Bank Capital, Bank Concentration, and Risk-taking

Yu Yi*

February 21, 2025

Abstract

Do bank concentration and capital regulation affect loan quality and capital allocation? To answer this question, I develop a tractable dynamic model featuring heterogeneous risk-taking entrepreneurs and an imperfectly competitive banking sector. The model shows that in a more concentrated banking sector where the bank capital constraint is non-binding, increased concentration reduces default risks and generates an inverse U-shaped relationship between concentration and efficiency. Conversely, in less concentrated markets where capital requirements are binding, a lower concentration improves efficiency with negligible effects on default risks. Empirical evidence from U.S. data supports the model predictions: first, there is a non-monotonic relationship between local HHI and loan rates; second, local HHI is positively correlated with bank capital ratios. The model suggests that reducing bank concentration and raising capital requirements can boost efficiency without increasing entrepreneurial default risks.

Keywords: Bank Concentration, Bank Capital, Risk-shifting, Risk-taking, Allocative Efficiency

---

*Department of Economics, Nankai University (E-mail: ivan_yyi@nankai.edu.cn). I am extremely grateful for Shengxing Zhang for his continued guidance and support. I would like to thank Ben Moll, Ricardo Reis, and Wouter Den Haan for invaluable mentorship. I would like to thank Maarten De Ridder, Matthias Doepke, Jonathon Hazell, Kai Li, John Moore, Martin Oehmke, Mark Schankerman, John Van Reenen, Pengfei Wang, Russell Wong, Zhiwei Xu, Hongda Zhong, seminar participants at LSE, and other conference participants for their valuable comments. All errors are my own.

1 Introduction

Bank concentration has become a significant concern, especially after the 1994 Riegle-Neal Act and the subsequent wave of mergers and acquisitions in the 21st century (Corbae and D'Erasmo (2020), Drechsler et al. (2017)). Figure 1 shows the trend in the Herfindahl-Hirschman Index (HHI) and the share of deposits held by the top 4 banks, with the HHI increasing from approximately 0.05 in 1994 to 0.2 in 2022, and the top 4 deposit share increasing from around 30% to 50% during this period. This raises the natural question of the impact of a highly concentrated banking sector[1].

Do bank concentration and capital regulation affect loan quality (default risk) and capital allocation? My theory suggests that in highly concentrated banking sectors, where the bank capital constraint is non-binding, increased concentration reduces default risks and leads to an inverse U-shaped relationship between concentration and efficiency. Conversely, in less concentrated markets where capital requirements are binding, lower concentration enhances efficiency with negligible impact on default risks. To make these arguments, I develop a tractable dynamic model featuring heterogeneous risk-taking entrepreneurs and a finite number of bankers. Entrepreneurs with differing productivities[2] choose between prudent and gambling projects under limited liability, which incentivizes inefficient gambling (Hellmann et al. (2000)). Bankers compete à la Cournot (Van Hoose et al. (2010)) in both loan and deposit markets[3], subject to a minimum capital requirement[4].

The framework underscores two novel mechanisms through which bank concentration affects entrepreneurial default risks, loan quality, and the efficiency of capital allocation. First, in less concentrated banking sectors, increased concentration enables banks to extract higher rents by raising loan rates. This exacerbates moral hazard, as entrepreneurs facing higher borrowing costs are more likely to invest in gambling projects with high potential returns while shifting downside risks to banks. This finding is consistent with Boyd and De Nicolo (2005). Since gambling reduces efficiency, greater bank concentration also distorts capital allocation. However, in more concentrated banking sectors where capital constraints are non-binding, higher concentration reduces the effective loan rate, discouraging entrepreneurs from undertaking gambling projects. As a result, the allocation of bank lending to entrepreneurs is more efficient. This

[1]It is important to distinguish between bank competition and bank concentration, although the latter is often treated as a proxy for the former.

[2]The model is related to heterogeneous agent frameworks, such as Angeletos (2007), Kiyotaki and Moore (2019), and Moll (2014).

[3]See alternative models of imperfect bank competition in Drechsler et al. (2017), Lagos and Zhang (2022), and Head et al. (2022).

[4]The Basel Committee on Banking Supervision developed international regulatory capital standards through a series of capital accords and related publications, which have collectively been in effect since 1988. For more details, see https://www.federalreserve.gov/supervisionreg/basel/basel-default.htm

mechanism is summarized as the *risk-shifting mechanism*.

Second, higher bank concentration increases the net interest margin, which induces more self-financed entrepreneurs, who tend to be less productive but allocate their entire endowment to prudent projects. This is referred to as the *net margin mechanism*.

In the model, the prudent project provides a guaranteed return, while the gambling project offers a higher potential return but results in total loss if unsuccessful. The prudent project yields a higher expected return. Entrepreneurs operate under a linear production technology and face borrowing constraints. Bankers solve an *optimal contracting problem*, where entrepreneurs' actions (project choices) and types (productivities) are either unobservable or observable at a cost.

Bankers are subject to a minimum capital requirement, but whether this capital constraint is binding is endogenously determined by market concentration. In contrast to the literature that assumes exogenous capital requirements or perpetually binding capital constraints (Brunnermeier and Koby (2018), Li (2019), Repullo (2004)), I allow the capital constraint to be non-binding, thereby exploring the role of bank capital in shaping the effects of concentration on loan quality and efficiency. This approach aligns with empirical findings that banks voluntarily hold more capital than the regulatory minimum and independently adjust their capital ratios (Alfon et al. (2004), Flannery and Rangan (2008)). Theoretically, I show that in concentrated banking sectors, market power incentivizes banks to accumulate more capital, making the capital constraint non-binding. Empirically, I find a significant positive correlation between local bank concentration and the Tier 1 risk-based capital ratio: as the HHI increases from 0 to 1, the capital ratio rises by 3.2%. These results are consistent with Yi (2022) and Li and Song (2023).

Entrepreneurs have heterogeneous productivity. The efficiency of their investments depends on whether they are allocated bank loans and whether they take on risks. In equilibrium, entrepreneurs are classified into four types based on productivity: (1) borrowing entrepreneurs who undertake risky ventures, (2) borrowing entrepreneurs who engage in cautious activities, (3) lending entrepreneurs who provide credit, and (4) autarky entrepreneurs who remain financially inactive. Highly productive entrepreneurs borrow to operate at full capacity and choose optimal investment projects, whereas less productive entrepreneurs typically deposit their endowments in banks. Due to imperfect competition in the banking sector, a positive net margin between loan and deposit rates persists, prompting some entrepreneurs (autarky entrepreneurs) to withdraw from the credit market and use their initial endowments for production instead.

My theory has four key predictions. First, less productive borrowing entrepreneurs are more likely to engage in gambling projects, and higher loan rates amplify this tendency. Due to asymmetric information, bankers charge uniform repayment rates, resulting in more productive borrowers receiving a larger share of production returns. This incentivizes them to select prudent projects with higher expected profits. Conversely, less productive borrowing entrepreneurs tend to pursue gambling projects, benefiting from limited liability protection when their projects fail. Additionally, higher loan rates increase funding costs, thereby intensifying moral hazard on the part of borrowers and raising the share of gambling projects and risky loans through the extensive margin. In this context, the loan rate deteriorates loan quality, which is consistent with the findings of Boyd and De Nicolo (2005).

Second, the relationship between bank concentration and loan rates is non-monotonic. As bank concentration increases, banks gain market power, leading to higher loan rates. However, as the banking sector becomes more concentrated, the negative correlation between loan rates and loan quality weakens banks' incentives to further raise loan rates. Instead, banks accumulate capital beyond the minimum requirement, distorting resource allocation to the real sector and restricting borrowing capacity. This leads to more entrepreneurs entering the credit market, which lowers the equilibrium loan rate. Consequently, the relationship between concentration and loan rates initially turns negative. At very high levels of concentration, the cost of capital declines, causing loan rates to rise again[5]. Thus, the relationship between bank concentration and loan rates follows a non-monotonic pattern: initially positive, then negative, and positive once more at high levels of concentration. This finding is analogous to Petersen and Rajan (1995), who demonstrate that reduced bank competition can enhance credit access and lower interest rates (or exhibit a non-monotonic relationship) as a result of information asymmetries and bank-firm relationships.

Using data from RateWatch and the FDIC, I find a similar non-monotonic pattern between bank concentration and loan rates. I use the HHI as a proxy for local bank concentration, which ranges from 0 to 1. Specifically, when the HHI is below 0.6, higher bank concentration leads to significantly higher loan rates. However, as the HHI increases from 0.6 to 0.7, the loan rate significantly decreases by 0.024%. When the HHI exceeds 0.8, the relationship turns significantly positive again. These patterns persist when I use a shift-share instrument (Borusyak et al. (2021), Adão et al. (2019), and Schubert et al. (2024)) to identify exogenous variation in local bank concentration.

Third, the effect of bank concentration on entrepreneurial default risks depends on both the *risk-shifting mechanism* and *net margin mechanism*. The direction of the *risk-shifting mechanism* depends on whether the bank capital constraint is binding. When the bank capital constraint is binding, a more concentrated banking sector is associated with higher loan rates, which in turn increases entrepreneurial default risks. Conversely, when banks accumulate ex-

[5]The effective loan rate, defined as the product of the capital price and loan rate, is consistently negatively correlated with bank concentration when the capital constraint is non-binding.

cess capital beyond the minimum requirement, higher bank concentration reduces default risks. Additionally, the *net margin mechanism* suggests that as the banking sector becomes more concentrated, the spread between loan and deposit rates widens, leading to a higher proportion of autarky entrepreneurs. Although these entrepreneurs are less efficient, they tend to invest in prudent projects due to their reliance on internal financing. Consequently, as bank concentration increases, the *net margin mechanism* results in lower entrepreneurial default risks.

Fourth, the model has implications for how bank concentration affects the efficiency of capital allocation through the *risk-shifting mechanism* and *net margin mechanism*. Allocative inefficiency arises when resources are directed toward autarky entrepreneurs or gambling borrowers. When the capital constraint is binding, higher bank concentration exacerbates inefficiency through both mechanisms. However, when banks hold excess capital beyond the minimum requirement, higher bank concentration leads to allocative efficiency through *risk-shifting mechanism*, but inefficiency through the *net margin mechanism*.

To summarize, in a less concentrated banking sector where the bank capital constraint is binding, higher bank concentration is associated with lower efficiency and ambiguous default risks, as the two mechanisms exert opposing effects on entrepreneurs' risk-taking behavior. Calibration to U.S. data suggests that the effect of bank concentration on entrepreneurial default risks is negligible when the capital constraint is binding, due to the comparable magnitudes of the two mechanisms. In contrast, in more concentrated banking sectors with a non-binding capital constraint, higher bank concentration reduces entrepreneurial default risks and creates an inverse U-shaped relationship with output, as the two mechanisms have opposing effects on allocative efficiency.

The model suggests that reducing bank concentration is beneficial when the bank capital constraint is binding, as it enhances efficiency without significantly affecting entrepreneurial default risks. However, when the capital ratio exceeds the minimum requirement, achieving allocative efficiency without increasing risks may require both a reduction in bank concentration and an increase in the minimum capital requirement.

My paper reconciles the debate on the relationship between bank concentration and default risks. Despite extensive studies, the relationship remains contentious among both theorists and empirical researchers. Boyd and De Nicolo (2005) argue that lower bank concentration reduces lending rates and borrowing costs for entrepreneurs, thus encouraging less risky behavior. In contrast, Corbae and Levine (2018) show that lower bank concentration increases risk-taking, as shrinking profit margins erode franchise values. Their framework assumes that banks invest in assets with exogenous return distributions. Meanwhile, Martinez-Miera and Repullo (2010) identify a U-shaped relationship between bank concentration and stability when loan defaults are imperfectly correlated. Unlike these studies, I examine the role of bank capital and demonstrate that the effect of bank concentration on entrepreneurial default risks depends on whether capital constraints are binding. Additionally, I contribute to this literature by moving beyond representative household models and exploring the implications of resource allocation.

While some empirical studies support both views, the overall evidence on the relationship between banking market structure and risk-taking remains inconclusive. Some papers find a negative correlation between bank competition and stability (Hellmann et al. (2000); Beck et al. (2003); Agoraki et al. (2011); Tabak et al. (2012); Jiang et al. (2023); Carlson et al. (2022); Beck et al. (2013)). Others argue that increased bank competition can enhance economic stability (De Nicolò et al. (2004); Beck et al. (2006); Carlson and Mitchener (2009); Craig and Dinger (2013)). My theory bridges these two strands of evidence.

I also contribute to the literature on the impact of bank concentration on efficiency. Several studies, including Cetorelli and Gambera (2001), Jayaratne and Strahan (1996), Black and Strahan (2002), Diez et al. (2018), and Joaquim et al. (2019), suggest that lower bank concentration fosters competition, thereby improving credit access, allocative efficiency, and economic growth. However, banks finance risky projects with safe liabilities, with their counterparts' risk-taking motives as another pivotal driver of inefficiency. This paper extends this literature by examining how the *net margin mechanism* and *risk-shifting mechanism* interact to produce a non-monotonic effect of bank concentration on capital allocation.

The rest of the paper is organized as follows. In Section 2, I present the model environment. Section 3 characterizes the symmetric model equilibrium and discusses the implications of the *risk-shifting mechanism* and the *net margin mechanism*. Section 4 provides a quantitative calibration of the model, examining how these two mechanisms affect the relationship between bank concentration, efficiency, and risk. In Section 5, I provide micro-data evidence to test model predictions, and discuss the policy implications. Finally, Section 6 concludes the paper, while the appendix includes proofs and robustness checks.

2 Model Environment

I introduce an imperfectly competitive banking sector and entrepreneurs' risk-taking behaviors into a heterogeneous agent model in the spirit of Moll (2014). Consider a model economy with discrete time and infinite horizon, where time is indexed by $t = 0, 1, 2, \cdots$. It captures the credit structure of an economy consisting of a continuum of entrepreneurs, a fixed number of bankers, and a unit measure of capital suppliers. Entrepreneurs are short-lived, while bankers and capital suppliers are long-lived. At each period, bankers intermediate resources among a continuum of ex-ante heterogeneous entrepreneurs, and capital suppliers provide capital to both bankers and entrepreneurs.

2.1 Entrepreneurs

There is a continuum of short lived entrepreneurs, who are indexed by their productivity z. The productivity of entrepreneurs is assumed to follow an exogenous distribution $G(z)$ that is identically and independently distributed (*i.i.d.*). Entrepreneurs are risk neutral and thus maximize the expected consumption

$$E_{t-1}[c_t]$$

At period t, entrepreneurs are endowed with two production technologies: a prudent project and a gambling project. The prudent project yields a return of z per unit of capital, while the gambling project generates a return of αz with probability p, and zero otherwise. The success of the gambling project is subject to an idiosyncratic shock. Following Hellmann et al. (2000):

Assumption 1 $\alpha > 1$ *and* $\alpha p < 1$.

This assumption indicates that although the gambling project yields a higher return in the event of success, its expected return is lower than that of the prudent project. Entrepreneurs who choose gambling projects benefit from limited liability, ensuring that they lose nothing if the project fails, leading to some entrepreneurs opting to gamble in the equilibrium.

In the middle of each period, entrepreneurs choose between borrowing and lending. Borrowers cannot commit to repayments, and lenders face enforcement issues. Bankers, however, have the ability to both commit and enforce agreements, facilitating financial intermediation. Entrepreneurs can either borrow from bankers and repay at a predetermined loan rate (r_t^l) if their project succeeds or deposit funds in banks at a fixed deposit rate (r_t^d).

After production and financial transactions, entrepreneurs generate offspring. They consume a portion (s) of their net returns and invest the remainder as capital, which is equally distributed among the next generation of entrepreneurs. Unlike Moll (2014), there is no wealth heterogeneity within the same generation of entrepreneurs, ensuring that every entrepreneur receives a non-zero endowment, even without a positive inheritance from their parents. All entrepreneurs are endowed with a_t units of capital at the beginning of period t.

Additionally, entrepreneurs face a borrowing constraint

$$k_t \leq \lambda a_t, \ 1 < \lambda < \infty \tag{1}$$

The parameter λ measures financial market efficiency. A value of 1 indicates a complete credit market shutdown, leaving all entrepreneurs financially inactive, while an infinite λ represents a perfect financial market. A finite λ reflects market imperfections, constraining entrepreneurs' borrowing based on their initial endowment. Let $\theta_t = \frac{k_t}{a_t}$ denote the actual leverage ratio of entrepreneurs at period t.

