Finance, R&D Investment, and TFP Dynamics*

Zhiyuan Chen†

November 13, 2019

Click for the latest version

Abstract

This paper investigates the role of R&D investment in shaping the relationship between financial constraints and aggregate total factor productivity (TFP). I study a dynamic model in which R&D investment, which affects productivity evolution endogenously, is subject to financial constraints. I parameterize the model with production, innovation, and balance sheet data. The estimated model implies sizeable *static* TFP losses caused by capital misallocation and *dynamic* TFP losses from distorting R&D investment. The accumulation of internal funds reduces the static TFP loss gradually. In contrast, because R&D has a persistent effect on productivity, the dynamic TFP loss rises initially and declines later. Compared to a model with exogenous productivity, innovation investment makes firms less able to use self-financing to reduce TFP losses, and prolongs the transition. Endogenous productivity growth amplifies the gains in TFP and output from financial reform, and leads to a longer-lasting consequence from a credit crunch. Improving the pledgeability of intangible assets in China to be the US level reduces the static TFP loss only 0.4%, but the dynamic TFP loss by 7.1%.

JEL: E22, E23, G31, L11, O31, O33, O47

Keywords: Financial constraints; R&D investment; capital misallocation; intangible assets; TFP

*I am grateful to Jonathan Eaton and Mark Roberts for their continued guidance and support throughout this project. I have benefited from numerous insightful discussions with Jingting Fan, Kenneth Judd, and James Tybout. I thank Francisco Buera, Shoumitro Chatterjee, Paul Grieco, Marc Henry, Fernando Parro, Karl Schurter, Stephen Yeaple for their comments and suggestions, as well as Jie Zhang for sharing the Chinese firm-level dataset with me. I also thank many seminar participants at Trade&Development Reading Group and IO Brown Bag at Penn State, as well as EGSC at Washington University in St. Louis for their valuable comments and suggestions. All remaining errors are my own.

†PhD candidate, Department of Economics, The Pennsylvania State University. Email: zxc5109@psu.edu.

1 Introduction

Differences in total factor productivity (TFP) are an important contributor to disparities in income and development across countries. Underdeveloped financial markets can reduce aggregate TFP by distorting innovation decisions. Despite a large literature that has documented a stimulating effect of financial development on innovation investment and productivity growth,[1] financial frictions are often absent in R&D investment models.[2] This limitation prevents us from analyzing the interplay of financial friction, R&D investment, and TFP together. The objective of this paper, thus, is to evaluate the role of R&D investment in shaping the relation between financial constraints and aggregate TFP.

R&D activities have a persistent effect on productivity. When financial constraints restrict a firm's ability to undertake R&D investment, the negative consequence on productivity will be carried over time. Endogenous R&D investment complicates the relationship between financial constraints and aggregate TFP. First, introducing R&D investment can potentially amplify the TFP loss from financial constraints. Financial constraints, meanwhile, can cause (1) a *static* TFP loss by generating differences in the marginal product of capital across entrepreneurs– i.e., giving rise to a misallocation of capital, and (2) a *dynamic* TFP loss by constraining R&D investment and hindering productivity growth. Second, R&D investment can also influence the efficacy of self-financing in reducing TFP losses by increasing productivity and assisting the accumulation of (intangible) assets. As Moll (2014) points out, in the presence of financial constraints, the dynamism of productivity and assets is the key to understanding transition dynamics and the steady state of TFP.

How much does the R&D channel account for the productivity loss caused by financial constraints? How does R&D investment affect the efficacy of self-financing in the reduction of TFP losses? This paper attempts to shed light on each of these questions by using a quantitative model of R&D investment with financial frictions. I estimate the model using a panel of manufacturing firms in China, a country with relatively less developed financial markets. The parameterized model matches firms' size distribution as well as their decisions on R&D investment and asset accumulation. Through this model's lens, I first quantify static and dynamic TFP losses. I then analyze the transition dynamics and the steady state of the model. Finally, to gauge the importance of R&D investment for understanding the relationship between finance and TFP, I compare the results of the estimated model with a model's special case where the productivity process is exogenous.

I find that, on average, financial constraints cause a TFP loss of 29% through the R&D channel within the sample period. This is close to the static TFP loss (37%) caused by capital misallocation. To examine the robustness of our benchmark results, I have explicitly considered the heterogeneity in R&D costs, heterogeneity in cost-benefits structure of R&D across industries, and endogenous

[1] For the related empirical evidence, see Rajan and Zingales (1998); Chava et al. (2013); Gorodnichenko and Schnitzer (2013); Kerr and Nanda (2015), among others.

[2] See Klette and Kortum (2004); Eaton and Kortum (1999); Aw et al. (2011); Doraszelski and Jaumandreu (2013); Peters et al. (2017), among others.

uncertainty in R&D investment. The quantification of TFP losses is robust to all of these modifications. Re-calibrating the exogenous-productivity model predicts a similar degree of productivity loss, yet it fails to detect the dynamic TFP loss caused by distorted R&D decisions.

Over time, the accumulation of internal funds enables some firms to escape from financial constraints, which reduces TFP losses. In the model with R&D, over time, the static TFP loss declines more slowly because the endogenous productivity growth tends to entrap relatively more firms in financial constraints. As a results, in the steady state, static TFP loss is reduced to be 18.4% (16.7%) in the model with endogenous (exogenous) productivity. In contrast, because of the persistent impact of R&D on productivity, the dynamic productivity loss rises initially and falls ultimately as firms become wealthier. In the steady state, dynamic TFP loss is 20%, showing only a decrease of 9%. This suggests that it is more difficult for firms to undo the dynamic TFP loss by self-financing.

I explore several policy implications of the quantitative model. First, I consider a financial reform that relaxes the credit constraints permanently. I find that the boosting effect of financial reform on aggregate TFP and output is amplified when considering the endogenous response of R&D investment. Second, I study a credit crunch by tightening the financial constraints for one period. I show that the detrimental impact of a credit crunch on aggregate TFP and output tends to be longer-lasting when I account for the endogenous growth of productivity.

My model also sheds light on the impact of using intangibles as collateral on R&D investment and TFP. In practice, products of R&D investment, such as patents and trademarks, are used as collateral when firms borrow from financial institutions.[3] In the last counterfactual experiment, I investigate the impacts of allowing more intangibles to serve as collateral on R&D investment and TFP. I find that improving the pledgeability of intangible assets in China to be as the US level encourages R&D investment and reduces the static (dynamic) TFP loss by 0.4% (down to 7.1%). Employing a policy shock on patent pledge financing in China, I also provide causal evidence supporting that increasing the pledgeability of intangible assets enhances the R&D investment in China.

This paper is closely related to studies that use quantitative models to analyze firms' incentives for undertaking R&D investment (Eaton and Kortum, 1999, 2007; Aw et al., 2011; Doraszelski and Jaumandreu, 2013; Warusawitharana, 2015; Peters et al., 2017). In these predecessor models, R&D investment entails certain costs and leads to productivity growth in the future. By assuming that firms operate in a perfect financial system, these models provide no room for analyzing the impact of financial development on R&D investment and productivity dynamics. Meanwhile, there is growing reduced-form evidence that financial markets do play a key role in supporting R&D investment (Rajan and Zingales, 1998; Robb and Robinson, 2012; Chava et al., 2013; Gorodnichenko and Schnitzer, 2013; Nanda and Nicholas, 2014; Kerr and Nanda, 2015; Cornaggia et al., 2015; Varela, 2018). In gen-

[3]As for relevant empirical studies, Loumioti (2012) finds twenty-one percent of US-originated secured syndicated loans during 1996-2005 have been collateralized by intangibles. More specifically, More recently, Hochberg et al. (2018) document that start-ups with more redeployable patents as assets are able to receive more funds from investors. And Mann (2018) shows that patents that are pledged as collateral to help US firms raise more debt and spend more on R&D.

eral, these studies employ certain indicators to measure financial development and then link them to firms' innovation behavior. The current paper contributes to these studies by explicitly considering financial constraints in a structural model of R&D investment.

