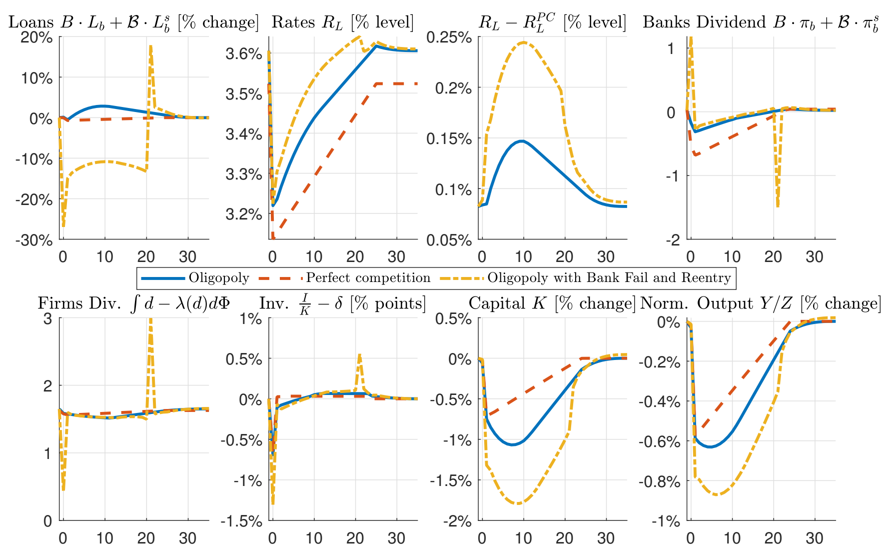
$$\pi_b^\chi = (1 - \zeta_t^\chi) \cdot \int \mathcal{I} \cdot r_l \cdot l_b^\chi \, \mathrm{d}j + r_M M_b^\chi - r_D D_b^\chi - F^\chi. \tag{32}$$

I calibrate the shock so that it is rational for one big bank to default on the equilibrium path at time $t = 0$. In particular, the shock is such that $\forall t : \zeta_t^s = 0$ (small banks) and $\zeta_0 = 0.016$ (big banks) and then it reverts back to 0 linearly with a persitence calibrated consistently with that of the TFP shock. In Subsection 6.4, I report the resulting equilibrium path which includes a new bank entry at time $t = 21$. This is in line with the events of the Great Recession, during which there were numerous bank failures, bailouts, as well as significant failures and mergers involving major banks (e.g., the bankruptcy of Lehman Brothers, the collapse of Washington Mutual, and the acquisition of Wachovia by Wells Fargo in 2008). Additionally, the Great Recession was marked by a general trend of consolidation in the banking market structure, accompanied by a sharp decline in new bank entries since 2009. In order to mimic the collapse in bank entry and the long run trend of market structure consolidation of the banking sector, in Subsection 6.5, I combine the shock of Subsection 6.4 with a permanent increase in the fixed cost to entry F_E, so that it is not profitable for a new bank to enter the credit market along the equilibrium path.

6.4 Bank Failure and Banks' Market Power

In this subsection, the fixed cost to entry F_E is held fixed throughout the dynamics, allowing one bank to enter the credit market along the equilibrium path. At $t = 0$, when the shock hits, one incumbent big bank decides to exit since $\tilde{V}_{b,0} < 0$ and the market structure changes temporarily from 3 big banks and 7 small banks to 2 big banks and 7 small banks. Note that the shock is calibrated such that exactly one big bank chooses to default; that is, when one bank exits at $t = 0$, the value functions of all remaining banks become positive.[23] Additionally, since all large banks are identical, any of them could potentially default. It is immaterial for the equilibrium dynamics which specific bank defaults because, due to symmetry, any default scenario would yield identical equilibrium outcomes. The equilibrium path is determined by the fact that, given it is optimal for exactly one bank to default—regardless of which one—it is also optimal for the remaining banks to continue operating.

Figure 4. Bank Default and Reentry, Financing, and Real Activity



Notes: This figure reports the transitional dynamics following the TFP shock reported in Figure C4 in Appendix C.1. At $t = 0$ (on impact) one incumbent big bank exits. At $t = 21$, it is profitable for one big bank to enter the market given that the fixed cost F_E held fixed at its stationary equilibrium value. X-axes report time t.

At $t = 21$, it is profitable for a new bank to enter the credit market given the fixed cost F_E has not changed. Hence, the market structure changes back from 2 to 3 big banks. Figure 4 reports the resulting dynamics. At the beginning of period $t = 0$, when one bank decides to exit, non-defaulting firms fully repay their outstanding loans to all banks, including

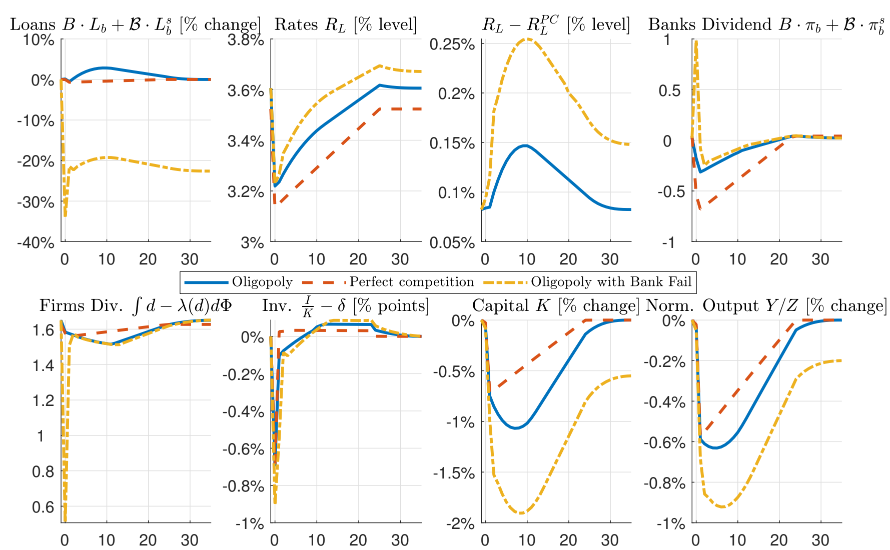
[23]This requires solving for the unexpected shock multiple times and checking the resulting value functions under different market structures.

the exiting bank. To prevent the destruction of assets, the defaulting bank liquidates its remaining assets, distributing them to the household as dividends. All loans are repaid prior to the change in market structure, as agents only trade one-period securities. Following the change, all agents make decisions in accordance with the new market structure. When a big bank fails, the surviving banks start to slowly extend more credit to firms in order to partially capture the market share of the defaulted bank. However, the speed of this adjustment is dampened by the decreased level of competition among surviving banks. The surviving banks' market power interacts with credit constraints, yielding a sharp drop in the aggregate volume of credit in the short run with interest rates that quickly increase after the initial drop. Moreover, banks lower credit supply in anticipation of the new bank's entry. The credit crunch propagates into the real economy, yielding a sharp and persistent drop in investment; hence physical capital and output. The entry of the new bank at $t = 21$ requires an initial equity injection from households (i.e., a negative bank dividend). The resulting expansion in credit supply increases loan availability and lowers interest rates, which raises firms' aggregate dividends. Investment in physical capital also temporarily rises, bringing capital and output back to their pre-default levels.

6.5 A Permanent Change in Bank Market Structure

In this subsection, I combine the shock of Subsection 6.4 with a permanent increase in the fixed entry cost F_E, so that it is not profitable for a new bank to enter the credit market along the equilibrium path.

Figure 5. Bank Default, Financing, and Real Activity



Notes: This figure reports the transitional dynamics following the TFP shock reported in Figure C4 in Appendix C.1. At $t = 0$ (on impact) one incumbent big bank exits. X-axes report time t.

When the market structure of the banking sector changes permanently, the model captures two ideas: (i) in the short term, the effects of market power of the surviving bank contributes to lowering the supply of credit to firms, further slowing down the economy and (ii) in the long run, the heightened banks' market power further contributes to amplifying and prolonging the recession. Figure 5 reports the resulting dynamics. Because of the general equilibrium effects and the reduced level of competition, in the long run, the economy stabilizes at a lower level of volume of credit, which results in less investment, capital, and output. Subsection 6.6 discusses the dynamic effects on the dispersion of loan rates (associated with the dispersion of marginal products of capital) and aggregate TFP.

6.6 Dispersion of Loan Rates and Aggregate TFP

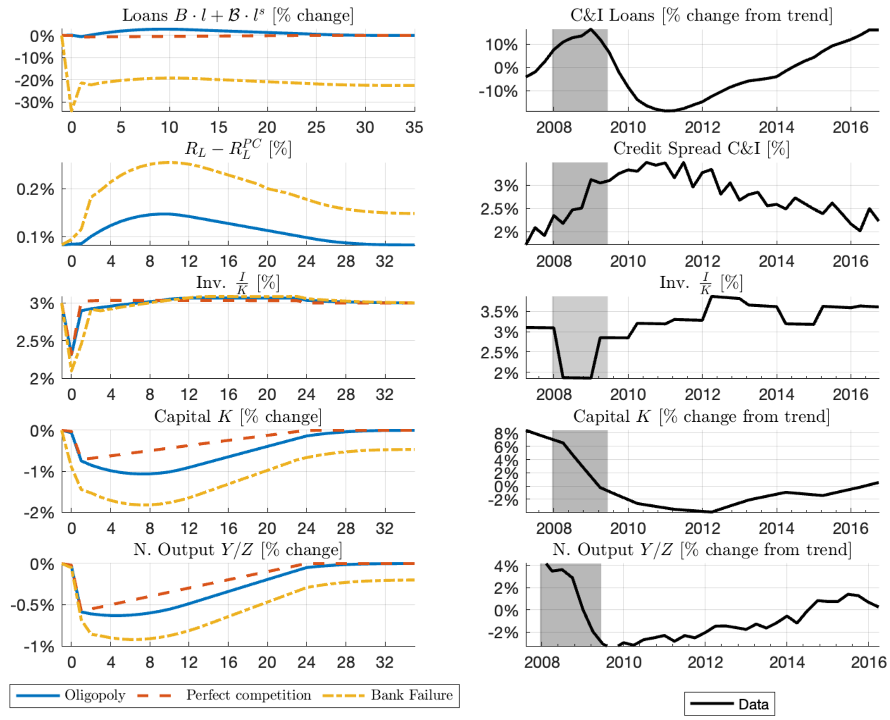
In this subsection, I study the effects on the dispersion of loan rates and aggregate TFP of the shock of Subsection 6.5. The dynamic financial oligopoly combined with heterogeneous firms generates firm-level endogenous financial frictions that create time-varying second moments, such as the dispersion of loan rates (directly linked to the dispersion of marginal products of capital) and aggregate TFP. The left panel of Figure C8 in Appendix C.1 reports the dynamic of the standard deviation of loan rates, expressed in percentage levels. The right panel of Figure C8 reports the associated aggregate TFP (calculated as the residual of an aggregate production $Y_t = \text{TFP}_t \cdot K_t^{\alpha}$). Both measures are calculated as difference from the perfectly competitive benchmark.

In agreement with empirical evidence (e.g., Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry, 2018 and David, Schmid, and Zeke, 2022), the model produces a dynamic with an increasing dispersion of loan rates during recessions; hence, an increasing dispersion of marginal productivity of capital. Figure C8 also suggests that the failure of a large bank induces a small but persistent decline in aggregate TFP.

6.7 Comparison with the Great Recession

Figure 6 shows that the model dynamics of Subsection 6.5 are broadly consistent with those of the Great Recession. For example, in the aftermath of the financial crisis C&I credit spreads increased at the peak by 1.5 percentage points (annualized). As shown in the second panel on the left, the model suggests that, at the peak, approximately 0.15 percentage points and 0.25 percentage points of these credit spreads are attributable to financial markups (calculated with and without bank default, respectively). Moreover, the model captures a drop in investment rate similar in magnitude to that of the Great Recession, and a persistent drop in capital (around 2% at the peak) and output (around 1% at the peak).

Figure 6. Model vs. the Great Recession



Notes: This figure compares the model dynamics of Subsection 6.5 with the data in proximity of the Great Recession. In particular, the right column reports: i) C&I loans (expressed in % deviation from the linear trend calculated from 1997:Q2 to 2017:Q2), ii) C&I credit spread, iii) quarterly investment rate, iv) physical capital (also expressed in % deviation from the linear trend calculated from 1997:Q2 to 2017:Q2), and v) normalized output (also expressed in % deviation from the linear trend calculated from 1997:Q2 to 2017:Q2). Data are linearly detrended. The panels on the right columns report the percentage change of the detrended series from the trend. X-axes report time t, expressed in quarters for the model (left column).

The resulting increase in banks' market power amplifies and prolongs the recession, which is suggestive that banks' market power may have played a significant role in amplifying and prolonging the crisis. The figure also depicts the dynamics of total loans. Around the onset of the Great Recession, total loans decline in the model following a bank failure, in line with the empirical evidence. Without bank reentry, total loans remain persistently depressed as the economy converges to a permanently lower level of total credit. In contrast, as shown by Figure C6 in Appendix C.1, when endogenous bank reentry occurs along the equilibrium path, banks reduce lending by less in response to a failure—anticipating future reentry—and the reentry itself restores total loans to their pre-crisis steady state. In this sense, bank reentry helps generate total lending dynamics that more closely align with those observed in the data.

7 Bank Market Power and Endogenous Firm Default

In this section, I extend the model to allow banks' market power to interact with firms' default decisions. Banks strategically internalize the effects of their choices on both current and future firm defaults, while also interacting strategically with other banks. This problem is particularly challenging, as it features a non-differentiable dynamic demand for loans, which renders the Markov Perfect equilibrium non-differentiable and precludes the use of the generalized Euler equations approach. I address this challenge using dynamic discrete choice methods, which yield smooth decision rules and restore differentiability. Similar methods have long been used in various areas of economics (e.g., McFadden, 1977), including industrial organization and, more recently, the sovereign default literature.[24] Following this approach, I introduce shocks to firms' outside option values, assumed to follow a logistic distribution. I then derive the generalized Euler equation that captures the interaction between banks' market power and firms' default behavior.

7.1 Firms

The value function $V_F(x, X)$ of a firm satisfies the following functional equation:

$$\tilde{V}_F(x, X) \ = \max_{\{l_b'\}_b^{\mathcal{B}'}, \{l_b^{s\prime}\}_b^{\mathcal{B}'}, k'} \ d - \lambda(d) + \mathbb{E}\left[\mathcal{I}' \cdot \mathcal{M}' \cdot V_F(x', X')|(x, X)\right], \tag{33}$$

where, at the beginning of each period, each firm decides whether to default or not $\mathcal{I}^d = \{0, 1\}$, so that the function $V_F(x, X)$ satisfies the following functional equation

$$V_F(x, X) \ = \max\left\{\tilde{V}_F(x, X), \epsilon\right\} = \max_{\mathcal{I}^d = \{0,1\}} \ \mathcal{I}^d \cdot \tilde{V}_F(x, X) + (1 - \mathcal{I}^d) \cdot \epsilon, \tag{34}$$

where the outside value shock ϵ follows a logistic distribution with mean zero and scale parameter $\zeta \bar{V}_F$, and $\bar{V}_F > 0$ is the average firm equity value in the stationary equilibrium used for normalization. All firms' constraints are identical to those specified in Section 4.

Use equation (34) to evaluate the expectation of the value function with respect to the outside value shocks:[25]

$$\mathbb{E}^{\epsilon}[V_F(x, X)] = \tilde{V}_F(x, X) + \zeta \bar{V}_F \cdot \ln\left(1 + e^{-\zeta^{-1} \frac{\tilde{V}_F(x,X)}{\bar{V}_F}}\right). \tag{35}$$

[24] Recent relevant contributions include Chatterjee and Eyigungor (2012); Dvorkin, Sánchez, Sapriza, and Yurdagul (2021); and Chatterjee, Corbae, Dempsey, and Ríos-Rull (2023).

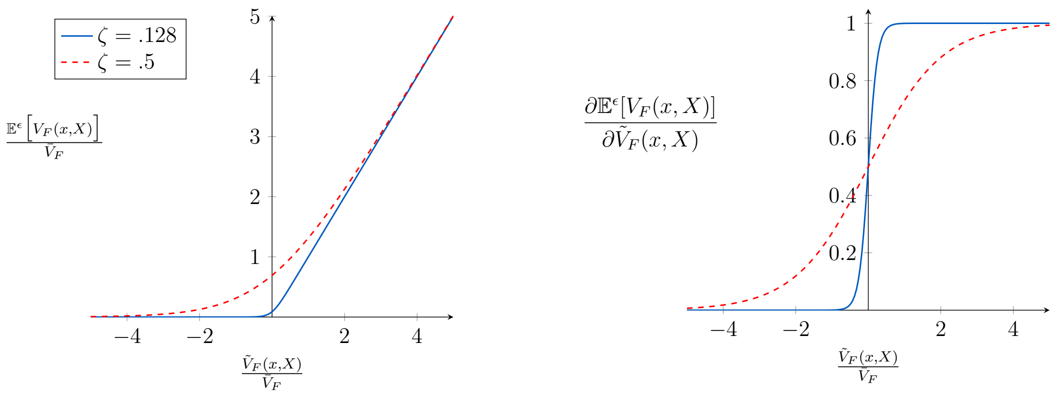
[25] See Appendix B.3 for mathematical details.

From equation (35) it follows that

$$\frac{\partial \mathbb{E}^{\epsilon}[V_F(x,X)]}{\partial \tilde{V}_F(x,X)} = \frac{1}{1+e^{-\zeta^{-1}\frac{\tilde{V}_F(x,X)}{\bar{V}_F}}} \equiv \tilde{\mathcal{I}}^d(\tilde{V}_F(x,X)), \tag{36}$$

is the logistic function which, intuitively, corresponds to the probability of a firm not defaulting ($\Pr[\mathcal{I}^d = 1|(x,X)]$) given all possible realizations of the exit value shocks.

Figure 7. Firm Expected Value with Logistic Shocks



Notes: This figure reports the firm expected value and its derivative for decreasing values of ζ.

