Bank Concentration, Product Market Competition and Firm Dynamics *

Lei Ye †

January 8, 2023

Please click here for the latest version

Abstract

Bank concentration in the U.S. has increased markedly since the Riegle-Neal Act implemented in 1994. This paper investigates the implications of bank concentration on product market competition, firm dynamics and aggregate economic performance. Empirically, I document that bank concentration is associated with less competition in product markets. It demotivates competition from new entrants on the extensive margin, and also strengthens market power of large incumbents at the cost of small ones on the intensive margin. I rationalize the empirical facts with an endogenous growth model that features imperfect bank competition, endogenous product market structure and interactions between banks and heterogeneous product market firms. Small firms are more reliant on bank financing than large firms, thus they are more heavily impacted by an interest rate spread increase that reflects the rise of bank market power in the bank concentration process. Changes in strategic competition among product market firms in response to the increased financing cost amplify the market reallocation towards large firms. Firm entry is deterred following the market structure change, further decreasing competition in product markets. In a calibration of the model to infer changes of the U.S. economy between the 1990s and the 2000s, bank concentration accounts for about 60 percent of the product market concentration, 20 percent of the entry and exit decline, and 60 percent of the enlarged markup ratio between large and small firms. The increased financing cost and the concentrated markets weaken firms' incentives to grow, thereby lowering long-run economic growth. Quantitative analysis suggests that bank concentration can explain about 3 percent of the U.S. productivity slowdown since the Act.

Keywords: bank concentration, industry competition, markups, firm dynamics, growth

JEL Codes: O40, E44, O30

*I am deeply indebted to my advisor Ping Wang for his continuous guidance and support. I would like to extend a special thanks to Yongseok Shin and Gaetano Antinolfi for their invaluable advice and comments. This paper also benefited from discussions with seminar participants at WUSTL Macro Study Group, Economics Graduate Student Conference (WUSTL), ES China meeting (CUHK ShenZhen) and ES Australia meeting (UQ). All errors are my own.

†Department of Economics, Washington University in St. Louis, MSC 1208-228-308, One Brookings Drive, St. Louis, MO 63130, USA. E-mail: l.ye@wustl.edu.

1 Introduction

Recent evidence suggests that U.S. product markets have experienced an increase in concentration levels over the past three decades. According to Grullon et al. (2019), market concentration has risen in more than 75% of the U.S. industries, and the average increase in concentration levels has reached 90%. At the same time, the U.S. banking industry has become increasingly concentrated after the implementation of the 1994 Riegle-Neal Act which removed many of the restrictions on banks' branching across state lines. The large banks have significantly increase their share in the banking system and the commercial loan markup has remarkably grown since the late 1990s (Figure 1a - 1b). While literature has studied common factors that simultaneously drive the concentration in the product markets and the banking industry, little attention has been paid to the causal relationship, where bank concentration can potentially lead to product market concentration.

(a) Bank Concentration (b) Commercial Loan Spread

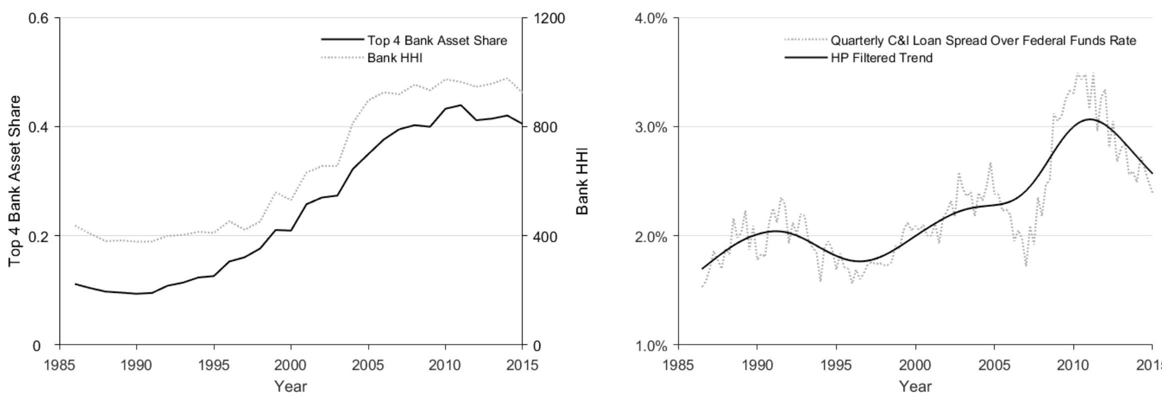


Figure 1: Bank Concentration and Commercial Loan Spread

Notes: The left panel displays two measures on bank concentration level: top four bank asset share and HHI of the largest fifty banks. The calculations are based on bank asset data from Call Report managed by Federal Deposit Insurance Corporation (FDIC). The right panel displays the quarterly commercial and industrial loan spread over federal funds rate and its HP-filtered trend. The commercial loan spreads are obtained from the Federal Reserve Board database.

In this paper, I study (i) whether the bank concentration truly causes the product market concentration, and in this process, (ii) how does bank concentration affect firm dynamics and aggregate economic performance. I find that bank concentration indeed leads to product market concentration. Regarding the firm dynamics, bank concentration deters firm entry from outside, and enables large incumbents to further expand their power over small incumbents from inside; regarding the aggregate economic performance, bank concentration negatively affects the productivity growth in the product markets. I also quantitatively measure the importance of bank concentration in driving these trends. It is found that bank concentration explains 59.13% of the observed product market concentration in the U.S. economy between the 1990s and the 2000s. In addition, bank concentration contributes to 24.56% in the total decline of the firm entry, 64.41% markup increase of large incumbents over small incumbents, and 2.35% decline in the productivity growth.

I propose the following mechanism to explain the above findings. Bank concentration reduces the competition in the banking system, which increases the interest rate spread for the product market firms, specifically, bank concentration increases the loan rate and decreases the deposit rate. During this process, small incumbents are tremendously hurt by the increased loan rate since they are more financially constrained and more reliant on the bank financing. Meanwhile, the reduced deposit rate motivates the large incumbents to switch to heavier capital investment instead, which allows them to gain stronger market power. The above two dynamics lead to concentration in the product markets. The rising dominance of large incumbents discourages firm entry from outside, which further exaggerates the product market concentration. Besides, the increased bank financing cost and the concentrated product market dis-incentivize firms to grow, thereby posting a negative impact on the economic growth.

To model the mechanism I propose, I develop an endogenous growth framework that incorporates imperfect bank competition in the spirit of Klein (1971) and Monti (1972), and strategic competition among heterogeneous product market firms along the lines of step-by-step quality ladder models (Aghion, Harris, et al. 2001). On the banking industry side, banks serve as financial intermediaries that collect deposits and issue loans to product market firms. The deposit market is in perfect competition, while the loan market is in Cournot competition. On the product market side, incumbent firms are heterogeneous in both asset and productivity, and engage in a two-stage strategic competition. Specifically, in the first stage of each period (i.e. production stage), incumbents involve in a Bertrand competition and set prices, which determines their market shares and corresponding profits. With the updated asset positions, in the second stage of each period (i.e. innovation stage), incumbents choose the amount of R&D investment out of their assets, which determines the productivity level in the next period. At the beginning of each period, entrants with heterogeneous initial assets choose their entry efforts, which determines if they can enter the market and replace one incumbent firm. The mechanism through which bank concentration affects the product market concentration is achieved via the imperfect competition in the loan market. Specifically, during the production stage, incumbent firms with insufficient asset can borrow from banks to cover the production cost. Incumbents with higher financing needs are more adversely affected by the increase of the financing cost related to the bank concentration. As a result, larger incumbents with better asset positions strengthen the advantage against smaller incumbents with insufficient assets. In other words, product markets become more concentrated as well. The product market concentration in turn deters firm entry.

I calibrate the model to the U.S. economy in the 1990s, targeting moments calculated based on data from DealScan, Compustat and Business Dynamics Statistics. Then I implement an exogenous increase in the bank concentration to its post-1990s' level and track the model responses. The model suggests that bank concentration could generate quantitatively important effects on product market structure and firm dynamics, including a sharp increase in markup ratio between large and small firms, remarkable declines in firm entry and exit rates, and a modest increase in product market concentration. The responses of the model to a concentrated banking system closely follows the changes of their empirical counterparts. Bank concentration drives 56.97%, 38.21% and 5.86% of observed changes in relative markups, firm entry rate and product market concentration, respectively. In addition, the calibrated model predicts a marginal decline in aggregate productivity growth, which accounts for about 2.40% of the productivity slowdown in the U.S. economy since the 1990s.

On top of bank concentration, the U.S. economy has experienced many fundamental changes between the 1990s and the 2000s that might have contributed to the rise in product market concentration, the decline in business dynamism, the shift of power balance to large firms, and the slowdown in productivity growth. Some of these changes in primitives include reduced research productivity, increased R&D cost, increased entry cost, and expanded market power of dominant firms. I look into bank concentration and these alternative explanations, and decompose the contribution of each channel in a counterfactual analysis to quantify their relative importance in driving the empirical regularities that the U.S. economy has been witnessing. Decomposition results suggest that 59.13% of the rise in product market concentration stems from bank concentration. Rising firm market power plays a equivalently important but opposite role, which explains 51.10% of the drop in market concentration. Bank concentration and firm market power are also the most important factors in driving the change in relative markup between large and small firms, with bank concentration accounting for 64.41% of the increase and firm market power leading to 35.89% of the decrease. Regarding the declines in firm entry and exit rates, higher entry cost is the most powerful force that drives the bulk of the changes. Bank concentration channel explains about 20.00% of the drop on the entry and exit margins. In terms of the change in productivity growth, the reduced research productivity channel and the increased R&D cost channel together explain almost the entire decline. In contrast to the technological factors that bring large impact on economic growth, bank concentration only contributes to 2.35% in the productivity slowdown.

Related Literature This paper is closely related to three strands of literature. The first strand of literature investigates the U.S. market concentration, business dynamism decline and productivity slowdown since the late 1990s. A set of recent papers document a broad increasing market concentration across U.S. industries since the late 1990s (Autor et al. 2017, Gutiérrez et al. 2017, Grullon et al. 2019, Autor et al. 2020, Barkai 2020). Measures of economic profits and markups have been on the rise along with the concentration process (Gutiérrez et al. 2017, R. E. Hall 2018, De Loecker, Eeckhout, et al. 2020, Eggertsson et al. 2021) while business dynamism has slowed down (Decker et al. 2016, Pugsley et al. 2019), reflecting increased average market power and reduced competition within the product markets. At the same time, productivity growth has been sluggish (Fernald 2014). The long-run productivity growth has been declining, except the IT-fueled boom in the late 1990s. These findings drew considerable attention as they likely indicate fundamental changes in the U.S. economy over the past decades. Many papers emphasize the increasing importance of information technology and intangible asset as a possible explanation (Aghion, Bergeaud, et al. 2019, De Ridder 2020, Corhay et al. 2020). Non-technological explanations include demographic changes (Karahan et al. 2019, Peters and Walsh 2020), regulation changes

(Gutiérrez et al. 2017, Grullon et al. 2019), or declining real interest rate (Chatterjee et al. 2019, Eggertsson et al. 2021, Liu et al. 2022). This paper contributes a new mechanism to the literature that links concentration in different markets, and shows that bank concentration has significant power in jointly explaining the observed trends in U.S. firm market power, business dynamism and productivity growth.

This paper also adds to the vast literature on financial development and economic growth. Economic research has focused intensely on the role played by financial market for real economic activities since first outlined in Schumpeter (1911). Levine (2005) provides a extensive review on the empirical and theoretical works exploiting the relationship between financial development and economic growth. Recent research has turned to the analysis of the specific market characteristics and the mechanisms through which finance affects real economic activity. This paper goes straight to the heart of this line of research by investigating the impact of competition in banking markets on the economic performance in the product markets, within an environment of imperfect bank competition that only a few studies has examined. From the empirical scope, this paper is closest to Cetorelli et al. (2006) which tests the relationship between competition in local U.S. banking markets and the market structure of non-financial sectors. They find that more vigorous banking competition is associated with more firms in operation and a smaller average firm size, which is suggestive of more active competition within non-financial sectors. This paper documents similar empirical regularities between bank competition and product market competition, and provides additional empirical evidence on the differential lending relationships between banks and heterogeneous firms to shed light on the potential mechanisms through which bank competition shapes the competition dynamics in the product markets. In terms of theoretical work, this paper is related to two papers that study the influences of bank concentration on economic growth within the endogenous growth framework. Deidda et al. (2005) explores the trade-off between economies to scale and economies to specialization in the provision of financial intermediation services within an AK endogenous growth model, and finds mixed relationship between between bank concentration and growth depending on the level of economic development. Diallo et al. (2018) introduces imperfect bank competition in the Schumpeterian growth paradigm, and demonstrates a negative effect of bank concentration on economic growth, the magnitude of which depends on the proximity to the world technological frontier. This paper builds on the Schumpeterian growth framework but, unlike Deidda et al. (2005) and Diallo et al. (2018) focusing on particular features of banking services, examines the differential effects of bank concentration on heterogeneous firms. The heterogeneous effects are found to be critical in shaping the product market structure and firm dynamics, and affect economic growth indirectly through its impacts on product market competition.

The theoretical framework in this paper is based on Schumpeterian growth models with step-by-step quality improvement in the tradition of Aghion, Harris, et al. (2001) and Aghion, Bloom, et al. (2005), and incorporates imperfect bank competition along the lines of Klein (1971) and Monti (1972). The class of Schumpeterian innovation-driven growth models have been widely used to study general issues related to firm competition, firm dynamics and aggregate growth for its analytical tractability yet rich model dynamics (Lentz et al. 2008, Akcigit and Ates 2019, Peters 2020, Ates et al. 2021). I build on this class of models by introducing ex-ante heterogeneous firms with external funding choices to adapt to the context of an economy with financial intermediation services and show that bank competition has significant effect on product market competition and economic growth. Klein (1971) and Monti (1972) provides simple yet powerful framework to investigate banking industry in the spirit of structure-conduct-performance paradigm, hence is widely adopted in studies about bank competition and its effects. I view my theory as extensions to the two classes of models separately, and a first trial to link them together.

The paper proceeds as follows. Section 2 presents empirical evidences on the link between U.S. bank concentration and product market concentration. Section 3 describes the theoretical model and its implications. Section 4 presents the model calibration and the quantitative investigation of U.S. bank concentration. Section 5 concludes.

2 Empirical Findings

This section presents empirical findings on the relationship between bank concentration and product market competition in the U.S. economy. Using information from commercial loan deals and firm financial statements, I document that bank concentration is associated with lower competition level in the product markets. I also examine the differential effects of bank concentration on different firms in the product markets. On the extensive margin, higher bank concentration is related with lower firm entry rate, implying lessened competition from new entrants to incumbents. On the intensive margin, bank concentration is linked with increasing market power of large incumbents at the cost of small ones, reflecting a shift of power balance to large companies within incumbents. I begin this section by describing the data, and then present empirical results.

2.1 Datasets and Variable Construction

The empirical analysis employs commercial loan deal data from *Thomson-Reuters LPC DealScan database (DealScan)*, fundamental firm accounting data from *Compustat Fundamentals Annual database (Compustat)*, and firm entry data from *Business Dynamics Statistics (BDS)* to construct a combined industry panel at a yearly frequency over the 1990-2015 period. The data sample consists of measures on bank credit concentration, product market competition, firm entry rate and industry characteristics for 57 industries defined with SIC two-digit codes. Descriptive statistics can be found in Table of Appendix A.1.

2.1.1 DealScan

The measure of bank concentration is constructed based on commercial loan deal data from DealScan. DealScan database provides comprehensive information on loan deals in the global commercial loan markets. For United States, this data covers around 75% of the value of all commercial loans.[1] The basic unit of observation in DealScan is a loan, alternatively referred to as a facility. Each loan observation contains rich information including borrower and lender identities, lender roles, facility amount, maturity, and loan rate. Compustat annual data is merged into the DealScan loan observations in order to complete borrowers and lenders' country information and industry classification.[2] In this study, I employ only lending relationships between U.S. banks and firms.[3] Observations are excluded if borrowers are agricultural firms (SIC codes 0100-0999), utilities firms (SIC codes 4900-4999), financial firms (SIC codes 6000-6999), and public administration and nonclassifiable firms (SIC codes 9000-9999). The final sample consists of 66867 loan observations to U.S. publicly listed firms covering 57 SIC two-digit level industries over the 1985-2015 period.