This paper also contributes to the literature that focuses on the role of endogenous productivity change in understanding the impact of various distortions, particularly tax distortions (e.g., Bhattacharya et al. (2013); Bento and Restuccia (2017); Da-Rocha et al. (2017)) and financial frictions (e.g., Mestieri et al. (2017); Vereshchagina (2018); Caggese (2019)).[4] Vereshchagina (2018) and Caggese (2019) are the two most relevant papers. Vereshchagina (2018) a variation of the Bewley-Ayiagary-Hugget model (Huggett, 1993) in which firms can invest in intangible capital to improve their future productivity in a deterministic way. In contrast, this paper extends a Hopenhayn firm dynamics model (Hopenhayn, 1992) by treating productivity as a controlled Markov process: future productivity is partly random and partly under the control of R&D investment. The current paper goes on to estimates its own model by focusing on dynamic decisions about R&D and wealth accumulation. It analyzes the effect that innovation investment has on the efficacy of self-financing for reducing TFP loss along the transitional path and in the steady state.

Caggese (2019) analyzes the role of innovative investment in understanding the impact of financial constraints on firm growth. He argues that incorporating radical as well as incremental innovations can explain the differences in the evolution of firm sizes over time between poor and rich countries. My paper differs from Caggese's work in several aspects. First, I focus on the influence of R&D investment on affecting the relationship between financial constraints and productivity. Second, I allow the output of R&D investment–the intangible assets–to be used as collateral. Third, I incorporate R&D into the productivity process a lá Aw et al. (2011). Therefore, R&D investment always faces a certain degree of uncertainty in my model, as with the radical innovation modelled by Caggese (2019).

The rest of this paper is organized as follows: In Section 2, I introduce the benchmark model and a decomposition of TFP losses. In Section 3, then, I present the data and estimation results. Section 4 displays the results of the quantitative analysis, while Section 5 analyzes several robustness checks. Section 6 concludes the paper.

2 The Benchmark Model

In this section, I introduce a heterogeneous-firm model in which firms finance their R&D investment out of cash flow; a firm's capital investment is restricted by a collateral constraint a lá Midrigan and Xu (2014) and Moll (2014). More importantly, I also introduce intangibles as a collateral for borrowing.

[4] More broadly, these studies originates from a large literature on financial frictions and economic development (see Jeong and Townsend (2007), Buera et al. (2011), Buera and Shin (2013), Midrigan and Xu (2014), and Moll (2014), among others). In most of such previous studies, the productivity process has been treated as exogenous. See also Buera et al. (2015) for an excellent review of this literature.

2.1 Setup

Production I consider an industry populated with a fixed number of firms, each producing a single variety. Firm i operating in period t uses labor l_{it}, capital k_{it}, and a constant-return-to-scale production technology to produce its output q_{it}:

$$q_{it} = \phi_{it} k_{it}^{\alpha} l_{it}^{1-\alpha} \tag{1}$$

where $0 < \alpha < 1$ is the capital's share. ϕ_{it} is the current state of technology. Labor is hired in a competitive market at the wage rate ω. Each firm has a constant-elasticity demand function:

$$q_{it} = p_{it}^{-\sigma}, \tag{2}$$

where σ is the demand elasticity and assumed to be greater than one. The revenue production function is

$$y_{it} = \left(\phi_{it} k_{it}^{\alpha} l_{it}^{1-\alpha}\right)^{\frac{\sigma-1}{\sigma}}, \tag{3}$$

which has decreasing returns to scale.

R&D and Productivity Following Aw, Roberts and Xu (2011) and Doraszelski and Jaumandreu (2013), I assume that R&D affects the firm's productivity process. Specifically, letting x_{it} be R&D investment, the log-productivity follows a controlled Markov process:

$$\ln(\phi_{it+1}) = \rho \ln(\phi_{it}) + \gamma \ln(x_{it} + 1) + \xi_{it+1}, \tag{4}$$

where γ governs the marginal effect of R&D investment on productivity; $\gamma > 0$ means that more R&D investment leads to a more favorable productivity distribution in the future. Note that $\ln(x_{it} + 1) = 0$ when $x_{it} = 0$, meaning that zero R&D investment generates no enhancement in productivity. Also note that $\partial \ln(\phi_{it+1})/\partial x_{it} = \gamma/(x_{it} + 1)$, implying that the rate of growth of firm productivity increase with x_{it} at a decreasing speed, showing no discontinuity at the extensive margin.[5] ρ is the persistence of the productivity.[6] ξ_{it+1} is an exogenous i.i.d shock that follows a normal distribution $\mathbf{N}\left(0,\, \sigma_{\xi}^2\right)$. σ_{ξ} measures the uncertainty facing R&D investment. A larger σ_{ξ} indicates a higher degree

[5] Another possible specification is $\ln(\phi_{it+1} = \rho \ln(\phi_{it}) + \gamma_0 \mathbb{I}(x_{it}) + \gamma_1 \ln(x_{it}+1) + \xi_{it+1}$, where $\mathbb{I}_{it}$ is an indicator function equals to one when x_{it} is positive and zero otherwise. Doraszelski and Jaumandreu (2013) have considered such a possibility. However, I do not find supporting evidence for this specification in our data set.

[6] Another interpretation of ρ is that it captures the depreciation of past R&D investment. To see this, note that I can rewrite this productivity process as

$$\phi_{it} = \exp\left[\sum_{s=0}^{t-1} \delta^s \gamma \ln(x_{it} + 1) + \sum_{s'=1}^{t} \delta^{s'} \xi_{is'}\right],$$

which means that the state-of-art technology summarizes all of the past R&D activities and exogenous shocks.

of uncertainty.[7]

Financial constraints The extent to which firms can use capital is determined by following constraint:

$$k_{it} \leq \frac{1}{1-\theta}a_{it} + \frac{\theta}{1-\theta}\phi_{it}^{\eta} \tag{5}$$

where $\theta \in [0, 1]$ captures the severity of borrowing constraint. A larger θ means a better financial environment where firms have better capacity to borrow. In particular, when $\theta = 1$, the capital constraint is never binding, which indicates a perfect financial system. ϕ_{it}^{η} summarizes the value of intangible assets (such as patents, trade marks, and other intellectual properties) used as collateral when a firm obtains external financing.[8] η is the elasticity between pledgeable intangible assets in response the measured productivity. A larger η means firms with relatively high productivity can borrow against more intangibles. I expect that $\eta > 0$ so that more intangible assets are available for more productive firms. Different from Midrigan and Xu (2014), who treat the intangible asset to be fixed over time, I allow it to be varying across firms and time. This more realistic setting allows us to analyze the potential effect of R&D investment on relaxing financial constraints by accumulating intangible assets.