Figure 7 illustrates the intuition behind equations (35) and (36). It shows how this smooth, differentiable formulation nests the textbook case—where shareholders' outside option is deterministically zero—and converges to it as the scale parameter ζ approaches zero. I provide a formal proof in Proposition II presented in Appendix B.3. The left panel plots the normalized expected equity value $\mathbb{E}^{\varepsilon}\left[V_F(x,X)\right]/\bar{V}_F$ from (35) against $\tilde{V}_F(x,X)/\bar{V}_F$. The right panel shows the corresponding solvency probability $\tilde{\mathcal{I}}^d$, i.e. the logistic CDF, as a function of $\tilde{V}_F(x,X)/\bar{V}_F$. In both panels the solid curve is $\zeta = .128$ (the calibrated value as described in the next subsection) and the dashed curve is $\zeta = .5$. Thanks to its "softmax" form, the mapping remains everywhere differentiable but collapses to the textbook indicator as $\zeta \to 0^+$. Equation (36) is handy to compute envelope conditions. The first-order condition with respect to k' is:

$$1-\lambda_d(d) = \mathbb{E}\left[\mathcal{M}'(X,X')\cdot\mathcal{I}'\cdot\tilde{\mathcal{I}}^d(\tilde{V}_F(x',X'))\cdot\Big(1+(1-\tau)(z'\alpha k'^{\alpha-1}-\delta)\Big)\cdot(1-\lambda_d(d'))|(x,X)\right], \tag{37}$$

where V_F' is shorthand notation for $V_F(x',X')$. The first-order condition with respect to $l_b^{\chi\prime}$ is:

$$1-\lambda_d(d) = \mathbb{E}\left[\mathcal{M}'(X,X')\cdot\mathcal{I}'\cdot\tilde{\mathcal{I}}^d(\tilde{V}_F(x',X'))\cdot(1+(1-\tau)r_l')\cdot(1-\lambda_d(d'))|(x,X)\right]. \tag{38}$$

Expressions (37) and (38) are key to derive generalized Euler equations which, in turns, are essential to characterize the banks behavior in a computationally tractable way.

7.2 Banks

The decision problem for a new entrant is identical to that in the baseline model. The decision problem for an incumbent bank now leads to modified generalized Euler equations that take into account the sensitivity of banks' decisions on firms' default decisions. Similarly to the baseline model, given other banks contracts $\{D'_{-b}, r'_D\}$ and $\{l'_{-b}(x,X), {l^s_{-b}}'(x,X), r'_l(x,X)\}$, a bank b best responds with a contract $\{D'_b, r'_D\}, \{l'_b(x,X), l^{s\prime}_b(x,X), r'_l(x,X)\}$ that satisfies the functional equations (18) and (19) subject to: (i) equation (17), (ii) the household's interest rate-quantity schedule jointly defined by equations (13) and (14), (iii) each firm's interest rate-quantity schedule which, in the extension, is jointly defined by equations (37) and (38). Each big bank's best response function satisfies the following generalized Euler equation:

$$1 = \mathbb{E}\left[\mathcal{I}' \cdot \mathcal{M}' \cdot \left(\mathcal{D}(x,X,x',X') + \tilde{\mathcal{I}}^d(x',X') \cdot \left(1 + \frac{l'_b}{R'_l}\frac{\partial R'_l}{\partial l'_b}\right)\right) \cdot R'_l | (x,X)\right]. \tag{39}$$

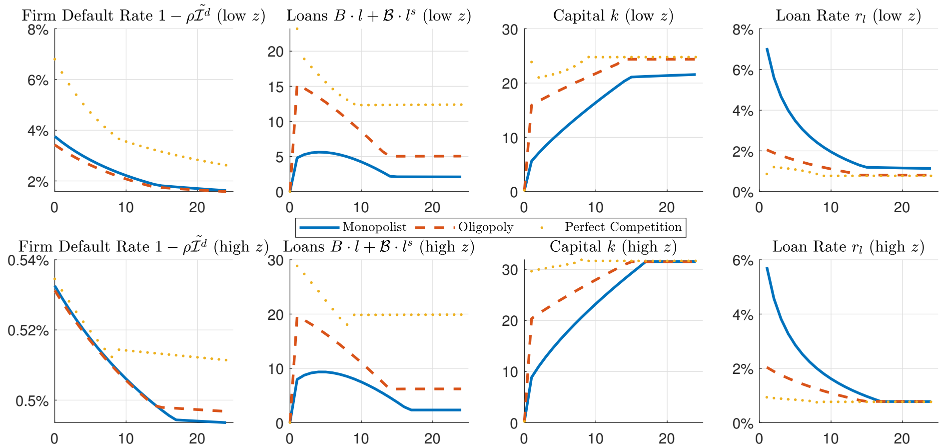
This equation further generalizes equation (23). For the sake of clarity, equation (39) is reported without the marginal recovery rate, which can be added similarly to equation (23). Also, a similar GEE arises for small banks, in the spirit of equation (24). The inverse elasticity can be computed with similar steps as in Subsection 4.4.2. When firms' default decisions are endogenous, a bank internalizes the impact that an additional unit of loan has on each firm's default decision according to

$$\forall z' : \quad \mathcal{D}(x,X,x',X') \equiv \frac{\mathrm{d}\tilde{\mathcal{I}}^d(\tilde{V}'_F)}{\mathrm{d}\tilde{V}'_F} \cdot \left[\tilde{V}_{F,l_b}(x',X') \cdot l'_b + \tilde{V}_{F,R_l}(x',X')\frac{\partial R'_l}{\partial l'_b} \cdot l'_b\right].$$

Note that when firms' default is treated exogenously, the term $\frac{\mathrm{d}\tilde{\mathcal{I}}^d(\tilde{V}'_F)}{\mathrm{d}\tilde{V}'_F}$ is zero and equation (39) collapses to equation (23). The term $\mathcal{D}(.)$ captures the strategic behavior of a bank that internalizes the marginal effects that an additional unit of loans has on the firm's future equity value and future default decisions. I set $\bar{\rho} = 1 - 0.21\%/2$ to maintain an effective level of default shocks in the stationary equilibrium. Hence, the new relevant parameter to calibrate is ζ, which I use to target an annual default rate of 0.84%.

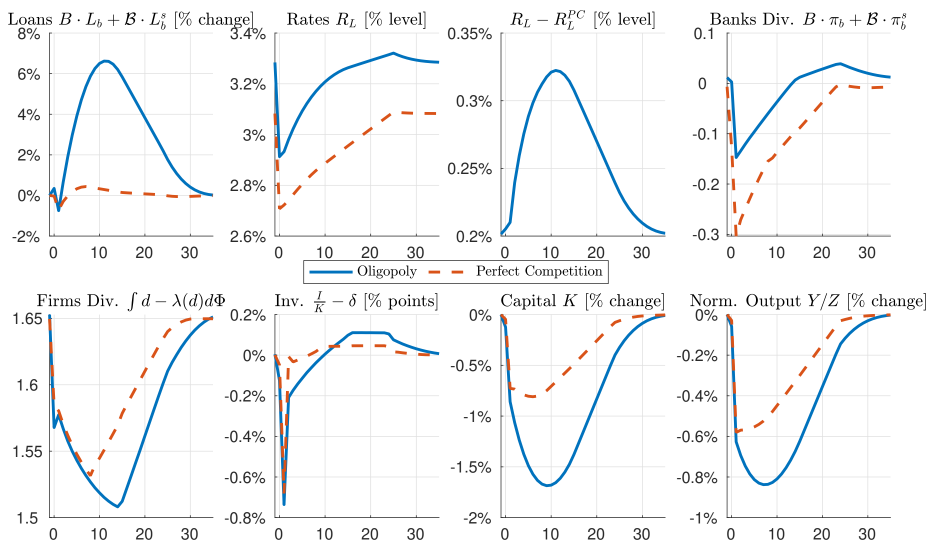
First, I calibrate the stationary equilibrium to target: (i) the profitability of the banking sector through the fixed entry cost F_E, (ii) the frequency of equity issuance through the parameter λ_0, (iii) the market leverage through the parameter k_0, (iv) the recovery rate through the parameter ν, (v) the standard deviation of investment rate through the parameter σ_z, and (vi) the default rate through the standard deviation of the outside value shocks ζ. All parameters resulting from the calibration process are reported in Table C2 in Appendix C.1, where the new parameter ζ is calibrated to be 0.128. All the corresponding aggregate stationary equilibrium moments are reported in Table C3 in Appendix C.1.

Figure 8. Stationary Equilibrium Policy Functions



Notes: This figure reports the stationary equilibrium policy functions when firms can endogenously default (the counterpart in the baseline model is Figure 2). X-axes report the firms' age.

Figure 9. TFP Shock, Financing, and Real Activity



Notes: This figure reports the transitional dynamics when firms can endogenously default (the counterpart in the baseline model is Figure 3), following the TFP shock reported in Figure C4 in Appendix C.1. X-axes report time t.

Figure 8 reports the stationary equilibrium policy functions, which are qualitatively similar to that of the baseline model. Additionally, as shown by Figure 8, the firms that are more financially constrained, and subject to banks market power, are also the ones with a higher default risk. Second, I hit the economy with the same unexpected shocks of Subsection 6.5 and compare the dynamics. Figure 9 reports the transitional dynamics when firms can endogenously default following the TFP shock reported in Figure C4 in Appendix C.1. Figure C12 in Appendix C.3 complements Figure 9 and also reports the endogenous response of the aggregate firms' default rate along the shock. The resulting dynamics are qualitative comparable to the baseline model. Quantitatively, the effects of bank market power are slightly amplified.

8 Conclusion

This paper introduces a new macroeconomic model that examines the role of banks' market power on firm dynamics, motivated by rising concentration and markups in the U.S. banking sector. Incorporating oligopolistic banks and firm heterogeneity, the model formalizes how banks' market power varies across firms with different characteristics, impacting firm growth, aggregate productivity, and output. The model shows that limited competition enables banks to charge firm-specific markups, disproportionately affecting productive financially constrained firms and dampening growth. Bank market power amplifies the effects of macroeconomic shocks, as banks extract higher markups during crises, thereby dampening capital accumulation, and intensifying declines in real activity. Following a major bank's failure, surviving banks use their increased market power to restrict credit supply, deepening the economic downturn. These findings suggest that policymakers should consider the interaction between banks' market power and credit constraints as an additional rationale for supporting large banks, beyond traditional concerns like preventing bank runs.

References

Asplund, M., and V. Nocke (2006): "Firm Turnover in Imperfectly Competitive Markets," *The Review of Economic Studies*, 73(2), 295–327.

Bernanke, B., and M. Gertler (1989): "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review*, 79(1), 14–31.

Black, S. E., and P. E. Strahan (2002): "Entrepreneurship and Bank Credit Availability," *The Journal of Finance*, 57(6), 2807–2833.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry (2018): "Really Uncertain Business Cycles," *Econometrica*, 86(3), 1031–1065.

Cetorelli, N., and M. Gambera (2001): "Banking Market Structure, Financial Dependence and Growth: International Evidence from Industry Data," *The Journal of Finance*, 56(2), 617–648.

Cetorelli, N., and P. E. Strahan (2006): "Finance as a Barrier to Entry: Bank Competition and Industry Structure in Local U.S. Markets," *The Journal of Finance*, 61(1), 437–461.

Chatterjee, S., D. Corbae, K. Dempsey, and J.-V. Ríos-Rull (2023): "A Quantitative Theory of the Credit Score," *Econometrica*, 91(5), 1803–1840.

Chatterjee, S., and B. Eyigungor (2012): "Maturity, Indebtedness, and Default Risk," *American Economic Review*, 102(6), 2674–99.

Chava, S., and A. Purnanandam (2011): "The effect of banking crisis on bank-dependent borrowers," *Journal of Financial Economics*, 99(1), 116–135.

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

Chodorow-Reich, G. (2013): " The Employment Effects of Credit Market Disruptions: Firm-level Evidence from the 2008–9 Financial Crisis *," *The Quarterly Journal of Economics*, 129(1), 1–59.

Clementi, G. L., and B. Palazzo (2016): "Entry, Exit, Firm Dynamics, and Aggregate Fluctuations," *American Economic Journal: Macroeconomics*, 8(3), 1–41.

Clymo, A., and A. Lanteri (2020): "Fiscal Policy with Limited-Time Commitment," *The Economic Journal*, 130(627), 623–652.

Cooley, T. F., and V. Quadrini (2001): "Financial Markets and Firm Dynamics," *American Economic Review*, 91(5), 1286–1310.

Cooper, R., and J. Haltiwanger (2006): "On the Nature of Capital Adjustment Costs," *The Review of Economic Studies*, 73(3), 611–633.

Corbae, D., and P. D'Erasmo (2021): "Capital Buffers in a Quantitative Model of Banking Industry Dynamics," *Econometrica*, 89(6), 2975–3023.

Covas, F., and W. J. den Haan (2011): "The Cyclical Behavior of Debt and Equity Finance," *American Economic Review*, 101(2), 877–99.

David, J. M., L. Schmid, and D. Zeke (2022): "Risk-adjusted capital allocation and misallocation," *Journal of Financial Economics*, 145(3), 684–705.

De Loecker, J., J. Eeckhout, and G. Unger (2020): "The Rise of Market Power and the Macroeconomic Implications*," *The Quarterly Journal of Economics*, 135(2), 561–644.

Demirgüç-Kunt, A., and M. S. Martinez Peria (2010): "A Framework for Analyzing Competition in the Banking Sector: An Application to the Case of Jordan," *World Bank Policy Research Working Paper No. 5499.*

Dempsey, K. (2025): "Capital Requirements with Non-Bank Finance," *The Review of Economic Studies*, Forthcoming.

Diamond, D. W., and R. G. Rajan (2005): "Liquidity Shortages and Banking Crises," *The Journal of Finance*, 60(2), 615–647.

Dvorkin, M., J. M. Sánchez, H. Sapriza, and E. Yurdagul (2021): "Sovereign Debt Restructurings," *American Economic Journal: Macroeconomics*, 13(2), 26–77.

Elenev, V., T. Landvoigt, and S. Van Nieuwerburgh (2021): "A Macroeconomic Model With Financially Constrained Producers and Intermediaries," *Econometrica*, 89(3), 1361–1418.

Freixas, X., and J. Rochet (1997): "Microeconomics of Banking, Volume 1," *The MIT Press, 1 edition.*

Gomes, J. F. (2001): "Financing Investment," *American Economic Review*, 91(5), 1263–1285.

Guerrieri, V., and G. Lorenzoni (2017): "Credit Crises, Precautionary Savings, and the Liquidity Trap*," *The Quarterly Journal of Economics*, 132(3), 1427–1467.

Hennessy, C. A., and T. M. Whited (2005): "Debt Dynamics," *The Journal of Finance*, 60(3), 1129–1165.

——— (2007): "How Costly Is External Financing? Evidence from a Structural Estimation," *The Journal of Finance*, 62(4), 1705–1745.

Holmstrom, B., and J. Tirole (1997): "Financial Intermediation, Loanable Funds, and the Real Sector," *The Quarterly Journal of Economics*, 112(3), 663–691.

Jamilov, R., and T. Monacelli (2025): "Bewley Banks," *The Review of Economic Studies*, Forthcoming.

Jermann, U., and V. Quadrini (2012): "Macroeconomic Effects of Financial Shocks," *American Economic Review*, 102(1), 238–71.

Khan, A., and J. K. Thomas (2013): "Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity," *Journal of Political Economy*, 121(6), 1055–1107.

Klein, P., P. Krusell, and J.-V. Ríos-Rull (2008): "Time-Consistent Public Policy," *The Review of Economic Studies*, 75(3), 789–808.

Klein, P., and J.-V. Ríos-Rull (2003): "Time-Consistent Optimal Fiscal Policy," *International Economic Review*, 44(4), 1217–1245.

Kreps, D. M., and J. A. Scheinkman (1983): "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes," *The Bell Journal of Economics*, 14(2), 326–337.

Lanteri, A. (2018): "The Market for Used Capital: Endogenous Irreversibility and Reallocation over the Business Cycle," *American Economic Review*, 108(9), 2383–2419.

Maskin, E., and J. Tirole (2001): "Markov Perfect Equilibrium: I. Observable Actions," *Journal of Economic Theory*, 100(2), 191–219.

McFadden, D. (1977): "Modelling the Choice of Residential Location," *Transportation Research Record*.

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

Midrigan, V., and D. Y. Xu (2014): "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, 104(2), 422–58.

Petersen, M. A., and R. G. Rajan (1995): "The Effect of Credit Market Competition on Lending Relationships," *The Quarterly Journal of Economics*, 110(2), 407–443.

Rajan, R. G., and L. Zingales (1998): "Financial Dependence and Growth," *American Economic Review*, 88(3), 559–586.

Saunders, A., A. Spina, S. Steffen, and D. Streitz (2024): "Corporate Loan Spreads and Economic Activity," *The Review of Financial Studies*, 38(2), 507–546.

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

APPENDIX (For Online Publication)

This online appendix is organized as follows. First, Appendix A is the computational appendix that contains details about the algorithms to solve for (i) the stationary equilibrium (ii) and the transitional dynamics (both for the baseline model and the extension). Second, Appendix B contains mathematical details. Third, Appendix C contains additional material.