2.2 Bankers

The banking sector is characterized by imperfect competition, with $i = 1, 2, \cdots, M$ long-lived bankers competing à la Cournot for loans (L_{it}) and deposits (D_{it})[6]. In the case where $M = 1$, the banking sector is a monopoly, while as M increases, it approaches perfect competition. At the beginning of each period, banker i holds equity capital N_{it}. Each bank invests all its funding in loans, classified as either safe or risky, with the fraction of risky loans denoted by v_t. Balance sheet identity of banker i then follows

$$L_{it} = D_{it} + N_{it} \tag{2}$$

The balance sheet items at the beginning of period t are summarized in Table 1. The banker i is risk-neutral and derives utility from dividend payouts

$$\sum_{t=0}^{\infty} \beta^t c_{it}^b$$

Each banker is assumed to fully diversify idiosyncratic risk and accumulates equity capital solely through retained earnings[7]. At the end of period t, bankers finance their dividend payouts and retained earnings with returns from their loan and deposit market activities

$$c_{it}^b + q_t N_{it+1} \leq \left[1 + r_t^l(L_t)\right](1 - v_t) q_t L_{it} + p\left[1 + r_t^b(L_t)\right] v_t q_t L_{it} - \left[1 + r_t^d(D_t)\right] q_t D_{it} \tag{3}$$

where the deposit and loan rates are implicitly determined by the aggregate deposit size D_t and loan size L_t. The right-hand side terms in equation (3) capture the income derived by banker i from issuing both safe loans $\left[1 + r_t^l(L_t)\right](1 - v_t) q_t L_{it}$ and risky loans $p\left[1 + r_t^b(L_t)\right] v_t q_t L_{it}$, subtracting the repayment to depositors $\left[1 + r_t^d(D_t)\right] q_t D_{it}$. The price of capital is denoted as q_t. To streamline equation (3), I define

$$p_t^e = (1 - v_t) \times 1 + v_t \times p, \tag{4}$$

[6]Imperfect bank competition follows the Cournot framework from Van Hoose et al. (2010), providing a straightforward yet effective way to analyze the banking structure between the extremes of monopoly ($M = 1$) and perfect competition ($M \to \infty$), that is, a complete spectrum of competitiveness. Under Cournot competition, these extremes mirror those found in Bertrand competition. However, additional frictions may be needed to capture an intermediate market structure under Bertrand competition.

[7]New equity issuance by new investors does not affect the model's main mechanisms.

which reflects the expected loan repayment probability, weighted by the proportion of safe and risky loans. Equation (3) then becomes

$$c_{it}^{b} + q_t N_{it+1} \leq q_t \left\{ \left[1 + r_t^l(L_t)\right] p_t^e L_{it} - \left[1 + r_t^d(D_t)\right] D_{it} \right\}. \tag{3'}$$

I assume that bankers' screening technology is sufficiently inefficient that they remain uninformed about entrepreneurs' productivity and project choices, resulting in a uniform loan rate applied to all borrowers. Bankers protect themselves only through collateralization. This is consistent with the findings of Asriyan et al. (2021), indicating that collateralization and costly screening are substitutes. The implementation of more efficient screening technology mitigates and shifts entrepreneurs' risk-taking behavior.

There is a a minimum capital requirement limits the maximum loan capacity that bankers can provide

$$N_{it} \geq \kappa L_{it} \tag{5}$$

where κ measures the flexibility of the minimum capital requirement. This requirement mandates that a fraction κ of bank loans must be backed by capital. Established by the Basel Committee on Banking Supervision in the late 1990s, this framework aims to regulate market risk and ensure that banks maintain adequate capital to absorb financial losses. The capital constraint is represented as a simplified minimum requirement for the capital to risk-weighted asset ratio. Incorporating the requirement on capital to risk-weighted asset ratio does not change the model's fundamental mechanisms.

2.3 Capital Supplier

There is a continuum of capital suppliers, who are endowed with $\overline{K}$ units of capital. At the end of each period t, they provide capital to entrepreneurs and bankers in a perfectly competitive capital market. Capital suppliers are assumed to lack storage technology so that they rationally choose to be hand-to-mouth.

3 Equilibrium Characterization

In this section, I characterize the model equilibrium and subsequently use the model to discuss how bank concentration affect entrepreneurs' risk taking behaviours through two mechanisms: the *net margin mechanism* and the *risk-shifting mechanism*.

3.1 Entrepreneurs' Side

For simplification, the time index is omitted when it causes no confusion. I first derive the equilibrium conditions under which entrepreneurs choose to invest in gambling projects. While these entrepreneurs benefit from limited liability, they also face lower expected returns, which compels them to borrow in order to take advantage of gambling opportunities. Given the linear production function, borrowing gambling entrepreneurs will consistently reach their borrowing limit.

The incentive compatibility condition requires that gambling borrowers secure a higher expected return compared to self-financed entrepreneurs and prudent borrowers. The benefits of of borrowing entrepreneurs engaged in gambling projects surpass those who are financially inactive when

$$p\left[\alpha z\lambda - q(1+r^l)(\lambda-1)\right]a \geq za.$$

The left-hand side of the inequality indicates that, with a probability of p, borrowing entrepreneurs engaged in gambling projects receive a return from production $\alpha z\lambda a$ net of loan repayment $q(1+r^l)(\lambda-1)a$.

Assumption 2 $\lambda\alpha p > 1$.

Assumption 2 ensures the presence of gambling in the equilibrium. Under this condition, there exists a lower bound on productivity, denoted as z_2, above which entrepreneurs are willing to take risks

$$z \geq \frac{(\lambda-1)p}{\lambda\alpha p - 1}q(1+r^l) \equiv z_2 \tag{6}$$

Moreover, the benefits of of borrowing entrepreneurs who engaged in gambling projects exceed those borrowing entrepreneurs who engaged in prudent projects:

$$p\left[\alpha z\lambda - q(1+r^l)(\lambda-1)\right]a \geq z\lambda a - q(1+r^l)(\lambda-1)a,$$

Investing in the prudent project may be attractive due to its higher expected return, that is, $\alpha p < 1$. Conversely, borrowers may prefer gambling projects, as they face a repayment probability less than 1. Due to the inefficiency of the gambling project, the incentive compatibility condition establishes an upper bound on productivity, denoted by z_3:

$$z \leq \frac{(\lambda-1)(1-p)}{\lambda(1-\alpha p)}q(1+r^l) \equiv z_3 \tag{7}$$

When both incentive compatibility conditions are satisfied, borrowing entrepreneurs will choose the gambling project. These conditions can be simultaneously met in equilibrium if and only if $z_2 < z_3$, leading to the following assumption:

Assumption 3 $\frac{(\lambda-1)p}{\lambda\alpha p-1} < 1$.

Assumption 3 posits that entrepreneurs with productivity levels between z_2 and z_3 will choose to invest in gambling projects. This assumption depends on three key parameters: α, p, and λ. The likelihood of Assumption 3 being satisfied increases with higher values of these parameters. Intuitively, the benefits of gambling grows with an increase in the excess return α or the success probability p. Moreover, when asset pledgeability λ rises, less entrepreneurs borrow, allowing bankers to charge higher loan rate. Therefore, more entrepreneurs take risks.

Given the deposit and loan rates, Proposition 1 outlines the financial decisions of entrepreneurs.

Proposition 1 *There are three productivity cutoffs z_1, z_2 and $\overline{z}_3$, such that*

- *The capital demand for individual entrepreneur is:*

$$k = \begin{cases} \lambda a & z \geq z_2 \\ a & z_1 \leq z \leq z_2 \\ 0 & z \leq z_1 \end{cases}$$

- *Entrepreneurs with productivity between z_2 and $\overline{z}_3$ invest in the gambling project, while those with $z > \overline{z}_3$ or $z_1 < z < z_2$ invest in the prudent project.*

Here, $z_1 = q(r^d + 1)$, $\overline{z}_3 = \min\{z_3, z_{max}\}$.

The cutoff property relies on the linearity of the production function. As outlined in Proposition 1, entrepreneurs' optimal capital demand is at corners: entrepreneurs with productivity levels below z_1 demand zero capital, while those with productivity levels exceeding z_2 demand the maximum amount permitted by their borrowing constraint. Entrepreneurs with productivity levels between z_1 and z_2 demand an amount equal to their initial wealth.

This capital demand structure differentiates two types of marginal entrepreneurs. For those at the productivity level z_1, the return of each additional unit of investment, $\frac{z}{q}$, equals the opportunity cost of not depositing that amount in the bank, $r^d + 1$. In contrast, entrepreneurs at the productivity level z_2 are indifferent between self-financing and borrowing for gambling projects. Under Assumption 3, since $z_2 < q(1 + r^l)$, entrepreneurs at the productivity level z_2 are incentivized to take risks.

Therefore, entrepreneurs with productivity levels below z_1 are termed lending entrepreneurs, those surpassing z_2 are termed borrowing entrepreneurs, and those in between are labeled autarky entrepreneurs. Lending entrepreneurs find production unprofitable due to low productivity and thus deposit their entire endowment in banks. Borrowing entrepreneurs, with higher productivity, borrow up to the borrowing limit to finance production. Autarky entrepreneurs arise from imperfect competition in the banking sector, where bankers charge a positive margin between loan and deposit rates. This discourages entrepreneurs with intermediate productivity from either borrowing or lending, rendering them financially inactive. Autarky entrepreneurs rely on their own funds for production and choose the prudent project, which offers a higher expected return.

Borrowing entrepreneurs face the decision of choosing between the prudent and gambling projects. While gambling offers a lower expected return, the repayment probability is less than one. Proposition 1 shows that those with productivity above $\overline{z}_3$ will choose the prudent project, whereas those within $z_2 < z < \overline{z}_3$ will engage in gambling.

Since bankers cannot observe entrepreneurs' productivity or project choices, they set a uniform loan rate for all borrowers. Entrepreneurs with productivity above $\overline{z}_3$ capture a larger share of production returns and thus invest in the prudent project. Conversely, those within $z_2 < z < \overline{z}_3$ benefit more from limited liability and gamble. Specifically, entrepreneurs with productivity equal to $q(1 + r^l)$ gain no return from the prudent project but receive a positive return if they gamble and the project succeeds.

Loan contracts, whether risky or safe, depend on borrowing entrepreneurs' investment preferences. In the extreme scenario where $\overline{z}_3 > z_{max}$, all loans are risky. In the next section, I will discuss how the equilibrium behaves in scenarios featuring solely risky loans and a mix of both risky and safe loans.

Entrepreneurs' financial and intertemporal decisions generate endogenous loan demand and deposit supply, along with a law of motion for aggregate entrepreneurial capital demand, as described in Lemma 1. Notably, since each generation of entrepreneurs are born with identical initial wealth, the aggregate capital demand aligns with individual capital demand.

Lemma 1 *Aggregate quantities* $\{L_t, D_t, a_{t+1}\}$ *satisfy:*

$$L_t = [1 - G(z_{2t})]\,(\lambda - 1)a_t \tag{8}$$

$$D_t = G(z_{1t})a_t \tag{9}$$

$$\begin{aligned} q_t a_{t+1} =& s\Big\{ \int_{z_{min}}^{z_{1t}} q_t(1 + r_t^d)dG(z_t) + \int_{\overline{z}_{3t}}^{z_{max}} \left\{\lambda[z_t - q_t(1 + r_t^l)] + q_t(r_t^l + 1)\right\} dG(z_t) \\ &+ \int_{z_{1t}}^{z_{2t}} z_t dG(z_t) + p\int_{z_{2t}}^{\overline{z}_{3t}} \left[\alpha\lambda z_t - (\lambda - 1)q_t(r_t^l + 1)\right] dG(z_t)\Big\}a_t \end{aligned} \tag{10}$$

Equation (8) reveals that the aggregate loan demand is determined by three factors: the proportion of borrowing entrepreneurs, the amount each borrows, and their initial capital. Similarly, deposit supply, as described in Equation (9), depends on the total initial capital of lending entrepreneurs. Additionally, Equation (10) outlines the law of motion for aggregate entrepreneurial capital, where the next generation's net wealth, $q_t a_{t+1}$, relies on their saving rate s and net return. Specifically, this net return is a weighted average across depositors, borrowing entrepreneurs in the prudent project, autarky entrepreneurs, and borrowing entrepreneurs engaged in gambling.

3.2 Bankers' Side

Banker i chooses the deposit and loan quantities to determine the optimal deposit and loan rates.

$$1 + r^d + D_i \frac{\partial r^d}{\partial D_i} = \mu_i \tag{11}$$

$$p_e[(1 + r^l) + L_i \frac{\partial r^l}{\partial L_i}] + (1 + r^l) L_i \frac{\partial p_e}{\partial r^l} \frac{\partial r^l}{\partial L_i} = \mu_i + \kappa \chi_i \tag{12}$$

where $q\mu_i$ represents the Lagrangian multiplier on the balance sheet identity, and $q\chi_i$ is the multiplier on the bank capital constraint. Equation (11) shows that the deposit rate depends on the elasticity of deposit supply and the balance sheet multiplier. Conversely, Equation (12) highlights that loan issuance tightens both the balance sheet and capital constraints, with the marginal benefit of issuing loans contingent on the elasticity of loan demand and the sensitivity of the loan repayment probability to loan size.

Let the aggregate loan demand elasticity be $\epsilon^l = -\frac{\partial \ln L}{\partial \ln(1+r^l)}$, and the aggregate deposit supply elasticity be $\epsilon^d = \frac{\partial \ln D}{\partial \ln(1+r^d)}$. The market share of loans and deposits held by banker i are denoted by s_i^l and s_i^d, respectively. The expressions for the deposit and loan rates then become

$$1 + r^d = \frac{\epsilon^d}{\epsilon^d + s_i^d} \mu_i \tag{13}$$

$$p_e(1 + r^l) = \frac{\epsilon^l}{\epsilon^l - s_i^l[1 + \frac{\partial \ln p_e}{\partial (1+r^l)}]} (\mu_i + \kappa \chi_i) \tag{14}$$

Equations 13 and 14 demonstrate that the optimal deposit and loan rates involve a markdown or markup over the marginal benefit or cost, depending on the banker i's market share. In a perfectly competitive market, where s_i^l and s_i^d approach zero, no markup or markdown exists. However, the negative correlation between loan rate and the expected repayment probability, as demonstrated in Proposition 2, introduces an additional term in Equation (14), reducing the expected markup.

Proposition 2 (Loan Rate and Entrepreneurial Default Risks) *Assume* $\frac{zg(z)}{1-G(z)}$ *is increasing*[8]*, we have:*

$$\frac{\partial v}{\partial r^l} \geq 0 \ \& \ \frac{\partial p_e}{\partial r^l} \leq 0$$

where the equality holds when $\overline{z}_3 = z_{max}$*.*

Proposition 2 suggests that higher loan rates result in an increased proportion of risky loans, a lower expected probability of loan repayment, and elevated entrepreneurial default risks in partial equilibrium. Specifically, rising loan rates reduce the net returns for borrowing entrepreneurs, prompting more of them to pursue riskier investments, thereby increasing the share of risky loans. This finding is consistent with Boyd and De Nicolo (2005), where they find that banks with greater market power charge higher loan rates and exhibit higher risk levels. My model differs in two key aspects. First, by incorporating entrepreneurs' risk-taking incentives into a heterogeneous agent framework, it allows for an analysis of their impact on resource allocation. Second, as the following section will show, the relationship between bank concentration and entrepreneurial default risk is shaped by a non-monotonic relationship between bank concentration and loan rates. I term the effect of bank concentration on risks through loan rate the *risk-shifting mechanism*." Notably, if $\overline{z}_3 = z_{max}$, resulting in all loans becoming risky ($v = 1$, $p_e = p$), the *risk-shifting mechanism* is deactivated.

The optimal condition for bank capital is given by

$$q_t = \beta q_{t+1}(\mu_{it+1} + \chi_{it+1}), \tag{15}$$

where accumulating one unit of bank capital today at a cost of q_t relaxes the balance sheet identity and bank capital requirement by multipliers μ_{it+1} and χ_{it+1}, respectively, in the following period.

3.3 Steady State Equilibrium

In this section, I delve into the general equilibrium, emphasizing the symmetric equilibrium throughout the paper.

Definition 1 (Symmetric Equilibrium) *A Symmetric Equilibrium in the economy consists of a sequence of policy function of bankers' decisions* $\{c_{it}^b, N_{it+1}, D_{it}, L_{it}\}_{t=0}^{\infty}$*, a sequence of aggregate quantities* $\{a_{t+1}, D_t, L_t\}_{t=0}^{\infty}$*, a sequence of interest rates* $\{r_t^l, r_t^d\}_{t=0}^{\infty}$*, and a sequence of price* $\{q_t\}_{t=0}^{\infty}$ *such that:*

[8] This is a mild assumption that most distributions satisfy.

1. *Entrepreneurs and bankers maximize expected lifetime utility given prices and interest rates;*

2. *Bankers choose the same quantities for all assets and liabilities;*

3. *Market clearing conditions for*

   - *loan market:* $\sum_{i=1}^{M} L_{it} = L_t$*;*
   - *deposit market:* $\sum_{i=1}^{M} D_{it} = D_t$*;*
   - *capital market:* $\sum_{i=1}^{M} N_{it} + a_t = \overline{K}$*.*

The concept of symmetry suggests that in equilibrium, there is no heterogeneity among bankers. To begin with, it is useful to explore how the symmetric equilibrium in a perfectly competitive banking sector differs from the competitive benchmark scenario, characterized by the absence of risk taking behaviors—specifically, when $\alpha = p = 1$.