Firm's problem and static choices Each firm is owned by an entrepreneur whose objective is to maximize its life-time utility:

$$\mathbf{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{c_{it}^{1-\epsilon} - 1}{1-\epsilon},$$

where c_{it} is the entrepreneur's consumption and ϵ denotes the inverse of elasticity of inter-temporal substitution. He is subject to a budget constraint

$$c_{it} + \mathbb{I}\left(x_{it}\right) f + C\left(x_{it}\right) + a_{it+1} \leq y_{it} - wl_{it} - \left(r+\delta\right) k_{it} + \left(1+r\right) a_{it}, \tag{6}$$

where $\mathbb{I}\left(x_{it}\right)$ is an indicator function of x_{it} which equals one when x_{it} is positive and zero otherwise. Unlike existing R&D investment models, I assume that the financing of R&D costs faces credit market imperfections. Hall and Lerner (2010) argue that the nature of R&D investment–intangible outcome and high degree of uncertainty, makes it more costly for innovators to use external financing for R&D activities. Consistent with their finding, I assume that R&D investment can only be financed using the firm's internal cash flow. The innovation investment is modelled as a two-step process. First, an entrepreneur needs to pay a fixed cost f to find a new research idea.[9] By introducing a fixed cost

[7] In the benchmark model, R&D investment does not alter the conditional variance of the future productivity. In an alternative specification, I relax this assumption. See more details in the section of extensions and robustness.

[8] In Appendix B, I provide a micro foundation for the chosen specification of intangible assets.

[9] Peters et al. (2017) and Chen (2019) have documented the persistence of R&D activities and distinguished between start up costs and maintenance costs. For computational tractability, I impose that the start up costs are equal to maintenance costs.

for innovation investment, the model captures an important data feature that only a small fraction of firms undertake R&D investment. The entrepreneur also needs to build research labs and hire research teams to implement the innovative idea. $C(x_{it})$ represent these expenses. Following a large body of investment literature, I use a quadratic form for the R&D investment:

$$C(x_{it}) = \frac{d}{2} x_{it}^2. \tag{7}$$

Firms take the interest rate and wage rate as given, therefore I can obtain optimal choices of labor and capital by solving the following constrained optimization problem:

$$\max_{l_{it},\, k_{it}} \{y_{it} - wl_{it} - (r+\delta)k_{it}\}$$
$$s.t.\ k_{it} \leq \frac{a_{it}}{1-\theta} + \frac{\theta}{1-\theta}\phi_{it}^{\eta}$$

The first-order condition delivers that

$$l_{it} = \frac{(1-\alpha)(\sigma-1)\, y_{it}}{\sigma w} \tag{8}$$

$$k_{it} = \frac{\alpha(\sigma-1)\, y_{it}}{\sigma R(a_{it},\, \phi_{it})} \tag{9}$$

and MRPK is given by

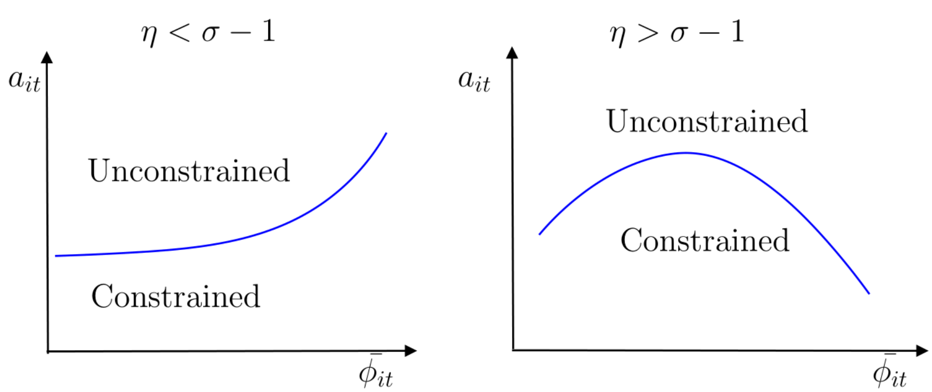
$$R_{it} = \max\left\{ r+\delta,\ \underbrace{\frac{\alpha}{\bar{m}}\left(\frac{1-\alpha}{\bar{m}w}\right)^{\frac{1-\alpha}{\bar{m}+\alpha-1}} \left[\phi_{it}\left(\frac{a_{it}}{1-\theta} + \frac{\theta\phi_{it}^{\eta}}{1-\theta}\right)^{1-\bar{m}}\right]^{\frac{1}{\bar{m}+\alpha-1}}}_{R(a_{it},\phi_{it})} \right\} \tag{10}$$

where $\bar{m} = \sigma/(\sigma-1)$ is the markup. There are two regimes in which MRPK are determined by different factors. When the financial constraint is not binding, $R(a_{it}, \phi_{it})$ is equal to the market's capital price $(r+\delta)$. In contrast, when the financial constraint is binding, the cost of capital $R(a_{it},\, \phi_{it})$ is jointly determined by a_{it}, ϕ_{it}, as well as parameters including θ and η. In particular, the costs of using capital is non-increasing in net worth a_{it}, meaning that wealthier firms tend to face a lower shadow costs of using capital conditional on the productivity, thus less likely to be constrained. Because the pledge-able intangible assets ϕ_{it}^{η} enter into the financial constraint, the relationship between the capital costs and productivity depends on model parameters. To see it more clearly, for any level of wealth a_{it}, let's define a cut-off productivity $\bar{\phi}(a_{it})$ as

$$R\left(a_{it},\, \bar{\phi}(a_{it})\right) = r+\delta \tag{11}$$

The relationship between $\bar{\phi}(a_{it})$ and a_{it} is affected by the value of η. In Figure 1, I illustrate how $\bar{\phi}(a_{it})$ varies with a_{it} in cases when $\eta < \sigma - 1$ and $\eta > \sigma - 1$, respectively. The left panel shows that when $\eta < \sigma - 1$, $\bar{\phi}$ is increasing with a_{it}. Given the level of net worth, firms that are more productive are more likely to have a binding financial constraint. In contrast, there is an inverted U relationship between a_{it} and $\bar{\phi}(a_{it})$ when $\eta > \sigma - 1$. In the region where $\bar{\phi}(a_{it})$ is increasing with a_{it}, the bifurcation is similar to the situation when $\eta < \sigma - 1$. However, when $\bar{\phi}(a_{it})$ is decreasing with a_{it}, conditional on net worth, less productive firms are more likely to be constrained because of a lack of intangibles to be used as collateral. In summary, if $\bar{\phi}(a_{it})$ increases (decreases) with a_{it}, it describes the highest (lowest) productivity above (below) which the financial constraints are binding. Note that when $\eta = 0$ the model degenerates into the case in which intangible assets are fixed across periods. It is also easy to see that $\bar{\phi}(a_{it})$ is always increasing in θ, meaning that relatively fewer firms are constrained in a better financial environment.[10]

Figure 1: Relationship between net worth and cut-off productivity for different values of η



Note:When $\bar{\phi}_{it}$ is increasing (decreasing) with a_{it}, more (less) productive firms tend to be more constrained.