A Computational Appendix

In this section, I describe the algorithms to solve both the stationary equilibrium and the dynamics with the MIT shocks. I highlight the novel methodology to solve for both General Equilibrium and strategic interactions. Since banks optimize over the optimal choices of the firms, solving this problem using value function iteration (VFI) would require to nesting two value function iterations inside each other and iterating on the nested value function system given guesses for the aggregate dynamics. Moreover, accounting for strategic interactions with value function iteration would require to solving this system of two nested VFIs given other banks' strategies and finding the fixed point of the resulting policies. This brute force approach is clearly not viable. To avoid this, I use projection methods jointly on the generalized Euler equations (23) and the loan firms optimality conditions (15) and (16). Hence, I leverage the fact that the elasticities can be calculated applying the implicit function theorem as described in Subsection 4.4.2. Note that several aggregate quantities are not only contained in the discount factors, but also in the elasticities of the generalized Euler equations (see Subsection 4.4.2). In order to account for strategic interactions and solve the generalized Euler equations (which can also be interpreted as best response functions), I impose ex-post symmetric strategies between banks after calculating the elasticity as described in Subsection 4.4.2; hence, I proceed to calculate the root of the resulting equation, given current state variables, as any other Euler equation.[26]

The algorithms to solve for the stationary equilibrium and the transitional dynamics in the extension of Section 7 are similar to those of the baseline model except that they require additional guesses to account for the firms' default policy functions and their derivatives, consistently with the equations presented in Section 7. This also implicitly requires to guess the distribution of the firms over age and productivity since, in the extension, it is not fixed ex-ante but depends on the equilibrium firms' default decisions.

A.1 Oligopolistic Stationary Equilibrium

Here are the main steps to solve for the *oligopolistic stationary equilibrium* (see Definition 4.1). Create grids $\mathcal{K} = [0, k_1, ..., \bar{k}]$, $\mathcal{L} = [0, l_{b,1}, ..., l_{b,\bar{L}}]$, and $\mathcal{Z} = [z_0, z_1, ..., \bar{z}]$. For each value of z creates a correspondent probability $P_z(z)$ consistently with the discretization of a

[26] This is equivalent to finding the fixed-point of the banks' strategies.

Gaussian probability distribution function with mean 1 and variance σ_z^2. Initialize the policy functions for investment, loan, and interest rate to the solution of the corresponding steady-state model without firms heterogeneity; i.e., $\forall(k, l_b, z) \in \mathcal{K} \times \mathcal{L} \times \mathcal{Z}, \quad l_b'(k, l_b, z) = l_b^*(z)$, $k'(k, l_b, z) = k^*(z)$, $R_l'(k, l_b, z) = R_l^*(z)$. Create an iterator j and set $j = 0$; hence, proceed as follows.

1. Guess the numbers of big and small banks $B^j = 1$ and $\mathcal{B}^j = 1$, an aggregate consumption $C^j = \int \tilde{d}\, \mathrm{d}\Phi + \sum_b^{B^j} \pi_b + \sum_b^{\mathcal{B}^j} \pi_b^s$ (e.g., use the steady-state consumption calculated without firm heterogeneity).[27]

2. Create an iterator w and set $w = 0$.

   (a) Start with guessed policy functions $k'^w(k, l_b, z)$, $l_b'^w(k, l_b, z)$, and $R_l'^w(k, l_b, z)$, whose derivatives are contained in the elasticity η_l' of equation (23). For every z, solve over age the firms' first-order conditions (15) and (16), and the generalized Euler equation (23), given the guessed policy functions. The elasticity η_l' of equation (23) is calculated according to equations (26), (27), (28), (29) and the condition of symmetry among bank's strategies $l_1'^{w+1} = \ldots = l_b'^{w+1} = \ldots = l_B'^{w+1}$, which is imposed ex-post. The elasticity η_l' contains not only the derivatives of the aforementioned policy functions but also the distribution values inside the derivative of aggregate consumption. Given that the distribution is fixed over age, with $\rho = \bar{\rho}$, simply use the probability density function over age $\phi(\text{age}, z)$ given by

   $$\phi(\text{age}, z) = P_z(z) \frac{\rho^{\text{age}}}{\sum_{\text{age}=0}^{\bar{N}} \rho^{\text{age}}}.$$

   (b) Start from age $= 0$ and simulate the policy functions up to age $\bar{N}$.

   (c) Project the simulated policy functions over age (representing the life cycle of a firm) onto (k, l_b, z) to determine $k'^{w+1}(k, l_b, z)$, $l_b'^{w+1}(k, l_b, z)$, and $R_l'^{w+1}(k, l_b, z)$.

   (d) If the policy functions converged (i.e., $\max(\sup|k'^{w+1} - k'^w|, \sup|l_b'^{w+1} - l_b'^w|, \sup|R_l'^{w+1} - R_l'^w|) < \epsilon$) proceed to step 3. Otherwise, set $w = w + 1$ and restart from step 2.

3. Compute the implied aggregate consumption C^{j+1} according to equation (11).

4. If the aggregate consumption converged (i.e., $|C^{j+1} - C^j| < \epsilon$) and the number of banks B^j an $\mathcal{B}^j$ is such that there is no incentive for an additional bank to enter the market as per equations (30) (these are two equations one for the big and one for the small banks), the program terminates. Otherwise, set $j = j + 1$, update $B^{j+1} = B^j + 1$ and $\mathcal{B}^{j+1} = \mathcal{B}^j + 1$ in order to satisfy equations (30) and restart from step 2. Use a

[27] Note that the aggregate dividend $\tilde{D}^j$ is contained in the elasticity η_l of equation (23).

quasi-Newton method to correct the guess of the aggregate consumption C^j, given the implied aggregate dividend C^{j+1}.[28]

A.2 Transitional Dynamics

The economy is initially in its stationary equilibrium when all agents discover a sudden change in a model parameter at $t = 0$. In order to compute the equilibrium dynamics, I need to find sequences of: (i) aggregate consumption $\{C_t\}_{t=0}^T$, and (ii) firms distributions $\{\phi_t(x)\}_{t=0}^T$; such that the representative household maximizes its utility, all markets clear in each period and the firms distributions evolve according to: (i) the firms' policy functions, (ii) the incumbent banks generalized Euler equations and (iii) the idiosyncratic default shocks. First, compute the two stationary equilibria associated with the configuration of parameters before and after the shock, as described previously.[29] Second, create an iterator j and set $j = 0$; hence, proceed as follows. Note that if there is a big (or small) bank (or banks) default on impact, i.e. $t = 0$, than the transitional dynamics could include a new big (or small) bank (or banks) entry on the equilibrium path (this depends on on the fixed entry cost). This requires to repeat this procedure several times to find the date t that satisfies the entry condition (30). In principle, this requires to guess a sequence $\{B_t\}_{t=0}^T$ and $\{\mathcal{B}_t\}_{t=0}^T$. In practice, since the idiosyncratic big bank shock described in Subsection 6.3 is calibrated to induce one and only one big bank to default, I know that only one big bank entry will occurs on the equilibrium path, given that the fixed entry cost is held at its stationary equilibrium value and there are no permanent parameters changes, eventually all shocks fade away and the economy will converge to the initial stationary equilibrium pre-default. Hence, I can solve the transitional dynamics several times till I find the new bank entry's date. Similarly, solving for the transitional dynamics with a big bank exit on the equilibrium path, it requires to solve this transitional dynamics several times till the shocks in question induce the value functions of the big banks to fall below zero on impact. Hence, re-solve with $B - 1$ big banks and check that the value functions of the big banks is now positive on impact, given the rest of the equilibrium is solved consistently with all equilibrium conditions (including a potential bank entry at some time t on the equilibrium path).

Given exogenous sequences for all shocks, create an iterator t and set $t = T - 1$ and proceed as follows.

1. Guess a sequence of aggregate consumption $\{C_t^j\}_{t=0}^T$.[30]

2. Guess an entire path of policy functions $\{k'^{t,j}(k, l_b, z),\ l_b'^{t,j}(k, l_b, z),\ R_l'^{t,j}(k, l_b, z)\}_{t=0}^{T-1}$.

[28] Or simply update the guessed consumption using the implied consumption with a dampening parameter, similarly to the update of the policy functions.

[29] If there are not permanent change to the parameters, the two stationary equilibria coincides.

[30] T should be long enough, so that after the shock the economy converges to its long-run stationary equilibrium. In this paper, I use T=40 quarters.

This guesses are needed since the derivatives of the policy functions are required to solve for the policy functions themselves. A potential guess is given by the ending stationary equilibrium policy functions.

3. Solve the policy functions backward from $t = T - 1$ to $t = 0$ given the guesses. The policy functions at $t = T$, are the ones associated with the ending stationary equilibrium, previously calculated. At each time t proceed similarly to before.

   (a) Given all guesses and the ending stationary equilibrium policy functions solve over time and age (and for every z) the firms' first-order conditions (15) and (16), and the generalized Euler equation (23). As before, the elasticity η_l' of equation (23) is calculated according to equations (26), (27), (28), (29) and the condition of symmetry among bank's strategies $l_1'^t = ... = l_b'^t = ... = l_B'^t$, which is imposed ex-post. At each point t, the elasticity η_l' contains not only the derivatives of the aforementioned policy functions but also the distribution values inside the derivative of aggregate consumption. Given that the distribution is fixed over age and time (given a level of productivity), use the probability density function over age, which is now time-varying

$$\phi_t(\text{age}, z) = P_z(z) \frac{\rho_t^{\text{age}}}{\sum_{\text{age}=0}^{\bar{N}} \rho_t^{\text{age}}}.$$

4. Now, start from $t = 0$ and iterate forward up to $t = T$. At each time t, start from age $= 0$ and simulate the time t policy functions up to age $\bar{N}$. This yields a mapping between age and (k, l_b, z).

5. For each time t, compute the implied aggregate consumption C_t^{j+1} according to equation (11) and the implied path of policy functions $\{k'^{t,j+1}(k, l_b, z), l_b'^{t,j+1}(k, l_b, z), R_l'^{t,j+1}(k, l_b, z)\}_{t=0}^{T-1}$.

6. If the sequences for aggregate consumption converged; i.e.,

$$\sup\{|C_t^{j+1} - C_t^j|\}_{t=0}^T < \epsilon,$$

   and if the sequences of policy functions converged; i.e.,

$$\max(\sup\{|k_t'^{j+1} - k'^j|\}_{t=0}^T, \sup\{|l_{b,t}'^{j+1} - l_{b,t}'^j|\}_{t=0}^T, \sup\{|R_{l,t}'^{j+1} - R_{l,t}'^j|\}_{t=0}^T) < \epsilon,$$

   the program terminates. Otherwise, set $j = j+1$, update the guessed with a dampening parameter, and restart from step 3.

A.3 Oligopolistic Stationary Equilibrium with Banks' Market Power Interacting with the Firms' Endogenous Default Decisions

Here are the main steps to solve for the *oligopolistic stationary equilibrium* when firms make endogenous default decisions (see Section 7). Create grids $\mathcal{K} = [0, k_1, ..., \bar{k}]$, $\mathcal{L} = [0, l_{b,1}, ..., l_{b,\bar{L}}]$, and $\mathcal{Z} = [z_0, z_1, ..., \bar{z}]$. For each value of z creates a correspondent probability $P_z(z)$ consistently with the discretization of a Gaussian probability distribution function with mean 1 and variance σ_z^2. Initialize the policy functions for investment, loan, and interest rate to the solution of the corresponding steady-state model without firms heterogeneity; i.e., $\forall(k, l_b, z) \in \mathcal{K} \times \mathcal{L} \times \mathcal{Z}$, $\quad l_b'(k, l_b, z) = l_b^*(z)$, $k'(k, l_b, z) = k^*(z)$, $R_l'(k, l_b, z) = R_l^*(z)$. Create an iterator j and set $j = 0$; hence, proceed as follows.

1. Guess the numbers of big and small banks $B^j = 1$ and $\mathcal{B}^j = 1$, an aggregate consumption $C^j = \int \tilde{d}\, \mathrm{d}\Phi + \sum_b^{B^j} \pi_b + \sum_b^{\mathcal{B}^j} \pi_b^s$ (e.g., use the steady-state consumption calculated without firm heterogeneity).

2. Create an iterator w and set $w = 0$.

   (a) Start with guessed policy functions $k'^w(k, l_b, z)$, $l_b'^w(k, l_b, z)$, and $R_l'^w(k, l_b, z)$, whose derivatives are contained in the elasticity η_l' of equation (23). Also guess the value function of each firm at every age; i.e., $\tilde{V}^{F,w}(age, z)$.[31] A guess for the value function is needed to find the distribution over age, which depends on default and its values are contains in the Generalized Euler equation elasticity. For every z, solve over age the firms' first-order conditions (37) and (38), and the generalized Euler equation (39), given the guessed policy and value functions. The elasticity η_l' of equation (39) is calculated using the implicit function theorem similarly to equations (26), (27), (28), (29) and the condition of symmetry among bank's strategies $l_1'^{w+1} = ... = l_b'^{w+1} = ... = l_B'^{w+1}$, which is imposed ex-post. Similarly to the case without firms' endogenous default, the elasticity η_l' contains not only the derivatives of the aforementioned policy functions but also the distribution values inside the derivative of aggregate consumption. Differently from the case without firms' endogenous default, the distribution over age is not just fixed by parameters values but also depends on the endogenous default decisions. Hence, start from $age = 0$ and simulate using the guessed value function to determine the default decisions up to an age $\bar{N}$ in order to find $\phi(\text{age}, z)$. These values are needed in the calculations of η_l' of equation (39).

   (b) Start from age $= 0$ and simulate the policy functions up to age $\bar{N}$. Calculate the implied value function $\tilde{V}^{F,w+1}(age, z)$.

[31] A good guess is the corresponding steady-state value function without firms heterogeneity.

(c) Project the simulated policy functions over age (representing the life cycle of a firm) onto (k, l_b, z) to determine $k'^{w+1}(k, l_b, z)$, $l_b'^{w+1}(k, l_b, z)$, and $R_l'^{w+1}(k, l_b, z)$. Recall that these projected policy functions are needed to compute the derivative of the policy function inside the inverse elasticity of the Generalized Euler equation. Note that there is no need to project the $\tilde{V}^{F,w+1}(age, z)$ into (k, l_b, z) since this is not needed, the derivatives of the default decisions do shop inside the inverse elasticity of the Generalized Euler equation but can be partially computed analytically using the envelope theorems and expressed as functions of the derivatives of $k'^{w+1}(k, l_b, z)$, $l_b'^{w+1}(k, l_b, z)$, and $R_l'^{w+1}(k, l_b, z)$.

(d) If the policy functions converged (i.e., $\max(\sup|k'^{w+1} - k'^{w}|, \sup|l_b'^{w+1} - l_b'^{w}|, \sup|R_l'^{w+1} - R_l'^{w}|) < \epsilon$) and the value function converged (i.e., $\max(\sup|\tilde{V}^{F,w+1} - \tilde{V}^{F,w}|) < \epsilon$) proceed to step 3. Otherwise, set $w = w + 1$ and restart from step 2.

3. Compute the implied aggregate consumption C^{j+1} according to equation (11).

4. If the aggregate consumption converged (i.e., $|C^{j+1} - C^{j}| < \epsilon$) and the number of banks B^j an $\mathcal{B}^j$ is such that there is no incentive for an additional bank to enter the market as per equations (30) (these are two equations one for the big and one for the small banks), the program terminates. Otherwise, set $j = j + 1$, update $B^{j+1} = B^j + 1$ and $\mathcal{B}^{j+1} = \mathcal{B}^j + 1$ in order to satisfy equations (30) and restart from step 2. Use a quasi-Newton method to correct the guess of the aggregate consumption C^j, given the implied aggregate dividend C^{j+1}.[32]

A.4 Transitional Dynamics with Banks' Market Power Interacting with the Firms' Endogenous Default Decisions

The economy is initially in its stationary equilibrium when all agents discover a sudden change in a model parameter at $t = 0$. In order to compute the equilibrium dynamics, I need to find sequences of: (i) aggregate consumption $\{C_t\}_{t=0}^{T}$, and (ii) firms distributions $\{\phi_t(x)\}_{t=0}^{T}$; such that the representative household maximizes its utility, all markets clear in each period and the firms distributions evolve according to: (i) the firms' policy functions, (ii) the incumbent banks generalized Euler equations and (iii) the idiosyncratic default shocks. First, compute the two stationary equilibria associated with the configuration of parameters before and after the shock, as described previously.[33] Second, create an iterator j and set $j = 0$; hence, proceed as follows. In the extension, all shocks are calibrated so that there is no bank default on the equilibrium path (hence, not even a bank entry). Similarly to the algorithm without firms' endogenous default, it is possible to repeat this procedure several

[32]Or simply update the guessed consumption using the implied consumption with a dampening parameter, similarly to the update of the policy functions.

[33]If there are not permanent change to the parameters, the two stationary equilibria coincides.

times in order to have a bank default and/or a bank entry on the equilibrium path. This requires to check that all entry and exit conditions are satisfied at each time t.

Given exogenous sequences for all shocks, create an iterator t and set $t = T - 1$ and proceed as follows.

1. Guess a sequence of aggregate consumption $\{C_t^j\}_{t=0}^T$.[34]

2. Guess an entire path of policy functions $\{k'^{t,j}(k, l_b, z),\ l_b'^{t,j}(k, l_b, z),\ R_l'^{t,j}(k, l_b, z)\}_{t=0}^{T-1}$. Also guess an entire path of value functions over firm age; i.e., $\{\tilde{V}_t^{F,j}(age, z)\}_{t=0}^{T-1}$. All these guesses are needed since the derivatives of the policy functions are required to solve for the policy functions themselves. The guess for the value function is required to find the distribution, which is contained inside the inverse elasticity of Generalized Euler equation. A potential guess is given by the ending stationary equilibrium policy and value functions.

3. Solve the policy functions backward from $t = T - 1$ to $t = 0$ given the guesses. The policy functions at $t = T$, are the ones associated with the ending stationary equilibrium, previously calculated. At each time t proceed similarly to before.

   (a) Given all guesses and the ending stationary equilibrium policy functions solve over time and age (and for every z) the firms' first-order conditions (37) and (38), and the generalized Euler equation (39), given the guessed policy and value functions. The elasticity η_l' of equation (39) is calculated using the implicit function theorem similarly to equations (26), (27), (28), (29) and the condition of symmetry among bank's strategies $l_1'^{w+1} = ... = l_b'^{w+1} = ... = l_B'^{w+1}$, which is imposed ex-post. At each point t, the elasticity η_l' contains not only the derivatives of the aforementioned policy functions but also the distribution values inside the derivative of aggregate consumption. Given that the distribution depends on the firms' default decisions, at each t, start from $age = 0$ and simulate using the guessed value functions to determine the default decisions up to an age $\bar{N}$ in order to find $\phi_t(\text{age}, z)$.