Corollary 1 *Assume* $\alpha p \lessapprox 1$*. When* $M \to \infty$ *and* $\kappa = 0$*, there exists a positive net margin (*$r^l > r^d$*) and a non-zero fraction of autarky entrepreneurs (*$z_2 > z_1$*), where*

$$1 + r^l = \frac{1 + r^d}{p} \tag{16}$$

$$z_2 = qp(1 + r^l)\frac{\lambda - 1}{\lambda \alpha p - 1} > q(1 + r^d) = z_1 \tag{17}$$

In the absence of risk-taking incentives, the equilibrium in the model with a perfectly competitive banking sector closely resembles the framework presented in Moll (2014), except for the exclusion of labor. In that model, a single cutoff determines the roles of creditors and lenders, with no positive margins. However, when there is a slight deviation ($\alpha p \lessapprox 1$) from the benchmark case, bankers introduce a positive spread between the loan and deposit rates, interpreted as a risk premium. Moreover, due to the inefficiency of the gambling project ($\alpha p < 1$), a segment of autarky entrepreneurs arises. This risk-taking behavior is undesirable not only because of the inefficiency of the gambling project but also due to the misallocation of resources to unproductive producers.

The following analysis explores the effects of bank concentration on entrepreneurial risk-taking. Specifically, risk-taking is measured by the amount of capital allocated to the gambling project, denoted as risky capital (rc_t) at time t. The equilibrium level of risky capital can be represented as:

$$rc = v[\overline{K} - v_a(\overline{K} - N)] \tag{18}$$

where v_a represents the proportion of autarky entrepreneurs. The relationship between bank concentration and entrepreneurial risk-taking is characterized in Proposition 3.

Proposition 3 (Bank Concentration and Entrepreneurial Risk-taking)

$$\frac{\partial rc}{\partial M} = \left[(1-v_a)\overline{K} + v_a N\right] \frac{\partial v}{\partial M} - v(\overline{K} - N)\frac{\partial v_a}{\partial M} + v_a v \frac{\partial N}{\partial M} \tag{19}$$

The first term in Equation (19) reflects the impact of bank concentration on entrepreneurial risk-taking through the share of risky loans, known as the *"risk-shifting mechanism"*. Proposition 2 shows that entrepreneurs facing higher borrowing costs are more likely to invest in gambling projects with high potential returns while transferring downside risks to banks. However, the model leaves the sign of $\frac{\partial v}{\partial M}$ ambiguous, given the uncertain relationship between bank concentration and loan rates.

The second term in Equation (19) represents how bank concentration influences risk-taking through the proportion of autarky entrepreneurs. As bank concentration increases, the net interest margin expands, leading to a rise in autarky entrepreneurs, described as the *"net margin mechanism"*. This mechanism, supported by Yi (2022) under the assumption of uniform productivity distribution without risk-taking motives, will be demonstrated quantitatively in Section 4. Autarky entrepreneurs, relying on personal funds, generally invest in safer projects, generating a negative relationship between bank concentration and entrepreneurial risk-taking.

The third term in Equation (19) captures the effect of bank concentration on entrepreneurial risk-taking through bank capital. Higher bank capital reduces entrepreneurial initial capital, thereby lowering autarky entrepreneurs' investment and heightening entrepreneurial risk-taking. However, this effect is quantitatively outweighed by the *net margin mechanism*, as will be shown in Section 4.

Overall, the relationship between bank concentration and risk depends on how concentration affects loan rates. If a highly concentrated banking sector that drives up loan rates, bank concentration would increase risk through the *"risk-shifting mechanism"*, while simultaneously reducing risk via the *"net margin mechanism"*. The net effect remains uncertain, depending on the relative magnitude of these two mechanisms. Conversely, if higher bank concentration leads to lower loan rates, both mechanisms would produce a negative correlation between bank concentration and risk.

Role of Bank Capital. Given bankers' limited information about entrepreneurs' productivity and investment preferences, their decisions in the loan market are restricted to two key instruments: bank capital and loan quantity (or loan rate). In a highly concentrated banking sector, the positive relationship between loan rate and entrepreneurial default risk discourages bankers from raising loan rates excessively. Instead, they may adopt a more prudent approach by increasing bank capital to enhance loan repayment prospects. Higher bank capital reduces the endowment available to borrowing entrepreneurs, increasing the share of borrowers and lowering z_2, which subsequently exerts downward pressure on the loan rate. Thus, bank capital plays a pivotal role in influencing how bank concentration affects risk via the loan rate, known as the *"risk-shifting mechanism"*. The subsequent section will quantitatively analyze the role of bank capital, focusing on equilibrium regions where the bank capital constraint is either binding or non-binding.

4 Quantitative Analysis

This section begins with parameter calibration of the model, followed by a quantitative analysis of how bank concentration affects risks through the *"risk-shifting mechanism"* and the *"net margin mechanism"*. The analysis considers two scenarios: one where only risky loans are present ($\overline{z}_3 = z_{max}$) and another where both safe and risky loans coexist ($\overline{z}_3 < z_{max}$). Quantifying these mechanisms enables a detailed examination of their distinct effects and provides insights into the effect of bank concentration on allocative efficiency.

4.1 Calibration

I select parameter values to reflect key characteristics of the U.S. economy from 1994 to 2020. This calibration primarily focuses on productivity distribution $G(z)$, bank concentration $\frac{1}{M}$, and the asset pledgeability parameter λ.

Bank concentration ($\frac{1}{M}$) is measured using the average Herfindahl-Hirschman Index (HHI) of the U.S. banking sector over this period. The HHI is defined as:

$$HHI = \sum_{i=1}^{M} {s_i^d}^2 = \sum_{i=1}^{M} (\frac{1}{M})^2 = \frac{1}{M} \tag{20}$$

The second equality follows that each banker holds a market share of $1/M$ in the deposit market. To evaluate bank concentration, I apply the method from Drechsler et al. (2017), calculating the national-level HHI as a weighted average of branch-level HHIs, using branch deposits as weights. Based on Equation (20), M is estimated to be approximately 7.45.[9]

I assume the distribution of productivity follows a bounded Pareto distribution (Melitz (2003)), whose property satisfies the assumption in Proposition 2. The bounded Pareto distribution is characterized by by the shape parameter γ, the upper bound z_{max}, and the lower

[9] It should be noted that M is an integer in the model economy. However, to maintain calibration precision, approximations like 7 or 8 are not used. In comparative statics, M values are treated as integers.

bound z_{min}, where z_{min} is normalized to 1. Parameters z_{max} and γ are then calibrated to reflect the observed dispersion of productivity and markups in the U.S. economy over the sample period.

Following Hsieh and Klenow (2009), the disparity between the 75th and 25th percentiles of Total Factor Productivity Ratio (TFPR) is 0.53[10]. The cumulative density distribution function of $log(z)$ for a bounded Pareto distribution is expressed as $\frac{z_{min}^{-\gamma}-e^{-\gamma z}}{z_{min}^{-\gamma}-z_{max}^{-\gamma}}$[11]. Based on this, z_{max} is estimated to be around 3. The shape parameter γ is set to 1.5, consistent with an average markup of approximately 20%, as suggested by Liu and Wang (2014)[12].

Based on the value of M, the asset pledgeability parameter λ is chosen to match the bank capital to asset ratio in the U.S. between 2001 and 2017. A higher value of λ indicates a more efficient financial market, leading to a higher bank capital to asset ratio. According to FRED, the average U.S. regulatory capital-to-risk-weighted assets ratio during this period was approximately 13.71%. Based on the calibrated M value, λ is thus set at approximately 15.

Following Basel III guidelines, the parameter κ is used to derive the implied capital policy requirement. Basel III sets a minimum total capital ratio of 8%, with an additional capital conservation buffer that raises the minimum capital requirement to 10.5% of risk-weighted assets. Since the benchmark model excludes a risk-based capital constraint, κ is initially set at 0.08. However, including risk-based capital constraints does not significantly affect the model's core mechanisms.

In this model, each period represents one year. Following Gali and Monacelli (2005) and Christiano et al. (2005), the discount factor β is calibrated to 0.96, implying a steady-state annual risk-free rate of approximately 4%. Entrepreneurs are assumed to be more patient, with a discount factor $s = 0.98$, following Gentry and Hubbard (2000). Aggregate capital capacity, $\overline{K}$, is normalized to 1. The parameter α is calibrated to 1.05, reflecting the average difference between the loan rate and the federal funds target rate in the U.S. Table 2 provides a summary of the calibration for all parameters.

[10] Hsieh and Klenow (2009) differentiate between TFPQ (using plant-specific price deflators) and TFPR (using industry-level price deflators). However, with the price of consumption goods normalized, TFPQ and TFPR are equivalent in this model.

[11] Assume there is a random variable X which follows a bounded Pareto distribution with parameter L, H and γ, where γ denotes the shape parameter, L denotes the minimum, and H denotes the maximum. Define $Y = log(X)$. The cumulative distribution function (cdf.) of X is $F_X(x) = Pr(X \leq x) = \frac{L^{-\gamma}-x^{-\gamma}}{L^{-\gamma}-H^{-\gamma}}$. Therefore, the $c.d.f.$ of Y is $F_Y(x) = Pr(Y \leq x) = Pr(log(X) \leq x) = Pr(X \leq e^x) = \frac{L^{-\gamma}-e^{-\gamma x}}{L^{-\gamma}-H^{-\gamma}}$. Correspondingly, the probability distribution function of Y is $\frac{\gamma e^{-\gamma x}}{L^{-\gamma}-H^{-\gamma}}$.

[12] It should be noted that I introduce an imperfect competition in the banking sector, which results in an increase in the markup. As a result, the required value of γ is not as high as in Liu and Wang (2014).

4.2 Equilibrium with only Risky Loans

In a scenario where all borrowing entrepreneurs favor gambling projects, the outcome is that all loans become risky. This extreme case is formally described as follows

Corollary 2 *Assume $\alpha p \lessapprox 1$. All loans are risky in the equilibrium.*

The result follows directly from Equation (7), where, as αp nears 1, z_3 approaches infinity. Gambling projects, despite incurring higher costs due to their lower expected returns compared to prudent projects, become more attractive as the expected return gap between the two types of projects narrows. Therefore, when $\alpha p \lessapprox 1$, all borrowing entrepreneurs are incentivized to pursue gambling projects.

To ensure that all loans are risky in equilibrium, I set $p = 0.9$. Figure 3 presents the comparative statics regarding bank concentration, measured by the number of bankers. A higher concentration in the banking sector corresponds to fewer bankers. Panel (b) confirms that all loans are indeed risky under this setup. Panel (a) demonstrates that in a highly concentrated banking sector, banks hold surplus capital exceeding the minimum requirement, while in a less concentrated sector, the minimum capital requirement becomes binding. This finding aligns with the empirical evidence in Section 5.1, which shows a positive relationship between bank concentration and capital holdings, with U.S. banks often accumulating capital above regulatory minimums. The underlying intuition is straightforward: greater concentration in the banking sector reduces deposit rates and the overall deposit supply. Since both deposits and equity capital are liabilities on the bank's balance sheet, a more concentrated sector raises the bank capital ratio due to a substitution effect between these two funding sources.

Panel (c) of Figure 3 illustrates how increased bank concentration affects the capital allocated to gambling projects, showing a consistently negative correlation between bank concentration and entrepreneurial default risks, regardless of whether the capital constraint is binding. In cases where all loans are risky, the *risk-shifting mechanism* becomes inapplicable, leaving the *net margin mechanism* and the bank capital channel, $\frac{\partial N}{\partial M}$, as the primary influences on the relationship between bank concentration and risk-taking. Notably, the bank capital channel is strictly dominated by the *net margin mechanism*, allowing for the exclusion of the bank capital channel from this point forward. A more concentrated banking sector, as shown in Panels (e) and (f) of Figure 3, leads to a higher interest spread imposed by bankers, which increases the proportion of autarky entrepreneurs. These entrepreneurs, using their initial endowments for production, are more inclined to select prudent projects. As a result, a strong negative correlation between bank concentration and risk emerges.

Allocative inefficiency resulting from bank concentration can be attributed to the *net margin mechanism*, as shown in Panel (d) of Figure 3. As bank concentration increases, both the net margin and the proportion of autarky entrepreneurs rise, leading to a greater allocation of capital toward inefficient autarky entrepreneurs through the extensive margin.

4.3 Equilibrium with Both Safe and Risky Loans

In a more general scenario where loans can be either risky or safe, the *risk-shifting mechanism*, outlined in Proposition 2, interacts with the *net margin mechanism* to influence the relationship between bank concentration and risk, as well as allocative efficiency. To enable both gambling and prudent projects to coexist in steady-state, I set $p = 0.7$ in this section.

Figure 4 first illustrates a non-monotonic relationship between bank concentration and the loan rate. Initially, as bank concentration rises, so does the loan rate, due to heightened market power and reduced elasticity of loan demand. However, a negative correlation between the loan rate and the expected probability of loan repayment causes bankers, in the general equilibrium, to internalize entrepreneurs' responses and avoid excessively high rates. This results in an unexpected negative correlation between bank concentration and the loan rate when there are approximately 4 to 8 banks.

In a highly concentrated banking sector, the correlation between bank concentration and the loan rate becomes positive again, driven by the general equilibrium effect through the price of capital. As bank concentration increases, the demand for capital declines, causing a decrease in the capital price (q), which in turn drives up the loan rate.

As bank concentration increases, the observed decline in the loan rate (Figure 4) corresponds with banks accumulating capital beyond the minimum requirement, as shown in panel (a) of Figure 5. In concentrated banking sectors, bankers are incentivized to exceed minimum capital ratios due to a substitution effect between deposits and capital, a motivation further reinforced by the inverse relationship between loan rates and expected loan repayment probabilities. Higher bank capital decreases the endowment available to borrowing entrepreneurs, raising the proportion of such borrowers and lowering z_2, collectively driving down the loan rate.

Panel (b) of Figure 5 further illustrates the relationship between bank concentration and the fraction of risky loans. When the bank capital constraint is binding, increased bank concentration corresponds to a greater proportion of risky loans. Conversely, in cases where the capital constraint is non-binding, the fraction of risky loans decreases as bank concentration intensifies. Specifically, under a binding capital constraint, banks with greater market power tend to set higher loan rates, thereby increasing the proportion of risky loans through the *risk-shifting mechanism*. However, when banks hold excess capital beyond the required minimum, increased concentration results in a lower effective loan rate, $q(1 + r^l)$, leading to a smaller fraction of risky loans.

In the general equilibrium, the relationship between the proportion of risky loans v and the loan rate r^l is not one-to-one. At high levels of bank concentration, equilibrium effects through the capital price q result in an increase in the loan rate. Importantly, the fraction of risky loans is fully determined by $q(1 + r^l)$. Beyond the effect of the loan rate, an increase in the capital price further raises the proportion of risky loans by elevating external funding costs for entrepreneurs. As illustrated in Figure 6, there is a one-to-one relationship between the effective loan return rate and the fraction of risky loans.

4.3.1 Bank Concentration and Entrepreneurial Default Risk

Panel (c) of Figure 5 illustrates the relationship between bank concentration and entrepreneurial default risks. When the bank capital constraint is non-binding, risky capital shows a negative correlation with bank concentration. However, under a binding constraint, this correlation becomes ambiguous.

In cases where the capital constraint binds, an increase in bank concentration raises loan rates. While the *risk-shifting mechanism* suggests a positive correlation between bank concentration and risky capital, this is not clearly observed in panel (c) due to the countervailing effect of the *net margin mechanism*, the second term in Equation (19). As bank concentration increases, the proportion of autarky entrepreneurs—who invest exclusively in prudent projects—also rises, leading to a decline in risky capital and mitigating the expected positive relationship with bank concentration. The calibrated parameters, aligned with U.S. data, indicate that these mechanisms nearly offset each other, rendering the correlation between bank concentration and risks negligible. However, when banks' capital ratios exceed the minimum requirement, both mechanisms jointly contribute to a negative correlation between bank concentration and risk.

4.3.2 Bank Concentration and Allocative Efficiency

Panel (d) of Figure 5 shows a non-monotonic relationship between bank concentration and allocative efficiency, shaped by the interplay of the *net margin mechanism* and the *risk-shifting mechanism*.

According to the *net margin mechanism*, reduced bank concentration improves output by narrowing the loan-deposit rate spread, thereby limiting capital allocated to autarky entrepreneurs, typically less efficient producers. Empirical evidence from Joaquim et al. (2019) supports this, indicating that aligning lending spreads with the global average could increase Brazilian output by 5%. Moreover, when the bank capital constraint is non-binding, higher bank concentration can lower the effective loan rate, thereby reducing the share of risky loans through the *risk-shifting mechanism*. Assumption 1, which imposes lower expected payoffs for gambling projects, mitigates the adverse effects of bank concentration on allocative efficiency. Consequently, the relationship between bank concentration and efficiency follows an inverse U-shape when the capital constraint is non-binding.

Interestingly, the negative correlation between bank concentration and output reverses when the number of banks is around 6 to 8, indicating a local efficiency optimum. Given that the calibrated number of banks (M) is set at 7.45 in the preceding analysis, this result carries significant quantitative policy implications.

5 Discussions

In this section, I present empirical evidence supporting the model's predictions. The results are twofold: first, there is a non-monotonic relationship between bank concentration and loan rates, with the patterns precisely consistent with the model. Second, higher bank concentration correlates with an increased bank capital ratio. Additionally, I discuss regulatory interventions on bank capital as potential strategies for enhancing allocative efficiency and mitigating risks.