2.2 Value Functions and Equilibrium

Value functions Now I omit firm and time subscripts and formulate the entrepreneur's problem in a recursive form. The state variables are (a, ϕ). Aggregate variables are assumed to be constant and exogenous to an individual firm. An entrepreneur makes two dynamic decisions: asset accumulation and R&D investment. The firm's recursive problem is given by

$$V(a, \phi) = \max_{a', x} \left\{ \frac{c^{1-\epsilon}}{1-\epsilon} + \beta \mathbf{E} V(a', \phi') \right\} \tag{12}$$

[10]See the math appendix for related proofs.

subject to a budget constraint:

$$c + \mathbb{I}(x) f + \frac{d}{2} x^2 + a' = \frac{1}{\sigma} y\,(a,\,\phi) + (1+r)\,a \tag{13}$$

where $y\,(a,\,\phi)$ is the firm's revenue generated by optimal labor and capital choices. The firm's revenue is given by

$$y\,(a,\,\phi) = \left[\frac{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}{\bar{m} w^{1-\alpha}}\right]^{\frac{1}{\bar{m}-1}} \phi^{\sigma-1} R\,(a,\,\phi)^{\alpha(1-\sigma)}\,, \tag{14}$$

The model has extensive and intensive margins of R&D investment due to the fixed cost of undertaking R&D investment. This generates kinks in the value function. Let $V^0\,(a,\,\phi)$ ($V^1\,(a,\,\phi)$) be the value function when $x = 0$ ($x > 0$). Then $V\,(a,\,\phi)$ can be expressed as

$$V\,(a,\,\phi) = \max\left\{V^0\,(a,\,\phi)\,,\, V^1\,(a,\,\phi)\right\} \tag{15}$$

Note that V^0 and V^1 can be written recursively as

$$V^0\,(a,\phi) = \max_{a'}\left\{\frac{[c\,(a,\phi,a')]^{1-\epsilon}}{1-\epsilon} + \beta\int_{\mathbf{R}} V\,(a',\,\phi')\,Q_0\,(\phi,\,d\phi')\right\}, \tag{16}$$

$$V^1\,(a,\phi) = \max_{a',\,x}\left\{\frac{[c\,(a,\,\phi,\,a',\,x)]^{1-\epsilon}}{1-\epsilon} + \beta\int_{\mathbf{R}} V\,(a',\,\phi')\,Q_x\,(\phi,\,d\phi')\right\}, \tag{17}$$

where the consumption levels are:

$$c\,(a,\,\phi,\,a') = \frac{1}{\sigma} y\,(a,\,\phi) + (1+r)\,a - a' \tag{18}$$

$$c\,(a,\,\phi,\,a',\,x) = c\,(a,\,\phi,\,a') - f - \frac{d}{2} x^2 \tag{19}$$

$Q_x\,(\phi,\,\cdot)$ ($Q_0\,(\phi,\,\cdot)$) denotes the transition kernel of the stochastic productivity process when R&D investment is positive (zero). At the extensive margin, a firm invests in R&D if and only if

$$V^1\,(a,\,\phi) > V^0\,(a,\,\phi) \tag{20}$$

Conditional on that the firm finds it optimal to invest in R&D, the intensive margin of R&D is characterized by (17). Because financial constraints lower firm's current profits, it will also have a negative effect on the firm's innovation investment. It is easy to verify that the R&D investment is non-decreasing with θ.

Equilibrium We are interested in both the transition dynamics and steady state of our model. To this end, we consider a partial recursive equilibrium which we formally define as below.

Definition 1. A *partial recursive equilibrium* is a set of value functions $V(a, \phi)$, policy functions $a'(a, \phi)$, $x(a, \phi)$ such that given $P = Q = 1$:

1. $V(a, \phi)$ solves the firm's Bellman equation;

2. $a'(a, \phi)$ and $x(a, \phi)$ are optimal decision rules for net worth accumulation and R&D investment, respectively.

Since the partial recursive equilibrium does not require that state variables are in stationary distributions, we can analyze firm dynamics over time by simulating the model forward. To analyze the long-run effects of R&D investment in affecting the relation between finance and aggregate TFP, I also consider the steady-state equilibrium of the model.

Definition 2. A *steady-state equilibrium* is a partial recursive equilibrium such that joint distribution of (a_{it}, ϕ_{it}) is invariant over time.

Given the joint distribution of (a, ϕ), the aggregate TFP, output, and inputs are determined. In the appendix, I discuss the existence of the steady state of the model.

2.3 Exogenous productivity

I now briefly discuss a special case of the benchmark model with endogenous productivity: the exogenous productivity. The benchmark model degenerates into a model of exogenous productivity when we impose $\gamma = 0$ or R&D costs to be infinity. Because the benefits from R&D investment is realized through improving productivity, these conditions immediately imply that no firm would undertake R&D investment. Let $W(a, \phi)$ be the value function when no R&D investment is undertaken, I can write the firm's recursive problem as

$$W(a, \phi) = \max_{a'} \left\{ \frac{c(a, \phi, a')^{1-\epsilon}}{1-\epsilon} + \beta \int_{\mathbf{R}} W(a', \phi') Q_0(\phi, d\phi') \right\} \tag{21}$$

subject to the budget constraint

$$c(a, \phi, a') = \frac{1}{\sigma} y(a, \phi) + (1 + r)a - a', \tag{22}$$

and the exogenous productivity evolution rule:

$$\ln(\phi_{it+1}) = \rho_1 \ln(\phi_{it}) + \xi^1_{it+1}. \tag{23}$$

When the productivity is exogenous, the only dynamic decision is asset accumulation. The equilibrium concepts I have just described for the endogenous productivity model can be readily applied here.

2.4 Aggregation and TFP losses

Let $\mathbb{N}$ be the set of active producers, the measure of which is N. The sectoral output is

$$Q_t = TFPQ_t K_t^{\alpha} L_t^{1-\alpha} \tag{24}$$

where the aggregate physical productivity $TFPQ_t$ can be expressed as[11]

$$TFPQ_t = \frac{\left[\int_{i\in\mathbb{N}} R_{it}^{\alpha(1-\sigma)} \phi_{it}^{\sigma-1} di\right]^{\frac{1}{\sigma-1}+\alpha}}{\left[\int_{i\in\mathbb{N}} R_{it}^{\alpha(1-\sigma)-1} \phi_{it}^{\sigma-1} di\right]^{\alpha}} \tag{25}$$

In the absence of financial constraints ($\theta = 1$), R_{it} is common to all firms and equals to $r + \delta$. This entails an efficiency allocation of capital, which implies an efficient level of aggregate TFP:

$$TFPQ_t^e = \left(\int_{i\in\mathbb{N}} \phi_{it}^{\sigma-1} di\right)^{\frac{1}{\sigma-1}} \tag{26}$$

R&D investment affects the evolution of productivity. By distorting R&d investment, financial constraints also have an influence on an individual firm's fundamental productivity. To analyze the TFP loss from underinvestment in R&D, I need to consider the endogenous evolution of fundamental productivity in the counterfactual scenario when $\theta = 1$.

I propose a counterfactual experiment to analyze the impact of financial constraints on aggregate TFP. I choose an initial period, t_0, and treat productivity and net worth in this period as the fundamental state. The chosen fundamental state is the starting point of our analysis. Given the initial productivity ϕ_{it_0}, let's define ϕ_{it}^* as the productivity in the counterfactual scenario in which no borrowing constraint exists (see Panel A of Figure 2). In this case, the MRPK is equalized for all firms. The associated *best* aggregate TFP is

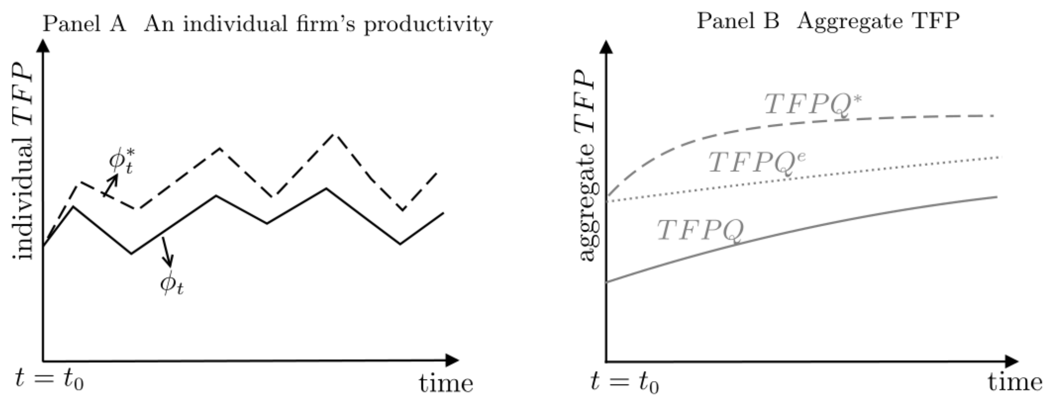
$$TFPQ_t^* = \left[\int_{i\in N} (\phi_{it}^*)^{\sigma-1}\right]^{\frac{1}{\sigma-1}} \tag{27}$$

I can decompose the actual aggregate TFP as

$$TFPQ_t = TFPQ_t^* - (TFPQ_t^* - TFPQ_t^e) - (TFPQ_t^e - TFPQ_t).$$

[11]See Appendix A for the derivations.