4. Now, start from $t = 0$ and iterate forward up to $t = T$. At each time t, start from age $= 0$ and simulate the time t policy functions up to age $\bar{N}$. This yields a mapping between age and (k, l_b, z).

5. For each time t, compute the implied aggregate consumption C_t^{j+1} according to equation (11) and the implied path of policy functions $\{k'^{t,j+1}(k, l_b, z),\ l_b'^{t,j+1}(k, l_b, z),\ R_l'^{t,j+1}(k, l_b, z)\}_{t=0}^{T-1}$. Also compute an implied path for the value functions $\{\tilde{V}_t^{F,w+1}(age, z)\}_{t=0}^{T-1}$.

[34] T should be long enough, so that after the shock the economy converges to its long-run stationary equilibrium. In this paper, I use T=40 quarters.

6. If the sequences for aggregate consumption converged; i.e.,

$$\sup\{|C_t^{j+1} - C_t^j|\}_{t=0}^T < \epsilon,$$

and if the sequences of policy functions converged; i.e.,

$$\max(\sup\{|k_t'^{j+1} - k'^j|\}_{t=0}^T, \sup\{|l_{b,t}'^{j+1} - l_{b,t}'^j|\}_{t=0}^T, \sup\{|R_{l,t}'^{j+1} - R_{l,t}'^j|\}_{t=0}^T) < \epsilon,$$

and if the sequence of value functions converged; i.e.,

$$\max(\sup\{|\tilde{V}_t^{F,j+1} - \tilde{V}_t^{F,j}|\}_{t=0}^T) < \epsilon,$$

the program terminates. Otherwise, set $j = j+1$, update the guessed with a dampening parameter, and restart from step 3.

B Mathematical Appendix

This appendix contains the mathematical details of the proposition in the simple model, as well as those pertaining to the extension of the baseline model.

B.1 Stylized Model: Proposition

This section contains the proposition related to statements 1–10 discussed in Section 3.

Proposition I

Assume that the distribution $\phi(x_0)$ is such that there is a non-zero measure of financially constrained firms:[a]

$$\mathcal{P} = \int \mathbf{1}[d_0(x_0, k_1^*, l_{b,1}^*) \geq 0] \,\mathrm{d}\Phi(x_0) < 1.$$

A higher number of banks (i.e., a higher B) has the following effects:

1. aggregate loans per bank $\int l_{b,1}^* \,\mathrm{d}\Phi$ decreases;
2. average loan interest rate $\int R_{l,1}^* \,\mathrm{d}\Phi$ decreases;
3. aggregate physical investment $\int k_1^* - (1-\delta)k_0 \,\mathrm{d}\Phi$ increases;
4. aggregate expected returns $\int \mathbb{E}_0 [d_1^*] / p_0^* \,\mathrm{d}\Phi$ decreases;
5. aggregate loans $\int \sum_b^B l_{b,1}^* \,\mathrm{d}\Phi$ increases;
6. aggregate leverage $\int \sum_b^B l_{b,1}^* / k_1^* \,\mathrm{d}\Phi$ increases;
7. aggregate TFP $\int k_1^{*\alpha} \,\mathrm{d}\Phi / \left(\int k_1^* \,\mathrm{d}\Phi\right)^\alpha$ increases;
8. variance of capital $\int k_1^{*2} \,\mathrm{d}\Phi - (\int k_1^* \,\mathrm{d}\Phi)^2$ decreases;
9. variance of loan interest rates $\int R_{l,1}^{*2} \,\mathrm{d}\Phi - (\int R_{l,1}^* \,\mathrm{d}\Phi)^2$ decreases;
10. variance of expected returns $\int (\mathbb{E}_0 [d_1^*] / p_0^*)^2 \,\mathrm{d}\Phi - \left(\int \mathbb{E}_0 [d_1^*] / p_0^* \,\mathrm{d}\Phi\right)^2$ decreases.

[a] For sub-points 7, 8, 9, and 10, I assume that the mass of financially constrained firms $1 - \mathcal{P}$ are all ex-ante identical.

For ease of notation, in the proofs, I denote $R_{l,1}$ simply with R_1.

B.1.1 Proofs

Proof. Statements 1,2 and 3. Note that the number of banks matters only for the financially constrained firms, so that each integral can be rewritten as follow

$$\frac{\partial}{\partial B} \int x^* \,\mathrm{d}\Phi = \int \frac{\partial x^*}{\partial B} \cdot \mathbf{1}[d_1^* < 0] \,\mathrm{d}\Phi + \int \frac{\partial x^*}{\partial B} \cdot \mathbf{1}[d_1^* \geq 0] \,\mathrm{d}\Phi = \int \frac{\partial x^*}{\partial B} \cdot \mathbf{1}[d_1^* < 0] \,\mathrm{d}\Phi,$$

where x^* is a place holder for $l_{b,1}^*$, R_1^*, p_0^* and k_1^*. Hence, a sufficient condition to establish the sign of $\frac{\partial}{\partial B} \int x^* \,\mathrm{d}\Phi$ is to determine the sign of $\frac{\partial x^*}{\partial B} \cdot \mathbf{1}[d_1^* < 0]$. Total differentiation of the optimality conditions (5), (7) and (6) of the financially constrained firms yields the following linear system

$$\begin{bmatrix} \kappa_1 & \kappa_2 & \frac{\partial R_1^*}{\partial l_{b,1}} \cdot \kappa_2 \\ -\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2} & 1 & 0 \\ -1 & \rho\beta & B \end{bmatrix} \begin{bmatrix} \frac{\partial k_1^*}{\partial B} \\ \frac{\partial R_1^*}{\partial B} \\ \frac{\partial l_{b,1}^*}{\partial B} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -l_{b,1}^* \end{bmatrix}.$$

Note first that equation (7), for firms with $d_0 < 0$, implies $R_1^* = \frac{1-\lambda_0 d_0}{\rho\beta} > \frac{1}{\rho\beta}$ for $\lambda_0 > 0$. Hence, the GEE implies $\frac{\partial R_1^*}{\partial l_{b,1}} < 0$ for financially constrained firms. Hence, by concavity of the production function and since $0 < \alpha < 1$, $\kappa_1 = \frac{\partial R_1^*}{\partial l_{b,1}} \frac{\lambda_0(\alpha-2)}{\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-1}} < 0$. It also follows that $\kappa_2 = \frac{1}{l_{b,1}^*}\left(\frac{\lambda_0}{\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}} - \rho\beta\right) < 0$. The determinant of the matrix is therefore:

$$\mathcal{D} = \frac{\partial R_1^*}{\partial l_{b,1}}\kappa_2 + \kappa_1 B + \kappa_2\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}B - \frac{\partial R_1^*}{\partial l_{b,1}}\kappa_2\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}\rho\beta.$$

Direct inversion yields:

$$\begin{bmatrix} \frac{\partial k_1^*}{\partial B} \\ \frac{\partial R_1^*}{\partial B} \\ \frac{\partial l_{b,1}^*}{\partial B} \end{bmatrix} = \begin{bmatrix} \frac{\partial R_1^*}{\partial l_{b,1}}\kappa_2 l_{b,1}^* \\ \frac{\partial R_1^*}{\partial l_{b,1}}\kappa_2\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}l_{b,1}^* \\ -l_{b,1}^*(\kappa_1 + \kappa_2\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}) \end{bmatrix} \cdot \mathcal{D}^{-1}.$$

Note that if $\kappa_1 + \kappa_2\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2} > 0$, then we can conclude that: $\frac{\partial k_1^*}{\partial B} > 0$, $\frac{\partial R_1^*}{\partial B} < 0$ and $\frac{\partial l_{b,1}^*}{\partial B} < 0$. This is equivalent to showing:

$$\frac{1-\rho\beta R_1^*}{\rho\beta l_{b,1}^*}\frac{\lambda_0(\alpha-2)}{\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-1}} + \frac{\lambda_0}{l_{b,1}^*} - \frac{1}{l_{b,1}^*}\rho\beta\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2} > 0.$$

Note that equation (5) of the optimalities of the constrained firm can be rewritten as

$$\lambda_0\frac{1-\rho\beta R_1^*}{\rho\beta l_{b,1}^*}\frac{1}{\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-1}k_1^{*-1}} = \rho\beta\frac{1-\rho\beta R_1^*}{\rho\beta l_{b,1}^*} + \lambda_0.$$

Using this equivalence *the want to show* can be rewritten as

$$\begin{aligned} &\frac{1-\rho\beta R_1^*}{\rho\beta l_{b,1}^*}\frac{\lambda_0(\alpha-2)}{\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-1}} + \frac{\lambda_0}{l_{b,1}^*} - \frac{1}{l_{b,1}^*}\rho\beta\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2} \\ &= \rho\beta\frac{1-\rho\beta R_1^*}{\rho\beta l_{b,1}^*}(\alpha-2)k_1^{*-1} + \lambda_0(\alpha-2)k_1^{*-1} + \frac{\lambda_0}{l_{b,1}^*} - \frac{1}{l_{b,1}^*}\rho\beta\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2} > 0. \end{aligned}$$

Multiply everything by $l_{b,1}^* > 0$, to get:

$$(1-\rho\beta R_1^*)(\alpha-2)k_1^{*-1} + \lambda_0(\alpha-2)\frac{l_1^*}{k_1^*} + \lambda_0 - \rho\beta\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2} > 0.$$

Hence, use equation (6) of the optimalities of the constrained firms to back out an expression for $l_{b,1}^*$ in function of R_1^* and k_1^*, and rewrite

$$(1-\rho\beta R_1^*)(\alpha-2)k_1^{*-1} + (\alpha-2)\frac{1-\rho\beta R_1^* - \lambda_0(z_0k_0^\alpha + (1-\delta)k_0 - k_1^*)}{Bk_1^*} + \lambda_0 - \rho\beta\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2} > 0.$$

The left-hand side can be rearranged as

$$\begin{aligned}&(\alpha-2)\frac{(1-\rho\beta R_1^*)(B+1) - \lambda_0(z_0k_0^\alpha + (1-\delta)k_0 - k_1^*)}{Bk_1^*} + \lambda_0 - \rho\beta\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}\\ &= (\alpha-2)\left[(1-\rho\beta R_1^*)(B+1) - \lambda_0(z_0k_0^\alpha + (1-\delta)k_0 - k_1^*)\right] + \lambda_0Bk_1^* - \rho\beta\alpha(\alpha-1)\mathbb{E}_0[z_1]Bk_1^{*\alpha-1}.\end{aligned}$$

Divide by $(\alpha-2) < 0$ (changing sign because it is always negative), the previous *want to show* is equivalent to show

$$\underbrace{(1-\rho\beta R_1^*)(B+1)}_{<0 \text{ when } d_0<0} - \underbrace{\lambda_0(z_0k_0^\alpha + (1-\delta)k_0)}_{>0} + \lambda_0k_1^* \underbrace{\frac{\alpha-2+B}{\alpha-2}}_{<0 \text{ if } B>1} - \underbrace{\rho\beta\alpha\frac{\alpha-1}{\alpha-2}\mathbb{E}_0[z_1]Bk_1^{*\alpha-1}}_{>0} < 0.$$

Consider two cases. If $B > 1$ (oligopoly), this last inequality is always satisfied. For $B = 1$ (monopoly), the inequality collapses to

$$\underbrace{(1-\rho\beta R_1^*)2}_{<0 \text{ when } d_0<0} - \underbrace{\lambda_0(z_0k_0^\alpha + (1-\delta)k_0)}_{>0} + \underbrace{(\lambda_0k_1^* - \rho\beta R_1^*)}_{<0}\underbrace{\frac{\alpha-1}{\alpha-2}}_{>0} < 0.$$

Finally, rearrange the Euler $\rho\beta R_1^* = 1 - \lambda_0 d_0$ to get

$$\lambda_0k_1^* - \rho\beta R_1^* = \lambda_0(z_0k_0^\alpha + (1-\delta)k_0 + l_{b,1}^*) - 1,$$

which yields the result

$$\underbrace{(1-\rho\beta R_1^*)2}_{<0 \text{ when } d_0<0} + \left(\frac{\alpha-1}{\alpha-2} - 1\right)\underbrace{\lambda_0(z_0k_0^\alpha + (1-\delta)k_0)}_{>0} + \underbrace{(\lambda_0 l_{b,1}^* - 1)}_{<0}\underbrace{\frac{\alpha-1}{\alpha-2}}_{>0} < 0.$$

□

Proof. Statements 4, 5 and 6. Following a similar logic as the one of equation (9) the proof focuses in studying the signs of the optimal choices of the financially constrained firms.

Hence, all equations that follow refer to those firms such that $d_0(k_0, z_0, k_1^*, l_{b,1}^*) < 0$.

For the expected return of the shares, equating the Euler equations for loans and for the price of shares provides the following non-arbitrage condition $\mathbb{E}_0\left[\frac{d_1^*}{p_0^*}\right] = R_1^*$. Hence, for financially constrained firms $\frac{\partial}{\partial B}\mathbb{E}_0\left[\frac{d_1^*}{p_0^*}\right] = \frac{\partial R_1^*}{\partial B} < 0$, which is always negative by previous result.

Before studying the effect of the number of banks on leverage, first note that the effect on total debt is ambiguous:

$$\frac{\partial}{\partial B} B \cdot l_{b,1}^* = B\frac{\partial l_{b,1}^*}{\partial B} + l_{b,1}^*.$$

As shown previously, as the number of banks increases $l_{b,1}$ decreases. Plugging the formula for $\frac{\partial l_{b,1}^*}{\partial B}$ found previously can resolve this ambiguity:

$$\frac{\partial}{\partial B} B \cdot l_{b,1}^* = l_{b,1}^*\left(1 - B\frac{\kappa_1 + \kappa_2\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}}{\mathcal{D}}\right) = l_{b,1}^*\left(1 - \frac{1}{1 + \frac{\frac{\partial R_1^*}{\partial l_{b,1}}\kappa_2\left(1-\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}\rho\beta\right)}{B\kappa_1 + B\kappa_2\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}}}\right),$$

since $l_{b,1}^* > 0$ and $\frac{\frac{\partial R_1^*}{\partial l_{b,1}}\kappa_2\left(1-\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}\rho\beta\right)}{B\kappa_1 + B\kappa_2\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}} > 0 \implies \frac{\partial}{\partial B} B \cdot l_{b,1}^* > 0$.

In order to prove that the leverage increases with the number of banks, it remains to show that the following inequality is always satisfied for the financially constrained firms:

$$\begin{aligned}
\frac{\partial}{\partial B}\frac{B \cdot l_{b,1}^*}{k_1^*} &= \left(B\frac{\partial l_{b,1}^*}{\partial B} + l_{b,1}^*\right)\frac{1}{k_1^*} - \frac{B \cdot l_{b,1}}{k_1^{*2}} \cdot \frac{\partial k_1^*}{\partial B} \\
&= \frac{l_{b,1}^*}{k_1^*}\left(1 - \frac{B\kappa_1 + B\kappa_2\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2} - B\frac{l_{b,1}^*}{k_1^*}\kappa_2\frac{\partial R_1^*}{\partial l_{b,1}}}{\mathcal{D}}\right) > 0.
\end{aligned}$$

Since $l_{b,1}^*/k_1^* > 0$, $\mathcal{D} > 0$ and $\kappa_2\frac{\partial R_1^*}{\partial l_{b,1}} > 0$, this is equivalent to show:

$$-B\frac{l_{b,1}^*}{k_1^*}\kappa_2\frac{\partial R_1^*}{\partial l_{b,1}} < \frac{\partial R_1^*}{\partial l_{b,1}}\kappa_2 - \frac{\partial R_1^*}{\partial l_{b,1}}\kappa_2\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}\rho\beta \iff -B\frac{l_{b,1}^*}{k_1^*} < 1 - \alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}\rho\beta,$$

which is always true since $-B\frac{l_{b,1}^*}{k_1^*}$ is always negative and $1-\alpha(\alpha-1)\mathbb{E}_0[z_1]k_1^{*\alpha-2}\rho\beta$ is always positive.

□

Proof. Statements 7, 8, 9 and 10. For statements 7, 8, 9 and 10, I assume that the mass of financially constrained firms $1-\mathcal{P}$ are all ex-ante identical. For TFP, the *want to show* is

$$\frac{\partial}{\partial B}\frac{\mathbb{E}\left[k_1^{*\alpha}\right]}{\left(\mathbb{E}\left[k_1^*\right]\right)^\alpha} = \frac{\alpha\mathbb{E}\left[k_1^{*\alpha-1}\frac{\partial k_1^*}{\partial B}\right]}{\left(\mathbb{E}\left[k_1^*\right]\right)^\alpha} - \alpha\frac{\mathbb{E}[k_1^{*\alpha}]\mathbb{E}\left[\frac{\partial k_1^*}{\partial B}\right]}{\left(\mathbb{E}\left[k_1^*\right]\right)^{\alpha+1}} > 0.$$

This is equivalent to show that

$$\mathbb{E}\left[k_1^{*\alpha-1}\frac{\partial k_1^*}{\partial B}\right]\mathbb{E}\left[k_1^*\right] - \mathbb{E}[k_1^{*\alpha}]\mathbb{E}\left[\frac{\partial k_1^*}{\partial B}\right] > 0.$$

Which is again equivalent to

$$\begin{aligned}
&k_1^{*\alpha-1}\frac{\partial k_1^*}{\partial B}(1-\mathcal{P})(k_1^*(1-\mathcal{P})+\bar{k}\mathcal{P}) - (k_1^{*\alpha}(1-\mathcal{P})+\bar{k}^\alpha\mathcal{P})\frac{\partial k_1^*}{\partial B}(1-\mathcal{P}) > 0\\
&\iff k_1^{*\alpha-1}(k_1^*(1-\mathcal{P})+\bar{k}\mathcal{P}) - (k_1^{*\alpha}(1-\mathcal{P})+\bar{k}^\alpha\mathcal{P}) > 0\\
&\iff k_1^{*\alpha}(1-\mathcal{P}) + k_1^{*\alpha-1}\bar{k}\mathcal{P} - k_1^{*\alpha}(1-\mathcal{P}-\bar{k}^\alpha\mathcal{P} > 0\\
&\iff k_1^{*\alpha-1}\bar{k}\mathcal{P} - \bar{k}^\alpha\mathcal{P} > 0\\
&\iff k_1^{*\alpha-1} > \bar{k}^{\alpha-1}.
\end{aligned}$$

Since $k_1^* < \bar{k}$, the last inequality is always verified.