From this commercial loan sample, I construct the measure of bank concentration, $Bank\ HHI_{it}$, as a Herfindahl–Hirschman Index (HHI) capturing the bank credit concentration at the industry-year level. $Bank\ HHI_{it}$ is defined as the sum of squared bank market shares, and a bank's market share is its total loan volume granted to an industry out of the aggregate loan volume this industry received. Both the bank's and the industry's loan volume are calculated by summing up the facility amount of related loans in the DealScan sample. Following Giannetti et al. (2019) and Saidi et al. (2021), I measure the bank's and the industry's loan volume over the previous five years, since the average loan maturity in the DealScan data is approximately five years. Due to the five-year lookback window, the bank concentration measure covers the period over 1990-2015.

2.1.2 Compustat

Compustat contains financial statement information for all U.S. publicly listed firms, and is used to construct measurement for product market competition in this study. Specifically, I use average firm markup in an industry, $Markup_{it}$, as the measure of competition level in this product market. Firm markup is estimated following the production approach in De Loecker and Warzynski (2012), as the the product of the output elasticity of a variable input in production and the ratio of the firm's sales to its expenditure on that input.[4] The output elasticity is obtained by estimating a translog production function using the procedure proposed in De Loecker and Warzynski (2012).

[1]According to Carey et al. (1999), DealScan covers between 50% and 75% of the value of all U.S. commercial loans during the early 1990s. From 1995 onwards, DealScan's coverage increases to include a greater fraction.

[2]Borrowers and lenders' information from Compustat is merged with DealScan data using the link tables provided in Chava et al. (2008) and Schwert (2018). Chava et al. (2008) provides files that link DealScan borrowing company with Compustat identifiers. In their work, Compustat data is merged with DealScan loan information by matching company names and loan origination dates in DealScan to company names and corresponding active dates in the CRSP historical header file. Schwert (2018) connects the most active lenders in DealScan with the identifiers of their bank holding companies in Compustat. This link includes all lenders that acted as lead arrangers on at least 50 loans or at least $10 billion in volume, as well as their related subsidiaries.

[3]Lending relationship is identified by focusing on lead arrangers of loan facilities. For a loan with multiple lead arrangers, the facility amount is split equally among them.

[4]The advantages of using the production approach to estimate firm markups in this context are twofold. First, the production approach does not impose any assumption on the market structure or competition. So the method can be consistently applied to different industries with differing characteristics. Second, firm markups are estimated based on a single variable input without imposing any particular substitution elasticity with respect to other variable or fixed inputs. As long as the one variable input is observed from data, markups can be estimated consistently.

Procedure details can be found in Appendix A.2. Firms' sales are directly observed on their financial statements from Compustat. Following De Loecker, Eeckhout, et al. (2020), the variable input expenditure is measured by the Compustat item "Cost of Goods Sold" (COGS), which is a sum of expenses directly attributable to the production of goods sold, including material costs, wage bills, energy and so on.

To control for the effects of heterogeneous industry characteristics on the bank concentration and product market competition relationship, I also construct a series of industry-level control variables using the Compustat data. Details on control variables are illustrated in regression specifications.

2.1.3 BDS

The U.S. Census Bureau's Business Dynamics Statistics provides annual measures of establishment openings and closings by firm size, age, industry and other statistics on U.S. business dynamics. I employ the establishment entry data by industry, $Entry_{it}$, to investigate the relationship between bank concentration and firm entry in the U.S. economy.

2.2 Bank Concentration and Product Market Competition

The main focus of this section is to investigate the relationship between bank concentration and product market competition in the U.S. economy. For this purpose, I conduct the following panel regression with industry and year fixed effect:

$$Markup_{it} = \beta_0 + \beta_1 Bank\ HHI_{it} + \gamma' X_{it} + \delta_i + \eta_t + \epsilon_{it}, \tag{1}$$

where $Markup_{it}$ denotes the average firm markup in industry i in year t measuring the competition level in this product market. $Bank\ HHI_{it}$ is the loan-based HHI capturing the bank credit concentration at the industry-year level. δ_i and η_t are the industry fixed effect and year fixed effect, respectively. To control for industry characteristics which could have impacts on average firm markup, the variable vector X_{it} includes $\ln(Asset)_{it}$, $\ln(Debt)_{it}$, $Leverage_{it}$ and $Ind\ HHI_{it}$. $\ln(Asset)_{it}$ is the log of sum of firms' assets in industry i in year t, which controls for the industry size. $\ln(Debt)_{it}$ is the log of sum of firms' long term debt plus debt in current liabilities, which controls for the industry's total loan demand. $Leverage_{it}$, defined as the ratio of debt to debt plus equity, controls for the the industry's financing structure and potential influence from the equity financing. $Ind\ HHI_{it}$ is the industry HHI based on firm sales which controls for product market structure.

The panel regression results are displayed in the left columns of Table 1. Column 1 shows a positive and statistically significant coefficient between bank credit concentration measure and average firm markup, suggesting that higher credit concentration is associated with lower competition level in the product markets. Column 2 reports the result after controlling for industry

Table 1: Bank Concentration and Average Firm Markup

	Panel Regression		IV	
	(1)	(2)	(3)	(4)
Bank HHI	0.114***	0.124***	1.286***	1.142***
	(0.034)	(0.036)	(0.360)	(0.292)
ln(Asset)		0.002		−0.079***
		(0.014)		(0.029)
ln(Debt)		0.037***		0.134***
		(0.013)		(0.032)
Leverage		−0.114***		−0.206***
		(0.042)		(0.058)
Ind HHI		0.244***		0.111*
		(0.049)		(0.070)
Constant	1.014***	0.672***	0.265	0.040
	(0.040)	(0.090)	(0.234)	(0.210)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
First-stage F-statistic	-	-	22.305	31.473
R^2	0.807	0.813	0.768	0.774
Obs	1466	1466	1466	1466

Notes: This table reports the panel regression results of average firm markup on bank credit concentration at the industry-year level (column 1 and 2), and the IV estimation results which uses cumulative bank merger events as an instrument variable for bank credit concentration measure (column 3 and 4). The sample period is 1990 to 2015. Industries are defined with the SIC two-digit codes, with agriculture, utilities, finance, public administration and nonclassifiable industries excluded. Nominal terms are deflated by CPI. Assets and debts are in thousands of U.S. dollars. Standard errors are in parentheses. *,**,*** denote significance at the 10%, 5% and 1% level, respectively.

heterogeneity. The coefficient between credit concentration and average firm markup is robust to the inclusion of industry characteristics.

Crucially, by including industry leverage as a control variable, I address the possibility that the effect of bank credit on product market competition could be systematically overestimated as firms can use equity as alternative source of funding. This problem is especially important when running with Compustat data, as firms covered in Compustat are all publicly listed. I also consider the impact that venture capital, another potential funding source for businesses, could exert on the estimation. Appendix A.3 provides robustness checks addressing the venture capital issue. Since venture capital investment cannot be observed from Compustat, I tackle this problem in an indirect way by removing observations related to high technology firms (e.g. information technology or biotechnology) and young firms that got publicly listed for three years and under, as these firms are likely to be targets of venture capitalists. The finding that higher bank credit concentration correlates with lower product market competition holds true when I control for equity and venture capital factors. The robustness to both equity financing and venture capital suggests that bank credit matters for product market competition, above and beyond the role of alternative financing sources.

As the U.S. economy has undergone market concentration in most industries since the 1990s (Grullon et al. 2019, Autor et al. 2020), there is a serious concern that the rising concentration in the banking industry and growing market power in the product markets are caused by common underlying factors, e.g. lax enforcement of antitrust policies or declining population. To address this endogeneity issue and to obtain some causal inference on the effect of bank concentration on product market competition, I conduct an instrument variable (IV) estimation. Following Favara et al. (2017) and Saidi et al. (2021), I use the cumulative number of bank mergers as an instrument variable for $Bank\ HHI_{it}$.[5] Only bank mergers involving acquirers and targets with non-zero market shares in a given industry are counted, since the credit concentration at the industry level would otherwise be unaffected by the bank mergers.

The IV regression results shown in Table 1 column 3 and 4 are in line with the baseline panel regression, which point to a positive and significant effect of bank concentration on product market markup. The first-stage F-statistics are suggestive of strong correlation between the instrument variable and the regressor $Bank\ HHI_{it}$. For exclusion restriction to hold here, it is required that bank merger events affect average firm markup only through credit concentration. This condition is generally true as the average industry only accounts for a small fraction of a bank's commercial loan portfolio, hence mergers are unlikely to occur because bank business strategies are leaning towards particular industries.

2.3 Bank Concentration and Differential Effects

The section explores the differential effects of bank concentration across firms in the product markets, with the purpose to see how the bank concentration process reshapes the competition and firm dynamics in the product markets. The growth literature has suggested that a substantial fraction of changes in market competition and aggregate productivity growth is accounted by the reallocation of resources among firms. Firm entry and incumbent reallocation are especially critical components of the competition and productivity reallocation dynamics. Therefore, I break the differential effect examination into two parts: the effect of bank concentration on entrants and the effect on the reallocation among incumbents.

2.3.1 Bank Concentration and Firm Entry

To see how bank concentration influences firm entry, I conduct the following panel regression:

$$Entry_{it} = \beta_0 + \beta_1 Bank\ HHI_{it} + \gamma' X_{it} + \delta_i + \eta_t + \epsilon_{it}, \tag{2}$$

where $Entry_{it}$ is the firm entry rate in industry i in year t, defined as the ratio of new establishments to the total establishments in this industry and year. $Bank\ HHI_{it}$ is the loan-based HHI measuring bank credit concentration. X_{it} is the control variable vector. δ_i and η_t are the industry fixed effect and year fixed effect, respectively. Like in the previous regression, I also conduct an IV

[5] Schwert (2018) provides a set of hand-collected bank merger events in which acquirers can be linked to Compustat.

estimation where the cumulative number of bank merger events is used as the instrument variable for $Bank\ HHI_{it}$.

Table 2 presents the estimation results on bank concentration and firm entry. Column 1 shows a significant negative relationship between firm entry rate and bank credit concentration under the panel regression. The coefficient estimate is stable when controlling for industry characteristics as in column 2. The IV estimation results are shown in column 3 and 4, which are in line with the panel regression results. The estimation results suggest that bank credit concentration is associated with less firm entry, thus implying lessened competition from potential entrants.

Table 2: Bank Concentration and Firm Entry Rate

	Panel Regression		IV	
	(1)	(2)	(3)	(4)
Bank HHI	−0.011***	−0.013***	−0.076***	−0.070***
	(0.004)	(0.004)	(0.028)	(0.026)
Constant	0.159***	0.154***	0.200***	0.188***
	(0.004)	(0.011)	(0.019)	(0.018)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry Control Variables	No	Yes	No	Yes
First-stage F-statistic	-	-	25.855	38.231
R^2	0.793	0.796	0.742	0.765
Obs	1404	1388	1404	1388

Notes: This table shows the estimation results of firm entry rate on bank credit concentration measure at the industry-year level. The sample period is 1990 to 2015. Industries are defined with the SIC two-digit codes, with agriculture, utilities, finance, public administration and nonclassifiable industries excluded. Control variables include log of industry asset, log of industry debt, industry leverage defined as the ratio of debt to debt plus equity, and sale-based industry HHI. Standard errors are in parentheses. *,**,*** denote significance at the 10%, 5% and 1% level, respectively.

2.3.2 Bank Concentration and Incumbent Reallocation

Now turn to the impacts of bank concentration on incumbents. As firm size is a crucial factor that affect firms' ability to borrow, the following analysis focuses on the differential effects on markups by incumbent firms of different sizes. In this study, firm size is measured by firm's total asset. Firms are sorted into asset quintiles for each industry in each year (first quintile denotes firms with the least asset and fifth quintile the most), hence five size classes are defined. To analyze the sensibility of firms to credit concentration for each size class, I conduct the following estimation which allow the coefficients of bank concentration to vary by firm size class:

$$Markup_{ist} = \beta_0 + \sum_{s=1}^{5} \beta_s \left(Bank\ HHI_{it} \times Firm\ Size_s\right) + \gamma' X_{it} + \delta_i + \eta_t + \epsilon_{ist}, \tag{3}$$

where $Markup_{ist}$ denotes the average firm markup of size class $s \in \{1, 2, .., 5\}$ in industry i in year t, measuring the market power that a given group of firms have. $Firm\ Size_s$ is the firm size dummy variable which equals 1 if the group of firms are in the s-th quintile of asset distribution and 0 otherwise. $Bank\ HHI_{it}$ is the loan-based HHI measuring bank credit concentration at the industry-year level. X_{it} is the control variable vector. δ_i and η_t are the industry fixed effect and year fixed effect, respectively.

The impacts of bank credit concentration vary across firm size as shown in Table 3. Rise in bank concentration reduces the average markup of firms in the first three small size classes ($Firm\ Size_1$ - $Firm\ Size_3$). Conversely, an increase in average firm markup is observed in the other two large size classes ($Firm\ Size_4$ - $Firm\ Size_5$) with increase in credit concentration. These relationships between bank credit concentration and firm size are robust to the inclusion of industry-level control variables, as well as the IV estimation where the cumulative number of bank merger events is used as the instrument variable for $Bank\ HHI_{it}$. The regression results suggest that bank concentration increases the market power of large incumbents at the cost of small businesses, reflecting a reallocation of economic activities and power balance towards large firms in the product markets. Moreover, the estimation coefficients increase monotonically with firm size. The monotonicity indicates that firm size is indeed an important factor in firm-bank lending

Table 3: Bank Concentration and Average Firm Markup by Size Categories

	Panel Regression		IV	
	(1)	(2)	(3)	(4)
Bank HHI × Firm Size$_1$	−0.265***	−0.263***	−0.407**	−0.441***
	(0.024)	(0.024)	(0.169)	(0.170)
Bank HHI × Firm Size$_2$	−0.123***	−0.123***	−0.029	−0.063
	(0.024)	(0.024)	(0.168)	(0.170)
Bank HHI × Firm Size$_3$	−0.037	−0.034	0.064	0.030
	(0.024)	(0.024)	(0.169)	(0.170)
Bank HHI × Firm Size$_4$	0.057**	0.059**	0.293*	0.259*
	(0.024)	(0.025)	(0.169)	(0.169)
Bank HHI × Firm Size$_5$	0.167***	0.170***	0.306*	0.272*
	(0.023)	(0.023)	(0.168)	(0.169)
Constant	0.905***	0.657***	0.850***	0.621***
	(0.020)	(0.049)	(0.108)	(0.125)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry Control Variables	No	Yes	No	Yes
First-stage F-statistic	-	-	21.302	21.346
R^2	0.473	0.478	0.445	0.450
Obs	7091	7090	7091	7090

Notes: This table reports the the differential effects of bank credit concentration on average markups by incumbent firms with different sizes. Firm size groups are quintiles of the cross-sectional asset distribution within each industry for each year. Industries are defined with the SIC two-digit codes, with agriculture, utilities, finance, public administration and nonclassifiable industries excluded. Control variables include log of industry asset, log of industry debt, industry leverage defined as the ratio of debt to debt plus equity, and sale-based industry HHI. Standard errors are in parentheses. *,**,*** denote significance at the 10%, 5% and 1% level, respectively.

relationships, and further has an impact on firm performance and economic reallocation dynamics. This observation guides me to develop a theoretical model that introduces firm heterogeneity along the size dimension to govern firm borrowing decisions, in addition to productivity heterogeneity as suggested by the literature of heterogeneous firms to be the main determinant of firm performance.

3 Model

Motivated by the empirical findings, I develop an endogenous growth model that rationalizes the effects of bank concentration on product market competition. In the model, banks serve as the only financial intermediaries which collect deposits and issue loans to firms in the product markets. Along the lines of Klein (1971) and Monti (1972), I model the bank side in a context of perfect competition for deposits while imperfect Cournot competition for loans. The imperfect competition assumption in the loan market yields a positive relationship between the degree of bank concentration and the loan spread faced by product market firms. The product market is modelled in the spirit of the step-by-step quality ladder framework (e.g. Aghion, Harris, et al. 2001, Aghion, Bloom, et al. 2005, Akcigit and Ates 2019). I introduce into this framework the possibility for firms to get external bank financing in their production process, which allows me to capture the effects of bank concentration on the product market competition through its influences on the borrowing decisions by heterogeneous product market firms. Incumbent firms in the product markets compete over price to obtain market power and invest in innovative activity to improve their productivity. There is also an outside pool of entrants engaging in research activity to enter the market and replace the laggard incumbents. An appealing feature of product market side is that a endogenous markup distribution is generated as a result of the different relative productivity levels, firm sizes and market competition across the product markets. Putting the two sides together, this model illustrates the interplay between bank competition and product market competition, which critically shapes firms' incentive for growth and, therefore, growth in the economy as a whole.