Figure 2: Decomposition of TFP loss



This further implies the total TFP loss as a share of $TFPQ_t^e$ can be computed as[12]

$$\begin{aligned} \textit{Total TFP Loss} =& \frac{TFPQ_t^* - TFPQ_t}{TFPQ_t^e} \\ =& \underbrace{\frac{TFPQ_t^* - TFPQ_t^e}{TFPQ_t^e}}_{\text{dynamic TFP loss}} + \underbrace{\frac{TFPQ_t^e - TFPQ_t}{TFPQ_t^e}}_{\text{static TFP loss}} \end{aligned} \tag{28}$$

The equation above decomposes the total productivity loss into two components (See Panel B of Figure 1 for three statistics of $TFPQ$). The static TFP loss is computed as in Hsieh and Klenow (2009), which measures the impact of capital misallocation in reducing the aggregate TFP. The dynamic TFP loss distortions in R&D investment caused by financial constraints. The proposed method of analyzing the TFP loss can be performed starting from any chosen states. This flexibility enables us to consider TFP dynamics and sources of its loss on the transitional path. Below I provide a characterization of these two components of TFP loss.

Static productivity loss The model features endogenous productivity evolution. By simulating the model, I notice that the marginal distributions of net worth and productivity are very close to log-normal distributions. I find it convenient to provide a characterization of the static TFP loss under

[12] I choose to use $TFPQ^e$ as the scaling variable to avoid the scale difference when comparing the results to that of exogenous productivity model.

the assumption that a_{it} and ϕ_{it} follow a joint log-normal distribution:

$$\begin{bmatrix} \log(a_{it}) \\ \log(\phi_{it}) \end{bmatrix} \sim \mathbf{N}\left(\begin{bmatrix} \mu_a \\ \mu_\phi \end{bmatrix}, \begin{bmatrix} \sigma_a^2 & \tilde{\rho}\sigma_a\sigma_\phi \\ \tilde{\rho}\sigma_a\sigma_\phi & \sigma_\phi^2 \end{bmatrix} \right),$$

where μ_a and μ_ϕ denote the mean of net worth and productivity, respectively. σ_a^2 and σ_ϕ^2 are the variance of net worth and productivity, separately. $\tilde{\rho}$ is the correlation between net worth and productivity. Under the assumption that productivity and MRPK are log-normally distributed, Midrigan and Xu (2014) and Ek and Wu (2018) have provided analytical expressions of aggregate TFP, showing that financial constraints reduce aggregate TFP. However, with the assumption that firms face borrowing constraints, the distributional assumption on productivity and MRPK is unlikely true because obviously MRPK is bounded below by $r + \delta$. By assuming that net worth and productivity follow a joint log-normal distribution, I provide a better approximation of the TFP loss. Employing the law of large numbers, I can express the aggregate TFP under efficient capital allocation as:

$$TFPQ_t^e = N^{\frac{1}{\sigma-1}} e^{(\sigma-1)\mu_\phi + \frac{\sigma-1}{2}\sigma_\phi^2} \tag{29}$$

Clearly, the efficient level of aggregate TFP is increasing with the mean and variance of the productivity distribution. For the empirical relevance, we consider $\eta < \sigma - 1$ which implies that more productive firms are more likely to be financially constrained. In this case, the fraction of constrained firms can be calculated as

$$\zeta = \int_0^\infty \int_{\bar{\phi}(a)}^\infty dG(a, \phi),$$

where $G(a, \phi)$ represent the joint log-normal distribution of (a, ϕ).[13] When ζ is relatively small, it can be shown that the aggregate TFP can be approximated as

$$TFPQ = \Upsilon_0^{\frac{1}{\sigma-1}} N^{\frac{1}{\sigma-1}} e^{(\sigma-1)\mu_\phi + \frac{\sigma-1}{2}\sigma_\phi^2} \tag{30}$$

where Υ_0 is defined as

$$\begin{aligned} \Upsilon_0 &\equiv \int_0^\infty \int_0^{\bar{\phi}(a)} \phi^{\sigma-1} dG(a, \phi) \\ &= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \Phi\left(h(z)\right) dz, \end{aligned} \tag{31}$$

[13]See the math appendix for related math derivations.

in which $h(z) = \frac{\bar{v}(z)-\tilde{v}(z)}{\sqrt{1-\tilde{\rho}^2}\sigma_\phi}$, and

$$\bar{v}(z) = \ln\left(\bar{\phi}(e^{\mu_a+(\sigma-1)\sigma_a\sigma_\phi+\sigma_a z})\right)$$
$$\tilde{v}(z) = \mu_\phi + (\sigma-1)(1+\tilde{\rho}-\tilde{\rho}^2)\sigma_\phi^2 + \tilde{\rho}\sigma_\phi z.$$

$\Phi(\cdot)$ is the CDF of standard normal distribution. Since $\Phi\left(\frac{\bar{v}(z)-\tilde{v}(z)}{\sqrt{1-\tilde{\rho}^2}\sigma_\phi}\right) < 1$, we know that $\Upsilon_0 < 1$. This implies that $TFPQ \leq TFPQ^e$. Therefore the static TFP loss can be computed as

$$\textit{Static TFP loss} \approx 1 - \Upsilon_0^{\frac{1}{\sigma-1}} \tag{32}$$

It is easy to show that the static TFP loss will be increasing with μ_ϕ and decreasing with μ_a. This means that both the levels of productivity and net wealth will affect the evolution of static productivity loss. The impact of the level of productivity on static TFP loss is the key for understanding the role that R&D investment plays in altering the dynamics static TFP loss. When μ_ϕ is fixed over time, allowing firms to accumulate wealth will reduce the static TFP loss unambiguously. However, when a firm can invest in R&D to increase its productivity, μ_ϕ may also be growing over time, which can potentially exacerbates the capital misallocation. This implies that endogenously productivity growth may undermine the role of self-financing in reducing TFP losses over time. In addition, Υ_0 is increasing with respect to θ because a larger θ implies a higher $\bar{\phi}(a)$, meaning less firms have binding financial constraints. Lastly, it is easy to verify that the TFP loss is increasing with ζ: a larger fraction of constrained firms leads to lower aggregate TFP.