5.1 Supporting Evidence

This section introduces new empirical evidence from U.S. data to investigate the correlation between bank concentration and loan rates. While the paper does not explicitly explore the well-studied impact of bank concentration on risks (Jiang et al. (2017); Carlson et al. (2022)), it focuses on empirically testing whether the direction of the *risk-shifting mechanism* is influenced by a non-monotonic relationship between bank concentration and loan rates, as well as a positive interaction between bank concentration and bank capital.

5.1.1 Data Description

The analysis incorporates data from three distinct sources: (i) the Federal Deposit Insurance Corporation (FDIC) provides the Summary of Deposits, (ii) U.S. Call Reports from the Federal Reserve Bank of Chicago offer bank balance sheet information, and (iii) branch-level rate data is obtained from RateWatch. Each dataset is detailed as follows:

Deposit Quantity. The FDIC dataset spans all U.S. bank branches from 1984 to 2021, furnishing information on branch characteristics, ownership details, and deposit quantities at the county level. The FDIC bank identifier is used to link this dataset with other relevant datasets.

Bank Balance Sheet. U.S. Call Reports from the Federal Reserve Bank of Chicago cover the period from March 1994 to March 2021. These reports provide quarterly balance sheet details for all U.S. commercial banks, encompassing information on assets, deposits, various loan types, equity capital, etc. The FDIC bank identifier is utilized to match this dataset with FDIC data.[13]

RateWatch. RateWatch data encompasses monthly loan rates at the branch level, covering the period from 2001 to 2021. The analysis focuses on one of the most frequently observed loan types, namely auto loans with a 72-month maturity (Wang (2024)), to address potential challenges associated with observed and unobserved heterogeneity among various loan products.[14] In particular, the analysis focuses on branches that actively participate in setting the loan rate.

Following the methodology detailed in Drechsler et al. (2017), bank concentration is measured using HHI. Initially, HHI is computed at the county-year level by summing the squared deposit market share of each branch within a county for each year, as specified in Equation (20). Subsequently, to obtain a bank-level HHI, the weighted average HHI is calculated for all branches associated with the same bank institution, employing branch deposit sizes as weights. The main analysis focuses on the periods between 2010 and 2021 to eliminate the structural changes in bank capital regulations in 2009. This leaves 133,360 observations in the main sample. The summary statistics are shown in Table 3 and Table 4.

5.1.2 Bank Concentration and Loan Rate Revisited

Aligned with the theoretical model's predictions, the correlation between bank concentration and the loan rate is expected to be non-monotonic. This section seeks to empirically validate this prediction through a fixed effect regression model, examining the impact of branch-level HHI on loan rates. The analysis initiates with the estimation of the following regression equation:

$$LoanRate_{kt} = \sum_{i=1}^{10} \beta_i HHI_{c(k)t} * \mathbb{1}(HHI_{c(k)t} \in (\frac{i-1}{10}, \frac{i}{10}]) + \alpha_{j(k)} + \alpha_t + \alpha_{s(k)t} + \epsilon_{jt} \tag{21}$$

where $LoanRate_{kt}$ denotes the loan rate for branch k at quarter t, $\alpha_{j(k)}$ represents the fixed effect [15] linked to branch k affiliated with institution j, α_t denotes the time fixed effect, $\alpha_{s(k)t}$

[13] In the Appendix, a local polynomial smoothing technique will be employed to demonstrate the non-monotonic correlation between bank concentration and loan rate at the bank level.

[14] According to the model, borrowers represent entrepreneurs who use external resources for production. This paper specifically focuses on auto loans, given their prevalence in the dataset. It is worth noting that households are not the exclusive borrowers of automobiles; some firms also procure automobiles for business purposes. In the appendix, an additional analysis will be performed using business loans as a robustness check.

[15] $\alpha_{j(k)}$ is the identifier assigned to financial institutions by the Federal Reserve. To account for missing bank identifiers in some observations, I include institution type as an additional control. The main results remain robust, as shown in Figures 9 and 10, even when observations with missing bank IDs are excluded.

accounts for the state-time fixed effect, and $HHI_{c(k)t}$ signifies the branch-level HHI for branch k at county c and quarter t. $\alpha_{s(k)t}$ is included in the regression to address concerns related to varying deregulation policies across states[16]. I cluster the standard error at bank level. The primary coefficients of interest in the regression, denoted by β_i for $i = 1, 2, \cdots, 10$, capture the heterogeneous impact of bank concentration on loan rates across deciles of bank concentration. Specifically, a positive β_i implies a positive correlation between HHI and loan rates when HHI falls within in the $i-1$th to ith decile. In line with the model's predictions, β_i is anticipated to exhibit both positive and negative values.

Figure 7 presents a visual representation of the results, accounting for bank fixed effects, time fixed effects, and state-year fixed effects. Additional specifications are detailed in Table 5. βs are positive and significant coefficients within the HHI range of $(0, 0.6]$, indicating a positive correlation between bank concentration and loan rates. This aligns with the model's prediction that banks with higher market power tend to charge elevated loan rates. However, as HHI increases, the coefficient becomes negative and significantly at the 1% level. Specifically, β_7 is -0.24, signifying that as HHI advances from 0.6 to 0.7, the loan rate decreases by 0.024%. This negative correlation corresponds with the model's prediction and can be ascribed to the interplay of the *risk-shifting mechanism* and non-binding capital constraints: heightened loan rates prompt greater engagement in gambling projects by entrepreneurs. Banks, cognizant of entrepreneurs' decisions, avoid excessively high loan rates, even within a highly concentrated banking sector. Instead, they accumulate excess capital beyond the minimum capital requirement. Furthermore, the positive and significant values of β_9 and β_{10} manifest the general equilibrium effect through the price of capital. Overall, the observed non-monotonic relationship between HHI and loan rates concurs with the model's equilibrium predictions.

Robustness of the Non-linearity. To test the robustness of the non-linear relationship between bank concentration and loan rates, I perform several additional exercises. First, in Figure 8, I use local polynomial smoothing between branch-level HHI and loan rates demeaned at the yearly level. Second, in Table 6, I conduct a third-degree polynomial regression of loan rates on branch-level HHI. Third, I rerun Equation (21) excluding observations with missing bank IDs, with results shown in Figures 9 and 10. The non-linear pattern between bank concentration and loan rates persists across all these specifications. In the Appendix B, I show the non-linear relationship persists when business loans are used.

[16] Interstate branching was restricted in the U.S. until the enactment of the Riegle-Neal Act in 1994. The Dodd-Frank Act of 2010 aimed to mitigate risks associated with financial institutions. Nevertheless, states retain authority to use provisions in the IBBEA to regulate out-of-state entry. Rice and Strahan (2010) devised a bank deregulation index, ranging from 0 to 4, with 4 denoting states with the most stringent entry requirements for out-of-state banks.

Identification: Bank Concentration

When estimating the effect of bank concentration on loan rates, endogeneity issues may bias the estimated coefficients on the HHI. The direction of the bias is ambiguous: an increase in bank concentration could reflect the expansion of a highly large banks, which would result in higher bank concentration (expected to increases loan rate) but also lower firms' willingness to borrow from banks (expected to decrease loan rate).

We therefore instrument for local bank concentration, creating an instrumental variable which leverages differential local branch-level exposure to large national bank growth, in a strategy which builds on the shift-share 'Bartik' approach. Our strategy is based on the facts that (a) increases in local bank concentration are often driven by individual large bank growing, (b) these banks usually operate across many counties, (c) local banking sectors are differentially exposed to different large banks, and (d) the deposit growth of these large banks nationally is likely orthogonal to economic conditions in a specific local counties.

I denote HHI for bank j that has a branch k in county c and time t as

$$HHI_{j,c,t} = \sum_k \sigma^2_{j,k,c,t} = \sum_k \sigma^2_{j,k,c,t-1} \frac{(1+g_{j,k,c,t})^2}{(1+g_{j,c,t})^2}$$

where $\sigma_{j,k,c,t} = \frac{d_{j,k,c,t}}{\sum_{k=1}^N d_{j,k,c,t}}$, the deposit $d_{j,k,c,t}$ of branch k in county c as a share of total deposit $\sum_{k=1}^N d_{j,k,c,t}$ in county c. $g_{j,k,c,t}$ denotes the growth rate of deposit for branch k, while $g_{c,j,t}$ denotes the growth rate of overall deposit in county c.

Following Schubert et al. (2024), I instrument for the deposit growth for bank j in county c with the national deposit growth of that bank j , leaving out the county c, which I denote as $\tilde{g}_{j,t}$. The instrument for HHI, $Z_{j,c,t}$, is therefore

$$Z_{j,c,t} = \sum_k \sigma^2_{j,k,c,t-1} \frac{(1+\tilde{g}_{j,t})^2}{(1+\tilde{g}_{j,c,t})^2} \tag{22}$$

where $\tilde{g}_{j,c,t} = \sigma_{j,k,c,t-1}\tilde{g}_{j,t}$ is the predicted local growth rate in deposit, as predicted from the national (leave-one-out) growth by each bank k. This instrument features plausibly exogenous 'shocks' (a function of banks' national deposit growth), and possibly endogenous exposure 'shares' (the last-period local deposit shares of each of those banks).[17]

In Table 7, I re-estimate Equation (21) to examine the non-linear relationship between the HHI and loan rates, confirming a consistent pattern. Specifically, the effect of bank concentration on loan rates is negative at the 1% significance level when the HHI lies in its 7th decile. When the HHI increases from 0.6 to 0.7, the loan rate decreases by 0.11% in column 3, where

[17] The shift-share IV has a significantly positive effect on HHI, as the first stage regression is significant.

bank, time, and state-year fixed effects are controlled.

5.1.3 Bank Concentration and Bank Capital

Consistent with the model predictions and the findings in Table 5, the relationship between bank concentration and loan rates exhibits a non-monotonic pattern due to non-binding capital constraints in highly concentrated banking sectors. Specifically, bankers accumulate more capital when the banking sector is more concentrated. To empirically verify these model predictions, the following regression is estimated:

$$\text{Bank Capital}_{j(k)t} = \beta_1 HHI_{c(k)t-1} + \alpha_{j(k)} + \alpha_t + \alpha_{s(k)t} + \epsilon_{jt} \tag{23}$$

where Bank Capital$_{i(k)t}$ is the Tier 1 risk based capital ratio for bank j at quarter t, $\alpha_{j(k)}$ denotes the fixed effect associated with branch k belonging to institution j, α_t represents the quarter fixed effect, $\alpha_{s(k)t}$ accounts for the state-year fixed effect, and $HHI_{c(k)t-1}$ signifies the HHI for $c(k)$ at time $t-1$. The main coefficient of interest, β_1, captures the effect of bank concentration on bank capital. A shift-share IV is used to control for endogeneity, and standard errors are clustered at the bank level.

Table 8 presents the estimation results, revealing a significantly positive effect of bank concentration on bank capital. This indicates that bankers hold more capital in more concentrated banking sectors. Specifically, in column 6, when HHI increases from 0 to 1, the Tier 1 risk-based capital ratio rises by 3.2%. It should be noted, however, that this result is not precisely consistent with the model predictions, as it is challenging to determine whether bank capital constraints are binding in the data.

5.2 Policy Implications

Based on the model, the bank capital constraint plays a central role in affecting the dynamics between bank concentration, entrepreneurial default risks, and allocative efficiency. This section explores the policy implications of promoting efficiency while simultaneously minimizing risk.

A key policy question is whether reducing barriers to bank competition alone would be sufficient to enhance efficiency in a low-risk environment. When the bank capital constraint is binding, decreasing bank concentration can improve efficiency with minimal impact on entrepreneurial default risks due to the offsetting effects of the *risk-shifting mechanism* and *net margin mechanism*. However, in cases where banks hold capital above the minimum requirement, a trade-off between efficiency and defaults risks arises. To achieve greater efficiency without increasing risk, policies should target both reduced bank concentration and higher minimum capital requirements. Higher bank capital not only broadens the range where the capital constraint is binding and risk becomes less sensitive to bank concentration but also limits entrepreneurial risk by lowering loan rates.

Figure 5 illustrates that reducing bank concentration could lead to decreased efficiency and increased risk when the number of banks is between approximately 6 and 8. Thus, an immediate reduction in bank concentration may not always produce the desired outcomes, as local maximum efficiency occurs around 7 banks. Policymakers should therefore be patient in lowering barriers to bank competition, even if this entails a temporary welfare loss. This finding is particularly relevant, given that the calibrated number of banks in the U.S., based on the HHI, is 7.45.

5.3 Exogenous Variation of Bank Concentration

In the baseline model, the effect of bank concentration, represented by the exogenous variable M, on entrepreneurial default risks (rc) depends on whether the bank capital constraint is binding. However, in practice, bank concentration is endogenously influenced by factors such as switching costs, regulations, mergers and acquisitions, and broader market conditions.

This section addresses the endogeneity of the number of bankers, M, by allowing entry costs to vary. Specifically, entry into the banking sector requires a fixed payment, denoted by τ. In the steady state of a symmetric equilibrium, the free entry condition can be expressed as

$$\frac{1}{1-\beta} c_{it}^{b} = \tau,$$

which equates the lifetime utility derived from consumption with the entry cost. To understand how entry costs influence bank concentration, it is useful to examine the relationship between bankers' consumption and the implied number of bankers. As shown in Figure 11, there is a negative correlation between bankers' consumption and the number of bankers. Consequently, as entry costs increase, fewer bankers enter the industry, resulting in a more concentrated banking sector.

Given the monotonic relationship between bankers' consumption and the number of bankers, extending the model to incorporate an endogenous M closely parallels the baseline model. Thus, in the baseline model, the number of bankers M can be viewed as an exogenous representation of bank concentration that corresponds to varying entry costs.

6 Conclusion

In this paper, I develop a tractable dynamic model to analyze the impact of bank concentration and capital regulation on entrepreneurial default risks and allocative efficiency. The model predicts that in highly concentrated banking sectors with non-binding capital constraints, greater concentration reduces default risks and generates an inverse U-shaped relationship between concentration and efficiency. In contrast, in less concentrated markets where capital requirements are binding, lower concentration improves efficiency with minimal effects on default risks. I identify two key mechanisms—*risk-shifting mechanism* and *net margin mechanism*—that drive these outcomes. These findings highlight the importance of incorporating bank capital levels in future empirical research on the relationship between bank concentration, risk, and efficiency.

This paper focuses on idiosyncratic risks, which are often associated with financial fragility. However, introducing aggregate risk into the model could provide a more comprehensive analysis of financial distress, bank runs, and similar phenomena. To maintain tractability, the model is expressed in real terms. With its rich heterogeneity, this framework could also be valuable for exploring monetary policy effects in the presence of price rigidity. I leave these extensions for future research.

Tables

Assets	Liabilities
Safe loans $(1 - v_t)L_{it}$	Deposits D_{it}
Risky loans $v_t L_{it}$	Equity capital N_{it}

Table 1: Banker's Balance Sheet

Parameters	Values	Description
β	0.96	Risk-free interest rate*
λ	15	Bank capital to asset ratio*
M	7.45	Average HHI between 1994-2020*
z_{max}	3	Hsieh and Klenow (2009)*
z_{min}	1	Normalized to 1
γ	1.5	Markup of 20%*
s	0.98	saving rate
κ	0.08	Basel III regulations*
α	1.05	Excess return of gambling project*
p	0.7	Success probability of gambling project
$\overline{K}$	1	Normalized to 1

Table 2: Calibrated Parameter Values

Notes: * indicates that the parameter is calculated to match moments from data

Variable	Mean	SD	Median
County-level HHI	0.110	0.127	0.070
Top 4 Deposit Share	0.450	0.225	0.417
Branch-level loan rate	4.27%	0.013	4%

Table 3: Summary Statistics. Source: FDIC and RateWatch, 2010Q1-2021Q1.

Variable	Mean	SD	Median
Bank-level HHI	0.146	0.129	0.114
Tier 1 capital to risk weighted asset ratio	17.2%	0.118	14.2%
Total capital to risk weighted asset ratio	18.3%	0.117	15.3%

Table 4: Summary Statistics. Source: Call Report, 2010Q1-2021Q1.