The model is of a closed economy in discrete time. The economy is populated by four types of agents: a representative household, a representative final good firm, firms producing differentiated intermediate goods, and banks. This section presents the model details by first describing the problem of each type of agent, and then analyzing the equilibrium.

3.1 Representative Household

The economy admits an infinitely-lived representative household that derives logarithmic utility from consumption over a final good:

$$U = \sum_{t=0}^{\infty} \beta^t \ln(C_t), \tag{4}$$

where C_t denotes the consumption at time t, and $\beta \in (0, 1)$ is the discount factor. The price of the final good is normalized to $P_t = 1$ in each period. The budget constraint of the representative household reads as:

$$C_t + B_{t+1} = w_t L_t + (1 + r_t^D) B_t + A_t^F - A_t^E + \Pi_t^B, \tag{5}$$

where B_t denotes the households' saving in the form of bank deposit that earns a deposit interest at the rate r_t^D. L_t denotes labor, which is supplied inelastically to firms in the intermediate goods markets. Labor supply is normalized such that $L_t = 1$ for any period and there is no population growth. In each period, labor earns a wage rate w_t that is determined endogenously to clear the labor market. Households own all the firms and banks in the economy, thus all firm values and bank profits accrue to the households. This includes a sum of intermediate goods firm values A_t^F, a lump-sum endowment provided to potential entrants in intermediate markets A_t^E, and a sum of bank profits Π_t^B. Specifically, $A_t^F = \int_0^1 (\sum_{j \in \{1,2\}} V_{jit} + V_{eit}) di$ with V_{jit} being the value of an incumbent intermediate goods firm j in market i at time t, and V_{eit} being the value of an potential entrant that can replace incumbents in market i. And $\Pi_t^B = \sum_{s=1}^S \pi_{st}^B$, where π_{st}^B denotes the profit of bank s at time t. These value functions, endowments and profits are explained in greater detail when it comes to intermediate goods firms' and banks' problems.

Optimal household decisions yield the standard Euler equation:

$$g_t = \beta(1 + r_t^D), \tag{6}$$

where g_t is the growth rate of aggregate consumption.

3.2 Representative Final Good Producer

The final good is produced competitively using inputs from a continuum of intermediate markets. The representative final good firm operates a Cobb-Douglas technology given by:

$$\ln Y_t = \int_0^1 \ln y_{it} di, \tag{7}$$

where y_{it} denotes the amount of intermediate good $i \in [0, 1]$ used at time t.

The optimization of the representative final good producer generates the following demand schedule for the intermediate good $i \in [0, 1]$:

$$y_{it} = \frac{Y_t}{p_{it}}, \tag{8}$$

where p_{it} is the price of intermediate good i. The demand schedule implies that expenditures are equalized across intermediate goods.

3.3 Intermediate Goods Producers

The continuum of intermediate markets i spans the interval $[0, 1]$. Each intermediate market features competition between two active incumbents that produce differentiated goods and invest to innovate over existing products, plus a potential entrant that engage in innovation activities to enter the market. Firm competition unfolds in three stages: entry, production and innovation. Every period starts with an *entry stage*, in which the potential entrant invests to attempt a new product and replaces the technological follower if the innovation is successful. The two incumbents, or the technological leader and the entrant in case of successful entry and replacement, then enter a *production stage*, in which they produce and compete over price. Following production is an *innovation stage*, in which the two firms invest in innovation activity to improve product qualities. This section describes the static price competition and the dynamic innovation decisions by incumbents, as well as the entry decisions by potential entrants.

3.3.1 Market Structure and Price Setting

Each intermediate market has two active incumbents $j \in \{1, 2\}$, which produce imperfect substitutes with differing qualities. The intermediate market output is a CES aggregation over outputs of the two incumbents:

$$y_{it} = \left[(q_{1it} y_{1it})^{\frac{\sigma-1}{\sigma}} + (q_{2it} y_{2it})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \quad \sigma \in [1, \infty) \tag{9}$$

where q_{jit} is the product quality of firm j in market i at time t, y_{jit} is firm j's output, and σ is the elasticity of substitution between product varieties within an intermediate market.[6] I refer to firm 1 as the technological leader (follower) in market i if $q_{1it} > q_{2it}$ ($q_{1it} < q_{2it}$). Firms are neck-and-neck if $q_{1it} = q_{2it}$. Initial product qualities are normalized such that $q_{ji0} = 1$.

The CES aggregator implies that the demand schedule facing the incumbent $j \in \{1, 2\}$ in intermediate market i is:

$$y_{jit} = \frac{q_{jit}^{\sigma-1} p_{jit}^{-\sigma}}{q_{1it}^{\sigma-1} p_{1it}^{1-\sigma} + q_{2it}^{\sigma-1} p_{2it}^{1-\sigma}} \cdot Y_t, \tag{10}$$

where the market price index is given by:

$$p_{it} = \left(q_{1it}^{\sigma-1} p_{1it}^{1-\sigma} + q_{2it}^{\sigma-1} p_{2it}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}. \tag{11}$$

The two incumbents in an intermediate market engage in Bertrand competition, that is, incumbents set prices to maximize profits, taking as given all the input and output prices and the

[6] q_{jit} can be alternatively interpreted as firm j's productivity, as the intermediate market output is homogeneous of degree one in either the qualities or the quantities of the intermediate goods firms' output. Therefore, I use product quality and productivity interchangeably.

price set by its competitor.[7] Given the demand function (10) above, profit maximization yields the following optimal pricing rule for each intermediate market incumbent:[8]

$$p_{jit} = \frac{\eta_{jit}}{\eta_{jit} - 1} c'_{jit}(y_{jit}, a_{jit}), \tag{12}$$

where η_{jit} is the price elasticity of demand for incumbent j. It can be shown that the elasticity takes the form $\eta_{jit} = \sigma + (1-\sigma)s_{jit}$, with $s_{jit} = \frac{p_{jit}y_{jit}}{p_{1it}y_{1it}+p_{2it}y_{2it}} = \frac{q_{jit}^{\sigma-1}p_{jit}^{1-\sigma}}{q_{1it}^{\sigma-1}p_{1it}^{1-\sigma}+q_{2it}^{\sigma-1}p_{2it}^{1-\sigma}}$ being incumbent j's market share in market i. $c'_{jit}(y_{jit}, a_{jit})$ is the marginal cost for incumbent j to produce output y_{jit} with net asset a_{jit}. The net asset and marginal cost will be illustrated in more detail when it comes to firms' production process in Section 3.3.2.

The optimal pricing rule implies a markup μ_{jit} for each incumbent which takes the following form:

$$\mu_{jit} = \frac{p_{jit}}{c'_{jit}(y_{jit}, a_{jit})} = \frac{\sigma + (1-\sigma)s_{jit}}{\sigma + (1-\sigma)s_{jit} - 1}. \tag{13}$$

The markup is increasing in the incumbent's market share and this can be easily seen in two extreme cases where the market share goes to 0 or 1. When market share approaches 0, the incumbent producer is atomistic with respect to the market and charges a markup of $\frac{\sigma}{\sigma-1}$, which is standard in the CES structure of monopolistic competition. When market share is getting close to 1, the incumbent weighs only the elasticity across markets and charges a markup that goes to infinity. An incumbent's markup is determined endogenously as a result of strategic interaction with its competitor. This duopolistic Bertrand competition setting provides a useful ground to analyze the relationship between firm competition and markup dynamics in the economy.

3.3.2 Production and Marginal Cost

Each incumbent produces the output for its variety using both capital and labor according to a Cobb-Douglas technology:

$$y_{jit} = k_{jit}^{\alpha} l_{jit}^{1-\alpha}, \tag{14}$$

where k_{jit} denotes the capital used, l_{jit} is the labor employed, and α is the capital share which is assumed to be common across intermediate goods markets.

---

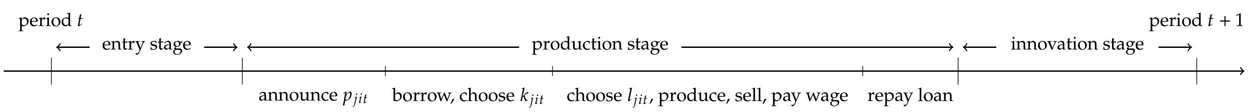
[7] One can alternatively assume Cournot competition between the two incumbents. Cournot competition yields an optimal markup given by $\mu_{jit} = \frac{p_{jit}}{c'_{jit}(y_{jit})} = \frac{\sigma}{(\sigma-1)(1-s_{jit})}$, which is also increasing in the incumbent's market share. Moreover, Cournot competition does not change the property that production profit only depends on the quality gap between the competing firms, not on the two quality levels.

[8] The model is with infinite horizon, and the pricing game is repeated infinitely. The folk theorem logic implies that collusive outcomes can arise as equilibrium. In my analysis, I assume firms are non-cooperative and only focus on the natural benchmark equilibrium, namely the repetition of the static game.

At the start of the production process, each incumbent decides how much capital to use. Capital investment can be financed by two funding sources, namely, internal fund and bank loan. Internal fund refers to the net asset that a firm accumulated from previous periods. When short in internal fund, the incumbent has access to bank loans to finance capital. Bank loan is more expensive than internal fund as banks charge a loan interest at the rate r_t^L on top of the principal. At the end of the production stage, after revenue is realized and capital gets depreciated with rate $\delta \in (0,1)$, the incumbent repays the loan and the interest. Importantly, the incumbent's borrowing is subject to a collateral constraint as in Moll (2014):[9]

$$k_{jit} \leq \theta a_{jit}, \quad \theta \in [1, +\infty). \tag{15}$$

where a_{jit} denotes the net asset of incumbent j in market i at the production stage of period t. The collateral constraint puts a limit on the incumbent's leverage ratio k_{jit}/a_{jit}, which captures the common intuition that the amount of capital available to a firm is limited by its net asset. In the other case where internal fund is sufficient to cover capital investment, the incumbent saves the residuals in the form of bank deposit and earns an interest at the rate r_t^D, where $r_t^D < r_t^L$. Capital, net asset, bank loan and bank deposit are all in units of the numeraire final good. After deciding its capital investment, the incumbent hires labor and starts producing. Wages are paid after realizing revenue and before repaying bank loan. The following timeline summarizes an incumbent's actions and decisions in the production stage.



Given the production process and input prices, cost minimization yields the factor demands for an incumbent with net asset a_{jit} and desired output level y_{jit}:

$$k_{jit}(y_{jit}, a_{jit}) = \begin{cases} \left(\frac{\alpha}{1-\alpha}\frac{w_t}{\delta+r_t^D}\right)^{1-\alpha} y_{jit}, & \text{if } k^D \leq a_{jit} \\ a_{jit}, & \text{if } k^D > a_{jit} \text{ and } k_u^L \leq a_{jit} \\ \left(\frac{\alpha}{1-\alpha}\frac{w_t}{\delta+r_t^L}\right)^{1-\alpha} y_{jit}, & \text{if } k^L > a_{jit} \text{ and } k_u^L \leq \theta a_{jit} \\ \theta a_{jit}, & \text{if } k^L > \theta a_{jit} \end{cases}, \tag{16}$$

[9]This formulation of collateral constraint is of analytical convenience and is isomorphic to the setup where firms own capital and issue debt to finance capital investment as in Buera and Moll (2015). Specifically, one can assume that capital is financed by both internal fund and debt: $k_{jit} \leq a_{jit} + d_{jit}$, where d_{jit} denotes the firm's debt with bank. The debt is collateralized by firm capital such that: $d_{jit} \leq \frac{\theta-1}{\theta} k_{jit}$. Rearranging the two equations above yields $k_{jit} \leq \theta a_{jit}$.

and

$$
l_{jit}(y_{jit}, a_{jit}) = \begin{cases} \left(\frac{1-\alpha}{\alpha}\frac{\delta+r_t^D}{w_t}\right)^{\alpha} y_{jit}, & \text{if } k^D \leq a_{jit} \\ \left(\frac{y_{jit}}{a_{jit}^{\alpha}}\right)^{\frac{1}{1-\alpha}}, & \text{if } k^D > a_{jit} \text{ and } k_u^L \leq a_{jit} \\ \left(\frac{1-\alpha}{\alpha}\frac{\delta+r_t^L}{w_t}\right)^{\alpha} y_{jit}, & \text{if } k^L > a_{jit} \text{ and } k_u^L \leq \theta a_{jit} \\ \left[\frac{y_{jit}}{(\theta a_{jit})^{\alpha}}\right]^{\frac{1}{1-\alpha}}, & \text{if } k^L > \theta a_{jit} \end{cases}, \tag{17}
$$

where $k^D = \left(\frac{\alpha}{1-\alpha}\frac{w_t}{\delta+r_t^D}\right)^{1-\alpha} y_{jit}$ is the optimal capital input if the incumbent finances it with internal fund only, and $k^L = \left(\frac{\alpha}{1-\alpha}\frac{w_t}{\delta+r_t^L}\right)^{1-\alpha} y_{jit}$ is the optimal capital level if the incumbent use both internal fund and bank loan as funding sources. The factor demands are piecewise functions with four regions. In the first region where $k^D \leq a_{jit}$, the incumbent is able to cover the desired capital level with its net asset, thus the factor demands take the standard form implied by the cost minimization for Cobb-Douglas production function. In the second region where $k^D > a_{jit}$ and $k^L \leq a_{jit}$, the incumbent is constrained in capital input, but hiring more labor to increase output is cheaper compared to increasing capital investment through bank loan. Hence, the incumbent sets capital equal to the maximum level that can be afford by its net asset and adjusts labor input at the margin. In the third region where $k^L > a_{jit}$ and $k^L \leq \theta a_{jit}$, the incumbent finances its optimal capital input with external fund from banks, and the marginal cost of capital increases from r_t^D to r_t^L. In the last region where $k^L > \theta a_{jit}$, capital investment is constrained even with bank loans. The incumbent, therefore, chooses capital at its maximum θa_{jit} and increase labor input to reach the desired output quantity. The derivation of factor demands is in Appendix B.1.

The optimal choices of capital and labor imply that the marginal cost of production for an incumbent with net asset a_{jit} and output y_{jit} takes the following form:

$$
c'_{jit}(y_{jit}, a_{jit}) = \left(\frac{\delta + r_t^D + \kappa_{jit}}{\alpha}\right)^{\alpha} \left(\frac{w_t}{1-\alpha}\right)^{1-\alpha}, \tag{18}
$$

where

$$
\kappa_{jit} = \begin{cases} 0, & \text{if } k^D \leq a_{jit} \\ \frac{\alpha}{1-\alpha} w_t \left(\frac{y_{jit}}{a_{jit}}\right)^{\frac{1}{1-\alpha}} - \delta - r_t^D, & \text{if } k^D > a_{jit} \text{ and } k^L \leq a_{jit} \\ r_t^L - r_t^D, & \text{if } k^L > a_{jit} \text{ and } k^L \leq \theta a_{jit} \\ \frac{\alpha}{1-\alpha} w_t \left(\frac{y_{jit}}{\theta a_{jit}}\right)^{\frac{1}{1-\alpha}} - \delta - r_t^D, & \text{if } k^L > \theta a_{jit} \end{cases}.
$$

The marginal cost function consists of four regions corresponding to the four cases illustrated in the factor demands. Figure 2a shows the marginal cost as a function of net asset for a given level of output. In the first and third region in the graph, incumbents are constrained in capital and face downward sloping marginal cost curves as net asset relaxes. Incumbent firms in these two regions increase production solely by hiring more labor, which is subject to decreasing returns. In the second and fourth region where incumbents are unconstrained, the marginal cost remains constant as in the standard cost minimization for production function with constant return to scale.