Dynamic productivity loss Let's continue to analyze the dynamic TFP loss. Given the productivity fundamentals $\{\phi_{it_0}\}$ in period t_0, using the optimal R&D decision rule I can write down the productivity in the next period as follows:

$$\ln(\phi_{it_0+1}) = \rho\ln(\phi_{it_0}) + \gamma\ln(x(a_{it_0}, \phi_{it_0}; \theta) + 1) + \xi_{it_0+1}$$
$$\ln(\phi^*_{it_0+1}) = \rho\ln(\phi_{it_0}) + \gamma\ln(x(a_{it_0}, \phi_{it_0}; \theta^*) + 1) + \xi_{it_0+1}$$

The difference in future productivity is determined by the gap in R&D investment, which reflects different levels of financial constraints. To control the impact of exogenous productivity shock, I impose same productivity shocks in these two scenarios. Conditional on the initial states of period t_0, the one-period mean difference of these two productivity distributions is

$$\mathbf{E}_{t_0}\left[\ln(\phi^*_{it_0+1})\right] - \mathbf{E}_{t_0}\left[\ln(\phi_{it_{0+1}})\right] = \gamma\ln\left(\frac{x(a_{it_0}, \phi_{it_0}; \theta^*) + 1}{x(a_{it_0}, \phi_{it_0}; \theta) + 1}\right) \tag{33}$$

If financial constraints restrict R&D investment, it is obvious that $x(a_{it_0}, \phi_{it_0}; \theta^*) > x(a_{it_0}, \phi_{it_0}; \theta)$. This implies that financial constraints cause underinvestment in R&D. Looking forward, to quantify the accumulative effects of financial constraints on TFP loss, I keep track of the two different productivity trajectories. In particular, in period $t+s$, the mean difference in productivity is

$$\mathbf{E}_{t_0}\left[\ln(\phi^*_{it_0+s})\right] - \mathbf{E}_{t_0}\left[\ln(\phi_{it_0+s})\right] = \gamma \sum_{j=1}^{s} \rho^{s-j} \ln\left(\frac{x^*_{t_0+j-1}+1}{x_{t_0+j-1}+1}\right) \tag{34}$$

where $x^*_{t_0+j-1} = x(a^*_{it_0+j-1}, \phi^*_{it_0+j-1}; \theta^*)$ and $x_{t_0+j-1} = x(a_{it_0+j-1}, \phi_{it_0+j-1}; \theta)$ represent R&D investment in two different scenarios. When productivity follows a log-normal distribution, employing (26), the dynamic productivity loss can be expressed as:

$$\begin{aligned} \textit{Dynamic TFP loss} &= \frac{TFPQ^*_{t_0+s}}{TFPQ^e_{t_0+s}} - 1 \\ &= \exp\left\{(\sigma-1)\gamma \sum_{j=0}^{s} \rho^{s-j} \ln\left(\frac{x^*_{t_0+j-1}+1}{x_{t_0+j-1}+1}\right)\right\} - 1 \end{aligned} \tag{35}$$

Equation (35) has several implications. First, the dynamic TFP loss reflects the current and past efforts of R&D investment and piles up over time. Second, R&D investment in more distant periods tends to be less important for current productivity because of the depreciation rate ρ. These two observations are critical for the understanding of the change in dynamic TFP loss over time. As time evolves, poor firms accumulate more wealth and grow out of financial constraints. This narrows the per-period difference in R&D investment. In addition, because past disparities in R&D investment become less important, dynamic TFP loss tend to decrease eventually. Lastly, the dynamic productivity loss will also be affected by the initial state of the counterfactual experiment.

2.5 Empirical goals

Before I introduce the data and estimation strategy, I now lay out my empirical goals. To begin with, I am going to measure the firm-year level productivity using a rich firm-level data set. Along with the observed information on firm net worth, capital, and employees, I can parameterize the equation of capital constraint by matching the cross-section distribution of capital and labor. Then I estimate the productivity evolution equation. In particular, I follow Vereshchagina (2018) to separately estimate the endogenous productivity process with R&D and the exogenous productivity process without R&D. This helps us understand the role that R&D plays in shaping the relation between R&D and TFP. The last empirical objective is to determine the costs of R&D investment. I choose cost parameters such that the R&D investment and net worth decisions predicted by the structural model are consistent with the data.

Utilizing the parameterized model, I analyze the effects of R&D investment on the relation between financial development and aggregate TFP by varying θ. In the exogenous productivity model, the only channel through which financial development can improve aggregate efficiency is reducing misallocation. In contrast, in the endogenous productivity model R&D investment also responds to the financial development and enhances the fundamental productivity as well as aggregate TFP. The quantitative exercise allows us to quantitatively evaluate the strength of R&D channel through which financial constraints affect the productivity distribution and aggregate TFP.

Because some firms can accumulate wealth over time to overcome financial constraints, the consequence of financial frictions differs in the short- and long-run. This point is theoretically explored in Moll (2014). How does the option of R&D investment affect transition of TFP losses? Eventually, what is the impact of R&D investment on the efficacy of self-financing in cutting TFP losses? Simulations of the empirical model will help us answer these questions.

3 Data and Estimation

3.1 Data

I use the administrative income tax records from Chinese State Administration of Tax (SAT) from 2008-2011. The SAT is in charge of collecting taxes and auditing, similar to the IRS in the United States. The SAT in China maintains its own firm-level database of tax payments as well as other balance-sheet and financial statement information that is necessary for tax-relevant calculations. For the purpose of this study, I use the information to estimate productivity and calculate other relevant variables. I have obtained these tax records from 2008 to 2011. I have followed several cleaning procedures. First, I have deleted the duplicated observations within a year. For firm names that are repeated within a year, I use the tax id as their unique identifier. Second, I have deleted observations with abnormal values for interested variables. These include: (i) negative sales, debt, total asset, fixed asset; (ii) number of employees smaller than 10; (iii) birth year later than 2011 or earlier than 1900. The final data set I use for estimation is a balanced panel with 21,428 firms spanning over four years. In Appendix D, I have included details of data processing and the construction of relevant variables.

3.2 Parameterization

The estimation procedure is a two-step procedure. In the first step, I calculate productivity using the data with some externally determined parameters. Together with data on net worth, this allows us to (i) calibrate the parameters describing the financial constraints, and (ii) estimate the productivity evolution equation. The second step employs a Simulated Methods of Moments (SMM) estimator which requires solving the dynamic structural model and pin down the parameters characterizing the costs of R&D investment.

External parameterization I first parameterize several parameters by choosing their conventional values.These parameters and their sources of values are summarized in following table.[14] The risk-free interest rate is chosen to be 0.0575, which is average of real lending interest rates between 2008 and 2011. Because the discounting rate is not separately identified in my model, I refer to the existing literature. Midrigan and Xu (2014) choose a discounting factor to be 0.92, Gopinath et al. (2017) set the discounting rate to be 0.87, and David and Venkateswaran (2019) set the value to be 0.95. In the end, I choose the discounting rate is chosen to be 0.90, which is smaller than $1/(1+r) = 0.946$.[15] The depreciation rate is set as 10% as in Song et al. (2011). I follow Midrigan and Xu (2014) to use a log-utility function. The capital share α is consistent with David and Venkateswaran (2019). Lastly, the value of the substitution elasticity is from the survey by Head and Mayer (2014).