For the dispersion of capital, the *want to show* is

$$\frac{\partial}{\partial B}\mathbb{E}\left[(k_1^*-\mathbb{E}\left[k_1^*\right])^2\right] = \mathbb{E}\left[\frac{\partial}{\partial B}(k_1^*-\mathbb{E}\left[k_1^*\right])^2|d_0<0\right](1-\mathcal{P}) + \mathbb{E}\left[\frac{\partial}{\partial B}(\bar{k}_1-\mathbb{E}\left[k_1^*\right])^2|d_0\geq 0\right]\mathcal{P} < 0,$$

where $\mathcal{P}$ is the mass of the firms not financially constrained and $\bar{\kappa}$ is the optimal choice of capital of the non financially constrained firms. Hence:

$$\begin{aligned}
\frac{\partial}{\partial B}\mathbb{E}\left[(k_1^*-\mathbb{E}\left[k_1^*\right])^2|d_0<0\right] &= 2(k_1^*-k_1^*(1-\mathcal{P})-\bar{k}\mathcal{P})\left(\frac{\partial k_1^*}{\partial B} - \frac{\partial k_1^*}{\partial B}(1-\mathcal{P}) - \underbrace{\frac{\partial \bar{k}}{\partial B}}_{=0}\mathcal{P}\right)\\
&= 2\mathcal{P}(k_1^*-\bar{k})\frac{\partial k_1^*}{\partial B}\mathcal{P} < 0.
\end{aligned}$$

Note that the last inequality follows from the fact that $k_1^* < \bar{k}$, otherwise the mass of firms $1-\mathcal{P}$ would not be financially constrained. $\frac{\partial k_1^*}{\partial B} > 0$ is positive from the previous proof. Note that the second term is always negative

$$\mathbb{E}\left[\frac{\partial}{\partial B}(\bar{k}_1-\mathbb{E}\left[k_1^*\right])^2|d_0\geq 0\right] = 2\mathbb{E}\left[\underbrace{(\bar{k}_1-\mathbb{E}\left[k_1^*\right])}_{>0}\left(\underbrace{\frac{\partial \bar{k}_1}{\partial B}}_{=0} - \underbrace{\frac{\partial}{\partial B}\mathbb{E}\left[k_1^*\right]}_{>0}\right)|d_0\geq 0\right] < 0.$$

Furthermore, note that $R_1^* = \mathbb{E}_0[1 + \alpha z_1 k_1^{*\alpha-1} - \delta]$ and

$$\sigma(R_1^*) = \sigma^2\left(1 + \alpha\mathbb{E}_0[z_1]k_1^{*\alpha-1} - \delta\right) = \alpha^2\mathbb{E}_0^2[z_1]\sigma\left(k_1^{*\alpha-1}\right).$$

Hence:

$$\frac{\partial\sigma^2(R_1^*)}{\partial B} = \alpha^2\mathbb{E}_0^2[z_1]\frac{\partial\sigma^2\left(k_1^{*\alpha-1}\right)}{\partial B}$$
$$= \alpha^2\mathbb{E}_0^2[z_1]\left(2\mathcal{P}\underbrace{(k_1^{*\alpha-1} - \bar{k}^{\alpha-1})}_{>0}\underbrace{(\alpha-1)}_{<0}k_1^{*\alpha-2}\frac{\partial k_1^*}{\partial B}\mathcal{P} + 2\mathbb{E}\left[\underbrace{(\bar{k}_1^{\alpha-1} - \mathbb{E}\left[k_1^{*\alpha-1}\right])}_{<0}\left(\underbrace{\frac{\partial\bar{k}_1^{\alpha-1}}{\partial B}}_{=0} - \underbrace{\frac{\partial\mathbb{E}\left[k_1^{*\alpha-1}\right]}{\partial B}}_{<0}\right)|d_0 \geq 0\right]\right) < 0.$$

Equating the two Euler equations for the price of the shares of the firms and the price of the bonds yields: $\frac{\partial}{\partial B}\mathbb{E}\left[\frac{d_1^*}{p_0^*}\right] = \frac{\partial R_1^*}{\partial B} < 0$.

□

B.2 Firm Loan Level under MPE

This subsection points to a feature of the MPE for which an optimal level of loan remains well defined even in absence of equity issuance cost in presence of a tax shield and curvature in the household's utility function. If there are no states of the world in which firms issue equity, or expect to issue equity, then the firms' first-order conditions reduce to:[35]

$$1 = \rho\cdot\mathcal{M}'\cdot\left(1 + (1-\tau)\left(\alpha {k'}^{\alpha-1} - \delta\right)\right) \text{ and } 1 = \rho\cdot\mathcal{M}'\cdot(1 + (1-\tau)r_l') \text{ with } \mathcal{M}' = \beta\left(\frac{C'}{C}\right)^{-\gamma},$$

and, in general, these conditions are not sufficient to pin down each firm's loan level in the stationary equilibrium. Loan is determined in equilibrium via the banking behavior captured by the generalized Euler equation which, in this simplified case, reduces to:

$$1 = \rho\cdot\mathcal{M}'\cdot(1 + r_l')\cdot(1 + \eta_l') \quad \text{with} \quad \eta_l' = \frac{\partial R_l'}{\partial l_b'}\frac{l_b'}{R_l'} = -\frac{f_{k'}\cdot g_{l_b'}}{g_{k'} + f_{k'}\cdot g_{R_l'}}\frac{l_b'}{R_l'},$$

where, under these simplified conditions, the partial derivatives of Subsection 4.4.2 are:

$$f_{k'} = \alpha(\alpha-1){k'}^{\alpha-2}, \qquad g_{k'} = \rho\frac{\partial\mathcal{M}'}{\partial k'}\left((1-\tau){r_l}' + 1\right),$$
$$g_{l_b'} = \rho\frac{\partial\mathcal{M}'}{\partial l_b'}\left((1-\tau){r_l}' + 1\right), \qquad g_{R_l'} = \rho\frac{\partial\mathcal{M}'}{\partial R_l'}\left((1-\tau){r_l}' + 1\right).$$

[35]Assume $\lambda_d(d) = \lambda_d(d') = 0$, no idiosyncratic TFP heterogeneity ($z = z' = 1$), no aggregate TFP shock ($Z = Z' = 1$), and that only idiosyncratic firm default shocks $\mathcal{I}$ are operative, with $\mathbb{E}[\mathcal{I}'] = \rho$. Note that consumption C is determined by the household budget constraint (equation (11) in the paper).

Plug the partial derivatives into η_l' to get the following expression for η_l'

$$\eta_l' \;=\; -\,\frac{\alpha(\alpha-1)\,k'^{\alpha-2}\,\partial_{l_b'}\mathcal{M}'}{\partial_{k'}\mathcal{M}' \;+\; \alpha(\alpha-1)\,k'^{\alpha-2}\,\partial_{R_l'}\mathcal{M}'}\,\frac{l_b'}{R_l'}, \tag{40}$$

where I indicated the partial derivatives of $\mathcal{M}'$ with short-hand notation for simplicity. The derivatives of the discount factors $(\partial_{l_b'}\mathcal{M}', \partial_{k'}\mathcal{M}', \partial_{R_l'}\mathcal{M}')$:

$$\frac{\partial \mathcal{M}'}{\partial l_b'} = \mathcal{M}'\left[-\gamma\,\frac{1}{C'}\,\frac{\partial C'}{\partial l_b'} \;+\;\gamma\,\frac{1}{C}\,\frac{\partial C}{\partial l_b'}\right],\;\frac{\partial \mathcal{M}'}{\partial k'} = \mathcal{M}'\left[-\gamma\,\frac{1}{C'}\,\frac{\partial C'}{\partial k'} \;+\;\gamma\,\frac{1}{C}\,\frac{\partial C}{\partial k'}\right],\;\frac{\partial \mathcal{M}'}{\partial R_l'} = \mathcal{M}'\left[-\gamma\,\frac{1}{C'}\,\frac{\partial C'}{\partial R_l'} \;+\;\gamma\,\frac{1}{C}\,\frac{\partial C}{\partial R_l'}\right], \tag{41}$$

appear since banks are big in the economy and do not take aggregates as given. These are the key terms that allow to have well defined values for loans even in the stationary equilibrium, in absence of equity issuance cost. In the stationary equilibrium aggregate variables are constant, i.e. $C = C'$, and $\mathcal{M}' = \beta\left(\frac{C'}{C}\right)^{-\gamma} = \beta$ but, in general, the derivatives of the SDF $(\partial_{l_b'}\mathcal{M}', \partial_{k'}\mathcal{M}', \partial_{R_l'}\mathcal{M}')$ are non-zero, as long as $\gamma > 0$, as shown by equations (41). Hence, it is fundamental to have some curvature in the utility function to have well defined loan levels even in absence of equity issuance cost.

Hence, for every firm, the problem reduces to find a triplet (l_b', r_l', k'), given states and future policy functions (which enter inside the discount factor through the term C'), such that

$$\begin{aligned}
1 &= \rho\beta\left(1 + (1-\tau)\left(\alpha k'^{\alpha-1} - \delta\right)\right)\\
1 &= \rho\beta\left(1 + (1-\tau)r_l'\right)\\
1 &= \rho\beta(1 + r_l')\left(1 - \;\frac{\alpha(\alpha-1)\,k'^{\alpha-2}\,\partial_{l_b'}\mathcal{M}'}{\partial_{k'}\mathcal{M}' \;+\; \alpha(\alpha-1)\,k'^{\alpha-2}\,\partial_{R_l'}\mathcal{M}'}\,\frac{l_b'}{R_l'}\right),
\end{aligned}$$

and where η_l' is substituted out using equation (40). In this simplified setting, the first equation yields a closed-form expression for k', and the second provides one for r_l', both of which do not depend on firm characteristics. Given these values, l_b' is determined by the third equation, which generally depends on firm-specific idiosyncratic characteristics as long as $\partial_{l_b'}\mathcal{M}' \neq 0$. Note that for $\partial_{l_b'}\mathcal{M}'$ to be non-zero, the model requires either a tax shield on interest payments or a scrap value. In the absence of both of these elements (and assuming no equity issuance costs), the resource constraint would simply determine consumption as total output minus aggregate investment in physical capital. In this case, firm debt would be irrelevant for consumption, and consequently, $\partial_{l_b'}\mathcal{M}' = 0$ and l_b' undetermined.

To demonstrate this formally, consider the case with zero recovery rate and no small banks. Using the resource constraint (31), we compute the partial derivatives of C and C'

with respect to a given bank's loan l'_b. First,

$$\frac{\partial C}{\partial l'_b} = 0 \quad \text{and} \quad \frac{\partial C'}{\partial l'_b} = -\frac{\partial T'}{\partial l'_b},$$

since C does not depend on l'_b. Next, rewrite T as

$$T = \sum_{\text{age}=0}^{\infty} \Big(k^{\alpha} \; - \; \sum_{b=1}^{B} r_l l_b \; - \; \delta k \Big) \, \phi(\text{age}),$$

and $\phi(\text{age}) = \rho^{\text{age}}(1-\rho)$ is the firm distribution.[36] Since

$$\frac{\partial T'}{\partial l'_b} = -\,\tau\, r'_l\, \phi(\text{age}+1) \quad \text{it follows that} \quad \frac{\partial C'}{\partial l'_b} = \tau\, r'_l\, \phi(\text{age}+1).$$

Substituting into the derivative of the stochastic discount factor gives

$$\frac{\partial \mathcal{M}'}{\partial l'_b} = \mathcal{M}' \Big[-\,\gamma\, \frac{1}{C'}\, \tau\, r'_l\, \phi(\text{age}+1) \Big]. \tag{42}$$

Thus, both $\gamma > 0$ and $\tau > 0$ are required for $\partial \mathcal{M}'/\partial l'_b \neq 0$. The presence of the probability density of the firm distribution, $\phi(.)$, within equation (42) underscores the non-atomistic behavior of banks, which evaluate the marginal impact of a non-zero measure of firms with identical characteristics—summarized by firm age—on aggregate consumption. Banks internalize that if they extend the same loan amount l'_b to all such firms, their collective behavior affects aggregate consumption.

B.3 Model Extension: Mathematical Details

This subsection provides the mathematical details of the extension in Section 7, in which firms make endogenous default decisions based on exit value shocks drawn from a logistic distribution.

[36] Note that this is the post default firm distribution: after the mass of defaulting firms, $1-\rho$, has reentered at age $= 0$; indeed, $\phi(0) = 1-\rho$.

Proposition II: Firm's Problem with Logistic Exit Value Shocks

The expected equity valuation equation $\mathbb{E}^\epsilon[V_F(x,X)]$ can be rewritten as

$$\mathbb{E}^\epsilon[V_F(x,X)] = \tilde{V}_F(x,X) + \zeta\bar{V}_F \cdot \ln\left(1 + e^{-\zeta^{-1}\frac{\tilde{V}_F(x,X)}{\bar{V}_F}}\right), \tag{43}$$

with $V_F(x,X)$ given by (33). As $\zeta \to 0^+$, the firm's problem characterized by equation (34) with $\mathbb{E}^\epsilon[V_F(x,X)]$ given by (43), subject to all firm constraints specified in Section 7, converges to

$$V_F(x,X) \;= \max\left\{\tilde{V}_F(x,X), 0\right\}. \tag{44}$$

The proof is derived in the calculations shown below.

Before proceeding note that the following indefinite integral can be computed as

$$\int \epsilon \cdot \frac{\zeta^{-1}\bar{V}_F^{-1}e^{-{-1}\bar{V}_F^{-1}\epsilon}}{\left(1+e^{-\zeta^{-1}\bar{V}_F^{-1}\epsilon}\right)^2}d\epsilon = -\epsilon\frac{e^{-\zeta^{-1}\bar{V}_F^{-1}\epsilon}}{1+e^{-\zeta^{-1}\bar{V}_F^{-1}\epsilon}} - \zeta^{-1}\ln\left(1+e^{-\zeta^{-1}\bar{V}_F^{-1}\epsilon}\right) + \mathcal{C},$$

where $\mathcal{C}$ is an integration constant. This indefinite integral is handy to compute the expectation of the value function with respect to ϵ, which is:

$$\begin{aligned}
\mathbb{E}^\epsilon[V_F(x,X)] &= \mathbb{E}^\epsilon[\mathcal{I}^d \cdot \tilde{V}_F(x,X)] + \mathbb{E}^\epsilon[(1-\mathcal{I}^d)\cdot\epsilon] \\
&= \mathbb{E}^\epsilon[\mathcal{I}^d]\cdot\tilde{V}_F(x,X) + \int_{\tilde{V}_F}^{\infty} \epsilon \cdot \frac{\zeta^{-1}\bar{V}_F^{-1}e^{-\zeta^{-1}\bar{V}_F^{-1}\epsilon}}{\left(1+e^{-\zeta^{-1}\bar{V}_F^{-1}\epsilon}\right)^2}d\epsilon \\
&= \mathbb{E}^\epsilon[\mathcal{I}^d]\cdot\tilde{V}_F(x,X) + (1-\mathbb{E}^\epsilon[\mathcal{I}^{d'}])\cdot\tilde{V}_F(x,X) + \zeta\bar{V}_F\cdot\ln\left(1+e^{-\zeta^{-1}\frac{\tilde{V}_F(x,X)}{\bar{V}_F}}\right) \\
&= \tilde{V}_F(x,X) + \zeta\bar{V}_F \cdot \ln\left(1+e^{-\zeta^{-1}\frac{\tilde{V}_F(x,X)}{\bar{V}_F}}\right),
\end{aligned}$$

with

$$\mathbb{E}^\epsilon[\mathcal{I}^d] = \tilde{\mathcal{I}}^d = Pr[\mathcal{I}^d = 1|(x,X)] = Pr[\tilde{V}_F(x,X) > \epsilon] = \frac{1}{1+e^{-\zeta^{-1}\frac{\tilde{V}_F(x,X)}{\bar{V}_F}}}.$$

For small positive ζ, ϵ converges to its mean, which is 0. The probability of solvency approaches

$$\lim_{\zeta\to 0^+}\mathbb{E}^\epsilon[\mathcal{I}^d] \;=\; \begin{cases} 1 & \text{if } \tilde{V}_F(x,X) > 0 \\ 0.5 & \text{if } \tilde{V}_F(x,X) = 0 \\ 0 & \text{if } \tilde{V}_F(x,X) < 0. \end{cases} \tag{45}$$

Rewrite

$$\mathbb{E}^{\epsilon}[V_F(x,X)] = \mathbb{E}^{\epsilon}[\mathcal{I}^d] \cdot \tilde{V}_F(x,X) + \int_{\tilde{V}_F}^{\infty} \epsilon \cdot \frac{\zeta^{-1}\bar{V}_F^{-1} e^{-\zeta^{-1}\bar{V}_F^{-1}\epsilon}}{\left(1+e^{-\zeta^{-1}\bar{V}_F^{-1}\epsilon}\right)^2} d\epsilon,$$

and note that $\lim_{z\to 0^+} \int_{\tilde{V}_F}^{\infty} \epsilon \cdot \frac{\zeta^{-1}\bar{V}_F^{-1} e^{-\zeta^{-1}\bar{V}_F^{-1}\epsilon}}{\left(1+e^{-\zeta^{-1}\bar{V}_F^{-1}\epsilon}\right)^2} d\epsilon$ converges to zero for any $\tilde{V}_F$. Thus,

$$\lim_{\zeta\to 0^+} \mathbb{E}^{\epsilon}[V_F(x,X)] = \begin{cases} \tilde{V}_F(x,X) & \text{if } \tilde{V}_F(x,X) > 0 \\ 0 & \text{if } \tilde{V}_F(x,X) \leq 0. \end{cases}$$

■

B.4 Role of the Capacity Constraint Parameter κ

In the stationary equilibrium, κ functions primarily as a scaling device that governs the number of small banks. It governs how finely the model can approximate aggregate moments—particularly the profitability of the banking sector. For instance, in a model with only a few large banks, adding or removing a bank causes discrete jumps in aggregate outcomes, which limits the ability to precisely match the observed profitability. Introducing a fringe of small banks with binding capacity constraints allows the model to "smoothly interpolate" the degree of competition and match the data more accurately.