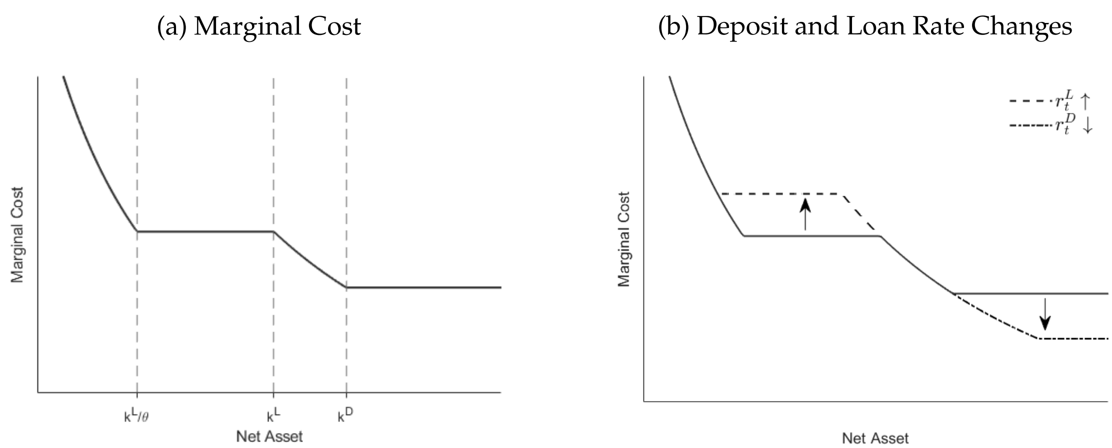


Figure 2: Marginal Cost Function

Banks influence the product markets through the deposit rate r_t^D and the loan rate r_t^L it imposes on the intermediate goods firms. As will be clear in the banks' problem in Section 3.4, an increase in the degree of bank concentration raises the loan spread $r_t^L - r_t^D$, reflecting a rise of bank market power. This change alters intermediate market firms' marginal cost, and thus production decisions. Figure 2b demonstrates how the marginal cost function is affected by the deposit rate and the loan rate, respectively. An increase in the loan rate raises incumbents' cost of external financing, leading to a shrink of firms turning to bank loan as funding source, and an increase on the marginal cost of firms using bank loans. As indicated by the dashed line, small firms are mostly impacted as they rely more heavily on bank loans. On the contrary, the deposit rate mainly influences the firms on the right side of the asset spectrum as displayed by the dash-dotted line. A reduction in the deposit rate brings down the opportunity cost of internal fund, reducing marginal costs and encouraging capital investment expansion by firms with large net assets. Inspecting the two effects together, an increase in loan spread benefits large firms while hurts small ones.

3.3.3 Innovation

After completing the production plans, incumbents can undertake innovative activities to improve their product qualities and increase consumer demand in future periods. When an innovation arrives in period t, it increases the innovating firm's product quality by a constant factor $\lambda > 1$ such that

$$q_{jit+1} = \lambda q_{jit}. \tag{19}$$

Summing up all the successful quality improvements that took place between period 0 and t, the number of which is denoted by N_{jit}, one can characterize an incumbent's current quality level as $q_{jit} = \lambda^{N_{jit}}$. Then the relative quality level between the two incumbent firms is given by

$$\frac{q_{1it}}{q_{2it}} = \frac{\lambda^{N_{1it}}}{\lambda^{N_{2it}}} = \lambda^{N_{1it}-N_{2it}} \equiv \lambda^{n_{it}}, \tag{20}$$

where $n_{it} \in \mathbb{Z}$ defines the technology gap between incumbent 1 and 2 in market i at time t. Hence, incumbent 1 is the technological leader (follower) in market i if $n_{it} > 0$ ($n_{it} < 0$). Firms are neck-and-neck if $n_{it} = 0$. I show in Appendix B.2 that the technology gap n_{it}, instead of the the quality levels q_{jit}, is one of the sufficient statistics to describe firm-specific payoffs.

Incumbents invest in R&D to improve product qualities. R&D investment is in units of final good, and R^X_{jit} units of final good generates an innovation with the arrival rate of $x_{jit} \in [0, 1]$. Following the Schumpeterian growth literature (e.g. Acemoglu et al. 2018, Akcigit and Kerr 2018), I consider a quadratic innovation cost function in the form of

$$R^X_{jit} = \frac{\nu_x}{2} x^2_{jit}, \tag{21}$$

where ν_x is the scale parameter of the incumbent's innovation cost function. Differing from capital investment that uses both internal fund and bank loan as funding sources, R&D investment is assumed to be only financed by firms' internal fund retained after production. This assumption is based on the intangible nature of the asset created in innovation activities as well as the associated high degree of uncertainty, which makes R&D investment lack of collateral that could be used as security for debt funding. It is supported by the empirical literature that firms tend to use internal funds over external ones when financing innovation projects (e.g. Himmelberg et al. 1994, B. H. Hall et al. 2009).

3.3.4 Entry and Exit

Before the firm production and innovation process, there is an entry stage in which a potential entrant in each market invests in R&D to enter the business. If the potential entrant generates a successful innovation, it enters the market, replaces the technological follower and starts producing. With probability $\phi \in [0, 1]$, the entrant's innovation is radical which allows it to catch up with the leader's quality. With the rest probability $1 - \phi$, the innovation is incremental such that it stays the same quality level with the initial follower. In case the new firm enters a neck-and-neck market, it replaces one of the two incumbents with equal probability. If the innovation fails, the potential entrant simply disappears.

Similar to incumbents' innovation, the potential entrant employs an innovation cost function that is quadratic in the innovation arrival rate:

$$R^Z_{it} = \frac{\nu_z}{2} z^2_{it} \tag{22}$$

where R_{it}^Z denotes R&D investment by entrant in market i at time t, z_{it} is the innovation arrival rate, and ν_z is the scale parameter of the entrant's innovation cost function.

Entrants' R&D investment are financed with their asset endowment. At the beginning of each period, entrants draw their initial endowment a_{eit} from a time invariant uniform distribution $U(\bar{a}_e - \frac{1}{2}, \bar{a}_e + \frac{1}{2})$ and invest in innovation activities. Entrants succeed in innovation bring the endowment left over into future businesses, while failing entrants disappear with all their assets. The initial endowments are transferred from the representative household to potential entrants in a lump-sum form A_t^E as mentioned in household's problem. The uniform distribution assumption implies that $A_t^E = \bar{a}_e$.

Incumbent firms that are not replaced by the entrant face a constant exit risk $\xi \in [0, 1]$ at the end of each period. When an incumbent gets hit by the exit shock, it is replaced by a new firm that takes over the product line with the same quality level. The wealth of the deceased firm is distributed partly to the newborn firm as asset endowment and the left simply disappears. The portion of asset distributed is denoted as $\tau \in (0, 1)$. This shock captures many reasons that leads to incumbent exit or reorganization but are not directly related to their production and innovation activities, for example, negative demand shocks, or adverse financial shocks.

3.3.5 Market Competition and Firm Decisions

Incumbent firms in the intermediate markets compete strategically in production activities and innovation tasks. Potential entrants form rational expectations over future competition and decide on entry efforts. In this section, I present value functions that summarize strategic interactions among intermediate goods firms and characterize firm-optimal production, innovation and entry decisions. An intermediate market is fully characterized by $(n_{it}, a_{1it}, a_{2it}, a_{eit})$, where n_{it} is the technology gap between incumbents, a_{1it}, a_{2it} are the two incumbents' net assets, respectively, and a_{eit} is the potential entrant's asset endowment. I define the value of an incumbent firm $j \in \{1, 2\}$ in market i at the beginning of period t as $V_{jit}(n_{it}, a_{1it}, a_{2it}, a_{eit})$, and the value of the potential entrant as $V_{eit}(n_{it}, a_{1it}, a_{2it}, a_{eit})$. Subscript i and t are dropped for brevity in the remainder of the paper. These firm values are examined backwards, starting with the actions and payoffs in the static price competition and the dynamic innovation investment, then moving to the entry decision.

Production Decisions For any intermediate market in the production stage, there are two potential sets of firms involved depending on the innovation status of the potential entrant: (1) the entrant fails the innovation and the current incumbents continue production; (2) the entrant succeeds and replaces the technological follower in following production. In either case, only two firms engage actively in the production activity, and the payoff-relevant variables in the production stage are the technology gap plus the two firms' net assets. I denote the two actively producing firms as u and v with $(u, v) \in \{(1, 2), (1, e), (e, 2)\}$. The two firms compete à la Bertrand, in which they determine prices simultaneously according to the optimal pricing rule (12), taking into account the opponent's choice. The firms then employ capital and labor to realize the production plan which are scheduled when they set the prices. Figure 3a - 3d depict a firm's decisions on capital, labor and output price in different payoff-relevant states $(n_{it}, a_{uit}, a_{vit})$, as well as the implied borrowing amount in the capital investment choices. Capital and labor choices follow the four-region scheme in the firm's own asset as explained in Section 3.3.2, and they decrease in the competitor's asset because the opponent holds more market share as its asset increases. The capital choice implies a hump-shaped borrowing policy in the firm's own asset, where the firm initially increases borrowing to expand production and gradually reduces the amount as its financial constraint relaxes. Meanwhile, borrowing decreases in the competitor's asset as firm u is getting into a worse competitive position and produces less. Output price declines in both the firm's and the competitor's assets and becomes constant when the two firms have sufficiently large assets.

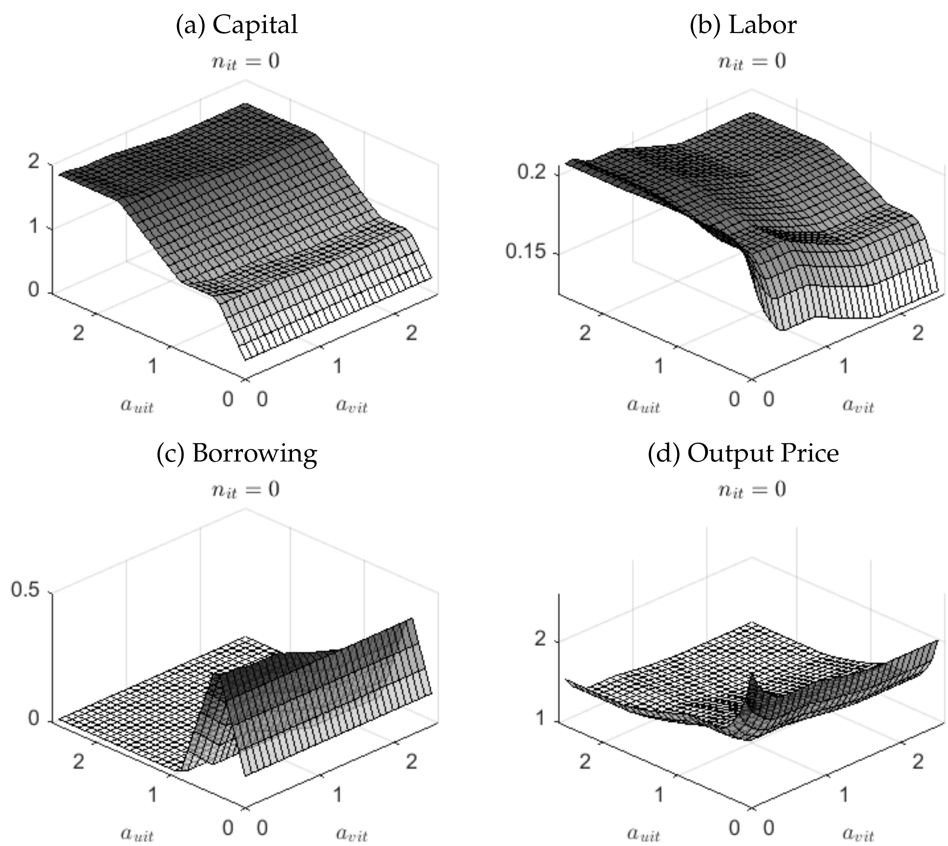


Figure 3: Production Decisions (Firm u)

Given the input choices and output price, the flow profit of firm $j \in \{u, v\}$ in the production stage can be characterized as:

$$\pi_{jit}(n_{it}, a_{uit}, a_{vit}) = p^*_{jit}(n_{it}, a_{uit}, a_{vit}) y^*_{jit}(n_{it}, a_{uit}, a_{vit}) - \int_0^{y^*_{jit}} c'(y_{jit}, a_{jit}) dy, \tag{23}$$

where $p^*_{jit}(n_{it}, a_{uit}, a_{vit})$ denotes the price that maximizes the profit of firm $j \in \{u, v\}$ and $y^*_{jit}(n_{it}, a_{uit}, a_{vit})$ is the corresponding output level implied by the demand schedule. Figure

4a illustrates how a firm's flow profit varies with the state variables. It shows that profit is strictly increasing in its own asset and decreasing in the competitor's asset. Markups are pinned down when firm set output prices. Figure 4b displays the a firm's markup in different payoff-relevant states. Markup depends positively on the firm's own asset and negatively on the competitor's, with a rate of change that identifies four regions in each dimension that corresponds to the four cases in the factor demand.

(a) Flow Profit (b) Markup

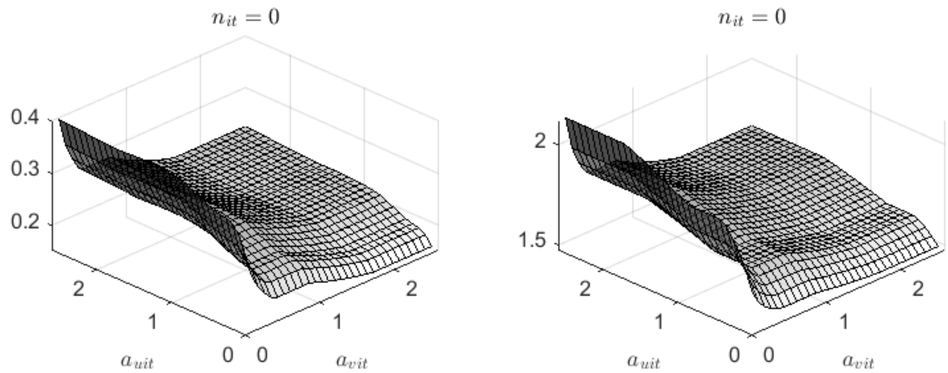


Figure 4: Production Profit and Markup (Firm u)

Innovation Decision The two actively producing firms continue competition in the innovation stage, each of which determines the optimal innovation arrival rate taking account of the opponent's innovation intensity. Define $V^P_{uvit}(n_{it}, a_{uit}, a_{vit})$ as the value of firm u at the start of the production stage when it competes with firm v, and $V^P_{vuit}(n_{it}, a_{uit}, a_{vit})$ the associated value of firm v. Subscript i and t are omitted for brevity in the following analysis. The set of value functions are given by:

$$
\begin{aligned}
V^P_{uv}(n_{it}, a_{uit}, a_{vit}) = & \max_{x_{uit}, a_{1it+1}} \pi_{uit}(n_{it}, a_{uit}, a_{vit}) - \frac{\nu_x}{2} x^2_{uit} \qquad (24) \\
& + \beta[(1-\eta)\cdot \mathbb{1}_{u=1} + 1 \cdot \mathbb{1}_{u=E}] \Big\{ [(1-x_{uit})(1-x_{vit}) + x_{uit}x_{vit}] \cdot \mathbb{E}_{a_e}\left[V_1(n_{it}, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] \\
& + x_{uit}(1-x_{vit}) \cdot \mathbb{E}_{a_e}\left[V_1(n_{it}+1, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] + (1-x_{uit})x_{vit} \cdot \mathbb{E}_{a_e}\left[V_1(n_{it}-1, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] \Big\}, \\
V^P_{vu}(n_{it}, a_{uit}, a_{vit}) = & \max_{x_{vit}, a_{2it+1}} \pi_{vit}(n_{it}, a_{uit}, a_{vit}) - \frac{\nu_x}{2} x^2_{vit} \qquad (25) \\
& + \beta[(1-\eta)\cdot \mathbb{1}_{v=2} + 1 \cdot \mathbb{1}_{v=E}] \Big\{ [(1-x_{uit})(1-x_{vit}) + x_{uit}x_{vit}] \cdot \mathbb{E}_{a_e}\left[V_2(n_{it}, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] \\
& + x_{uit}(1-x_{vit}) \cdot \mathbb{E}_{a_e}\left[V_2(n_{it}+1, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] + (1-x_{uit})x_{vit} \cdot \mathbb{E}_{a_e}\left[V_2(n_{it}-1, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] \Big\}, \\
s.t. \quad & a_{1it+1} = a_{uit} + \pi_u(n_{it}, a_{uit}, a_{vit}) - \frac{\nu_x}{2} x^2_{1it}, \quad a_{1it+1} > 0, \quad 0 \le x_{uit} \le 1, \\
& a_{2it+1} = a_{vit} + \pi_v(n_{it}, a_{uit}, a_{vit}) - \frac{\nu_x}{2} x^2_{2it}, \quad a_{2it+1} > 0, \quad 0 \le x_{vit} \le 1.
\end{aligned}
$$

The first term on the right-hand side of each value function represents the flow profit from production, and the second term is the R&D investment. The rest terms capture the continuation value as a result of changes in firms' technological positions. The bracket next to the discount factor displays the possibility of exogenous exit if the incumbent is not replaced by the entrant. The asset constraints describe the evolution of firms' net assets, which accumulate in the production stage as they receive profits and diminish in the innovation stage as they use internal fund to finance innovation.