VARIABLES	(1) OLS	(2) OLS	(3) OLS
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0, 0.1]$)	1.708***	0.768***	0.858***
	[0.238]	[0.186]	[0.196]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.1, 0.2]$)	0.709***	0.314***	0.310***
	[0.122]	[0.096]	[0.090]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.2, 0.3]$)	0.673***	0.420***	0.454***
	[0.086]	[0.087]	[0.078]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.3, 0.4]$)	0.430***	0.202**	0.221***
	[0.078]	[0.085]	[0.079]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.4, 0.5]$)	0.432***	0.299***	0.329***
	[0.084]	[0.095]	[0.100]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.5, 0.6]$)	0.134***	0.080	0.052
	[0.050]	[0.052]	[0.048]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.6, 0.7]$)	-0.304***	-0.236***	-0.235***
	[0.086]	[0.062]	[0.056]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.7, 0.8]$)	0.079	0.003	-0.027
	[0.052]	[0.051]	[0.070]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.8, 0.9]$)	0.457***	0.321***	0.449***
	[0.085]	[0.073]	[0.064]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.9, 1]$)	0.677***	0.544***	0.567***
	[0.141]	[0.099]	[0.162]
Observations	133,015	133,015	133,015
R-squared	0.669	0.678	0.684
Bank Fixed Effect	Yes	Yes	Yes
Institution Type Fixed Effect	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes
State-year Fixed Effect	No	No	Yes
State Fixed Effect	No	Yes	No

Table 5: Bank Concentration and Loan Rate (Auto Loan) with OLS

Notes: Table 5 shows the relationship between branch-level HHI and loan rate (Auto 6 years) within different deciles of HHI. The data is at the branch-quarter level and cover from 2010Q1 to 2021Q1. Rows 1-10 show the coefficients on the interaction term between HHI and the indicator of HHI within different deciles. The 10 coefficients reflect the heterogeneous effect of HHI on the loan rate within different deciles. From top to bottom, the coefficients are positive, negative, and then positive again, which indicates a non-monotonic relationship between bank concentration and the loan rate. The standard errors are clustered at the bank level. Compared to column 1, I additionally control for the state fixed effect in the second column and the state-time fixed effect in the third column. *** indicates significance at the 1% level; ** indicates significance at the 5% level; * indicates significance at the 10% level.

VARIABLES	(1) OLS	(2) OLS	(3) OLS
Branch-HHI	1.716***	1.027***	1.045***
	[0.238]	[0.156]	[0.185]
Branch-HHI^2	-5.588***	-3.409***	-3.499***
	[0.996]	[0.748]	[0.855]
Branch-HHI^3	4.621***	2.920***	3.062***
	[0.944]	[0.734]	[0.851]
Observations	133,015	133,015	133,015
R-squared	0.669	0.678	0.684
Bank Fixed Effect	Yes	Yes	Yes
Institution Type Fixed Effect	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes
State-year Fixed Effect	No	No	Yes
State Fixed Effect	No	Yes	No

Table 6: Bank Concentration and Loan Rate (Auto Loan) with third Polynomial

Notes: This table shows the non-linear relationship between branch-level HHI and loan rate (Auto 6 years) using the third order polynomial. The data is at the branch-quarter level and cover from 2010Q1 to 2021Q1. The standard errors are clustered at the bank level. Compared to column 1, I additionally control for the state fixed effect in the second column and the state-time fixed effect in the third column. *** indicates significance at the 1% level; ** indicates significance at the 5% level; * indicates significance at the 10% level.

VARIABLES	(1) IV	(2) IV	(3) IV
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0, 0.1]$)	2.811***	2.292***	2.280***
	[0.247]	[0.332]	[0.285]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.1, 0.2]$)	0.903***	0.686***	0.672***
	[0.127]	[0.113]	[0.141]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.2, 0.3]$)	0.667***	0.585***	0.625***
	[0.103]	[0.110]	[0.097]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.3, 0.4]$)	0.423***	0.345**	0.281*
	[0.137]	[0.156]	[0.146]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.4, 0.5]$)	1.062**	1.001*	1.367**
	[0.541]	[0.559]	[0.591]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.5, 0.6]$)	0.506**	0.444**	0.330
	[0.212]	[0.200]	[0.212]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.6, 0.7]$)	-1.266***	-1.015**	-1.115**
	[0.489]	[0.440]	[0.485]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.7, 0.8]$)	0.603	0.581	0.863
	[0.615]	[0.406]	[0.810]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.8, 0.9]$)	1.250	1.222	0.837
	[1.088]	[1.160]	[1.543]
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.9, 1]$)	0.571**	0.505*	0.439
	[0.249]	[0.297]	[0.477]
Observations	133,015	133,015	133,015
Bank Fixed Effect	Yes	Yes	Yes
Institution Type Fixed Effect	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes
State-year Fixed Effect	No	No	Yes
State Fixed Effect	No	Yes	No

Table 7: Bank Concentration and Loan Rate (Auto Loan) with IV Regression

Notes: This table shows the relationship between branch-level HHI and loan rate (Auto 6 years) within different deciles of HHI, using the shift share instrument a IV regression. The data is at the branch-quarter level and cover from 2010Q1 to 2021Q1. Rows 1-10 show the coefficients on the interaction term between HHI and the indicator of HHI within different deciles. The 10 coefficients reflect the heterogeneous effect of HHI on the loan rate within different deciles. From top to bottom, the coefficients are positive, negative, and then positive again, which indicates a non-monotonic relationship between bank concentration and the loan rate. The standard errors are clustered at the bank level. Compared to column 1, I additionally control for the state fixed effect in the second column and the state-time fixed effect in the third column. *** indicates significance at the 1% level; ** indicates significance at the 5% level; * indicates significance at the 10% level.

VARIABLES	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) IV	(6) IV
Branch-HHI	1.603	2.330*	2.377*	2.523	3.156	3.223*
	[1.016]	[1.248]	[1.268]	[1.761]	[1.980]	[1.827]
Observations	57,334	57,334	57,334	57,334	57,334	57,334
Bank Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Institution Type Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
State-year Fixed Effect	No	No	Yes	No	No	Yes
State Fixed Effect	No	Yes	No	No	Yes	No

Table 8: The Effect of Bank Concentration on Bank Capital

Notes: This table shows the relationship between branch-level HHI and bank capital. The data is at the branch-quarter level and cover from 2010Q1 to 2021Q1. The standard errors are clustered at the bank level. *** indicates significance at the 1% level; ** indicates significance at the 5% level; * indicates significance at the 10% level.

Figures

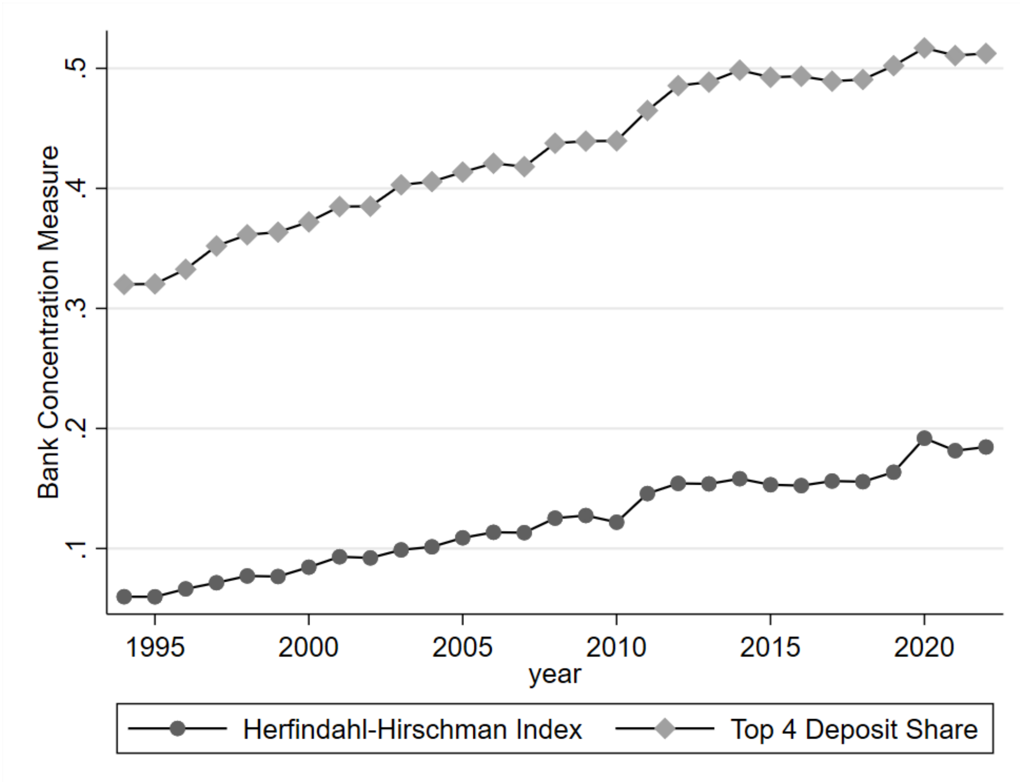


Figure 1: Trend of HHI and Top 4 Deposit Share

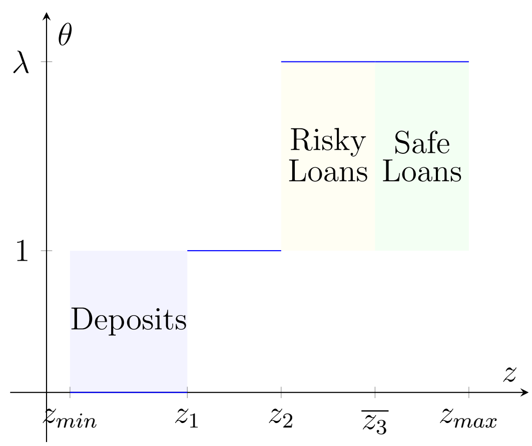
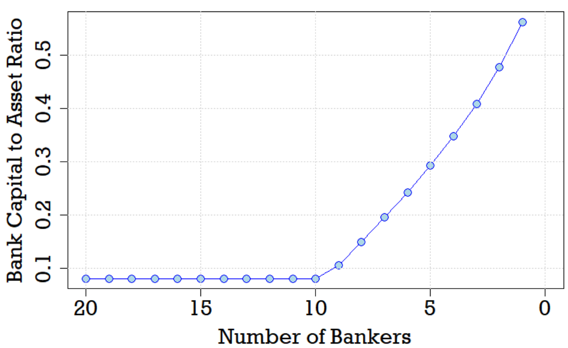
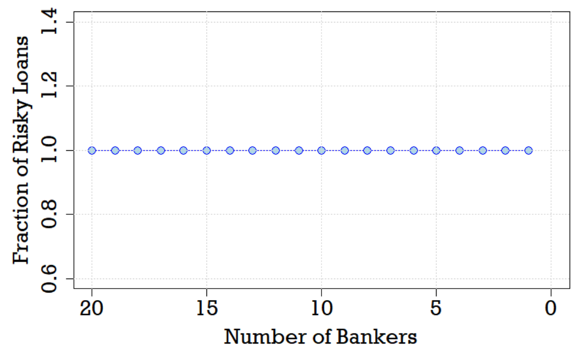


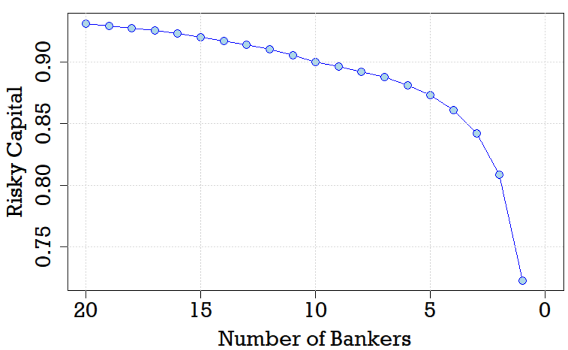
Figure 2: Entrepreneurs' Leverage and Project Choice



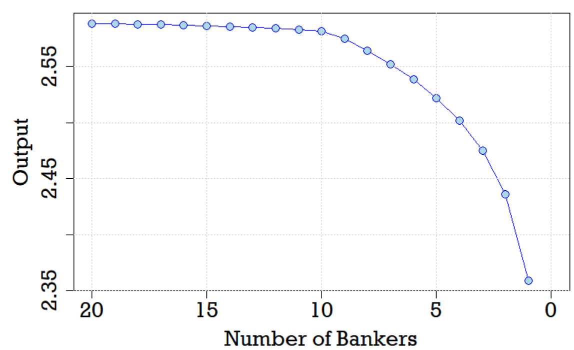
(a) Bank Capital to Asset Ratio



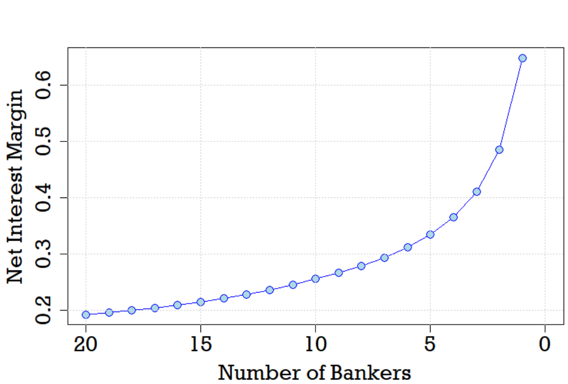
(b) Fraction of Risky Loan



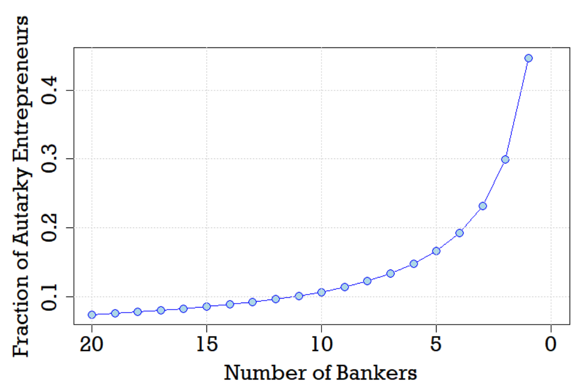
(c) Risky Capital



(d) Output



(e) Net Margin



(f) Fraction of Autarky Entrepreneurs

Figure 3: Comparative Statics of Model with only Risky Loan

Notes: This plot presents the relationship between bank concentration (number of bankers) and endogenous variables: bank capital to asset ratio, fraction of risky loan, risky capital, output, net margin, and fraction of autarky entrepreneurs, when all loans are risky. I focus on the comparative statics in the steady state.

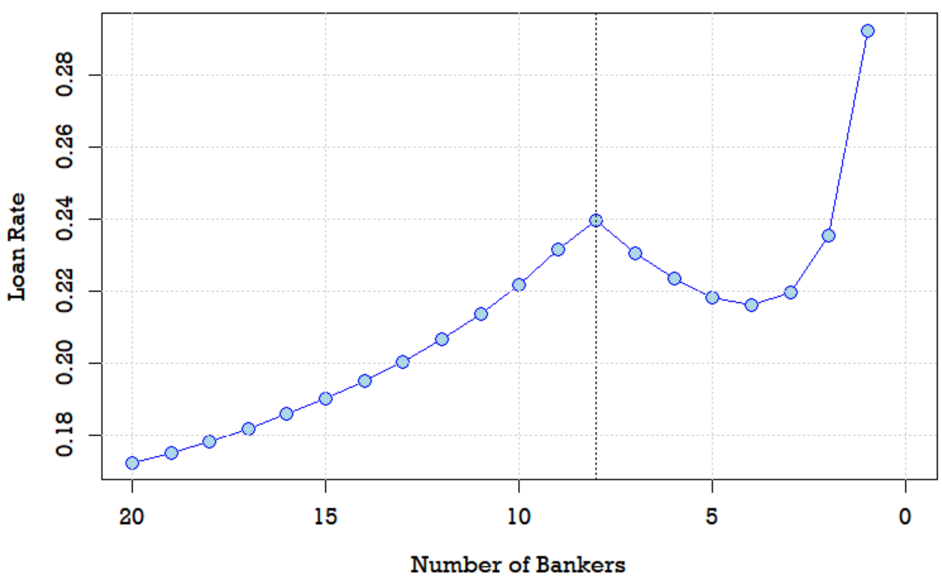
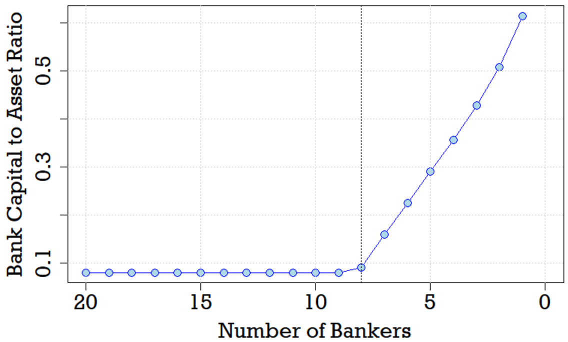
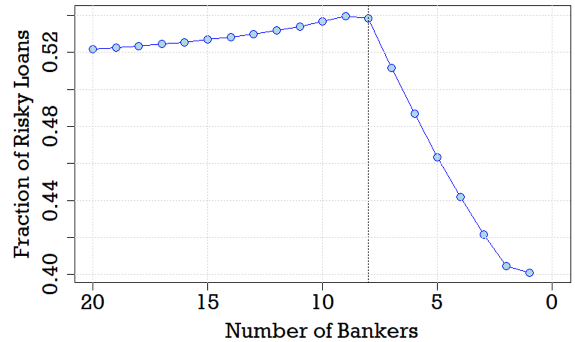


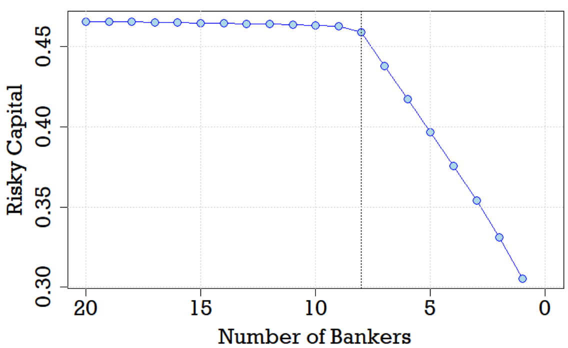
Figure 4: Bank Concentration (Number of Bankers) and Loan Rate