To illustrate, in the current calibration $\kappa = 0.1$, meaning each small bank can lend up to one-tenth as much as a big bank. With 7 small banks, this configuration contributes the equivalent of 0.7 large banks in aggregate capacity. Now suppose instead we set $\kappa = 0.000175$ and introduce 4,000 small banks to mimic the long tail of small banks in the data—this yields the same total capacity: $\mathcal{B}\kappa = 0.000175 \times 4000 = 0.7$.

This can also be seen algebraically by substituting the equilibrium condition $l_b^s = \kappa l_b$ into the expressions for investment and dividends (under symmetry and in stationary equilibrium):

$$\begin{aligned} \tilde{i} &= i + (B + \mathcal{B}\kappa)\, l_b - (B + \mathcal{B}\kappa)\, l_b', \\ d &= (1-\tau)\left[Zzk^{\alpha} - r_l\,(B + \mathcal{B}\kappa)\, l_b\right] + \tau\delta k - \tilde{i}. \end{aligned}$$

What matters for the stationary equilibrium is the aggregate capacity of small banks, summarized by the product $\mathcal{B}\kappa$, rather than the values of $\mathcal{B}$ or κ individually. This term $\mathcal{B}\kappa$ consistently appears as a multiplicative factor in the generalized Euler equation for the big banks—specifically in equations (27)–(29)—through its presence in the stochastic discount factor $\mathcal{M}$, its derivatives, and in the firm equity issuance terms. Recall that each generalized Euler equation is derived without imposing symmetry. Ex ante, each bank assesses the sensitivity of its own loan supply choice on the demand curve with respect to its own loan quantity, treating the strategies of all other banks as given. Once symmetry is imposed ex post, the term $\mathcal{B}\kappa$ always enters the generalized Euler equation of the big banks multiplicatively. Thus, as long as $\mathcal{B}\kappa$ is held constant, the stationary equilibrium remains unchanged.

However, while this rescaling leaves the stationary equilibrium unaffected, it does matter for out-of-steady-state dynamics. With 4,000 small banks and $\kappa = 0.000175$, each small bank is much smaller and more fragile. Given a macro shock of a certain size, it becomes more likely that one or more small banks would fail—potentially creating different transitional dynamics. No small bank fails along the transition paths under the current calibration ($\kappa = 0.1$, $\mathcal{B} = 7$). Hence, if the model were rescaled to feature 4,000 small banks—holding $\mathcal{B}\kappa$ constant in the stationary equilibrium—each bank would be smaller and therefore more likely to fail in response to the same aggregate shock. This could amplify the transition dynamics. In this sense, the current scaling can be seen as conservative.

Relatedly, given the structure of the model, the stationary equilibrium depends only on the total loan mass $B + \kappa\mathcal{B}$, rather than on the particular combination (with $B \geq 1$) of big and small banks. For example, defining bank market concentration as $B/(B + \kappa\mathcal{B})$, different concentration levels can be equivalently targeted by adjusting the composition of B and $\kappa\mathcal{B}$ while holding total capacity constant. This is because, under the capacity constraint, small banks do not behave strategically independently, but effectively act as scaled-down big banks.

B.5 Role of α on the Main Mechanism

As discussed in Section 5, higher values of the curvature of the profit function α tend to generate a higher bank profitability, ceteris paribus. Here, I discuss the role of the parameter α in shaping the mechanism of endogenous financial frictions discussed in the paper. Figure 10 illustrates the production function k^α and its derivative for two values of α. As α increases, firms with very low capital face a lower marginal product of capital (MPK), but the MPK falls less steeply with capital. Specifically, when $\alpha = 0.55$, the derivative curve lies below that for $\alpha = 0.34$ at very small values of k, but crosses and stays above it already starting from small values of k.

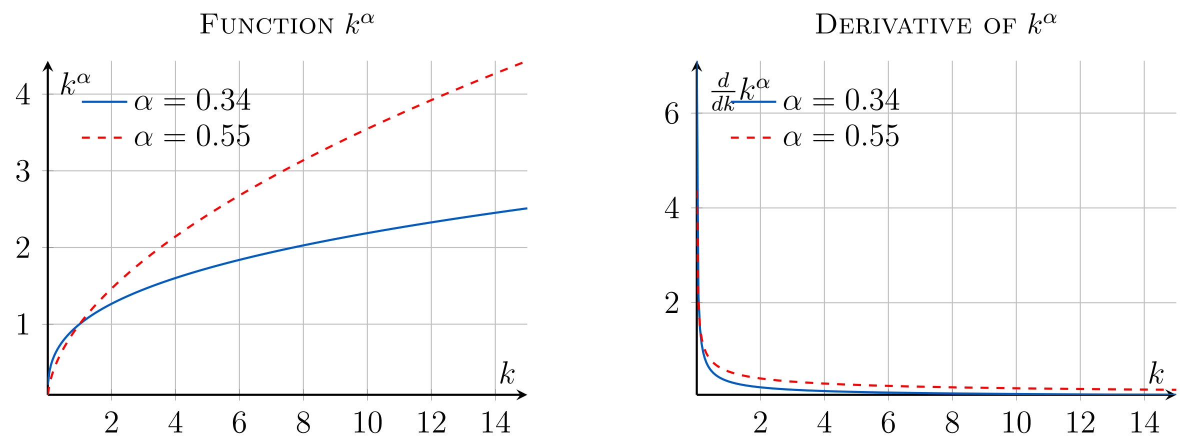


Figure 10. CAPITAL RESPONSE FUNCTIONS FOR DIFFERENT CURVATURES α

This implies that banks internalize a shift in marginal value: one dollar lent to a tiny firm yields less marginal return when α is higher, but becomes more valuable as the firm accumulates capital. As a result, banks with market power extract relatively smaller markups from the smallest firms (compared to the case with $\alpha = 0.34$) but shift toward higher markups (also compared to the case with $\alpha = 0.34$) as firms grow. Another way to visualize this is through the second derivative of the production function, shown in Figure 11. When $\alpha = 0.55$, the red dashed line lies below the blue line, pointing to a more negative inverse elasticity $\eta_l'(.)$ and thus more market power for the banks. This aligns with the generalized Euler equation idea, which evaluates the sensitivity of the marginal product of capital—see equation (9) in the simple model.

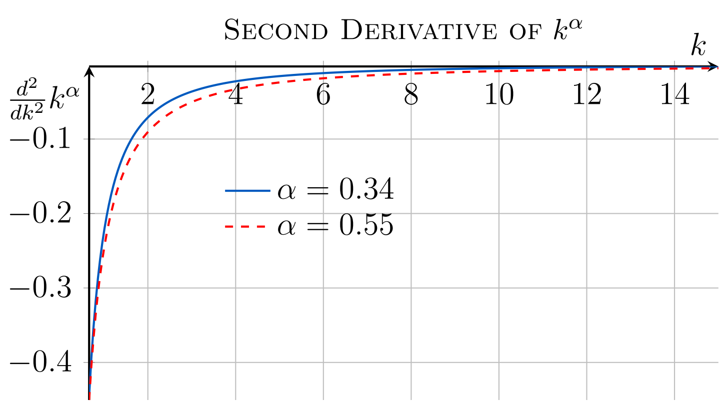


Figure 11. SECOND DERIVATIVE OF THE REDUCED FORM PROFIT FUNCTION k^α

This behavior of the profit function, in the range of α considered, substitutes financing costs across the firm distribution, reinforcing the paper's central mechanism.

C Additional Material

This section includes two parts: (i) the first part contains additional figures, and (ii) the second part contains additional empirics.

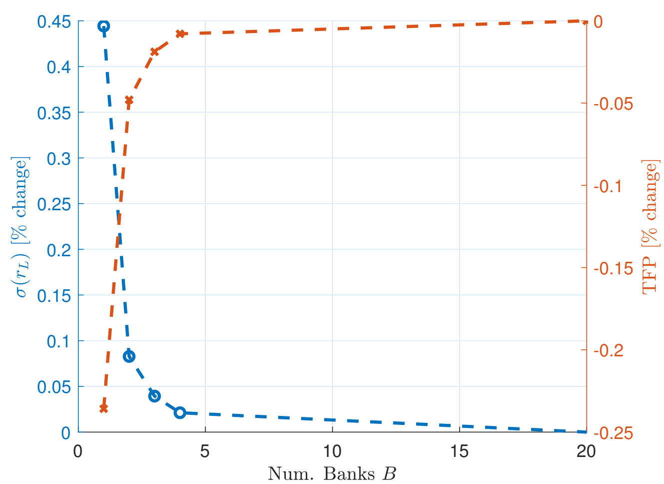
C.1 Additional Figures and Tables

Table C1. Parameter Values

		Parameter	Value	Target/Source
HHs	Time Discount	β	0.995	Match Deposit Rate
	CRRA	γ	1	Standard
Firms	Depreciation Rate	δ	0.03	Standard Quarterly Depreciation
	Curvature of the Profit Function	α	0.34	See Section 5
	Corporate Tax Rate	τ	0.197	Effective Corp. Tax (OECD Tax Database)
	Default Rate	$1-\bar{\rho}$	0.21%	Quarterly C&I Charge-off to Loan (FDIC)
	Equity Flotation Cost	λ_0	0.8	Internally calibrated (see Table I)
	Variance of Idio. Productivity	σ_z	0.08	Internally calibrated (see Table I)
	Starting Capital	k_0	0.237	Internally calibrated (see Table I)
Banks	Fixed Entry Cost	F_E	[0.831,1.21]	Internally calibrated (see Table I)
	Capacity Constraint Parameter	κ	0.1	Small-to-big assets ratio (see Figure C14)
	Recovery Rate	ν	0.084	Internally calibrated (see Table I)

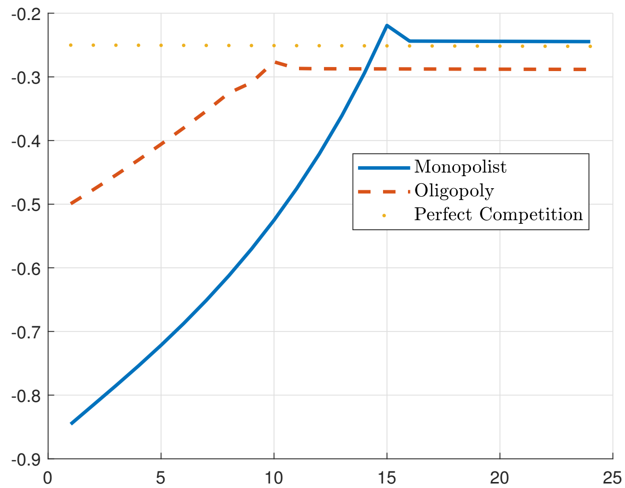
Notes: The table reports the parameter values.

Figure C1. Stationary Equilibrium and Credit Misallocation



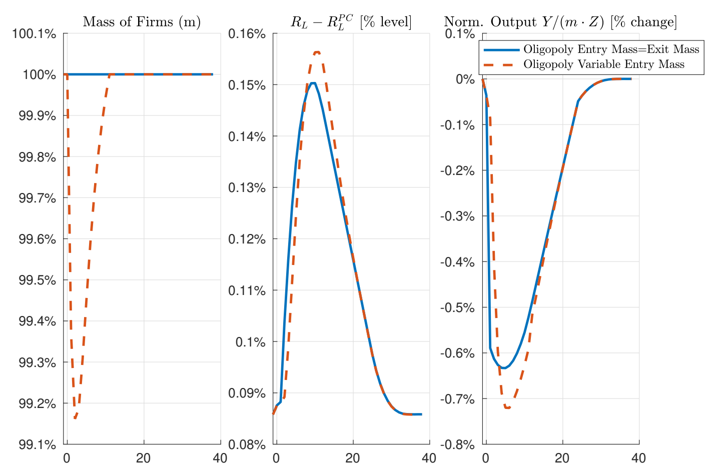
Notes: This figure reports the standard deviations of interest rates and TFP, each expressed as a ratio relative to the values observed in an economy with 20 banks. I vary the fixed cost of entry F_E, thereby allowing a gradual increase in the number of banks entering the financial intermediation market. The figure illustrates the mechanism: limited competition in the financial sector leads to a higher dispersion of interest rates and a lower aggregate productivity.

Figure C2. Stationary Equilibrium and Cross-Sectional Markups



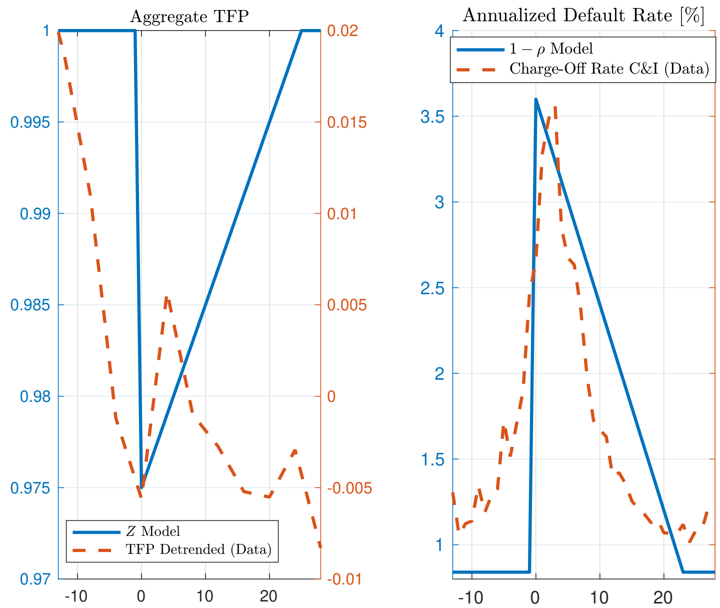
Notes: This figure reports the inverse elasticities $\eta'_L(x, X, x', X')$ appearing in the generalized Euler equations (23), evaluated for firms of different ages in the stationary equilibrium. These elasticities measure the sensitivity of the future loan interest rate with respect to the quantity of loans.The X-axes reports the firms' age.

Figure C3. TFP Shock, Financing, Real Activity, and Firms' Entry



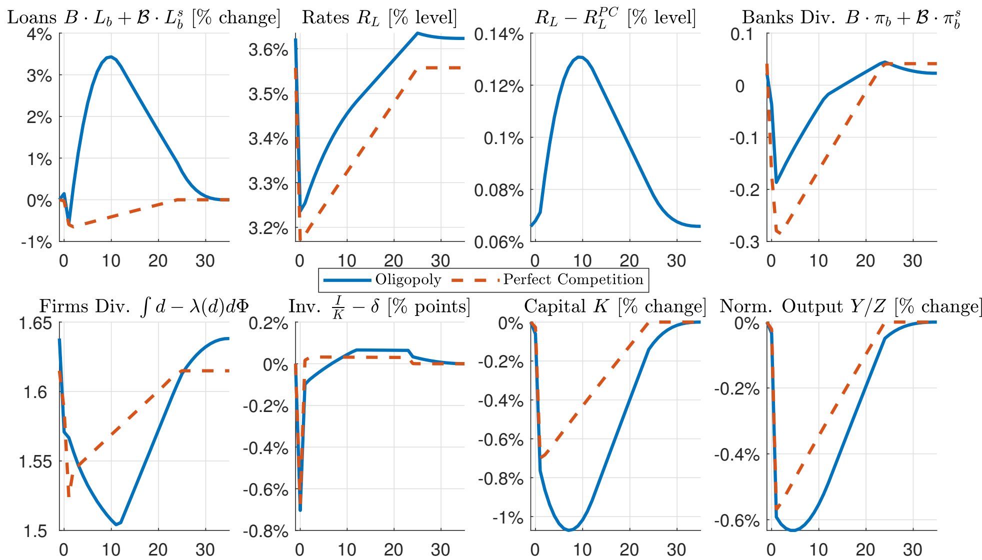
Notes: This figure reports the transitional dynamics of (i) the aggregate interest rate spread and (ii) the aggregate output (normalized to remove the exogenous component Z_t to isolate the effects on the endogenous component) $\int k_t^{\alpha}(x_t, X_t)\,\mathrm{d}\Phi_t$ following the TFP shock reported in Figure C4 in Appendix C.1 with and without the assumption *Entry Mass = Exit Mass*. The evolution over time of the aggregate interest rate spread is calculated as the difference between (i) the annualized aggregate interest rate $\int r_{l,t}(x_t, X_t)\,\mathrm{d}\Phi_t$ under the calibrated oligopoly and (ii) the annualized aggregate interest rate $\int r_{l,t}^{\mathrm{PC}}(x_t, X_t)\,\mathrm{d}\Phi_t^{\mathrm{PC}}$ under perfect competition. The right panel is divided period-by-period by the mass of firms m to eliminate the mechanical effect of GDP falling simply because the mass of firms falls. X-axes report time t.