The two competing firms decide on their innovation arrival rates simultaneously as best response to each other. I prove in Appendix B.3 that there exists at least one pair of innovation arrival rates (x_{uit}, x_{vit}) emerging as equilibrium.[10] Figure 5 displays a firm's innovation policy in different payoff-relevant states. It is shown that the optimal innovation rate crucially depends on the firm's technological position. When the firm is the technological follower in the market, the innovation rate increases with its own asset while decreases with the competitor's asset. The firm suppresses innovation effort when the competitor is in a better asset position. In the case of neck-and-neck market, the firm shows stronger innovation intensity than in the follower case, with the hope to escape from the competition and become the leader in the market. When the firm takes the technological leadership, it alters the innovation strategy to compete more intensively when the competitor owns more asset, the purpose of which is to hold its leading position.

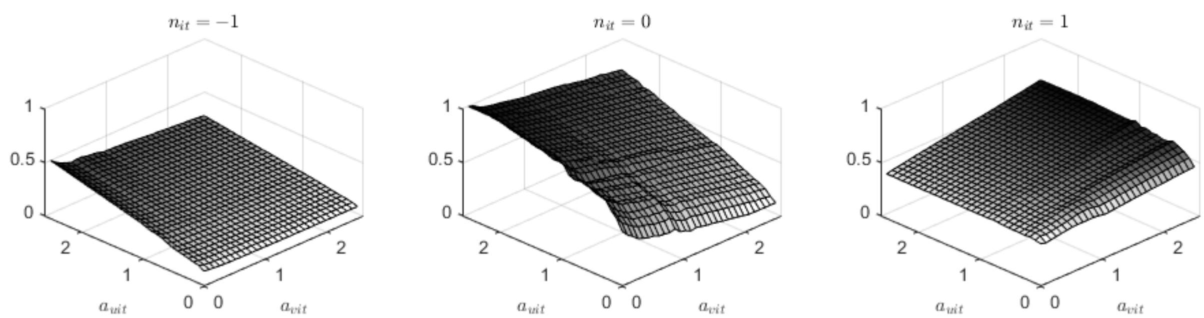


Figure 5: Innovation Rate (Firm u)

Entry Decision Entry occurs through successful innovation by potential entrants and it is directed at a specific intermediate good market. A successful entrant replaces the technological follower in the market, or drives either incumbent out of business with equal probability if it enters a neck-and-neck market. Hence the potential entrant forms expectation on future payoff and chooses over the innovation intensity, which can be equivalently interpreted as the firm's entry rate. When the technology gap between the two incumbents is positive, i.e. $n_{it} > 0$, the value functions for the

[10] For the case of multiple equilibria, I pick the pair of innovation arrival rates that delivers the highest firm values.

potential entrant and the two incumbents are given by:

$$\begin{aligned} &V_e(n_{it}, a_{1it}, a_{2it}, a_{eit}) \\ &\quad = \max_{z_{it}} -\frac{\nu_z}{2} z_{it}^2 + z_{it} \left[(1-\phi) V_{e1}^P(n_{it}, a_{1it}, a_{eit} - \frac{\nu_z}{2} z_{it}^2) + \phi V_{e1}^P(0, a_{1it}, a_{eit} - \frac{\nu_z}{2} z_{it}^2) \right], \end{aligned} \tag{26}$$

$$\begin{aligned} &V_1(n_{it}, a_{1it}, a_{2it}, a_{eit}) \\ &\quad = z_{it} \left[(1-\phi) V_{1e}^P(n_{it}, a_{1it}, a_{eit} - \frac{\nu_z}{2} z_{it}^2) + \phi V_{1e}^P(0, a_{1it}, a_{eit} - \frac{\nu_z}{2} z_{it}^2) \right] + (1 - z_{it}) V_{12}^P(n_{it}, a_{1it}, a_{2it}), \end{aligned} \tag{27}$$

$$V_2(n_{it}, a_{1it}, a_{2it}, a_{eit}) = (1 - z_{it}) V_{21}^P(n_{it}, a_{1it}, a_{2it}), \tag{28}$$

$$s.t. \qquad a_{eit} - \frac{\nu_z}{2} z_{it}^2 > 0.$$

Incumbent 1 is the technological leader in this case, and the potential entrant replaces incumbent 2 if it successfully enters the market. In the entrant's value function, the first term represents the innovation cost and the rest terms capture the continuation value when it enters the business, in which the entrant may bring a radical innovation or an incremental one. The incumbents' value functions summarize their continuation values as a result of the entry attempt. When the intermediate is neck-and-neck, i.e., $n_{it} = 0$, the value functions can be written as:

$$V_e(0, a_{1it}, a_{2it}, a_{eit}) = \max_{z_{it}} -\frac{\nu_z}{2} z_{it}^2 + \frac{z_{it}}{2} V_{e1}^P(0, a_{eit} - \frac{\nu_z}{2} z_{it}^2, a_{2it}) + \frac{z_{it}}{2} V_{e2}^P(0, a_{eit} - \frac{\nu_z}{2} z_{it}^2, a_{2it}), \tag{29}$$

$$V_1(0, a_{1it}, a_{2it}, a_{eit}) = \frac{z_{it}}{2} V_{1e}^P(0, a_{1it}, a_{eit} - \frac{\nu_z}{2} z_{it}^2) + (1 - z_{it}) V_{12}^P(0, a_{1it}, a_{2it}), \tag{30}$$

$$V_2(0, a_{1it}, a_{2it}, a_{eit}) = \frac{z_{it}}{2} V_{2e}^P(0, a_{eit} - \frac{\nu_z}{2} z_{it}^2, a_{2it}) + (1 - z_{it}) V_{21}^P(0, a_{1it}, a_{2it}), \tag{31}$$

$$s.t. \qquad a_{eit} - \frac{\nu_z}{2} z_{it}^2 > 0.$$

The value functions when an entrant aims for an intermediate market in the neck-and-neck state are similarly defined, except that either incumbent is equally likely to be replaced and any successful entry does not improve on the current product quality. When $n_{it} < 0$, incumbent 1 is the technological follower and it is driven out of business when entry occurs. The value functions are symmetric to the case where $n_{it} > 0$, with the two incumbents' roles flipped:

$$\begin{aligned} &V_e(n_{it}, a_{1it}, a_{2it}, a_{eit}) \\ &\quad = \max_{z_{it}} -\frac{\nu_z}{2} z_{it}^2 + z_{it} \left[(1-\phi) V_{e2}^P(n_{it}, a_{eit} - \frac{\nu_z}{2} z_{it}^2, a_{2it}) + \phi V_{e2}^P(0, a_{eit} - \frac{\nu_z}{2} z_{it}^2, a_{2it}) \right], \end{aligned} \tag{32}$$

$$V_1(n_{it}, a_{1it}, a_{2it}, a_{eit}) = (1 - z_{it}) V_{12}^P(n_{it}, a_{1it}, a_{2it}), \tag{33}$$

$$\begin{aligned} &V_2(n_{it}, a_{1it}, a_{2it}, a_{eit}) \\ &\quad = z_{it} \left[(1-\phi) V_{2e}^P(n_{it}, a_{eit} - \frac{\nu_z}{2} z_{it}^2, a_{2it}) + \phi V_{2e}^P(0, a_{eit} - \frac{\nu_z}{2} z_{it}^2, a_{2it}) \right] + (1 - z_{it}) V_{21}^P(n_{it}, a_{1it}, a_{2it}), \end{aligned} \tag{34}$$

$$s.t. \qquad a_{eit} - \frac{\nu_z}{2} z_{it}^2 > 0.$$

The potential entrant decides on its entry rate taking into account the future competitive

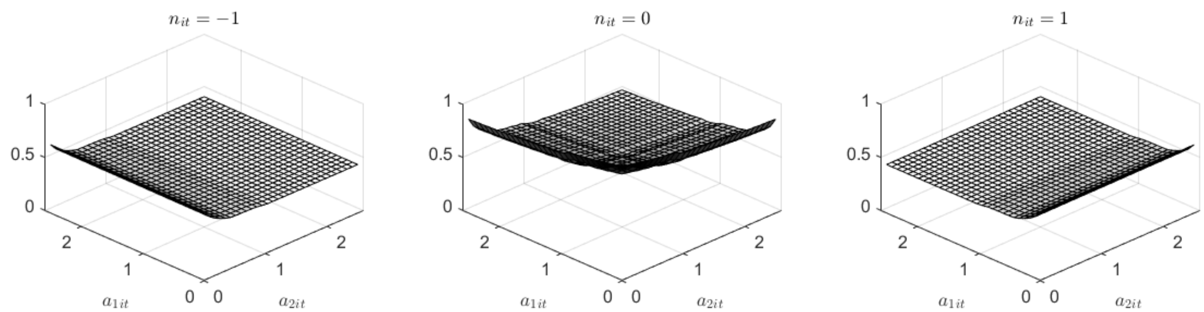


Figure 6: Entry Rate

pressure. Figure 6 shows the optimal entry policy for a potential entrant (fixed a_{eit}) when it faces intermediate markets with different payoff-relevant states $(n_{it}, a_{1it}, a_{2it})$. Entry effort is generally more intensive when the technological leader owns smaller amount of assets, as the entrant will face less fierce competition after entry.

3.4 Banks

The economy is populated by a finite number of banks indexed by $s \in \{1, 2, ..., S\}$. Banks collect deposits and issue loans to intermediate goods firms. The deposit market is assumed to be in perfect competition, while the loan market is modeled in a context of Cournot competition in which each bank determines its loan amount taking the loan quantities chosen by other banks as given. In each period, a bank chooses the deposit and loan amount to maximize its profit given by:

$$\pi^B_{st}(D_{st}, T_{st}) = r^L_t(T_t)T_{st} - r^D_t D_{st} - C(D_{st}, T_{st}), \tag{35}$$

where D_{st} is the amount of deposit bank s demands at time t, T_{st} is the amount of loan bank s offers at time t, and $T_t = \sum_{s=1}^{S} T_{st}$ denotes the aggregate loan quantity. The loan rate r^L_t is a function of T_t as each bank determines its loan amount taking into account the effect of its own choice on the loan market equilibrium through the aggregate loan quantity and the loan rate. $C(D_{st}, T_{st})$ is the bank's cost associated with the provision and management of deposits and loans. Following Klein (1971) and Monti (1972), I assume the cost function takes an additively separable form:

$$C(D_{st}, T_{st}) = \gamma^D_s D_{st} + \gamma^L_s T_{st}, \tag{36}$$

where $\gamma^D_s, \gamma^L_s \in [0, 1]$ are bank specific unit-cost parameters associated with the deposit and loan activities, respectively. I further assume banks have zero capital for simplicity, thus a bank faces the following balance sheet constraint in which its loan amount equals the deposit amount:

$$D_{st} = T_{st}. \tag{37}$$

Optimal bank decisions yield the following expression for the loan rate:[11]

$$\frac{r_t^L - r_t^D - \bar{\gamma}}{r_t^L} = \sum_{s=1}^{S} \left(\frac{T_{st}}{T_t}\right)^2 \cdot \frac{1}{\epsilon_T} \equiv h \cdot \frac{1}{\epsilon_T} \tag{38}$$

where $\bar{\gamma} = \sum_{s=1}^{S} \left[\frac{T_{st}}{T_t}(\gamma_s^D + \gamma_s^L)\right]$ is the weighted-average marginal cost of managing deposits and loans, $\sum_{s=1}^{S} \left(\frac{T_{st}}{T_t}\right)^2$ is the bank HHI measured by loan share and I define it as h for brevity, and $\epsilon_T = -\frac{\partial T_t/T_t}{\partial r_t^L/r_t^L}$ is the elasticity of aggregate loans. This condition shows that the Lerner index representing market power of banks is positively related with the bank concentration level. Higher concentration indicates an increase in banks' market power, as well as the loan spread faced by firms in the intermediate goods market.

3.5 Equilibrium

After describing the problems of the four type of agents, now it is ready to define the equilibrium of the model economy. My analysis focuses on the Balanced Growth Path (BGP) equilibrium of the model. A *BGP equilibrium* is defined as a set of allocations $\{C_t, B_t, Y_t, y_{it}, y_{jit}, k_{jit}, l_{jit}, x_{jit}, z_{it}, D_{st}, T_{st}\}_{i\in[0,1], j\in\{u,v\}, s\in\{1,2,...,S\}, t\in\mathbb{Z}}$, prices $\{p_{it}, p_{jit}, w_t, r_t^D, r_t^L\}_{i\in[0,1], j\in\{u,v\}, t\in\mathbb{Z}}$, and a distribution $f(n_{it}, a_{uit}, a_{vit})$ such that, for $\forall t$:

1. Given prices $\{w_t, r_t^D\}$, the representative household chooses $\{C_t, B_{t+1}\}$ to maximize utility (4) subject to the budget constraint (5);
2. Given price $\{p_{it}\}$, the representative final good producer chooses $\{y_{it}\}$ to maximize profit;
3. Given prices $\{w_t, r_t^D, r_t^L\}$, intermediate goods producers choose $\{p_{jit}, k_{jit}, l_{jit}\}$ to maximize flow profit in the production stage subject to the market demand (10);
4. Intermediate goods producers choose $\{x_{jit}, z_{it}\}$ to satisfy the Bellman equations (24) - (32);
5. Given price $\{r_t^D\}$, banks choose $\{D_{st}, T_{st}\}$ to maximize profits (35);
6. Y_t is as given in (7), and the final goods market clears;
7. Wage rate w_t clears the labor market;
8. Deposit rate r_t^D satisfies the household's Euler equation (6);
9. Loan rate r_t^L clears the loan market;
10. The distribution over technology gap and assets $f(n_{it}, a_{uit}, a_{vit})$ is stationary over time

[11]Individual bank's optimization yields an equation that expresses the optimal loan rate in terms of the bank's market share and the elasticity of loan demand: $\frac{r_t^L - r_t^D - (\gamma_S^D + \gamma_s^L)}{r_t^L} = -\frac{\partial r_t^L/r_t^L}{\partial T_t/T_t} \cdot \left(\frac{T_{st}}{T_t}\right)$. Weighting each bank's optimization condition by its loan share and then aggregating yields expression (38).

11. Aggregate variables grow at a constant rate

$$g = e^{X \ln \lambda} - 1, \quad \text{where } X = \int (x_{uit} + x_{vit} + \phi n_{it} z_{it}) f(n, a_{uit}, a_{vit}) d(n_{it}, a_{uit}, a_{vit}).$$

This completes the description of the model. I solve the model numerically and conduct comparative statics in the next section to shed light on the mechanisms through which bank concentration leads to less firm competition and more concentrated market structure in the product markets. The key of the mechanism lies in firms' heterogeneous responses to the loan spread change induced by bank concentration. Small firms rely more heavily on bank financing than large firms due to the financial constraint, thus they are more adversely impacted by a loan spread increase. Changes in strategic competition among product market firms in response to the increased financing cost amplify the market reallocation towards large firms. Firm entry is deterred following the market structure change as potential entrants expect more fierce competition from dominant firms after entry. The responses by incumbents and potential entrants bring down the competition dynamics in the economy and shift the market composition to more concentrated industries.

4 Quantitative Analysis

The ultimate aim in this paper is to understand the mechanisms by which bank concentration influences product market competition and firm dynamics, and to quantify the importance of bank concentration in driving the U.S. business dynamism. To this end, I calibrate the BGP equilibrium of the model to match data moments of the U.S. economy during 1994-2004 as a low bank concentration period ("1990s"). Using this baseline calibration, I impose a change on bank HHI to develop intuition on the channels through which bank concentration reshapes product market competition and firm dynamics. Then I re-calibrate the model parameters for the 2005-2015 period ("2000s") during which bank concentration level is high, in order to infer changes to the economy between the two periods and to quantify the relative importance of bank concentration among some key potential drivers of the changing U.S. business dynamism.