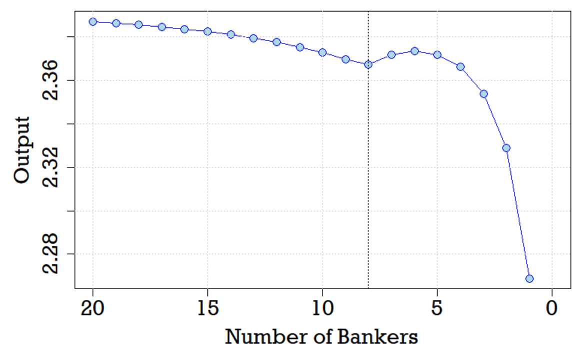
(a) Bank Capital to Asset Ratio



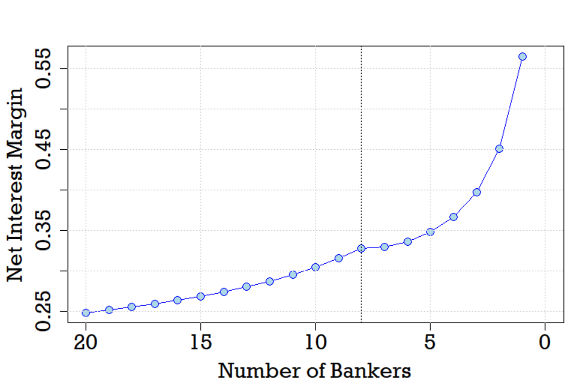
(b) Fraction of Risky Loan



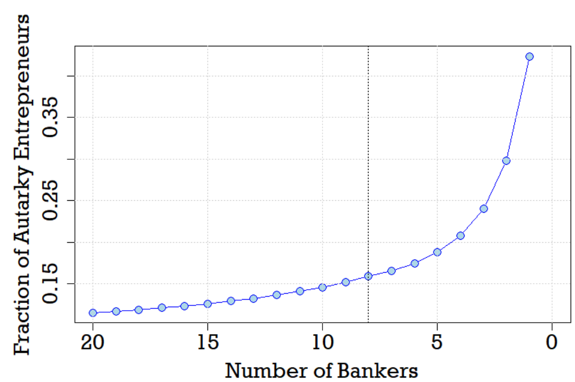
(c) Risky Capital



(d) Output



(e) Net Interest Margin



(f) Fraction of Autarky Entrepreneurs

Figure 5: Comparative Statics of Model with Both Safe and Risky Loan

Notes: This plot presents the relationship between bank competition (number of bankers) and endogenous variables: bank capital to asset ratio, fraction of risky loan, risky capital, output, net margin, and fraction of autarky entrepreneurs, when loans are either risky or safe. The vertial dashed line determines whether the capital requirement is binding. I focus on the comparative statics in the steady state.

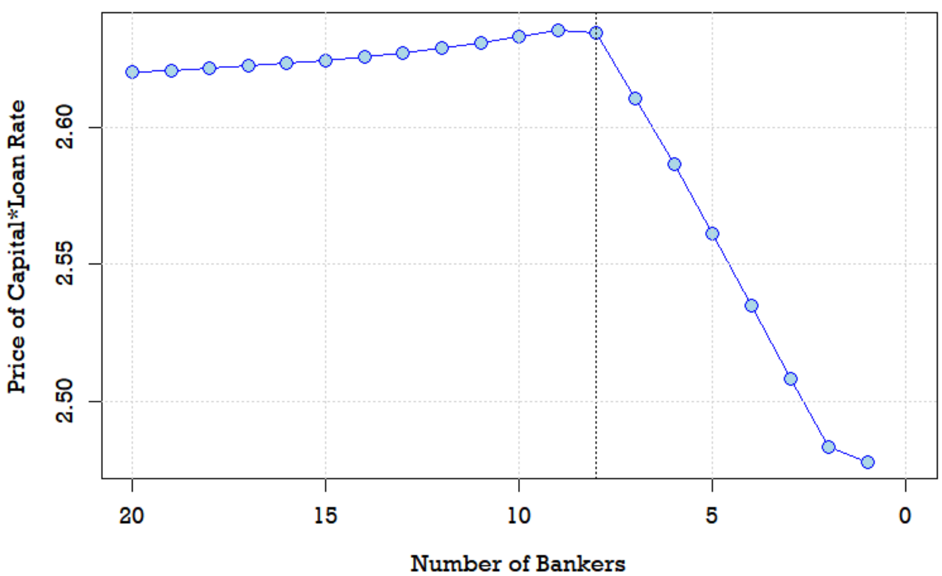


Figure 6: Bank Concentration (Number of Bankers) and Effective Loan Rate

Notes: This figure shows the effective return rate on loans $q(1 + r^l)$ in the general equilibrium, as a function of the number of banks. $q(1 + r^l)$ shapes exactly the same as the fraction of risky loan v.

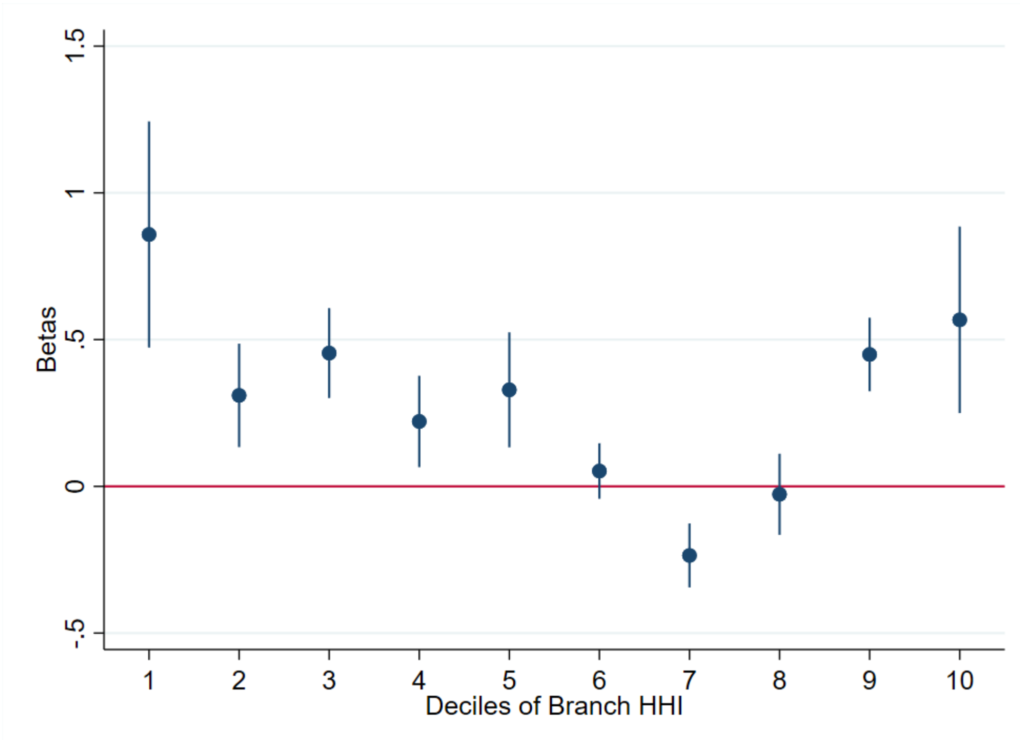


Figure 7: Branch-level HHI and Loan Rate (Auto Loan)

Notes: Figure 7 illustrates the relationship between HHI and loan rates across different levels of HHI. The dataset, spanning from 2010 to 2021, is sourced from RateWatch and FDIC. Bank and time fixed effects are controlled for in this analysis. The X-axis represents the ordinal deciles of branch-level HHI, while the Y-axis depicts the coefficients of interaction between HHI and the indicator of HHI being in different deciles (β_is in Equation 21). The figure displays pointwise estimates along with 5% confidence intervals. Notably, when HHI is extremely low or high, the pointwise estimate is significantly positive. However, when HHI falls within the 7th decile, the pointwise estimate turns significantly negative. Additional specifications will be detailed in Table 5.

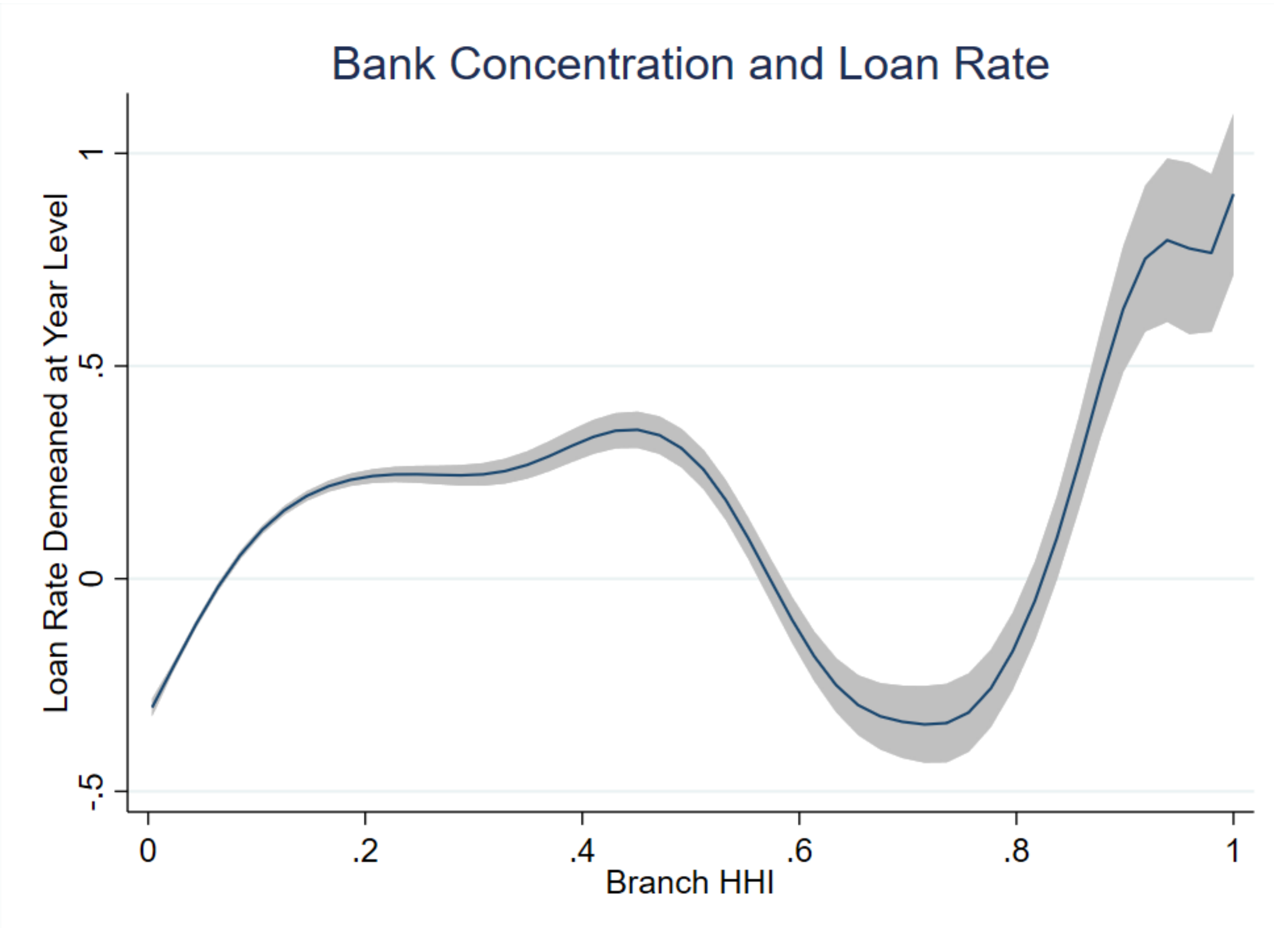


Figure 8: Local Polynomial Smoothing between Bank Concentration and Loan Rate (Auto Loan)

Notes: This figure shows the non-monotonic relationship between branch-level HHI and loan rate (Auto 72 loan). The data is at the branch-quarter level and cover from March 2010 to March 2021. The loan rate is demeaned at year level. A local polynomial smoothing is conducted between the demeaned loan rate and branch-level HHI. The shaded area represents the 95% confidence interval.

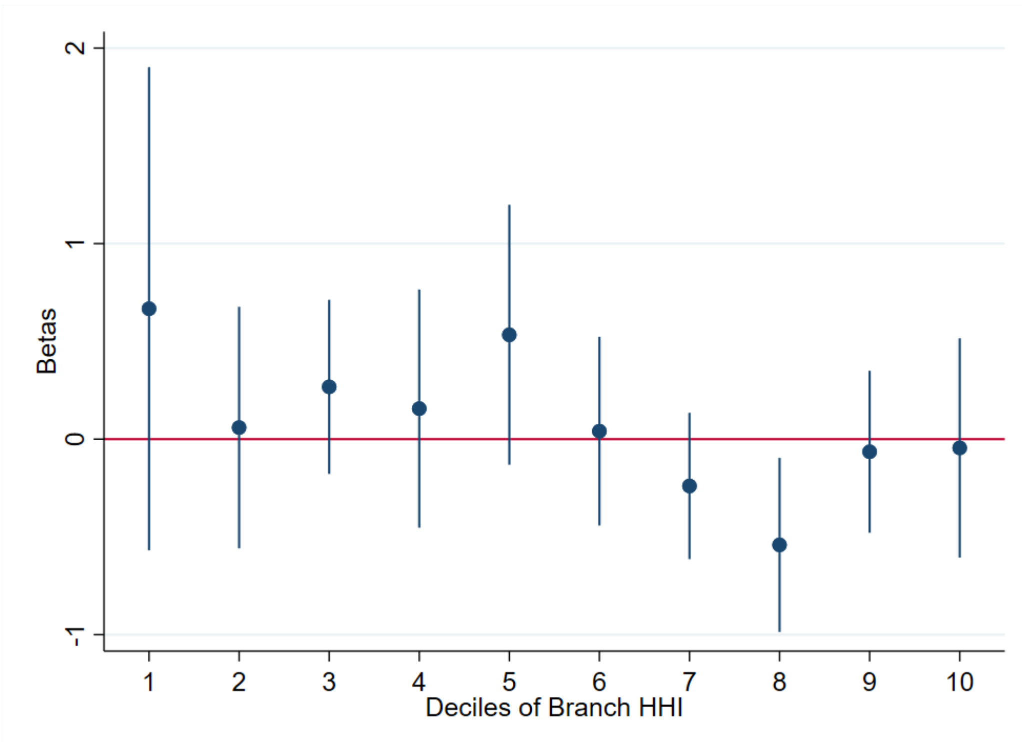


Figure 9: Branch-level HHI and Loan Rate (Auto Loan, Robustness)

Notes: Figure 7 illustrates the relationship between HHI and loan rates across different levels of HHI. The dataset, spanning from 2010 to 2021, is sourced from RateWatch and FDIC. I delete the observations with bank ID (rssdid) is labeled as 0. Bank, state-year, and time fixed effects are controlled for in this analysis. The X-axis represents the ordinal deciles of branch-level HHI, while the Y-axis depicts the coefficients of interaction between HHI and the indicator of HHI being in different deciles (β_is in Equation 21). The figure displays pointwise estimates along with 5% confidence intervals. Notably, when HHI is extremely low or high, the pointwise estimate is positive, although with large noise. However, when HHI falls within the middle decile, the pointwise estimate turns significantly negative.