Table 1: Baseline external parameters

Parameters	r	δ	β	ϵ	α	σ
Values	0.0575	0.10	0.90	1.00	0.50	5.00

Internal estimation Using the revenue production function (3), the physical productivity is estimated as:

$$\phi_{it} = \frac{(p_{it}y_{it})^{\bar{m}}}{k_{it}^{\alpha} l_{it}^{1-\alpha}} \tag{36}$$

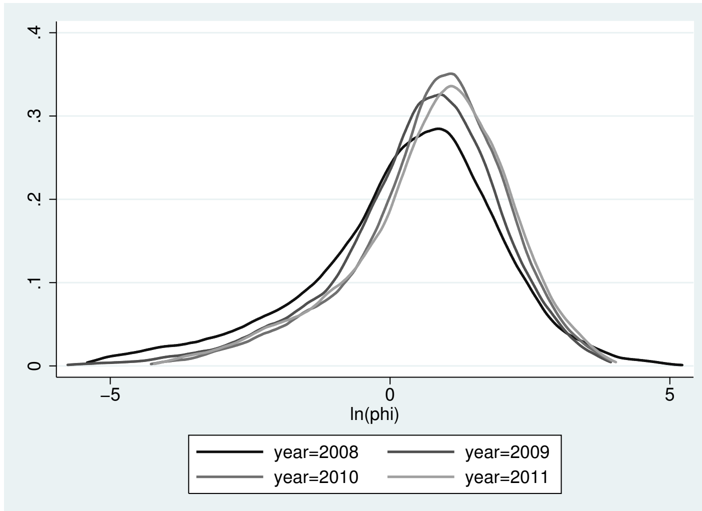
I use firm's value added to measure $p_{it}y_{it}$. k_{it} is measured using the firm's deflated fixed assets. I use wage bill to measure l_{it} in order to account for the unobserved differences in human capital composition in different firms. Figure 3 displays the kernel density of estimated productivity. The logged productivity ranges from -5 to 5, showing a large dispersion.

Note that θ, η, and w jointly affect the firm's choices of capital and labor. Because these choices are static, I calibrate them using the cross-section moments including averages of capital-to-net worth ratio, capital-to-productivity ratio, capital stock, and number of employees, as well as 0.25, 0.5, 0.75 percentiles of employees and capital. I choose (θ, η, w) to minimize the distance between the model-generated moments and these targeted moments. In Table 2, I display the value of targeted moments in the data and the calibrated model. The calibrated values for (θ, η, w) are presented in Table 3. $\theta = 0.324$ implies that only around 32% of the physical capital and intangible assets can be used as

[14] The capital share can potentially be backed out using the data if no distortions are imposed in the labor market. Note that by (8) I know that $\alpha = 1 - \bar{m}\omega l_{it}/y_{it}$. Therefore α can be estimated as $\hat{\alpha} = \frac{1}{NT}\sum_{n=1}^{N}\sum_{t=1}^{T}\left(1 - \frac{\bar{m}\omega l_{it}}{y_{it}}\right)$.

[15] The value of discounting rate will mainly affect the value of estimated R&D costs, it does not affect the quantitative results of analysis of aggregate TFP loss.

Figure 3: Kernel density of estimated productivity



collateral when the firm borrows from the financial institutions. This value is smaller than that is used in the literature. One possible reason is that the intangible asset can also be used as collateral in the model. $\eta = 0.513$ implies that the one percent improvement in the productivity will lead to 0.513 percent increase in the pledge-able collateral when using external financing. Note that $\eta < \sigma - 1$, implying that the cut-off productivity function ($\bar{\phi}(a_{it}, \theta)$) is increasing in a_{it}. This predicts that wealthier firms are less likely to be constrained in the empirical model.

After obtaining the productivity estimates, I can estimate the productivity process using a regression as follows:

$$\ln(\phi_{it+1}) = \rho \ln(\phi_{it}) + \gamma \ln(x_{it} + 1) + \mu_{jt} + \xi_{it+1} \tag{37}$$

where μ_{jt} denotes a three-digit industry-year fixed effect. μ_{jt} captures the factors that affecting the productivity evolution while not capture by our theoretical model. I apply OLS estimator to estimate this linear model. The variance of the error term σ_ξ^2 is estimated by the sample variance of the residuals. This step gives me estimates of $\left(\rho,\, \gamma,\, \sigma_\xi^2\right)$. I present the estimation results in the three middle columns of Table 3. I can see that the estimated endogenous productivity process has a persistence of 0.336. R&D investment shifts up the mean of the distribution of future productivity. In particular, one percent increase in R&D investment leads to around 0.056 percent increase in the mean of future productivity. The estimate of σ_ξ is 1.264, indicating that there is a relatively large dispersion in the exogenous shock to productivity. I use a similar method to back out the productivity process without R&D investment. In the absence of R&D investment in the productivity evolution, the estimated productivity process has a larger persistence of 0.349. This upward bias is mainly driven by the positive correlation between R&D investment and current productivity. Now the inferred dispersion of

Table 2: Targeted moments for internal calibration of (θ, η, w)

targeted moments	data	model
$\log(k_{it})$:		
mean	4.35	3.66
25th percentile	3.35	2.69
50th percentile	4.53	4.72
75th percentile	5.50	6.06
$\log(l_{it})$:		
mean	4.87	4.59
25th percentile	4.17	3.22
50th percentile	4.92	5.74
75th percentile	5.63	7.41
$\overline{\log(k_{it}/a_{it})}$	-0.51	-1.20
$\overline{\log(k_{it}/\phi_{it})}$	3.77	3.08

the productivity shocks turns to be 1.274, which is greater than that inferred from the endogenous productivity process.

Table 3: Parameters determined using the data

Estimation Method	Calibration		OLS			SMM	
	θ	η	ρ	γ	σ_ξ	f	d
Endogenous productivity	0.324	0.513	0.336***	0.056***	1.264	73.21***	0.073***
			(0.006)	(0.002)		(1.684)	(0.004)
Exogenous productivity	0.324	0.513	0.349***	0	1.274	n.a.	n.a.
			(0.006)				

Note: the OLS estimation contains a full set of industry-year fixed effects. For the OLS estimation, standard errors clustered at the 3-digit sectoral level are in the parenthesis. the *** indicates significance level at 1% significance level.

Finally, I use simulated methods of moments (SMM) to estimate the parameters for R&D costs, (f, d). I pool the data from 2008 to 2010. Given the observed net worth and productivity (a, ϕ), I solve the model to find the parameters to minimize the distance between model-generated optimal R&D investment and net worth accumulation policies and the data. Define the vector of moments for observation i in year t as:

$$\mathbf{m}_{it}(f, d) = \begin{pmatrix} \mathbb{I}(x_{it} > 0) \\ \ln(x_{it} + 1) \\ \ln(a_{it+1}) \end{pmatrix}_{\text{model}} - \begin{pmatrix} \mathbb{I}(x_{it} > 0) \\ \ln(x_{it} + 1) \\ \ln(a_{it+1}) \end{pmatrix}_{\text{data}}$$

The GMM estimator for (f, d) is obtained by minimizing following objective function:

$$L(f,d) = \frac{1}{2}\left[\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{m}_{it}(f,d)\right]' \mathbf{W}(f,d)\left[\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{m}_{it}(f,d)\right] \tag{38}$$

where the weighting matrix is $\mathbf{W}(f,d) = \left[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{m}_{it}\mathbf{m}_{it}'\right]^{-1}$. For a given pair of parameters (f, d), I solve the model, simulate the optimal R&D and wealth accumulating choices, and compute the objective function. Then I find parameters that minimize the objective function.To tackle the problem of possible local maximizers, I use the Markov Chain Monte Carlo (MCMC) estimator suggested by Chernozhukov and Hong (2003).[16] The estimation results shows that the fixed cost is around 7.32 million RMB (equivalent to around 1 million USD), which is significant at 1% significance level. The fixed costs help explain that a large fraction of firms do not participate in R&D investment. The intensive margin of R&D investment is characterized by the quadratic costs. The estimate of d is 0.073 and significant at 1% significance level.