4.1 Calibration for the 1990s

Fourteen structural parameters define the BGP equilibrium of the model economy: $\{\beta, \sigma, \alpha, \delta, \theta, \lambda, \nu_x, \nu_z, \phi, \bar{a}_e, \xi, \tau, \bar{\gamma}, h\}$. Six parameters are calibrated outside the model. On the household side, I take the discount factor $\beta = 0.97$. In combination with the calibrated growth rate of the economy, this discount factor results in a real deposit rate of 4.82%, close to the average real interest rate of 5.11% from the World Bank Open Data. On the firm side, I set the capital share to $\alpha = 1/3$ and the capital depreciation rate to $\delta = 0.06$, which are commonly adopted in the growth literature. The collateral constraint θ is calibrated externally to match the average liability to non-financial asset ratio of 0.69 for the U.S. non-financial sector as suggested in Buera and Nicolini (2020). The model implied firm liability to asset ratio is bounded above by $\theta - 1$, hence I set the collateral constraint to $\theta = 1.69$. For the probability of an entrant's innovation being radical, I take the value $\phi = 0.0423$, following the previous work by Akcigit and Ates (2019). The average entrant asset endowment is set to $\bar{a}_e = \frac{1}{2}$ for normalization. On the bank side, bank HHI is calculated to be $h = 0.0545$, based on the largest fifty banks' asset from the Call Report.

I jointly calibrate the rest of the parameters $\{\sigma, \lambda, \nu_x, \nu_z, \xi, \tau, \bar{\gamma}\}$ to a set of seven data targets that are informative about the key features of the model. The first target I include is the average firm markup, calculated based on Compustat data following the production approach in De Loecker and Warzynski (2012) and De Loecker, Eeckhout, et al. (2020). The average firm markup informs the calibration about the within-industry elasticity of substitution σ. The second target I use is the annual utilization-adjusted total factor productivity (TFP) growth rate obtained from the Federal Reserve Bank of San Francisco database (see Fernald 2014), which helps discipline the innovation step size λ. I also include the average aggregate R&D spending to GDP ratio to capture information on the incumbents' R&D cost scale parameter ν_x, for which the data are available from the Federal Reserve Economic Data (FRED). The next two data targets are obtained from the U.S. Census Bureau's Business Dynamics Statics (BDS). I employ the average firm entry rate in BDS to calibrate the scale parameter of the entrants' R&D cost function ν_z. I also use the average firm exit rate to inform about the probability of exit shock ξ. Firms in the model could be driven out of business due to the entry replacement or the exit shock. Hence, the model-generated firm exit rate consists of an endogenous part depending on the entrants' entry decisions and an exogenous part determined by the exogenous parameter ξ. The sixth data moment I target is the markup ratio between large and small firms. Large firms are defined as the ones in the top 40 percent of the asset distribution in each industry and year, and small firms are those in the bottom 60 percent. This asset-based threshold distinguishing large and small firms is chosen closely following the empirical findings in Section 2.3.2, which point to a negative effect of bank concentration on the average markup by firms in the bottom three quintiles and an opposite effect on firms in the top two asset quintiles. The target of small firm markup is informative about the fraction of asset distributed from a deceased firm to its replacing one, τ, in the event of firm exit. The idea is that net asset, on top of product quality, determines a firm's market share and thus markup. Replacing firms take over production with the initial quality level, hence the only factor governs their markup is the starting asset, which directly depend on parameter τ. As most of them start as the smaller firm in the intermediate market, the parameter τ influences the average markup of small firms and thus the markup ratio. The last target I include is the average loan spread calculated based on DealScan data. Given that the average bank HHI calculated based on top 50 banks' asset equals 0.0545 and the deposit rate plus the aggregate loan elasticity are endogenously determined within the model, the average loan spread helps back out the cost of managing deposits and loans $\bar{\gamma}$ according to equation (38).

For the seven parameters to be calibrated within the model, I compare model-generated moments to data targets, and choose parameter values that minimize a sum of the distance between model and data moments as in Acemoglu et al. (2018):

$$\sum_{k=1}^{7} \frac{|\text{model}(k) - \text{data}(k)|}{\frac{1}{2}|\text{model}(k)| + \frac{1}{2}|\text{data}(k)|}.$$

Table 4 summarizes the calibrated parameters and the fit of model to data. The model is fairly successful in matching key data moments, except for some overestimation in the R&D share of GDP. This can be attributed to the fact that in the model productivity growth is purely due to R&D, while in the reality productivity may be affected by other factors, e.g. human capital or management practices. The calibration results suggest that the condition of the model economy replicates well the state of the U.S. economy during the 1994-2004 period.

Table 4: List of Parameter Values for 1990s and Model Fit

	Description	Value	Target (Data Source)	Data	Model
A. Parameters Calibrated Within the Model					
σ	Elasticity of substitution within industry	6.098	Average firm markup (Compustat)	1.3984	1.3985
λ	Innovation step size	1.081	TFP growth (Fernald 2014)	1.68%	1.68%
ν_x	Incumbent innovation cost scale	1.519	R&D share of GDP (FRED)	2.66%	3.42%
ν_z	Entrant innovation cost scale	9.736	Firm entry rate (BDS)	10.52%	10.52%
ξ	Probability of exit shock	0.050	Firm exit rate (BDS)	10.03%	10.03%
τ	Fraction of asset distributed in firm exit	0.652	Large and small firm markup ratio (Compustat)	1.0480	1.0479
$\bar{\gamma}$	Cost of managing deposits and loans	0.006	Commercial Loan Spread (DealScan)	2.07%	2.07%
B. Parameters Calibrated Outside the Model					
β	Discount factor	0.97	Real interest rate (World Bank)		
α	Capital share	1/3	Set exogenously		
δ	Depreciation rate	0.06	Set exogenously		
θ	Collateral constraint	1.69	Buera and Nicolini (2020)		
ϕ	Probability of radical entrant innovation	0.0423	Akcigit and Ates (2019)		
$\bar{a}_e$	Average entrant asset endowment	0.50	Normalization		
h	Bank HHI	0.0545	Asset-based HHI of largest fifty Banks (Call Report)		

4.2 Understanding the Aggregate Impact of Bank Concentration

In this section, I conduct comparative statics in which a shock of higher bank concentration is introduced on the 1990s' BGP, and present the responses of model-based product market variables in order to understand the implications of bank concentration on product market behavior. This analysis also helps demonstrate the ability of bank concentration to account for some of the important empirical trends observed in the U.S. economy.

My approach is to impose a one-time shock on the bank HHI in equation (38), representing an exogenous change on the degree of concentration in the banking system. I consider a positive move in model bank HHI in line with the data counterpart, where the new bank HHI is set at 0.0942 as the average asset-based HHI of the top fifty banks during the 2005-2015 period. I track the model responses and Table 5 reports the results. It shows the model-generated changes in several important variables regarding aggregate economic performance, and compares them with the observed data counterparts.

Table 5: Counterfactual Results

	Data			Model			Sensitivity to Bank HHI	Explained data change
	1990s	2000s	Change	1990s	Counterfactual	Change		
Product market HHI	0.8130	0.8498	↑	0.5071	0.5093	↑	0.05%	5.86%
Firm entry rate	10.52%	9.37%	↓	10.52%	10.08%	↓	-0.11%	38.21%
Firm exit rate	10.03%	9.42%	↓	10.03%	9.82%	↓	-0.05%	34.21%
Large vs. small firm markup	1.0480	1.0670	↑	1.0479	1.0588	↑	0.27%	56.97%
Average firm markup	1.3984	1.4771	↑	1.3985	1.4003	↑	0.05%	2.38%
TFP growth	1.68%	0.33%	↓	1.68%	1.64%	↓	-0.01%	2.40%

Notes: This table reports model-generated change in variables of interest, and compares them with the empirical counterpart. Upward arrows indicate an increase in the variable of interest, and downward arrows indicate a decline. Product market HHI in data is calculated using Compustat statistics as the squared sales share of firms in the top 40 percent of the asset distribution plus the squared sales share of firms in the bottom 60 percent. Sensitivity analysis shows the percentage change in each model-generated variable in response to a 1% increase in bank HHI. Explained data change measures the share of the contribution from bank concentration to the total data change between 1990s and 2000s.

The model suggests that bank concentration could generate important effects on product market structure and firm dynamics, including a sharp increase in markup ratio between large and small firms, remarkable declines in firm entry and exit rates, and a modest increase in product market concentration. The model has remarkable success in generating reasonable variation in all margins with the directions suggested by the data. Comparing the model-generated changes with their empirical counterpart, bank concentration explains about 57%, 35% and 6% of observed changes in relative markups, firm entry and exit rate and product market concentration, respectively. In addition, the calibrated model predicts a marginal decline in aggregate productivity growth, which accounts for about 2% of the productivity slowdown in the U.S. economy since the 1990s.

The crux of these model-predicted effects lies in firms' heterogeneous responses to the loan spread change induced by bank concentration. Bank concentration raises the loan spread imposed on the product market firms, which reflects the rise of bank market power given the less competitive conditions in the loan market. On one hand, small firms incur higher cost on external financing and dampen debt choices as a result. While higher loan spread hurts all firms using bank loans, firms with fewer internal assets are more adversely impacted as they are more reliant on external credit in financing capital investment. As a supporting evidence, the average debt ratio between the small and the large firm in an intermediate market reduces from 23.81 in the 1990s BGP to 16.46 in the counterfactual exercise. On the other hand, the opportunity cost of internal funds decreases and large firms with sufficient assets expand capital investment as a response. The increase in loan spread favors large firms in production while hurts small firms, thus resulting in larger masses of product markets across relatively greater market share differences. This shift induces higher market concentration, wider gap between large and small firms' markups, and higher average firm markup in the product markets. This production reallocation lowers innovation incentives for all firms. Small firms reduce innovation activity with less hope to escape competition. Large firms innovate less intensively as they are facing lower competitive pressure. Overall innovation is dampened and thus aggregate productivity growth. Moreover, the process of new firms taking over production slows down as they expect lower chance to grow after entry, bringing down the overall dynamism decreases in the economy and reinforcing the dominate role of large firms.

4.3 Re-calibration for the 2000s

The U.S. economy has witnessed many fundamental changes since the late 1990s which result in a rise in product market concentration, a shift of power balance to dominant firms, a decline in business dynamism and a slowdown in productivity growth. Recent research has been making attempts to understand potential forces driving these trends. This paper contributes to this growing literature by providing an explanation from the bank concentration perspective. In order to investigate the roles of different channels suggested in the literature and to understand the strength of bank concentration in explaining the empirical trends, a decomposition analysis is conducted in this section. I calibrate the model to reflect the changes that the U.S. economy has been experiencing between the 1990s and the 2000s, and then decompose the contribution of each channel of interest to quantify their potential importance in driving the empirical regularities.

Five channels are considered within the scope of my theoretical framework. An drop in the innovation step size λ implies less quality improvement from each innovation, capturing the research output side hypothesis in Bloom et al. (2020) that the research productivity is declining. An increase in the incumbents' innovation cost scale parameter ν_x indicates that more R&D investment is needed to achieve the same innovation arrival rate, which matches the cost side of the hypothesis in Bloom et al. (2020) that ideas are getting harder to find. An increase in the entrants' innovation cost scale parameter ν_z captures various factors that could have resulted in increasing barriers to entry (Pugsley et al. 2019, Peters and Walsh 2020, Corhay et al. 2020) or the declining knowledge diffusion from incumbent firms to new entrants (Akcigit and Ates 2019). A decrease in the elasticity of substitution within industries σ implies increased market power of dominant firms as their produce more differentiated varieties, in line with De Loecker, Eeckhout, et al. (2020) and Edmond et al. (2018). Finally, an increase in bank concentration ratio h captures the bank concentration channel proposed in this paper.

The channels I consider have been moving simultaneously over the years. In order to correctly gauge the contribution of each channel to the observed trends, I first calibrate shock path for each channel that will jointly allow the model to replicate the salient feature of the U.S. economy in the 2000s. Then I shut down each channel one at a time on the 2000s' model equilibrium to quantify the contribution of each specific force. In the model calibration for the 2000s, I hold all the externally calibrated parameters other than bank HHI constant at their 1990s' level and re-calibrate the seven internal parameters $\{\sigma, \lambda, \nu_x, \nu_z, \xi, \tau, \bar{\gamma}\}$ to target data moments which are informative about the U.S. economy during the 2005-2015 period. Bank HHI h during the 2000s equals 0.0942 as the average asset-based HHI of the largest fifty banks.

Table 6 summarizes the calibrated parameters for the 2000s and the fit of model to data. Com-

Table 6: List of Parameter Values for 2000s and Model Fit

	Description	Value	Target (Data Source)	Data	Model
σ	Elasticity of substitution within industry	5.289	Average firm markup (Compustat)	1.4771	1.4771
λ	Innovation step size	1.072	TFP growth (Fernald 2014)	0.33%	0.89%
ν_x	Incumbent innovation cost scale	3.558	R&D share of GDP (FRED)	2.82%	2.71%
ν_z	Entrant innovation cost scale	12.993	Firm entry rate (BDS)	9.37%	9.37%
ξ	Probability of exit shock	0.049	Firm exit rate (BDS)	9.42%	9.42%
τ	Fraction of asset distributed in firm exit	0.140	Large and small firm markup ratio (Compustat)	1.0670	1.0668
$\bar{\gamma}$	Cost of managing deposits and loans	0.004	Commercial Loan Spread (DealScan)	2.63%	2.63%

paring the 1990s and the 2000s BGPs, a few observations stand out. The elasticity of substitution within industries decrease modestly, implying a greater exercising of market leader pricing power now than in the past, in line with the findings in De Loecker, Eeckhout, et al. (2020) and Barkai (2020). Consistent with Bloom et al. (2020), the innovation quality drops while the innovation cost increases. Ideas are getting harder to find after coming into the 2000s, and the effectiveness of R&D also declines. The entry cost increases substantially, resulting in a decline in entry rate of new firms as suggested in Decker et al. (2016) and Peters and Walsh (2020). The probability of exit shock drops but the change is quite muted. The decline in firm exit rate is mainly driven by the slowdown in firm entry and replacement.

4.4 Role of Bank Concentration and Other Channels

Now turn to the decomposition exercise. I base on the BGP in the 2000s and shut down each of the five channels one at a time. Shutting down a specific channel means that the parameter governing the particular margin remains constant at the 1990s' BGP level. Therefore, each experiment obtains the hypothetical results that would have arisen had the specific channel remained unchanged over time. The resulting differences between the hypothetical results and the 2000s' BGP provide a measure of the relative magnitude of the contribution by the specific channel in driving the model responses. Specifically, I express the contribution of a channel ω to the total model-generated deviation on variable M between the two periods as follows:

$$contribution_{\omega} = \frac{M(\omega_{2000s}, \Omega_{2000s}) - M(\omega_{1990s}, \Omega_{2000s})}{M(\omega_{2000s}, \Omega_{2000s}) - M(\omega_{1990s}, \Omega_{1990s})},$$

where $M(\omega_{t_1}, \Omega_{t_2})$ denotes the variable M is generated under model equilibrium with the parameter governing the particular channel ω set at t_1 level and other parameters Ω at t_2 level. Table 7 presents the contributions of the five channels to each of the variables of interest. For better comparison, I normalize the contributions to each variable of interest such that all channels sum up to one. The results decisively highlight that the bank concentration channel succeeds in capturing all the variable changes in correct directions. Other channels have different implications within my theoretical framework and account for a meaningful part of the changes only in a limited number of variables.

Table 7: Decomposition Results

	1990s vs. 2000s change	Higher bank concentration	Lower research productivity	Higher R&D cost	Higher entry cost	Higher firm market power	Other
		$h \uparrow$	$\lambda \downarrow$	$\nu_x \uparrow$	$\nu_z \uparrow$	$\sigma \downarrow$	
Product market HHI	↑	59.13%	-13.93%	-3.41%	-12.97%	-51.10%	122.28%
Firm entry rate	↓	24.56%	-2.91%	-0.84%	149.29%	-48.76%	-21.33%
Firm exit rate	↓	22.53%	-2.67%	-0.77%	136.93%	-44.72%	-11.29%
Large vs. small firm markup	↑	64.41%	-10.29%	-8.33%	-15.32%	-35.89%	105.41%
Average firm markup	↑	2.96%	-0.67%	-0.18%	-0.64%	92.21%	6.31%
TFP growth	↓	2.35%	15.93%	85.24%	4.21%	-8.69%	0.95%

Notes: This table reports the share of the contribution from the specific channel to the total model-generated deviation between the 1990s and the 2000s. The contributions are normalized such that all channels sum up to one. Upward arrows denote an increase in the variable of interest from the 1990s to the 2000s, and downward arrows indicate a decline. Positive values mean that the specific channel moves the model-generated variable in the same direction with the change between the 1990s and 2000s, and negative values indicate opposite moves. A value larger than 100% means that the difference from the 1990s BGP is larger than the model-generated total variation between the 1990s and the 2000s.