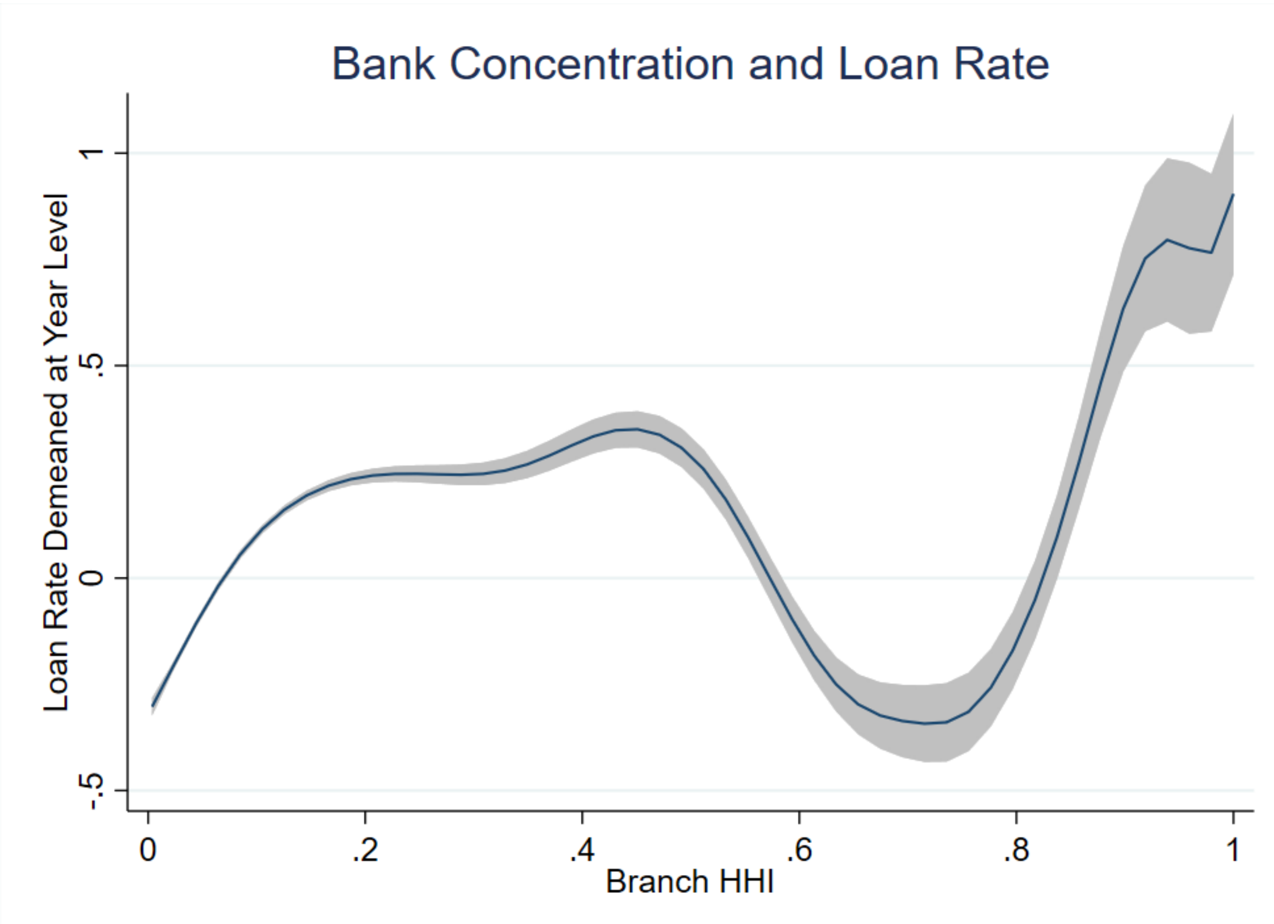


Figure 10: Local Polynomial Smoothing between Bank Concentration and Loan Rate (Auto Loan, Robustness)

Notes: This figure shows the non-monotonic relationship between branch-level HHI and loan rate (Auto 72 loan). The data is at the branch-quarter level and cover from March 2010 to March 2021. I delete the observations with bank ID (rssdid) is labeled as 0. The loan rate is demeaned at year level. A local polynomial smoothing is conducted between the demeaned loan rate and branch-level HHI. The shaded area represents the 95% confidence interval.

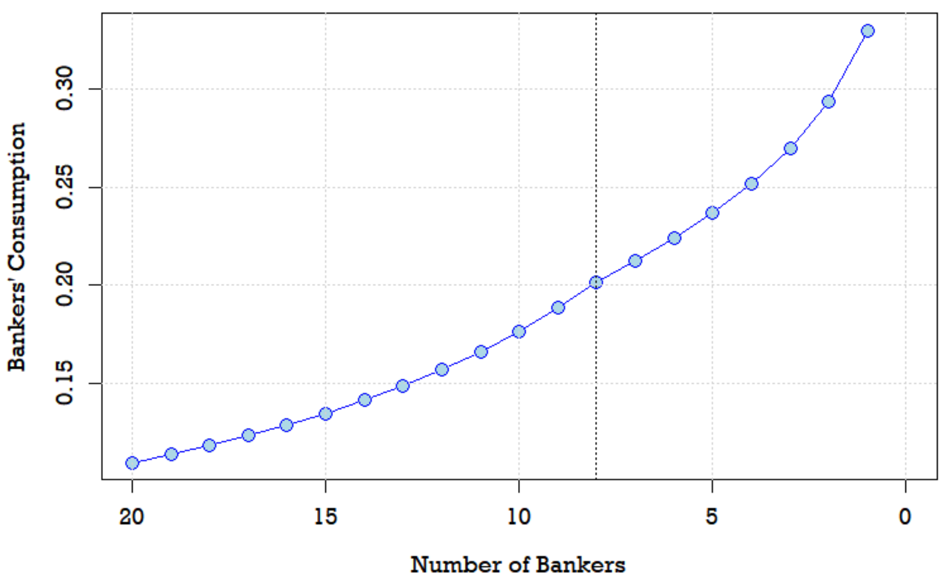


Figure 11: Bank Concentration (Number of Bankers) and Bankers' Consumption

References

Adão, R., Kolesár, M., and Morales, E. (2019). Shift-share designs: Theory and inference*. *The Quarterly Journal of Economics*, 134(4):1949–2010.

Agoraki, M.-E. K., Delis, M. D., and Pasiouras, F. (2011). Regulations, competition and bank risk-taking in transition countries. *Journal of Financial Stability*, 7(1):38–48.

Alfon, I., Argimon, I., and Bascuñana-Ambrós, P. (2004). *What Determines How Much Capital is Held by UK Banks and Building Societies?* Financial Services Authority London.

Angeletos, G.-M. (2007). Uninsured idiosyncratic investment risk and aggregate saving. *Review of Economic dynamics*, 10(1):1–30.

Asriyan, V., Laeven, L., and Martín, A. (2021). Collateral Booms and Information Depletion. *The Review of Economic Studies*, 89(2):517–555.

Beck, T., De Jonghe, O., and Schepens, G. (2013). Bank competition and stability: Cross-country heterogeneity. *Journal of financial Intermediation*, 22(2):218–244.

Beck, T., Demirguc-Kunt, A., and Levine, R. (2003). Bank concentration and crises.

Beck, T., Demirgüç-Kunt, A., and Levine, R. (2006). Bank concentration, competition, and crises: First results. *Journal of Banking & Finance*, 30(5):1581–1603.

Black, S. E. and Strahan, P. E. (2002). Entrepreneurship and bank credit availability. *The Journal of Finance*, 57(6):2807–2833.

Borusyak, K., Hull, P., and Jaravel, X. (2021). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, 89(1):181–213.

Boyd, J. H. and De Nicolo, G. (2005). The theory of bank risk taking and competition revisited. *The Journal of Finance*, 60(3):1329–1343.

Brunnermeier, M. K. and Koby, Y. (2018). The reversal interest rate. Technical report, National Bureau of Economic Research.

Carlson, M., Correia, S., and Luck, S. (2022). The effects of banking competition on growth and financial stability: Evidence from the bational banking era. *Journal of Political Economy*, 130(2):462–520.

Carlson, M. and Mitchener, K. J. (2009). Branch banking as a device for discipline: Competition and bank survivorship during the great depression. *Journal of Political Economy*, 117(2):165–210.

Cetorelli, N. and Gambera, M. (2001). Banking market structure, financial dependence and growth: International evidence from industry data. *The Journal of Finance*, 56(2):617–648.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of political Economy*, 113(1):1–45.

Corbae, D. and D'Erasmo, P. (2020). Rising bank concentration. *Journal of Economic Dynamics and Control*, 115:103877. St. Louis Fed -JEDC-SCG-SNB-UniBern Conference, titled "Disaggregate Data and Macroeconomic Models".

Corbae, D. and Levine, R. (2018). Competition, stability, and efficiency in financial markets. *Unpublished manuscript*.

Craig, B. R. and Dinger, V. (2013). Deposit market competition, wholesale funding, and bank risk. *Journal of Banking & Finance*, 37(9):3605–3622.

De Nicolò, G., Bartholomew, P., Zaman, J., and Zephirin, M. (2004). Bank consolidation, internationalization, and conglomeration: Trends and implications for financial risk. *Internationalization, and Conglomeration: Trends and Implications for Financial Risk*.

Diez, M. F., Leigh, M. D., and Tambunlertchai, S. (2018). *Global Market Power and its Macroeconomic Implications*. International Monetary Fund.

Drechsler, I., Savov, A., and Schnabl, P. (2017). The Deposits Channel of Monetary Policy. *The Quarterly Journal of Economics*, 132(4):1819–1876.

Flannery, M. J. and Rangan, K. P. (2008). What caused the bank capital build-up of the 1990s? *Review of Finance*, 12(2):391–429.

Gali, J. and Monacelli, T. (2005). Monetary policy and exchange rate volatility in a small open economy. *The Review of Economic Studies*, 72(3):707–734.

Gentry, W. M. and Hubbard, R. G. (2000). Entrepreneurship and household saving.

Head, A. C., Kam, T., Ng, S., and Pan, G. (2022). Money, credit and imperfect competition among banks. *Working Paper*.

Hellmann, T. F., Murdock, K. C., and Stiglitz, J. E. (2000). Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough? *American Economic Review*, 90(1):147–165.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Jayaratne, J. and Strahan, P. E. (1996). The finance-growth nexus: Evidence from bank branch deregulation. *The Quarterly Journal of Economics*, 111(3):639–670.

Jiang, L., Levine, R., and Lin, C. (2017). Does competition affect bank risk? Technical report, National Bureau of Economic Research.

Jiang, L., Levine, R., and Lin, C. (2023). Does competition affect bank risk? *Journal of Money, Credit and Banking*, 55(5):1043–1076.

Joaquim, G., van Doornik, B. F. N., Ornelas, J. R., et al. (2019). *Bank Competition, Cost of Credit and Economic Activity: Evidence from Brazil*. Banco Central do Brasil.

Kiyotaki, N. and Moore, J. (2019). Liquidity, business cycles, and monetary policy. *Journal of Political Economy*, 127(6):2926–2966.

Lagos, R. and Zhang, S. (2022). The limits of onetary economics: On money as a constraint on market power. *Econometrica*, 90(3):1177–1204.

Li, J. (2019). Imperfect banking competition and financial stability. Technical report, Technical report, Working Paper.

Li, Z. and Song, J. (2023). Deposit market power and bank risk-taking. *Working Paper*.

Liu, Z. and Wang, P. (2014). Credit constraints and self-fulfilling business cycles. *American Economic Journal: Macroeconomics*, 6(1):32–69.

Martinez-Miera, D. and Repullo, R. (2010). Does competition reduce the risk of bank failure? *The Review of Financial Studies*, 23(10):3638–3664.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Moll, B. (2014). Productivity losses from financial frictions: Can self-financing undo capital misallocation? *American Economic Review*, 104(10):3186–3221.

Petersen, M. A. and Rajan, R. G. (1995). The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics*, 110(2):407–443.

Repullo, R. (2004). Capital requirements, market power, and risk-taking in banking. *Journal of Financial Intermediation*, 13(2):156–182.

Rice, T. and Strahan, P. E. (2010). Does credit competition affect small-firm finance? *The Journal of Finance*, 65(3):861–889.

Schubert, G., Stansbury, A., and Taska, B. (2024). Employer concentration and outside options. *working Paper*.

Tabak, B. M., Fazio, D. M., and Cajueiro, D. O. (2012). The relationship between banking market competition and risk-taking: Do size and capitalization matter? *Journal of Banking & Finance*, 36(12):3366–3381.

Van Hoose, D. et al. (2010). *the Industrial Organization of Banking*. Springer.

Wang, O. (2024). Banks, low interest rates, and monetary policy transmission. *The Journal of Finance, forthcoming*.

Yi, Y. (2022). Bank concentration, bank capital and misallocation. *Working Paper*.

Appendices

A Proofs

Proof of Proposition 1. All the entrepreneurs are risk neutral and maximize their expected consumption today. Since the saving rate of entrepreneurs is exogenous given, consumption follows:

$$c_t = s_t \Pi_t$$

where Π_t denotes the net return of the generation t. The functional form of Π_t is different in the following 3 cases:

Case 1 *If the entrepreneurs choose to deposit part of their wealth ($k_t \leq a_t$), then*

$$\Pi_t = z_t k_t + q_t(r_t^d + 1)(a_t - k_t) = [z_t - q_t(r_t^d + 1)]k_t + q_t(r_t^d + 1)a_t \tag{24}$$

where q_t is the price of capital, r_t^d is the deposit rate and k_t is the capital that is used in production. Note that entrepreneurs who do not borrow will not invest in gambling projects. The reason for this is that they prefer projects with a higher expected return.

The above equation implies that k_t equals to 0 or a_t, which depends on whether the productivity is above $q_t(r_t^d + 1)$.

Case 2 *Suppose that the entrepreneur becomes a borrower and chooses the prudent project. Denote her leverage ratio as θ with $\theta \leq \lambda$, the net profit is then:*

$$\Pi_t = z_t \theta a_t - q_t(r_t^l + 1)(\theta - 1)a_t = [z_t - q_t(r_t^l + 1)]\theta a_t + q_t(r_t^l + 1)a_t \tag{25}$$

where r_t^l is the loan rate. Following the above equation, the value of θ equals to 1 or λ, which depends on whether the productivity is above $q_t(r_t^l + 1)$.

Case 3 *Suppose that the entrepreneur becomes a borrower while invests in the gambling project. Denote her leverage ratio as θ with $\theta \leq \lambda$, the net profit is then:*

$$\Pi_t = p\{\alpha z_t \theta a_t - q_t(r_t^l + 1)(\theta - 1)a_t\} = p\left\{[\alpha z_t - q_t(r_t^l + 1)]\theta a_t + q_t(r_t^l + 1)a_t\right\} \tag{26}$$

Following the above equation, the value of θ equals to 1 or λ, which depends on whether the productivity is above $\frac{q_t(r_t^l+1)}{\alpha}$. Since α is greater than 1, there is a region of productivity in which borrowing entrepreneurs might want to start a gambling project rather than a prudent one.

The remaining calculation is to identify the border of each case. If borrowing and gambling exists in the equilibrium, the benefit of doing so should dominate that of staying autarky, as well as borrowing and investing in the prudent project. The condition is derived in Equation (6) and (7) that:

$$\frac{(\lambda - 1)p}{\lambda \alpha p - 1} q(1 + r^l) = z_2 < z < z_3 = \frac{(\lambda - 1)(1 - p)}{\lambda(1 - \alpha p)} q(1 + r^l) \tag{27}$$

Further, $\frac{(\lambda-1)p}{\lambda\alpha p-1} > \frac{1}{\alpha}$ following Assumption 1. Therefore, under the condition implied by Equation (27), entrepreneurs will borrow up to the borrowing limits λ and invest in the gambling project.

By Assumption 3, $\frac{(\lambda-1)(1-p)}{\lambda(1-\alpha p)} > 1$ and entrepreneurs borrow and invest in the prudent project if and only if $z > z_3$. In an extreme when $z_3 > z_{max}$, there are no borrowing entrepreneurs who stay prudent in the equilibrium.

When $q(1+r^d) < z < z_2$, $k = a$, which means that entrepreneurs will use their internal finance to produce. When $z < q(1+r^d)$, $k = 0$, so that the entrepreneurs deposit all their money in banks. ■

Proof of Lemma 1. Equations (8) and (9) are directly obtained from Proposition 1, given that borrowing entrepreneurs borrow up to the borrowing limit and lending entrepreneurs deposit all their capital in banks.

For the lending entrepreneurs, their net return becomes:

$$\Pi_t = (r_t^d + 1)q_t a_t$$

For the borrowing entrepreneurs who invest in the prudent project, their net return becomes:

$$\Pi_t = \lambda(z_t - (r_t^l + 1)q_t)a_t + (r_t^l + 1)q_t a_t$$

For the borrowing entrepreneurs who gamble, their net return becomes:

$$\Pi_t = p\{\lambda(\alpha z_t - (r_t^l + 1)q_t)a_t + (r_t^l + 1)q_t a_t\}$$

For the autarky entrepreneurs, their net return becomes:

$$\Pi_t = z_t a_t$$

Given the constant saving rate, I have:

$$\begin{aligned} q_t a_{t+1} =& \beta\{\int_{z_{min}}^{z_{1t}} q_t(1+r_t^d)dG(z_t) + \int_{\overline{z}_{3t}}^{z_{max}} \left\{\lambda[(z_t - q_t(1+r_t^l)] + q_t(r_t^l+1)\right\} dG(z_t) \\ &+ \int_{z_{1t}}^{z_{2t}} z_t dG(z_t) + p\int_{z_{2t}}^{\overline{z}_{3t}} \left[\alpha\lambda z_t - (\lambda-1)q_t(r_t^l+1)\right] dG(z_t)\}a_t \end{aligned}$$

by simply aggregating all the entrepreneurs of different productivities. ■

Proof of Proposition 2. The Bellman equation for the banker i is:

$$V(N_{it}) = \max_{\{c_{it}^b, L_{it}, D_{it}\}} \{c_{it}^b + \beta V(N_{it+1})\}$$

subject to the balance sheet identity (2), the budget constraint (3) and the minimum capital requirement (5). The Lagrangian function for banker i becomes:

$$q_t\{(1+r_t^l)p_t^e L_{it} - (1+r_t^d)D_{it}\} - q_t N_{it+1} + q_t\mu_{it}(D_{it} + N_{it} - L_{it}) + q_t\chi_{it}(N_{it} - \kappa L_{it}) \quad (28)$$

by substituting the budget constraint into the utility function, where $q_t\mu_{it}$ is the multiplier of the bank's balance sheet identity. $q_t \chi_{it}$ is the multiplier of the bank capital constraint. Deriving the first order condition, I obtain Equations (11) and (12).