Figure C4. Shock



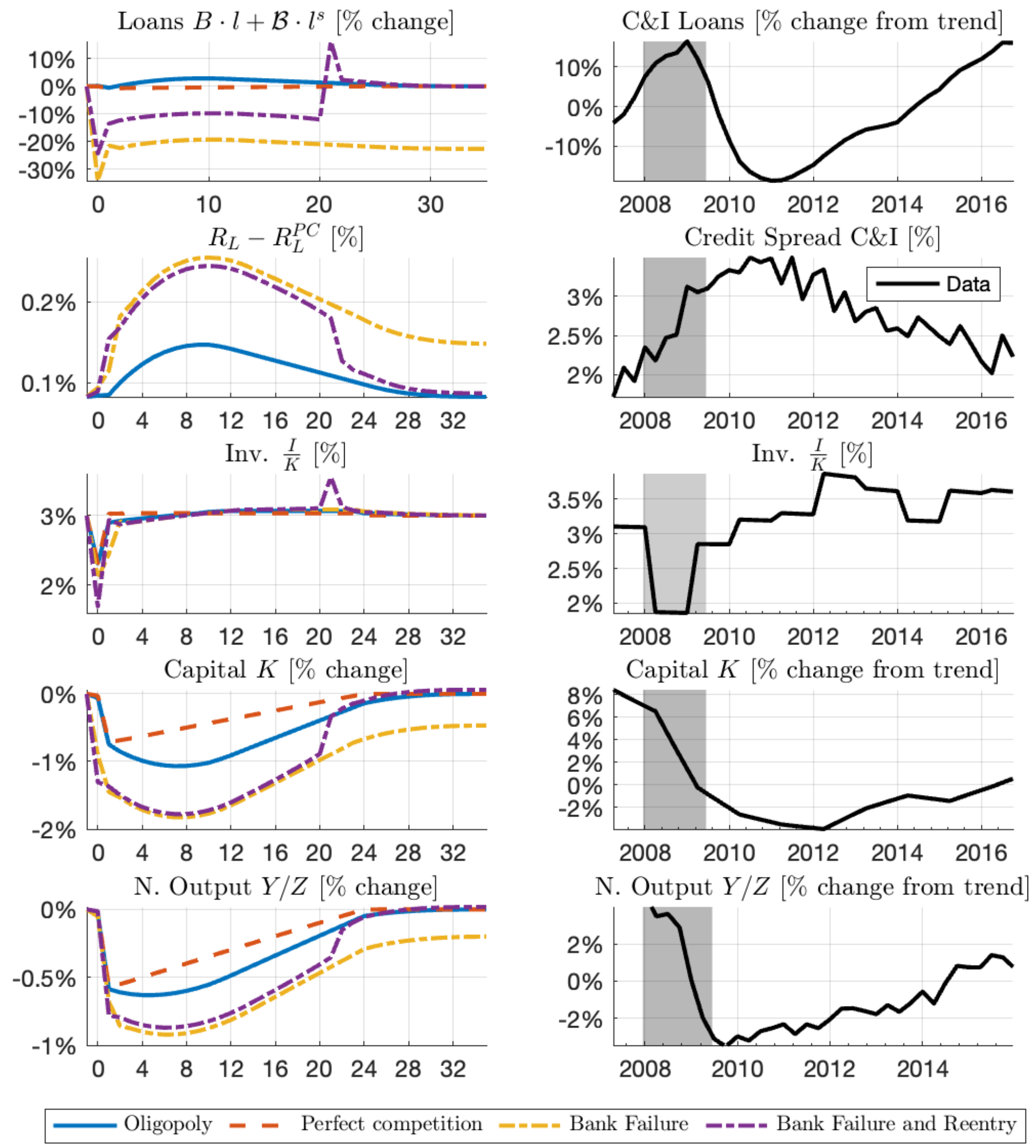
Notes: This shock is calibrated to a similar magnitude to that of the Great Recession. The X-axis reports time t, expressed in quarters both in the model and in the data. The data reported in the graph are from 2005:Q1 till 2014:Q3. TFP has been linearly detrended using data from 1997:Q2 to 2017:Q2. This same shock pushes one bank to default in the calibrated model. Following a sudden unexpected decrease in the aggregate TFP (the firms default probability $1-\rho$ at time $t=0$ decreases) the economy mean-reverts to its original level. After the unexpected shock, all agents can perfectly forecast the mean-reversion path.

Figure C5. TFP Shock, Financing, and Real Activity



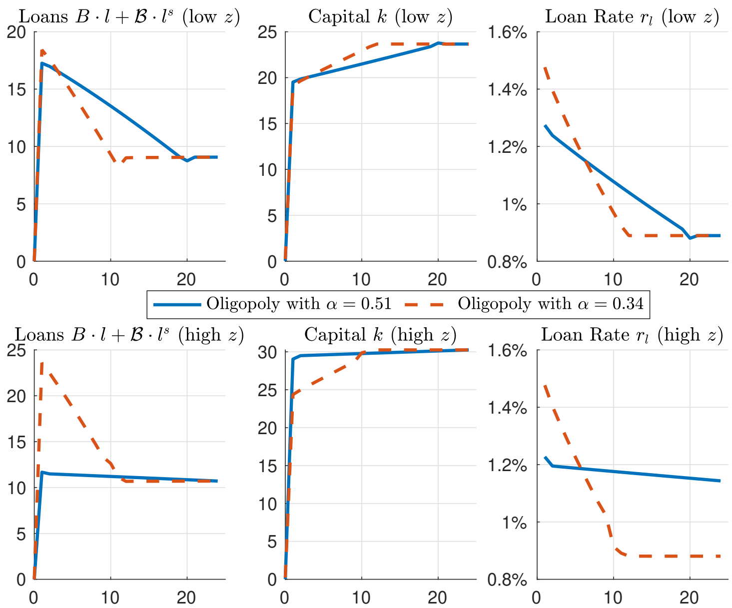
Notes: This figure reports the transitional dynamics without firm idiosyncratic TFP heterogeneity (the counterpart in the baseline model is figure 3), following the TFP shock reported in Figure C4. X-axes report time t.

Figure C6. Model vs. the Great Recession with Reentry



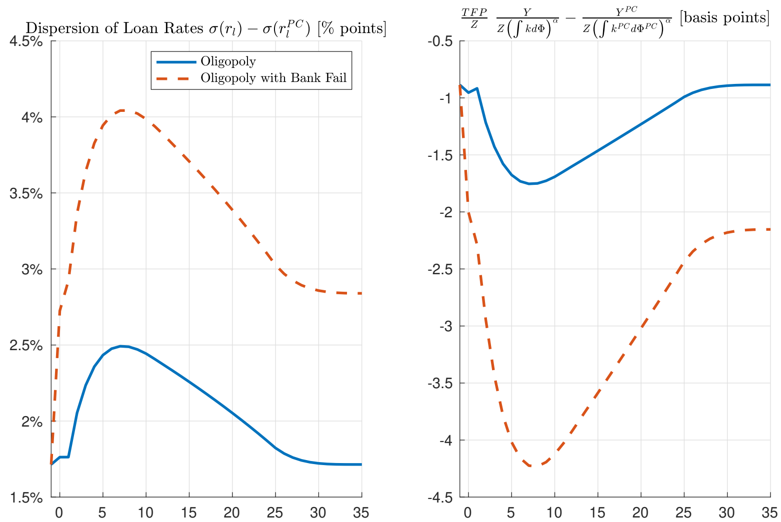
Notes: This figure compares the model dynamics of Subsection 6.5 with the data in proximity of the Great Recession. In particular, the right column reports: i) C&I loans (expressed in % deviation from the linear trend calculated from 1997:Q2 to 2017:Q2), ii) C&I credit spread, iii) quarterly investment rate, iv) physical capital (also expressed in % deviation from the linear trend calculated from 1997:Q2 to 2017:Q2), and v) normalized output (also expressed in % deviation from the linear trend calculated from 1997:Q2 to 2017:Q2). Data are linearly detrended. The panels on the right columns report the percentage change of the detrended series from the trend. X-axes report time t, expressed in quarters for the model (left column).

Figure C7. Stationary Equilibrium Policy Functions with Higher α



Notes: This figure shows the equilibrium loan quantity (left panel), physical capital (central panel), and the loan interest rate (right panel) for firms of different age and productivity, comparing $\alpha = 0.51$ to $\alpha = 0.34$ as calibrated in the stationary equilibrium of Section 5. The steady-state values of capital and loans for $\alpha = 0.51$ have been normalized by those for $\alpha = 0.34$ to make the graphs comparable. X-axes report the firms' age.

Figure C8. Bank Default and Dispersion of Loan Rates



Notes: This figure reports the transitional dynamics of the dispersion of loans' interest rates, calculated as the square root of $\int r_{l,t}^2(x_t, X_t)\,\mathrm{d}\Phi_t - (\int r_{l,t}(x_t, X_t)\,\mathrm{d}\Phi_t)^2$, and aggregate TFP, calculated as $\int k_t^{\alpha}(x_t, X_t)\,\mathrm{d}\Phi_t / (\int k_t(x_t, X_t)\,\mathrm{d}\Phi_t)^{\alpha}$, following the unexpected shock reported in Figure C4 in Appendix C.1. Both measures are calculated as percentage points differences with the perfectly competitive benchmark. X-axes report time t.

Table C2. Parameter Values in the Extension

		Parameter	Value	Target/Source
HHs	Time Discount	β	0.995	Match Deposit Rate
	CRRA	γ	1	Standard
Firms	Depreciation Rate	δ	0.03	Standard
	Curvature of the Profit Function	α	0.34	See Section 5
	Corporate Tax Rate	τ	0.197	Effective Corp. Tax (OECD Tax Database)
	Exogenous Default	$1-\bar{\rho}$	0.21%/2	See Section 7
	Scale Parameter of Exit Value Shocks	ζ	0.0245	Internally calibrated (see Table C3)
	Equity Flotation Cost	λ_0	0.7	Internally calibrated (see Table C3)
	Variance of Idio. Productivity	σ_z	0.07	Internally calibrated (see Table C3)
	Starting Capital	k_0	0.235	Internally calibrated (see Table C3)
Banks	Fixed Entry Cost	F_E	[0.21,0.81]	Internally calibrated (see Table C3)
	Capacity Constraint Parameter	κ	0.1	Small-to-big assets ratio (see Figure C14)
	Recovery Rate	ν	0.085	Internally calibrated (see Table C3)

Notes: The table reports the parameter values for for the extension with endogenous firm default.

Table C3. Stationary Equilibrium and Annualized Moments in the Extension

Targeted	**Description**	**Moment**	**Model**	**Data**
Yes	Profit/Revenue	$\pi_b / \int r_l(x,X) l_b d\Phi$	17.5%	16.2%
Yes	Freq. of Equity Iss.	$4\int (d(x,X)<0) d\Phi$	3.5%	4.2%
No	Capital to GDP	$K/(4Y)$	2.2	2.2
No	Investment to K	$4I/K$	14%	16%
No	Debt Adjust. to K	$\int 4(B\Delta l_b'(x,X) + \mathcal{B}\Delta l_b^{s\prime}(x,X))\,\mathrm{d}\Phi/K$	0.52%	0.62%
Yes	Market Leverage	$\int (Bl_b(x,X) + \mathcal{B}l_b^s(x,X))/V_F(x,X)\,\mathrm{d}\Phi$	32%	34%
No	Num. Big Banks	B	2	-
No	Num. Small Banks	$\mathcal{B}$	8	-
No	Std. MPK	$2\sqrt{\int (Zz\alpha k^{\alpha-1})^2 d\Phi - \left(\int Zz\alpha k^{\alpha-1} d\Phi\right)^2}$	0.16	0.68
Yes	Std. Investment Rate	$2\sqrt{\int (i/k)^2 d\Phi - \left(\int (i/k) d\Phi\right)^2}$	0.339	0.337
Yes	Recovery Rate	$\int \mathcal{II}^d(\nu k + k_0)/(\mathcal{II}^d Bl_b(x,X) + \mathcal{II}^d \mathcal{B}l_b^s(x,X)) d\Phi$	0.62	0.51
Yes	Default Rate	$\int 1 - \mathcal{II}^d d\Phi$	0.81%	0.84%

Notes: This table reports the targeted and untargeted aggregated annualized moments for the extension with endogenous firm default.

C.2 Different Functional Form for $\lambda(.)$

In this section, I consider a different functional form for the cost of equity issuance. In particular I add a fixed cost H of equity issuance such that:

$$\lambda(d) = \begin{cases} H + \lambda_0 \frac{d^2}{2} & \text{if } d \leq 0 \\ 0 & \text{if } d > 0 \end{cases}.$$

In order to preserve differentiability and being able to derive the GEE, I consider a smooth version, where H is approximated by a logit function:

$$\mathcal{H}(d;\kappa) = \frac{1}{1+\exp(-\kappa \cdot d)},$$

which, for high values of κ, approximates well the step function while preserving differentiability, as shown in figure C9.

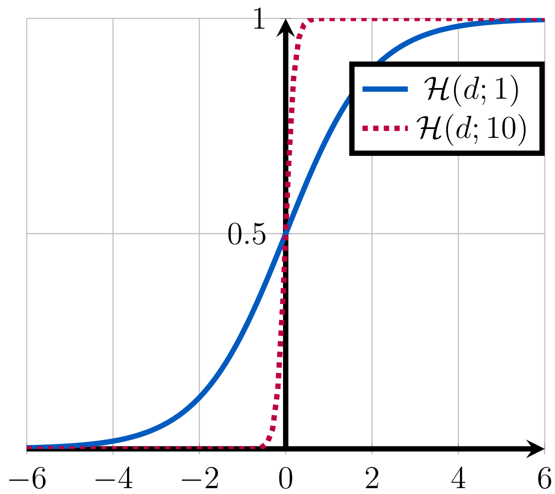


Figure C9. Sigmoid Function

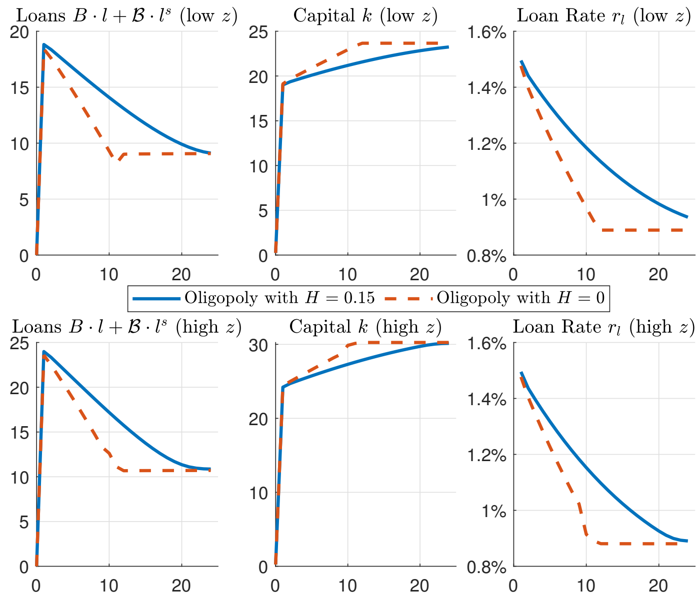
Notes: The plot of the function $\mathcal{H}(d;\kappa)$. Solid blue line: Plot of the sigmoid function (when $\alpha = 1$), typically used in the hidden layer of a neural network. Dashed purple line: Plot of the $\mathcal{H}(d;\kappa)$, when $\kappa = 10$. The higher is κ the more the $\mathcal{H}$ function acquires the shape of a step function.

In summary, I use the following differentiable function

$$\lambda(d) = \begin{cases} H \cdot \mathcal{H}(d;\kappa) + \lambda_0 \frac{d^2}{2} & \text{if } d \leq 0 \\ 0 & \text{if } d > 0 \end{cases}.$$

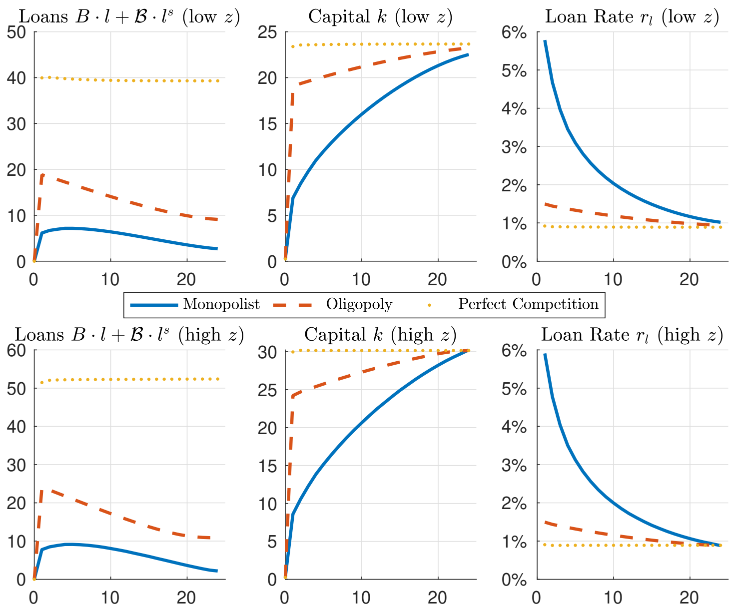
In order to inspect the impact of the fixed cost, I introduce a fixed cost $H = 0.15$ in the calibrated stationary equilibrium of Section 5, where $H = 0$. Results are reported in Figures C10 and C11. In particular, Figure C10 reports the comparison with the baseline stationary equilibrium and Figure C11 reports the comparison across different market structures of the banking sector. Figure C11 indicates that the firm life cycle exhibits qualitatively similar behavior across different bank market structures, consistent with the baseline. Figure C10 demonstrates that, when a fixed cost of equity issuance is introduced, firms take longer to reach their capital targets. As a result, banks extract rents for a longer period, as indicated by the interest rate panel, aligning with the notion that a steep fixed cost can locally amplify the "lack of outside options," thereby allowing banks to exploit this constraint.

Figure C10. Stationary Equilibrium Policy Functions with Equity Issuance Fixed Cost



Notes: This figure reports the equilibrium policies for loan quantity (left panel), physical capital (central panel), and loan interest rate (right panel), for firms of different age and productivity, when I introduce a fixed cost $H = 0.15$ in the equity issuance cost in the calibrated stationary equilibrium of Section 5 ($H = 0$). The fixed cost exacerbates the mechanism of endogenous financial friction, as captured by the central panel. X-axes report the firms' age.