The lower research productivity and the higher R&D cost play a similar qualitative role in determining the variables of interest since all incumbent firms expect less return from R&D investment which decreases incentives for innovation and dampens productivity growth. Laggard firms suppress innovation effort less compared to market leaders as they expect higher chance to attain the market leadership given the smaller quality gap or the lower innovation intensity by market leader. As laggard firms are likely to be the smaller one in the market, the relative markup between large and small firms decreases and market becomes less concentrated. The research productivity channel and the R&D cost channel can only capture the change in TFP growth in the correct direction.

The higher entry cost reduces firm entry and exit rates, the intuition of which is straightforward. Facing a diminished probability of being overtaken by new entrants, technological laggard firms grow into larger sizes. As laggard firms tend to be the smaller ones in product markets, this leads to a shrink in relative firm size between large and small firms. Markup ratio between large and small firms therefore drops. Product market concentration declines and average firm markup slightly decreases. When incumbent firms take less competition pressure from new firms, they respond by investing less in R&D. Aggregate productivity growth slows down as a result of the suppressed innovation effort by both entrants and incumbent firms. The entry cost channel captures the direction of changes in entry and exit rate and productivity growth, while lacks explaining power in variables regarding market concentration and power balance among competing firms.

The higher firm market power channel captures the increase in average firm markup while fails to match the direction of change in rest variables. This is because the model has the standard Schumpeterian feature that increased market power gives a greater incentive for competition. As the substitutability of products within an industry decreases, the firms' varieties are more differentiated and the market leader charges higher markup for the same level of quality differences. The increased profit given same quality gaps induces more innovation effort by laggard firms with the hope to take over the market leader, shrinking the markup gap between large and small firms. Firm entry is stimulated as entrants expect higher profitability after replacing current incumbents. Firm exit rate rises as a consequence and product markets become less concentrated. Since all firms try harder in innovation, growth rate increases.

Now look into the contributions from the five channels to each of the variables of interest. 59.13% of the rise in product market concentration stems from bank concentration. Rising firm market power leads to 51.10% of decline in concentration level, which roughly washes out the effect of bank concentration. The rest three channels all lead to less concentrated product markets, with a total contribution of 30.31%. Bank concentration and firm market power also play an important role in driving the change in relative markup between large and small firms, with bank concentration accounting for 64.41% of the increase and firm market power leading to 35.89% of the decrease. The other three channels together result in 33.94% of shrink in the large and small markup ratio. The decomposition suggests the greatest scope for higher entry cost in explaining the firm entry and exit decline. Higher entry cost alone accounts for 149.29% and 136.93% of the drop in firm entry and exit rate, respectively. Bank concentration explains about 20.00% of the decrease on the firm entry and exit, while firm market power leads to about 45.00% of increase on the two margins. The rest two channels have only marginal effects on firm entry and exit. The firm market power channel is the most important factor in the determination of the level of average firm markup, which explains 92.21% of the increase in it. Bank concentration only contributes 2.96% to the markup increase. In terms of the change in productivity growth, the research productivity channel and the R&D cost channel together explain almost the entire decline. Compared to these two technology-related channels that bring huge impact on economic growth, the decomposition results suggest only a very modest scope for the bank concentration channel in explaining the productivity slowdown.

5 Concluding Remarks

This paper studies the macroeconomic implications of concentration in the U.S. banking system following a decrease in regulation in the 1990s. I examine this process through the lens of an endogenous growth model. It features heterogeneous firms, variable product market structure and firm markups, realistic entry and exit dynamics and imperfect bank competition. The model suggests that bank concentration exerts differential effects on product market firms, shaping the competition dynamics in a way that favors large establishments and therefore leading to product market concentration. I structurally calibrate the model to match data on U.S. economy, and find that bank concentration causes quantitatively significant changes in variables reflecting the concentration level and firm dynamics in the product markets. Bank concentration accounts for 59.13% of the product market concentration, 24.56% of the firm entry rate decline, 22.53% of the firm exit rate decline and 64.41% of the enlarged markup ratio between large and small product market firms observed in the U.S. economy between the 1990s and the 2000s.

The findings in this paper are consistent with the idea that banks with market power erect an important financial barrier to entrants and small firms, perhaps in part to protect the profitability of their well-established borrowers (Cestone et al. 2003, Saidi et al. 2021). While the mechanism through which bank competition results in differentiated lending relationships remains to be explored in future studies, the insight that bank concentration has non-uniform impacts across firms updates the conventional wisdom that bank competition is either good or bad overall. Moreover, this paper finds only modest productivity loss in the bank concentration process. Quantitative analysis suggests that bank concentration explains 2.35% of the U.S. productivity slowdown since the late 1990s. This result provides an additional metric to weigh up the pros and cons of bank competition.

The policy implications associated with bank concentration are especially relevant. Bank market structure is a traditional policy variable that regulators across countries and over time attempt to influence. Low bank concentration and vigorous bank competition have obvious benefits including greater efficiency in the production of banking services, higher quality banking products and more innovation. This paper provides supporting evidence on these benefits of bank competition by evaluating the real effects on product markets. Nevertheless, financial stability is an important asset that comes with bank concentration. Less competition in the banking system could be beneficial as it increases bank profit margins and suppresses excessive risk taking (Beck et al. 2006). Preceding discussions are absent from the effect of bank concentration on financial stability as it involves different modeling approach regarding financial risks which is beyond the scope of this paper. Answers to trade-off between increased competition and financial stability will have policy implications in a wider economic context, which deserves further research.

References

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., & Kerr, W. (2018). Innovation, reallocation, and growth. *American Economic Review*, *108*(11), 3450–3491.

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., & Li, H. (2019). A theory of falling growth and rising rents. *NBER Working Paper*.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and innovation: An inverted-u relationship. *The Quarterly Journal of Economics*, *120*(2), 701–728.

Aghion, P., Fally, T., & Scarpetta, S. (2007). Credit constraints as a barrier to the entry and post-entry growth. *Economic Policy*, *22*(52), 733–779.

Aghion, P., Harris, C., Howitt, P., & Vickers, J. (2001). Competition, imitation and growth with step-by-step innovation. *The Review of Economic Studies*, *68*(3), 467–492.

Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, *60*(2), 323–351.

Akcigit, U., & Ates, S. T. (2019). What happened to u.s. business dynamism? *NBER Working Paper*.

Akcigit, U., & Kerr, W. (2018). Growth through heterogeneous innovations. *Journal of Political Economy*, *126*(4), 1374–1443.

Ates, S. T., & Saffie, F. E. (2021). Fewer but better: Sudden stops, firm entry, and financial selection. *American Economic Journal: Macroeconomics*, *13*(3), 304–356.

Atkeson, A., & Burstein, A. T. (2010). Innovation, firm dynamics, and international trade. *Journal of Political Economy*, *118*(3), 433–484.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Reenen, J. V. (2017). Concentrating on the fall of the labor share. *American Economic Review*, *107*(5), 180–185.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Reenen, J. V. (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, *135*(2), 645–709.

Barkai, S. (2020). Declining labor and capital shares. *The Journal of Finance*, *75*(5), 2421–2463.

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

Bloom, N., Jones, C. I., Reenen, J. V., & Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, *110*(4), 1104–1144.

Buera, F. J., Kaboski, J. P., & Shin, Y. (2011). Finance and development: A tale of two sectors. *American Economic Review*, *101*(5), 1964–2002.

Buera, F. J., & Moll, B. (2015). Aggregate implications of a credit crunch: The importance of heterogeneity. *American Economic Journal: Macroeconomics*, *7*(3), 1–42.

Buera, F. J., & Nicolini, J. P. (2020). Liquidity traps and monetary policy: Managing a credit crunch. *American Economic Journal: Macroeconomics*, *12*(3), 110–138.

Carey, M., & Hrycray, M. (1999). Credit flow, risk, and the role of private debt in capital structure. *Working Paper, Federal Reserve Board*.

Cestone, G., & White, L. (2003). Anti-competitive financial contracting: The design of financial claims. *The Journal of finance*, *58*(5), 2109–2141.

Cetorelli, N., & Strahan, P. E. (2006). Finance as a barrier to entry: Bank competition and industry structure in local u.s. markets. *The Journal of Finance*, *61*(1), 437–461.

Chatterjee, S., & Eyigungor, B. (2019). The firm size-leverage relationship and its implications for entry and business concentration. *Federal Reserve Bank of Philadelphia Working Paper*.

Chava, S., & Roberts, M. R. (2008). How does financing impact investment? the role of debt covenants. *The Journal of Finance*, *63*(5), 2085–2121.

Corbae, D., & D'Erasmo, P. (2020). Rising bank concentration. *NBER Working Paper*.

Corhay, A., Kung, H., & Schmid, L. (2020). Q: Risk, rents, or growth? *Working Paper*.

Cuciniello, V., & Signoretti, F. M. (2015). Large banks, loan rate markup, and monetary policy. *International Journal of Central Banking*, *11*(3), 141–177.

De Loecker, J., Eeckhout, J., & Unger, G. (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, *135*(2), 561–644.

De Loecker, J., & Warzynski, F. (2012). Markups and firm-level export status. *American Economic Review*, *102*(6), 2437–2471.

De Ridder, M. (2020). Market power and innovation in the intangible economy. *Working Paper*.

Decker, R. A., Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2016). Where has all the skewness gone? the decline in high-growth (young) firms in the u.s. *European Economic Review*, *86*, 4–23.

Deidda, L., & Fattouh, B. (2005). Concentration in the banking industry and economic growth. *Macroeconomic Dynamics*, *9*(2), 198–219.

Diallo, B., & Koch, W. (2018). Bank concentration and schumpeterian growth: Theory and international evidence. *The Review of Economics and Statistics*, *100*(3), 489–501.

Edmond, C., Midrigan, V., & Xu, D. Y. (2018). How costly are markups. *NBER Working Paper*.

Eggertsson, G. B., Robbins, J. A., & GetzWold, E. (2021). Kaldor and piketty's facts: The rise of monopoly power in the united states. *Journal of Monetary Economics*, *124*(Supplement), S19–S38.

Favara, G., & Giannetti, M. (2017). Forced asset sales and the concentration of outstanding debt: Evidence from the mortgage market. *The Journal of Finance*, *72*(3), 1081–1118.

Fernald, J. (2014). Productivity and potential output before, during, and after the great recession. *NBER Working Paper*.

Fernholz, R. T., & Koch, C. (2016). Why are big banks getting bigger? *Working Paper, Federal Reserve Bank of Dallas*.

Giannetti, M., & Saidi, F. (2019). Shock propagation and banking structure. *The Review of Financial Studies*, *32*(7), 2499–2540.

Grossman, G. M., & Helpman, E. (1991). Quality ladders in the theory of growth. *The Review of Economic Studies*, *58*(1), 43–61.

Grullon, G., Larkin, Y., & Michaely, R. (2019). Are us industries becoming more concentrated? *Review of Finance*, *23*(4), 697–743.

Gutiérrez, G., & Philippon, T. (2017). Declining competition and investment in the u.s. *NBER Working Paper*.

Hall, B. H., & Lerner, J. (2009). The financing of r&d and innovation. *NBER Working Paper*.
Hall, R. E. (2018). New evidence on the markup of prices over marginal costs and the role of mega-firms. *NBER Working Paper*.
Himmelberg, C. P., & Petersen, B. C. (1994). R&d and internal finance: A panel study of small firms in high-tech industries. *The Review of Economics and Statistics*, *76*(1), 38–51.
Janicki, H. P., & Prescott, E. S. (2006). Changes in the size distribution of u.s. banks: 1960-2006. *Federal Reserve Bank of Richmond Economic Quarterly*, *92*, 291–316.
Jiang, H., Zheng, Y., & Zhu, L. (2022). Entry barriers and growth: The role of endogenous market structure. *Working Paper*.
Karahan, F., Pugsley, B., & Sahin, A. (2019). Demographic origins of the startup deficit. *NBER Working Paper*.
Klein, M. A. (1971). A theory of the banking firm. *Journal of Money, Credit and Banking*, *3*(2), 205–218.
Klette, T. J., & Kortum, S. (2004). Innovating firms and aggregate innovation. *Journal of Political Economy*, *112*(5), 986–1018.
Lentz, R., & Mortensen, D. T. (2008). An empirical model of growth through product innovation. *Econometrica*, *76*(6), 1317–1373.
Levine, R. (2005). Finance and growth: Theory and evidence. *Handbook of Economic Growth*.
Liu, E., Mian, A., & Sufi, A. (2022). Low interest rates, market power, and productivity growth. *Econometrica*, *90*(1), 193–221.
Luttmer, E. G. J. (2007). Selection, growth, and the size distribution of firms. *The Quarterly Journal of Economics*, *122*(3), 1103–1144.
Midrigan, V., & Xu, D. Y. (2014). Finance and misallocation: Evidence from plant-level data. *The American Economic Review*, *104*(2), 422–458.
Moll, B. (2014). Productivity losses from financial frictions: Can self-financing undo capital misallocation? *American Economic Review*, *104*(10), 3186–3221.
Monti, M. (1972). Deposit, credit and interest rae determination under alternative bank objectives. *Mathematical Methods in Investment and Finance (North-Holland, Amsterdam)*, 431–454.
Olley, S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, *64*(6), 1263–1297.
Olmstead-Rumsey, J. (2020). Market concentration and the productivity slowdown. *Working Paper*.
Peters, M. (2020). Heterogeneous markups, growth and endogenous misallocation. *Econometrica*, *88*(5), 2037–2073.
Peters, M., & Walsh, C. (2020). Population growth and firm dynamics. *Working Paper*.
Pugsley, B. W., & Sahin, A. (2019). Grown-up business cycles. *The Review of Financial Studies*, *32*(3), 1102–1147.
Rajan, R. G., & Zingales, L. (1998). Financial dependence and growth. *The American Economic Review*, *88*(3), 559–586.
Saidi, F., & Streitz, D. (2021). Bank concentration and product market competition. *The Review of Financial Studies*, *34*(10), 4999–5035.

Schumpeter, J. A. (1911). *The theory of economic development*. Harvard University Press.
Schwert, M. (2018). Bank capital and lending relationships. *The Journal of Finance*, *73*(2), 787–830.
Syverson, C. (2017). Challenges to mismeasurement explanations for the us productivity slowdown. *Journal of Economic Perspectives*, *31*(2), 165–86.

A Data Appendix

A.1 Data Sample Descriptive Statistics

I construct a panel data sample consisting of measures on bank credit concentration, product market competition, firm entry rate and industry characteristics for 57 industries defined with SIC two-digit codes. Table 8 reports descriptive statistics on the constructed panel data sample. The sample period is 1990 to 2015.

Table 8: Descriptive Statistics

	Min	Median	Max	Mean	Std Dev	Obs
Average Firm Markup	0.78	1.29	3.05	1.36	0.31	1466
Bank HHI	0.08	0.22	1.00	0.26	0.14	1474
ln(Asset) (Asset in $Million)	0.18	10.55	14.26	10.44	1.72	1466
ln(Debt) (Debt in $Million)	-3.19	9.38	13.16	9.17	1.89	1466
Leverage	0.01	0.26	0.97	0.29	0.17	1466
Ind HHI	0.03	0.12	0.95	0.18	0.18	1466
Firm Entry Rate	0.03	0.10	0.24	0.10	0.03	1404

Notes: This table reports summary statistics on the panel data sample. Nominal terms, e.g. asset and debt, are deflated by CPI from FRED. Industries are defined with the SIC two-digit codes, with agriculture, utilities, finance, public administration and nonclassifiable industries excluded.

A.2 Firm Markup Estimation

According to De Loecker and Warzynski (2012), under a variety of pricing models, a firm's markup can be computed as a multiplication of the output elasticity of a variable input and the revenue share of the variable input:

$$\mu_{it} = \theta_{it}^{V} \frac{P_{it} Q_{it}}{P_{it}^{V} V_{it}},$$

where μ_{it} denotes firms i's markup at time t, θ_{it}^{v} is the output elasticity of the variable input, P_{it} is the firm's output price, Q_{it} is its output, P_{it}^{V} is the price of the variable input and V_{it} is the amount of input used. Therefore, estimating firm markups involves correctly gauging two key ingredients: the revenue share of the variable input $\frac{P_{it}^{V} V_{it}}{P_{it} Q_{it}}$ and the output elasticity of the variable input θ_{it}^{V}.

I rely on Compustat data from 1995-2015 to estimate firm markups. While Compustat only includes information on publicly listed U.S. firms, it provides substantial coverage of firms in the private sector over a long period of time. I restrict attention to firms incorporated in the U.S., and exclude observations in agriculture, utilities, finance, public administration and non-classifiable industries. The sample includes around 3000 firms per year, with this number varying over time. Following De Loecker, Eeckhout, et al. (2020), I use Compustat term "Cost of Goods Sold" (COGS) as the variable input cost of the firm and "Sales" (SALE) as the revenue. COGS bundles all expenses directly attributable to the production of the goods sold by the firm and includes materials and intermediate inputs, labor cost, energy and so on. The revenue share of the the variable input is thus calculated as $\frac{COGS}{SALE}$.