By definition, $v = \frac{G(z_3)-G(z_2)}{1-G(z_2)}$. I denote $\frac{(\lambda-1)p}{\lambda\alpha p-1}q = a_2$ and $\frac{(\lambda-1)(1-p)}{\lambda(1-\alpha p)}q = a_3$. Therefore:

$$\frac{\partial v}{\partial r^l} = \frac{[g(z_3)a_3 - g(z_2)a_2](1-G(z_2)) + (G(z_3)-G(z_2))g(z_2)a_2}{(1-G(z_2))^2} \tag{29}$$

The second element in the numerator is equivalent to $\{-[1-G(z_3)] + [1-G(z_2)]\}g(z_2)a_2$, so that Equation (29) becomes:

$$\begin{aligned}
\frac{\partial v}{\partial r^l} &= \frac{[g(z_3)a_3 - g(z_2)a_2](1-G(z_2)) + \{-[1-G(z_3)] + [1-G(z_2)]\}g(z_2)a_2}{(1-G(z_2))^2} \\
&= \frac{g(z_3)a_3(1-G(z_2)) - (1-G(z_3))g(z_2)a_2}{(1-G(z_2))^2} \\
&= \frac{1}{(1+r^l)(1-G(z_2))^2}(g(z_3)z_3(1-G(z_2)) - (1-G(z_3))g(z_2)z_2) \\
&= \frac{1}{g(z_3)g(z_2)z_3z_2(1+r^l)(1-G(z_2))^2}\left(\frac{(1-G(z_2))}{g(z_2)z_2} - \frac{(1-G(z_3))}{g(z_3)z_3}\right)
\end{aligned}$$

Since $\frac{zg(z)}{(1-G(z))}$ is increasing, $\frac{\partial v}{\partial r^l} \geq 0$. Further, p_e is a decreasing function of v so that $\frac{\partial p_e}{\partial r^l} \leq 0$.
■

B Robustness Checks with Business Loans

In this section, I will conduct regressions similar to those in Equation 21 and Equation 23 using secured business loans from the RateWatch dataset. The number of observations for secured business loans is 17,282, which is significantly smaller than the number of observations for auto loans. Given the limited dataset size, I estimate the following regression:

$$LoanRate_{kt} = \sum_{i=1}^{5} \beta_i HHI_{c(k)t} * \mathbb{1}(HHI_{c(k)t} \in (\frac{i-1}{5}, \frac{i}{5}]) + \alpha_{j(k)} + \alpha_t + \alpha_{s(k)t} + \epsilon_{jt}, \tag{30}$$

where the regression divides the entire sample into five equal parts and includes interaction terms between HHI and quintile indicators. Table 9 presents the results, showing that bank concentration has a positive and statistically significant impact on loan rates when branch-level HHI falls into the second or fifth quintile. Conversely, there is no significant association between bank concentration and the loan rate in other quintiles.

Based on the model predictions, the effect of bank concentration on loan rate is more likely to be significantly positive either when bank concentration is low or high. This model explains the positive correlation by considering the channel of the elasticity of loan demand, as well as the general equilibrium effect of the price of capital. Due to the *risk-shifting mechanism* in the model, however, the correlation between bank concentration and loan rate should be negative when the level of bank concentration is in between. The reason for not obtaining negative estimates in Table 9 might be that the dataset contains too much noise. There is a significant dispersion in the estimate due to the limited number of business loans. The correlation between bank concentration and the loan rate may be significantly negative if the quality of business loans is as good as that of auto loans.

Variables	(1) OLS	(2) OLS	(3) OLS
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0, 0.2]$)	-0.0142	-0.291	-0.256
	(0.649)	(0.698)	(0.669)
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.2, 0.4]$)	0.743*	0.657	0.734*
	(0.408)	(0.470)	(0.430)
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.4, 0.6]$)	0.216	0.173	0.121
	(0.361)	(0.373)	(0.433)
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.6, 0.8]$)	0.250	0.0852	-0.0002
	(0.181)	(0.295)	(0.391)
Branch-HHI*$\mathbb{1}$(Branch-HHI$\in (0.8, 1]$)	1.10***	1.12***	0.921**
	(0.363)	(0.399)	(0.424)
Constant	7.10***	7.13***	7.13***
	(0.0560)	(0.0626)	(0.0595)
Bank Fixed-effect	Yes	Yes	Yes
Quarter Fixed-effect	Yes	Yes	Yes
State Fixed-effect	No	Yes	No
State-time Fixed-effect	No	No	Yes
R-Squared	0.637	0.650	0.672
Observations	17,282	17,282	17,282

Table 9: Bank Concentration and Loan Rate (Business Loan)

Notes: Table 9 shows the relationship between branch-level HHI and loan rate (Secured Business Loan) within different quintiles of HHI. The data is at the branch-quarter level and cover from January 1994 to March 2021. Rows 1-5 show the coefficients on the interaction term between HHI and the indicator of HHI within different quintiles. The 5 coefficients reflect the heterogeneous effect of HHI on the loan rate within different quintiles. The standard errors are clustered at bank level. Compared to column 1, I additionally control for the state fixed effect in the second column and the state-time fixed effect in the third column. *** indicates significance at the 1% level; ** indicates significance at the 5% level; * indicates significance at the 10% level.

Figure B.1 shows the non-linear patterns between bank concentration and loan rate using the local polynomial smoothing.

C Evidence at Bank level

Using the FDIC and the Call Reports data, I examine the relationship between HHI and loan rate at the bank level. I calculate the loan rate by dividing the interest income over the loan size. What I do is running a local polynomial smoothing, and visualizing the non-linear correlation between the two variables in Figure C.1. There are four lines in each sub-figure, where the yellow line represents other personal loans, the green line represents commercial and industrial loans, the blue line represents the real estate loans, and the purple lines represents

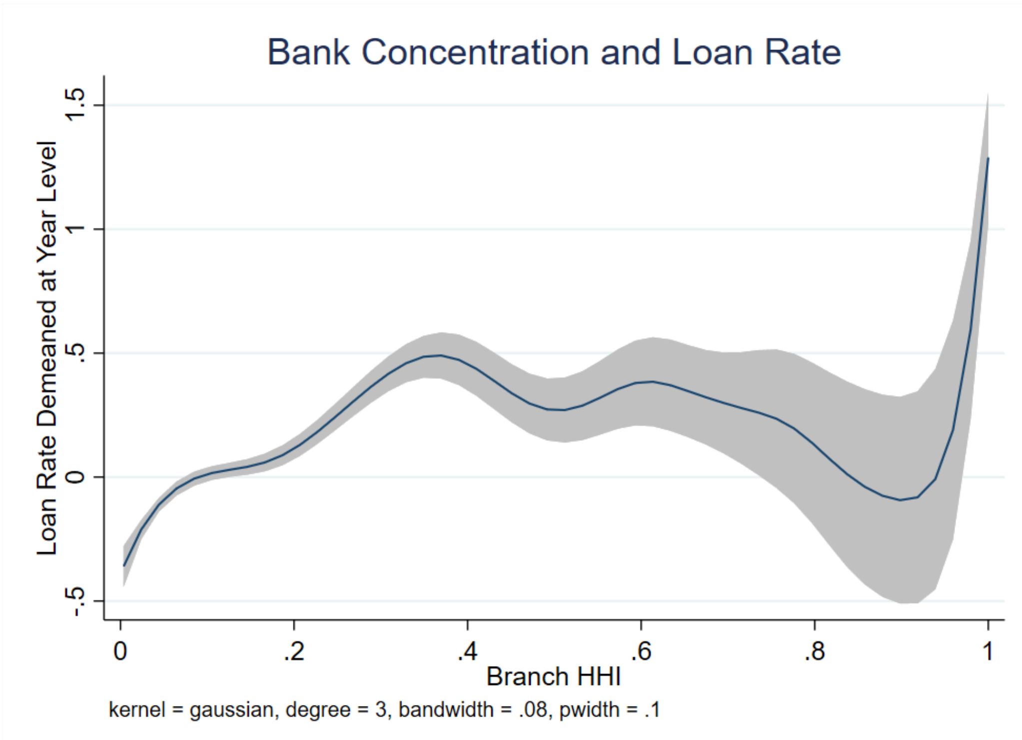
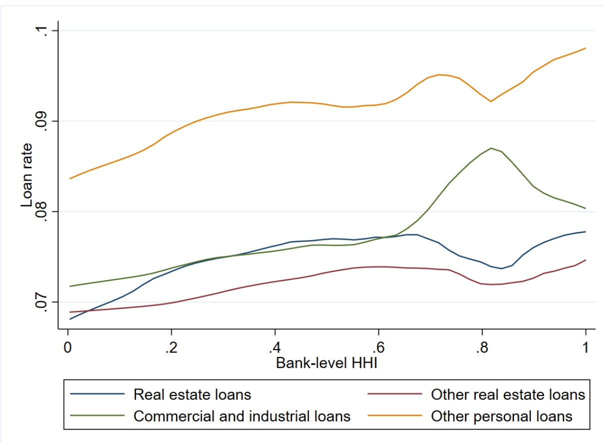


Figure B.1: Local Polynomial Smoothing between Bank Concentration and Loan Rate (Business Loan)

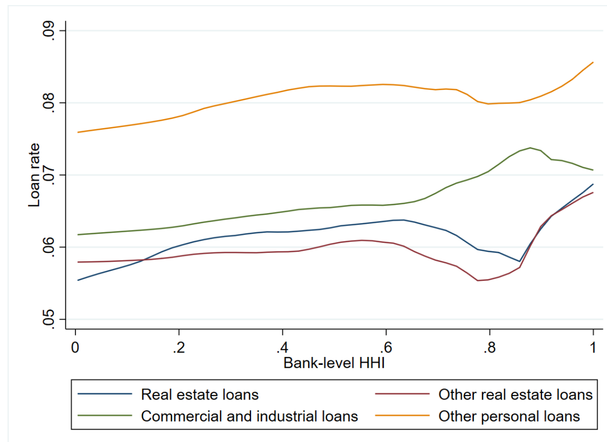
Notes: This figure shows the non-monotonic relationship between branch-level HHI and loan rate (business loan). The data is at the branch-quarter level and cover from January 1994 to March 2021. The loan rate is demeaned at year level. A local polynomial smoothing is conducted between the demeaned loan rate and branch-level HHI. The shaded area represents the 95% confidence interval.

other real estate loans. As illustrated in Figure C.2, these four loan types account for more than 80 percent of the total loan size.

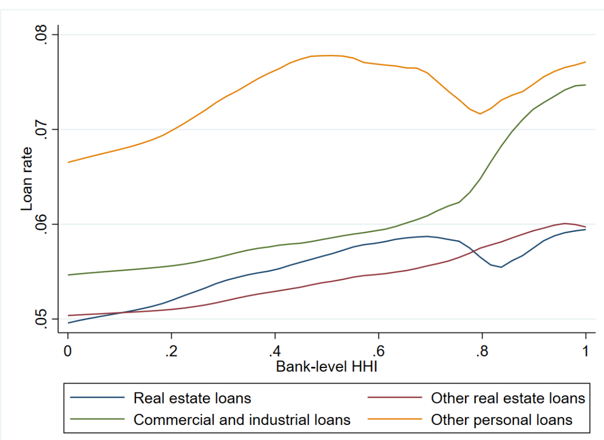
The four sub-figures capture the relationship between bank-level HHI and loan rate in years 2008, 2012, 2016, 2020, where I partially control for the time fixed effect. It is observed from the figure that the loan rate for personal loans is higher than for other loan types. Moreover, the correlation between bank-level HHI and loan rate is non-monotonic. When the bank concentration is high, there is a region where the correlation between bank concentration and loan rate is negative. The model predictions and branch-level evidence is consistent with this non-monotonicity. This could be attributed to the *risk-shifting mechanism* and the non-binding capital constraint.



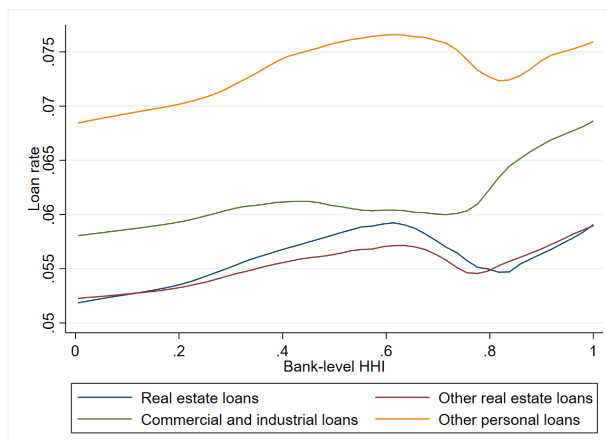
(a) Year 2008



(b) Year 2012



(c) Year 2016



(d) Year 2020

Figure C.1: Bank Concentration and Loan Rate at Bank-level

Notes: This plot presents the correlation between bank concentration and loan rate. There are four lines in the graph, where the yellow line represents other personal loans, the green line represents commercial and industrial loans, the blue line represents the real estate loans, and the purple line represents other real estate loans.

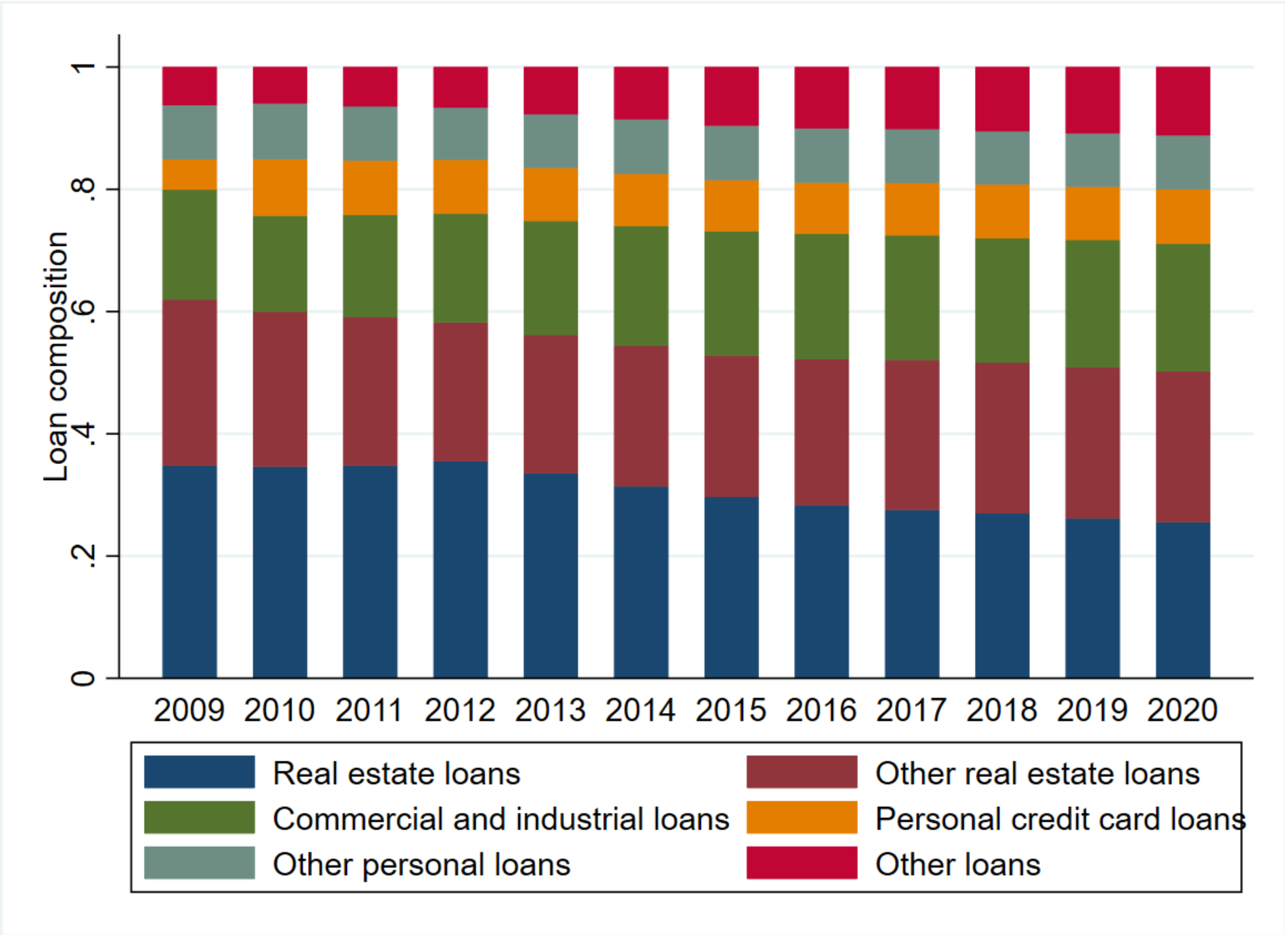


Figure C.2: Loan Composition in the U.S.