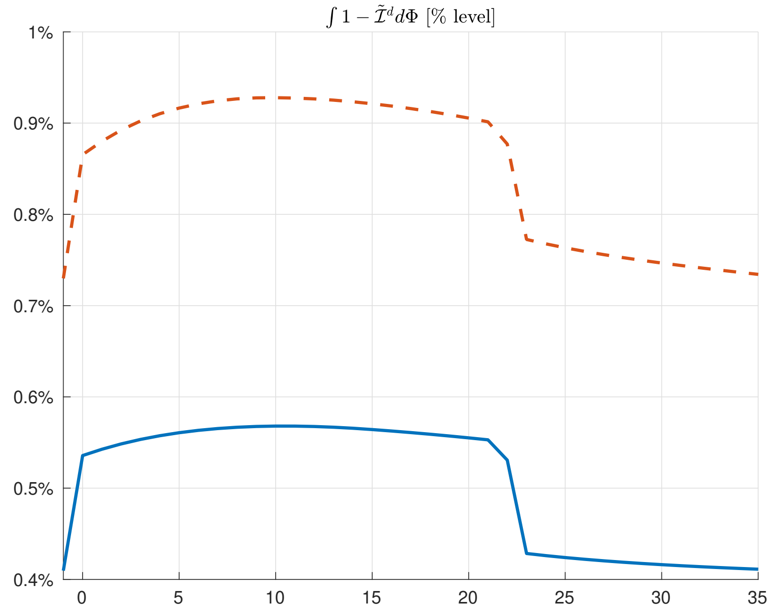
Figure C11. Stationary Equilibrium and Firms' Life Cycle with $H = 0.15$



Notes: This figure reports the equilibrium policies for loan quantity (left panel) and loan interest rate (right panel), along the life cycle of a firm where I perturb the calibrated stationary equilibrium of Section 5 with a fixed cost $H = 0.15$. The mechanism of *endogenous financial friction*, captured by the central panel, behaves in a qualitative similar way to the baseline when different stationary equilibria with different market structure of the banking sector are benchmarked to each others. X-axes report the firms' age.

C.3 Extension

Figure C12. Firm Endogenous Default Rate in the Extension



Notes: This figure reports the transitional dynamics of the annualized firm default rate in the extension. This figure complements Figure 9 in Section 7. X-axes report time t.

C.4 Data

Three different data sources were primarily used in the paper: (i) DealScan and Compustat linked database, (ii) Bank Financials from the Federal Financial Institutions Examination Council's (FFIEC) Central Data Repository for Public Data Distribution, and (iii) aggregate Federal Deposit Insurance Corporation (FDIC) data.

C.4.1 DealScan and Compustat Linked Database

I compile data on 1.027 million loans facilitated between borrowers and lenders from 1986 to 2012, with loan amounts ranging from a minimum of 36 million to a maximum of 81 billion. These loans were obtained from the DealScan database through Wharton Research Data Services (WRDS). Identifiers from this dataset were used to link borrower information with corresponding lender details, which were obtained from WRDS's Compustat database. The data is collected on an annual basis, with all financial information reported at the individual bank level. Tables C4, C5, and C6 present the DealScan, Compustat, and derived variables, respectively. The loans considered in the analysis include Revolver/Line Loans, Term Loans, and Revolver/Term Loans.

Table C4. Variable Mapping to Loan Report Data(DealScan)

Variable Name	Code	Year Start	Year End
Lender Identifier	gvkey	1986	2012
Interest Rate	allindrawn	1986	2012
Loan Amount	dealamount	1986	2012
Loan Maturity	maturity	1986	2012

Table C5. Variable Mapping to Loan Report Data(Compustat)

Variable Name	Code	Year Start	Year End
Lender Identifier	gvkey	1986	2012
Book Value	bookval	1986	2012

Table C6. Variable Mapping to Loan Report Data(Compustat)

Variable Name	
Net Working Capital Ratio	Working Capital/Total Assets
Profitabilities Ratio	Retained Earnings/Total Assets
Return on Total Assets	EBIT/Total Assets
Equity-to-Liability Ratio	Market Value of Equity/Book Value of Total Liabilities
Asset Turnover Ratio	Total Sales/Total Assets
Altman Z-Score	1.2*(Net Working Capital Ratio) + 1.4*(Profitability Ratio) + 3.3*(Return on Total Assets) + 0.6*(Eq-to-Liabilities Ratio) + 1.0*(Assets Turnover Ratio)

Table C7. Log(Credit Spread) Regression

	Dependent variable: All-In Drawn/100					
	(1)	(2)	(3)	(4)	(5)	(6)
log(Book Val)	-0.2535^{***}	-0.2430^{***}	-0.1059^{***}	-0.0105^{***}	-0.1026^{***}	-0.2447^{***}
	(0.0011)	(0.0012)	(0.0011)	(0.0017)	(0.0011)	(0.0021)
log(Deal Amount)	-0.0083^{***}	-0.0295^{***}	-0.0465^{***}	-0.0017	-0.0466^{***}	-0.0088^{***}
	(0.0013)	(0.0014)	(0.0013)	(0.0013)	(0.0013)	(0.0013)
Altman Z-Score	-0.0373^{***}	-0.0258^{***}	-0.0192^{***}	-0.0014^{***}	-0.0196^{***}	-0.0373^{***}
	(0.0004)	(0.0005)	(0.0004)	(0.0005)	(0.0004)	(0.0004)
Profitability		-0.0022^{***}	-0.0007^{***}	-0.0001	-0.0006^{***}	
		(0.0001)	(0.0001)	(0.0001)	(0.0001)	
Net Book Leverage		0.3841^{***}	0.1695^{***}	0.0163^{*}	0.1690^{***}	
		(0.0073)	(0.0066)	(0.0092)	(0.0066)	
Binary Secured			1.2820^{***}	1.0668^{***}	1.2759^{***}	
			(0.0030)	(0.0036)	(0.0030)	
Firm Age					-0.0010^{***}	
					(0.0001)	
C_5						1.5841^{***}
						(0.0456)
log(Book Val)$\times$ C_5						-0.0361^{***}
						(0.0071)
Constant	3.7986^{***}	4.0015^{***}	2.6295^{***}	1.3619^{***}	2.6369^{***}	3.4056^{***}
	(0.0212)	(0.0218)	(0.0200)	(0.0239)	(0.0200)	(0.0221)
Firm-Fixed Effects	No	No	No	Yes	No	No
Time-Fixed Effects	Yes	Yes	Yes	Yes	Yes	No
Observations	846854	843768	843768	843654	843768	846829
R^2	0.112	0.116	0.275	0.500	0.275	0.110
Adjusted R^2	0.112	0.116	0.275	0.497	0.275	0.110

Note: *p<0.1; **p<0.05; ***p<0.01

C.4.2 Bank Financials

I compile data on all banks registered in the U.S. from 1984 to 2022, using information from the last quarter of each year. The data is sourced from the Federal Financial Institutions Examination Council's (FFIEC) Central Data Repository for Public Data Distribution, to which banks submit their quarterly financials to the Federal Reserve. All financial data is reported at the individual bank level. Specifically, the bank financials were collected through the FFIEC's Central Data Repository for Public Data Distribution, covering all registered banks in the U.S. from 2001 to 2022, with annual reporting. Tables C8 and C9 present the balance sheet and income statement variables, respectively.

Table C8. Variable Mapping to Call Report Data (Balance Sheet)

Variable Name	Code	Year Start	Year End
Assets	TOTAL ASSETS	2001	2022
Commercial & Industrial Loans	C&I LOANS, C&I LOANS, U.S. ADDRESSEES	2001	2022
Deposits	TOTAL DEPOSITS	2001	2022
Federal Funds Purchased	SECURS PURCHSD UDR AGRMNTS TO RESELL	2001	2022
Federal Funds Sold	FEDERAL FNDS SOLD, FEDERAL FNDS SOLD IN DOMESTIC OFFICS	2001	2022
U.S. Treasury Securities	US TREAS SECS-AVL-FOR-SLE-FAIR VALUE	2001	2022
U.S. Agency Obligations	US OBLGS ISSD BY US GOV SPON AGC-FAI, US OBLGS ISSD BY US GOVT AGCS-FAIR	2001	2022

Table C9. Variable Mapping to Call Report Data (Income Statement)

Variable Name	Code	Year Start	Year End
Interest Income Loans	INTEREST AND FEES ON LOANS	2001	2022
Interest Expense Deposits	INTEREST EXPENSE ON SAVING DEPOSITS	2001	2022
Charge-Off	CHARGE-OFFS ON C&I LOANS	2001	2022
Total Non-Interest Income	TOTAL NONINTEREST INCOME	2001	2022
Total Non-Interest Expenses	TOTAL NON-INTEREST EXPENSES	2001	2022

Table C10. Bank-Level Credit spread and Banks Market Concentration

	Dependent variable:			
	$\log(R_{b,L,t} - R_{M,t})$			
	(1)	(2)	(3)	(4)
$\log L_{b,t}$	-0.058^{***}	-0.068^{***}	-0.155^{***}	-0.031^{***}
	(0.001)	(0.002)	(0.002)	(0.001)
Charge-off$_{b,t}$	0.123^{***}	0.106^{***}	0.097^{***}	0.108^{***}
	(0.004)	(0.004)	(0.004)	(0.002)
$C_{5,t}$			0.854^{***}	
			(0.015)	
Constant	-2.527^{***}	-2.412^{***}	-1.667^{***}	-2.831^{***}
	(0.009)	(0.018)	(0.022)	(0.016)
Bank-Fixed Effects	No	Yes	Yes	Yes
Time-Fixed Effects	No	No	No	Yes
Observations	261135	260236	260236	260236
R^2	0.025	0.178	0.188	0.707
Adjusted R^2	0.025	0.125	0.135	0.687
Note:	*p<0.1; **p<0.05; ***p<0.01			

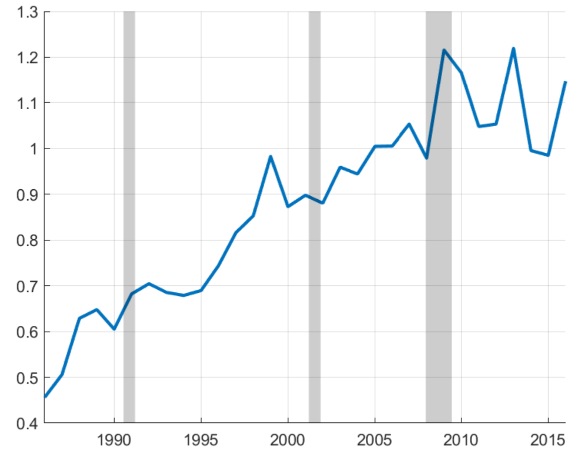
Notes: There is a strong significant positive correlation between credit spread, banks market concentration, and net charge-off rate. The correlation with the quantity of C&I Loans outstanding is significant and negative.

Table C11. Bank-Level Credit spread and Banks Market Concentration (Two-Way Clustering)

	Dependent variable:			
	$\log(R_{b,L,t} - R_{M,t})$			
	(1)	(2)	(3)	(4)
$\log L_{b,t}$	-0.058^{***}	-0.068^{***}	-0.155^{***}	-0.031^{***}
	(0.001)	(0.003)	(0.005)	(0.005)
Charge-off$_{b,t}$	0.123^{***}	0.106^{***}	0.097^{***}	0.108^{***}
	(0.037)	(0.029)	(0.026)	(0.025)
$C_{5,t}$			0.854^{***}	
			(0.023)	
Constant	-2.527^{***}	-2.412^{***}	-1.667^{**}	-2.831^{***}
	(0.011)	(0.032)	(0.047)	(0.057)
Bank-Fixed Effects	No	Yes	Yes	Yes
Time-Fixed Effects	No	No	No	Yes
Observations	261135	260236	260236	260236
R^2	0.025	0.178	0.188	0.707
Adjusted R^2	0.025	0.125	0.135	0.687
Note:	*p<0.1; **p<0.05; ***p<0.01			

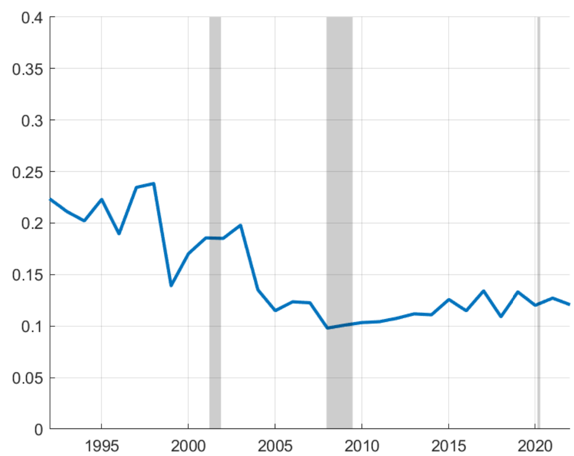
Notes: This table shows that the statistical significance of the coefficients presented in Table C10 is robust to two-way clustering of standard errors by bank and time.

Figure C13. Time Fixed Effects of Stylized Fact 3



Notes: This figure reports the time fixed effects γ_t of the main regression of stylized fact 3.

Figure C14. Ratio of the Tenth Bank by Size and the First Bank



Notes: This figure reports the ratio between the assets of the tenth bank and the first bank by size. X-axes report time in years.

Estimating the Bank Cost Structure. I estimate the marginal cost of producing a loan from an estimate of marginal net expenses, i.e. marginal non-interest expenses net of marginal non-interest income. Marginal non-interest expenses $\frac{\partial NIE^i_{\theta,t}}{\partial l^i_t}$ are derived from the trans-log function:[37]

[37]Non-interest expenses are calculated as total expenses minus the interest expense on deposits, the interest expense on federal funds purchased, and expenses on premises and fixed assets.

$$\log(NIE_t^i) = g_1\log(W_t^i) + \sigma_1\log(l_t^i) + g_2\log(q_t^i) + g_3\log(W_t^i)^2 + \sigma_2[\log(l_t^i)]^2 + g_4\log(q_t^i)^2$$
$$+\sigma_3\log(l_t^i)\log(q_t^i) + \sigma_4\log(l_t^i)\log(W_t^i) + g_5\log(q_t^i)\log(W_t^i) + \sum_{j=1,2} g_6^j t^j + g_{8,t} + g_9^i + \epsilon_t^i, \quad (46)$$

where $g_{8,t}$ indicates time fixed-effects, g_9^i represents bank-fixed effects, W_t^i are input prices (labor expenses over assets), l_t^i corresponds to real loans (one of the two bank i's outputs), q_t^i are safe securities (the second bank output), and t indicates a time trend. Similarly, marginal non-interest income $\frac{\partial NII_t^i}{\partial l_t^i}$ is derived from the trans-log function:

$$\log(NII_t^i) = \sigma_1\log(l_t^i) + g_1\log(q_t^i) + \sigma_2[\log(l_t^i)]^2 + g_2\log(q_t^i)^2$$
$$+\sigma_3\log(l_t^i)\log(q_t^i) + \sum_{j=1,2} g_3^j t^j + g_{4,t} + g_5^i + \epsilon_t^i. \quad (47)$$

Table C12. Derived Variables

Variable Name	
Cost of funds	(Int. exp. dep.+int. exp. Fed funds)/(dep.+Fed funds)
Mg. Net Exp.	Mg. non-int. exp. - mg. non-int. inc.
Int. return on loans	Int. income loans/loans
Safe securities	U.S. Treasury securities+U.S. Agency obligations
Markup	Int. return on loans/(cost of funds+mg. net exp.) - 1
Lerner Index	1 - (cost of funds+mg. net exp.)/int. return on loans

From equations (46) and (47), I calculate the marginal non-interest expenses and income:

$$\text{mg. non-int. exp.} \equiv \frac{\partial NIE_t^i}{\partial l_t^i} = \frac{NIE_t^i}{l_t^i}[\sigma_1 + 2\sigma_2\log(l_t^i) + \sigma_3\log(q_{it}) + \sigma_4\log(W_t^i)], \quad (48)$$

$$\text{mg. non-int. inc.} \equiv \frac{\partial NII_t^i}{\partial l_t^i} = \frac{NII_t^i}{l_t^i}[\sigma_1 + 2\sigma_2\log(l_t^i) + \sigma_3\log(q_{it})]. \quad (49)$$

Marginal net expenses (mg. net exp.) are thus computed as the difference between the marginal non-interest expenses and marginal non-interest income.

C.4.3 Credit spread and Banks Market Concentration with Aggregate Data

To conclude the analysis, I now use aggregate data. Table C13 reports the effects that the market share of the top 5 U.S. banks and the net charge-off rates have on C&I credit spreads. As a reminder, credit spreads are calculated as the difference between the weighted-average effective loan rate for all C&I Loans (R_L) and 3-Month T-bill rates (R_M). Two other measures are added to the analysis: (i) outstanding quantity of C&I Loans ($tn) and (ii) the weighted-average maturity for all C&I Loans. Each period t is a quarter between 1997Q2 and 2017Q2. The results of the following regression

$$R_{L,t} - R_{M,t} = \beta_0 + \beta_1 \times C_{5,t} + \beta_2 \times (1 - \rho_t) + \beta_3 \times L_t + \beta_4 \times M_t,$$

are reports in Table C13.

Table C13. Credit spread and Banks Market Concentration

	Dependent variable:		
	Commercial & Industrial Loan Rates Spreads over intended federal funds rate		
	(1)	(2)	(3)
Market share of top 5 banks (%)	0.040***	0.053***	0.056***
	(0.004)	(0.006)	(0.007)
Net Charge-Off Rate (%)	0.337***	0.295***	0.272***
	(0.051)	(0.051)	(0.059)
C&I Loans ($tn)		−0.391***	−0.345**
		(0.139)	(0.152)
Maturity			−0.121
			(0.157)
Constant	0.434**	0.384**	0.406**
	(0.183)	(0.177)	(0.179)
Observations	81	81	81
R^2	0.644	0.677	0.680
Adjusted R^2	0.635	0.664	0.663
Residual Std. Error	0.291 (df = 78)	0.279 (df = 77)	0.280 (df = 76)
F Statistic	70.500*** (df = 2; 78)	53.802*** (df = 3; 77)	40.292*** (df = 4; 76)

Note: *p<0.1; **p<0.05; ***p<0.01

Notes: There is a strong significant positive correlation between credit spread, banks market concentration, and net charge-off rate. The correlation with the quantity of C&I Loans outstanding is significant and negative, consistently with the model (the elasticity between loan and loan rate is negative).