To obtain a measure of the output elasticity of a variable input of production, I estimate a parametric production function for each industry-year. Specifically, I assume a Cobb-Douglas production function, with a variable input bundle and capital as input, for firm j in a two-digit SIC industry i in year t so that factor share vary across industries and time:

$$Y_{jit} = A_{jit} V_{jit}^{\theta_{it}^{V}} K_{jit}^{\theta_{it}^{K}},$$

where Y_{jit} is the firm's output and I use Compustat term SALE adjusted by industry-level input price deflator (PIRIC from FRED) to measure it. V_{jit} is the variable input and I rely on Compustat term COGS as a measure for it. COGS is also deflated to obtain real value at the firm level. K_{jit} is the firm's capital stock and the deflated Compustat term "Total Gross Property, Plant and Equipment (PPEGT) is used as a measure for it. Then I conduct the following regression for each two-digit industry:

$$y_{it} = \alpha + \theta_{it}^{V} v_{it} + \theta_{it}^{K} k_{it} + \delta_i + \eta_t + \epsilon_{it},$$

where lowercase letters denote variables in log. In the above regression specification, θ_{it}^{V} captures the industry-year specific variable output elasticity. I use the markup equation to obtain the firm markup from the estimated $\hat{\theta}_{it}^{V}$ and the inverse cost share $\frac{SALE}{COGS}$.

Note that, to calculate the average firm markup in each industry, I trim firm observations at the 1st and 99th markup percentiles to reduce the impact of extreme values which are common for ratios in firm panels drawn from accounting data.

A.3 Regression Robustness Check

I conduct three sets of robustness checks on the panel regression (1) testing the relationship between bank concentration and product market competition in the U.S. economy. The first robustness check removes observations during the Great Recession period 2007-2009 as bank market structure and loan spread changed dramatically in this period. The second and third robustness check deals with the venture capital (VC) issue as firms could have venture capital as alternative financing resource and the effect of bank credit on product market competition could be systematically overestimated due to this. The second robustness check excludes industries which are likely to be targeted by venture capitals, including electronics (SIC codes 3500-3599), information technology (SIC codes 7300-7399), biotechnology (SIC codes 8000-8099) and engineering, research and management (SIC codes 8700-8799). The third robustness check deletes firm observations which are publicly listed within three years as firms are likely to hold venture capital during the first years after IPO. The

Table 9: Bank Concentration and Average Firm Markup

	Remove 2007-2009		Remove VC Target Industries		Remove VC Target Firms	
	(1)	(2)	(3)	(4)	(5)	(6)
Bank HHI	0.094***	0.104***	0.109***	0.114***	0.162***	0.151***
	(0.036)	(0.037)	(0.360)	(0.038)	(0.041)	(0.042)
ln(Asset)		−0.004		−0.002		−0.009
		(0.015)		(0.016)		(0.017)
ln(Debt)		0.038***		0.037***		0.045***
		(0.014)		(0.014)		(0.015)
Leverage		−0.078***		−0.130***		−0.230***
		(0.045)		(0.045)		(0.053)
Ind HHI		0.202***		0.265*		0.440*
		(0.051)		(0.051)		(0.053)
Constant	1.012***	0.714***	1.030***	0.728***	0.983***	0.664***
	(0.042)	(0.093)	(0.042)	(0.095)	(0.047)	(0.103)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.810	0.814	0.776	0.784	0.751	0.769
Obs	1298	1298	1284	1284	1280	1279

estimation results are displayed in the Table 9.

B Model Appendix

B.1 Derivation of Factor Demand and Marginal Cost

An intermediate market firm produce its variety of output using both capital and labor. Capital can be financed by both internal fund and bank loan, and depreciates at the rate δ. Banks charge a loan interest at the rate r_t^L, and pay a deposit interest at the rate r_t^D if a firm has asset surplus and save it to banks. Capital, net asset, bank loan and bank deposit are all in units of final good. Labor input is determined after finalizing the capital investment, and the relative price for labor is w_t. Given the production process and input prices, the cost minimization problem can be formalized as

$$\begin{aligned}
&\min_{l_{jit},k_{jit}} w_t l_{jit} + \delta k_{jit} - r_t^D(a_{jit} - k_{jit}) \cdot \mathbb{1}_{k_{jit} \leq a_{jit}} + r_t^L(k_{jit} - a_{jit}) \cdot (1 - \mathbb{1}_{k_{jit} \leq a_{jit}}) \\
s.t. \quad & y_{jit} = k_{jit}^{\alpha} l_{jit}^{1-\alpha}, \\
& k_{jit} \leq \theta a_{jit},
\end{aligned}$$

where y_{jit} is given. There are four potential cases depending on a firm's desired capital level and its net asset status.

1. $\left(\frac{\alpha}{1-\alpha}\frac{w_t}{\delta+r_t^D}\right)^{1-\alpha} y_{jit} \leq a_{jit}$
   In this case, the firm can finance its desired capital investment solely by internal fund and save the residuals to banks in the form of deposits. The problem can be reduced to:

$$\min_{l_{jit},k_{jit}} w_t l_{jit} + \delta k_{jit} - r_t^D(a_{jit} - k_{jit})$$
$$s.t. \qquad y_{jit} = k_{jit}^{\alpha} l_{jit}^{1-\alpha},$$

   which yields the standard factor demands under Cobb-Douglas production function:

$$k_{jit}(y_{jit}) = \left(\frac{\alpha}{1-\alpha}\frac{w_t}{\delta + r_t^D}\right)^{1-\alpha} y_{jit}, \quad l_{jit}(y_{jit}) = \left(\frac{1-\alpha}{\alpha}\frac{\delta + r_t^D}{w_t}\right)^{\alpha} y_{jit}.$$

   The optimal choices of capital and labor imply the total cost and the marginal cost to produce y_{jit} are given by

$$c_{jit}(y_{jit}) = w_t l_{jit} + \delta k_{jit} - r_t^D(a_{jit} - k_{jit}) = \left(\frac{\delta + r_t^D}{\alpha}\right)^{\alpha}\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha} y_{jit} - r_t^D a_{jit},$$
$$c'_{jit}(y_{jit}) = \left(\frac{\delta + r_t^D}{\alpha}\right)^{\alpha}\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha}.$$

2. $\left(\frac{\alpha}{1-\alpha}\frac{w_t}{\delta+r_t^D}\right)^{1-\alpha} y_{jit} > a_{jit}$ and $\left(\frac{\alpha}{1-\alpha}\frac{w_t}{\delta+r_t^L}\right)^{1-\alpha} y_{jit} \leq a_{jit}$
   In this case, capital investment is constrained if the firm finances it only through internal fund. In order to increase output, hiring more labor is cheaper compared to increasing capital investment via bank loan. Thus, the firm chooses capital up to a_{jit} and adjusts labor on the margin:

$$k_{jit}(y_{jit}) = a_{jit}, \quad l_{jit}(y_{jit}) = \left(\frac{y_{jit}}{a_{jit}^{\alpha}}\right)^{\frac{1}{1-\alpha}}.$$

   This implies the total cost and the marginal cost to produce y_{jit} are

$$c_{jit}(y_{jit}) = w_t l_{jit} + \delta k_{jit} = w_t \cdot \left(\frac{y_{jit}}{a_{jit}^{\alpha}}\right)^{\frac{1}{1-\alpha}} + \delta a_{jit},$$
$$c'_{jit}(y_{jit}) = \frac{w_t}{1-\alpha}\left(\frac{y_{jit}}{a_{jit}}\right)^{\frac{\alpha}{1-\alpha}}.$$

3. $a_{jit} < \left(\frac{\alpha}{1-\alpha}\frac{w_t}{\delta+r_t^L}\right)^{1-\alpha} y_{jit} \le \theta a_{jit}$
   In this case, the firm can cover the desired capital level through the additional fund from banks. The cost minimization problem can be reduced to:

$$\min_{l_{jit},k_{jit}} w_t l_{jit} + \delta k_{jit} + r_t^L(k_{jit} - a_{jit})$$
$$s.t. \qquad y_{jit} = k_{jit}^{\alpha} l_{jit}^{1-\alpha},$$

   which yields the factor demands

$$k_{jit}(y_{jit}) = \left(\frac{\alpha}{1-\alpha}\frac{w_t}{\delta + r_t^L}\right)^{1-\alpha} y_{jit}, \quad l_{jit}(y_{jit}) = \left(\frac{1-\alpha}{\alpha}\frac{\delta + r_t^L}{w_t}\right)^{\alpha} y_{jit}.$$

   The factor demands imply the total cost and the marginal cost to produce y_{jit} are

$$c_{jit}(y_{jit}) = w_t l_{jit} + \delta k_{jit} + r_t^L(k_{jit} - a_{jit}) = \left(\frac{\delta + r_t^L}{\alpha}\right)^{\alpha}\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha} y_{jit} - r_t^L a_{jit},$$
$$c'_{jit}(y_{jit}) = \left(\frac{\delta + r_t^L}{\alpha}\right)^{\alpha}\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha}.$$

4. $\left(\frac{\alpha}{1-\alpha}\frac{w_t}{\delta+r_t^L}\right)^{1-\alpha} y_{jit} > \theta a_{jit}$
   In this case, capital is constrained even the firm has access to bank loan. Therefore, the firm sets capital at its maximum level and adjust labor input at the margin

$$k_{jit}(y_{jit}) = \theta a_{jit}, \quad l_{jit}(y_{jit}) = \left[\frac{y_{jit}}{(\theta a_{jit})^{\alpha}}\right]^{\frac{1}{1-\alpha}}$$

   This implies the total cost and the marginal cost to produce y_{jit} are

$$c_{jit}(y_{jit}) = w_t l_{jit} + \delta k_{jit} + r_t^L(k_{jit} - a_{jit}) = w_t\left[\frac{y_{jit}}{(\theta a_{jit})^{\alpha}}\right]^{\frac{1}{1-\alpha}} + (\theta\delta + \theta r_t^L - r_t^L)a_{jit},$$

$$c'_{jit}(y_{jit}) = \frac{w_t}{1-\alpha}\left(\frac{y_{jit}}{\theta a_{jit}}\right)^{\frac{\alpha}{1-\alpha}}.$$

Combining the four cases together completes the derivation of the factor demand and marginal cost functions. □

B.2 Proof of Relative Quality Being One State Variable

Relative quality refers to the ratio of product qualities between firm 1 and 2 in an intermediate goods market $\frac{q_{1it}}{q_{2it}}$. Below I prove that the firms' pricing strategy in the production stage depend on the relative quality, not the two quality levels, plus the two firms' net assets. For firm 1 in market i, plugging the final good firm's demand for good i into the market share expression and using the optimal pricing rule yields:

$$s_{1it} = \frac{p_{1it}y_{1it}}{p_{1it}y_{1it} + p_{2it}y_{2it}} = \frac{1}{1 + \left(\frac{q_{2it}}{q_{1it}}\right)^{\sigma-1}\left(\frac{p_{2it}}{p_{1it}}\right)^{1-\sigma}} = \frac{1}{1 + \left(\frac{q_{2it}}{q_{1it}}\right)^{\sigma-1}\left[\frac{\sigma+(1-\sigma)(1-s_{1it})}{\sigma+(1-\sigma)(1-s_{1it})-1} \cdot \frac{\sigma+(1-\sigma)s_{1it}-1}{\sigma+(1-\sigma)s_{1it}} \cdot \frac{c'(y_{2it},a_{2it})}{c'(y_{1it},a_{1it})}\right]^{1-\sigma}}.$$

The quality levels come into firms' problem from the demand side, hence independent of the marginal costs $c'(y_{1it}, a_{1it})$ and $c'(y_{2it}, a_{2it})$. The equation above suggests there is a mapping from relative quality to market shares and prices given a_{1it} and a_{2it}, hence independent of quality levels. Symmetric argument applies to s_{2it}. □

B.3 Proof of the Existence of Equilibrium in Innovation Competition

In the innovation stage, the two competing firms choose innovation arrival rates as best response to their competitors'. Hence the pair of innovation rates are determined simultaneously as the intersection of the two best response curves. Below I prove there exists at least one pair of innovation rates emerging as equilibrium. The proof relies on the fact that the best response curve for each firm has a positive intercept. Without loss of generality, I show this property for one firm. Symmetric argument applies to the other firm.

Given the competing firm v's innovation rate x_{vit}, firm u chooses its innovation arrival rate x_{uit} to maximize the firm value function in the innovation stage. The related value function reads as:

$$\begin{aligned}
V^P_{uv}(n_{it}, a_{uit}, a_{vit}) &= \max_{x_{uit}} \pi_{uit}(n_{it}, a_{uit}, a_{vit}) - \frac{\nu_x}{2}x^2_{uit} \\
&+ \beta[(1-\eta)\cdot \mathbb{1}_{u=1} + 1 \cdot \mathbb{1}_{u=E}]\Big\{[(1-x_{uit})(1-x_{vit}) + x_{uit}x_{vit}] \cdot \mathbb{E}_{a_e}\left[V_1(n_{it}, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] \\
&+ x_{uit}(1-x_{vit}) \cdot \mathbb{E}_{a_e}\left[V_1(n_{it}+1, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] + (1-x_{1it})x_{2it} \cdot \mathbb{E}_{a_e}\left[V_1(n_{it}-1, a_{1it+1}, a_{2it+1}, a_{eit+1})\right]\Big\}, \\
s.t. \quad & a_{1it+1} = a_{uit} + \pi_u(n_{it}, a_{uit}, a_{vit}) - \frac{\nu_x}{2}x^2_{1it}, \quad a_{1it+1} > 0, \quad 0 \le x_{uit} \le 1.
\end{aligned}$$

When $x_{vit} = 0$, the value function can be reduced to

$$\begin{aligned}
V^P_{uv}(n_{it}, a_{uit}, a_{vit}) = \max_{x_{uit}} \pi_{uit}(n_{it}, a_{uit}, a_{vit}) - \frac{\nu_x}{2}x^2_{uit} + \beta[(1-\eta)\cdot \mathbb{1}_{u=1} + 1 \cdot \mathbb{1}_{u=E}] \\
\Big\{(1-x_{uit}) \cdot \mathbb{E}_{a_e}\left[V_1(n_{it}, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] + x_{uit} \cdot \mathbb{E}_{a_e}\left[V_1(n_{it}+1, a_{1it+1}, a_{2it+1}, a_{eit+1})\right]\Big\}.
\end{aligned}$$

It is optimal for firm u to choose a positive innovation rate $x_{uit} > 0$ than going for $x_{uit} = 0$ if there exist a positive innovation rate such that the firm value is higher than the one when innovation rate equals zeros. I prove below that a stronger condition holds here. That is, for any positive innovation rate, the firm value is higher than the one with zero innovation rate. It requires the following inequality to be hold for $\forall x_{uit>0}$:

$$V^P_{uv}(n_{it}, a_{uit}, a_{vit})|_{x_{uit}>0} > V^P_{uv}(n_{it}, a_{uit}, a_{vit})|_{x_{uit}=0},$$

$$\frac{\nu_x}{2} x^2_{uit} > \beta[(1-\eta)\cdot \mathbb{1}_{u=1} + 1 \cdot \mathbb{1}_{u=E}] \cdot x_{uit} \cdot \left\{ \mathbb{E}_{a_e}\left[V_1(n_{it}, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] - \mathbb{E}_{a_e}\left[V_1(n_{it}+1, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] \right\}.$$

Since it holds for $\forall x_{uit} > 0$, it is equivalent to:

$$\beta[(1-\eta)\cdot \mathbb{1}_{u=1} + 1 \cdot \mathbb{1}_{u=E}] \cdot x_{uit} \cdot \left\{ \mathbb{E}_{a_e}\left[V_1(n_{it}, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] - \mathbb{E}_{a_e}\left[V_1(n_{it}+1, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] \right\} < 0$$

$$\mathbb{E}_{a_e}\left[V_1(n_{it}, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] - \mathbb{E}_{a_e}\left[V_1(n_{it}+1, a_{1it+1}, a_{2it+1}, a_{eit+1})\right] < 0$$

which obviously holds as firms gain from enlarging the product quality gap. The best response curve for a firm therefore has a positive intercept. Moreover, the best response curve is bounded above by one as innovation rate is a probability which cannot go beyond one. The two properties together ensures there exists at least one pair of innovation rates as best responses to each other emerging in the equilibrium. □