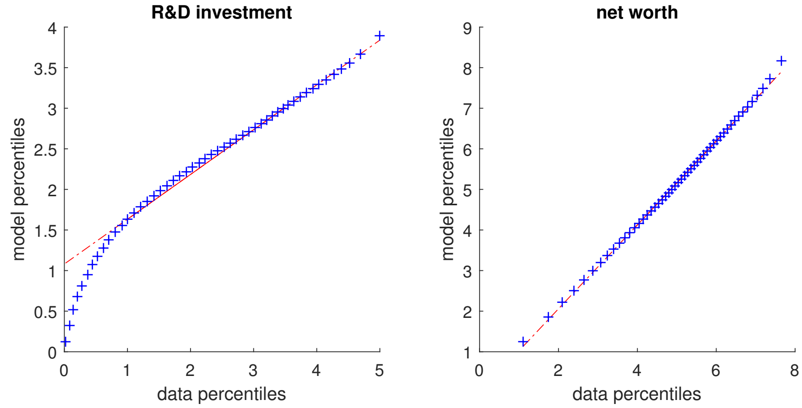
3.3 Model fit

The estimated model provides a good match for the observed R&D investment and net worth accumulation decisions. The estimated model predicts that 12% of firms undertake R&D investment, and the data shows that 15% of firms are active in R&D activities. In the left panel of Figure 4, I plot the percentiles of the simulated R&D investment distribution and future net worth distribution against those observed in the data. I can see that in either case, the fitted line is almost straight and lies along the 45-degree line, suggesting that the model simulated sample and the data sample have a similar distribution. For the R&D investment, I see that the model matches relatively worse in terms of lower percentiles of the R&D distribution, this may suggest that R&D costs may differ across firms. A larger variable cost may be able to generate small R&D investment. Later I will discuss an extension of the model with heterogeneity in R&D costs. Looking at the right panel of Figure 4, I see a very tight match between the model and the data. This shows that the model is successful in predicting the decision on net worth accumulation.

When comparing the results from the endogenous productivity model, it is important that these model can generate close predictions of observed outcomes. Though the exogenous model always predicts zero R&D investment, it produces very similar results of net worth accumulation rule and output. In the left panel of Figure 5, I plot the percentiles of model-predicted future net worth against that being observed in the data. It shows a tight match between the model and data in terms of the choice of net worth accumulation. In the right panel of Figure 5, I show the future output predicted by these two models are almost the same given the same states observed in the data. These results

[16]See Appendix F for a description of full computing procedures.

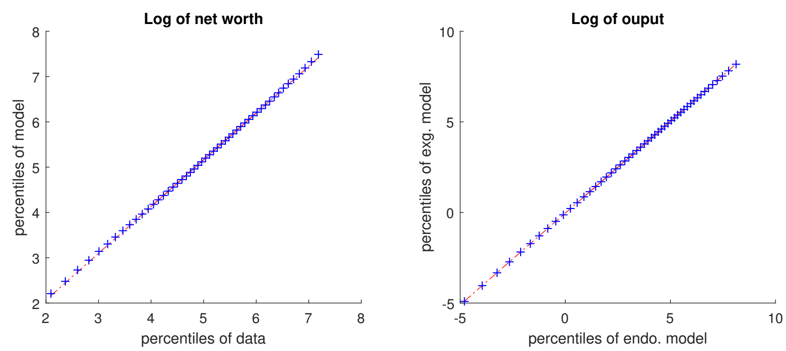
Figure 4: Model fit for endogenous productivity model



Note: the percentiles range from the 1st percentile to 99th percentile.

show that the exogenous productivity model and endogenous productivity model can match the net worth and output data equally well.

Figure 5: Model fit for exogenous productivity model



Note: the percentiles range from the 1st percentile to 99th percentile. The positive part of R&D investment is displayed.

3.4 Non-targeted moments

Now I investigate the model fit for the non-targeted moments. In particular, I explore several empirical implications from our quantitative model with R&D investment. I begin by discussing some firm-level implications of the model. These implications include cross-section correlations as well as firm-level dynamic choices. To evaluate the external validity of the model, I look at two groups of correlation moments. First, I look at the correlation between the log of MRPK ($\ln(R_{it})$) and two state variables ϕ and a. The MRPK is constructed as in Hsieh and Klenow (2009).[17] The second regression equation I am interested in is the relation between the dynamic decisions and the state variables. The closeness in the parameters between the model and the data indicates the model behaves well in terms of rationalizing the capital prices and the R&D decision. Moreover, these two regressions provides two main firm-level predictions that are testable using the firm-level data.

Implication 1 Conditional productivity, the firm's MRPK is negatively related to the firm's net worth. Conditional on firm's net worth, MRPK is negatively related to firm's productivity.

Implication 2 Conditional on productivity, both of firm's R&D investment and future net worth are positively related to firm's current net worth. Conditional on firm's net worth, R&D investment and future net worth are also positively related to firm's productivity.

Table 4: Correlation between dynamic choices and state variables

Dependent var.	$\ln(R_{it})$		$\ln(1+x_{it})$		$\ln(a_{it+1})$	
	model	data	model	data	model	data
$\ln(a_{it})$	-0.318***	-0.163***	0.216***	0.219***	0.914***	0.954***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.001)	(0.003)
$\ln(\phi_{it})$	0.824***	0.648***	0.242***	0.0829***	0.294***	0.0244***
	(0.001)	(0.003)	(0.0027)	(0.0028)	(0.0018)	(0.0013)
N	85268	85268	63989	63989	63989	63989
R^2	0.818	0.705	0.309	0.119	0.966	0.911

Note: Standard errors are in parentheses; *** $p < 0.01$.

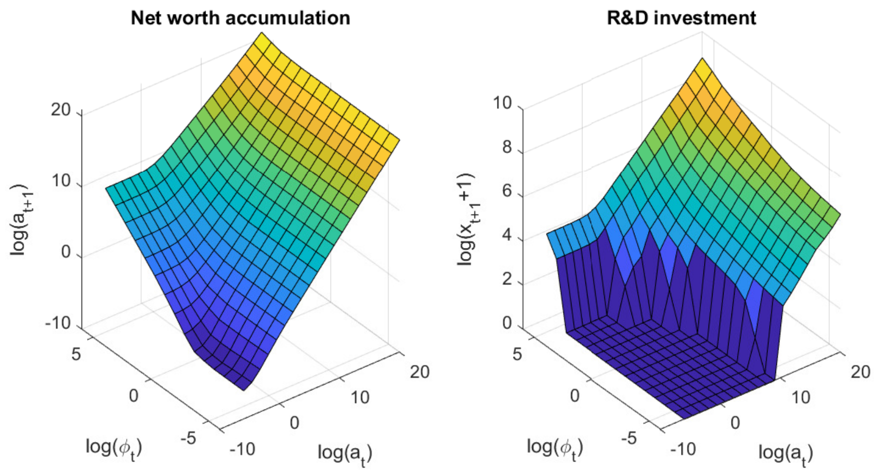
Notice that our model permits a possible negative relation between productivity and MRPK as η differs. The first implication is drawn from the static problem given the estimated parameters. Ceteris paribus, the estimated model predicts that richer producers are less likely to be financially constrained. More productive producers tend to be more financially constrained conditional on their net worth. In the second of Table 4, I do find such correlation patterns in the data. In the data, the partial correlation between $\ln(a_{it})$ and $\ln(R_{it})$ is -0.163, and the partial correlation between $\ln(\phi_{it})$ and $\ln(R_{it})$ is 0.648. This is close to what's being implied by the model.

Figure 6 illustrates the R&D decision and net-worth accumulation decision. The left panel depicts the optimal R&D investment. Conditional on current net worth, firms with higher productivity tend to invest in R&D investment, but less productive firms choose not to invest in R&D investment. When I fix the level of productivity, wealthier firms tend to invest more in R&D investment. Because in our model R&D can only be financed through internal cash flow,[18] the R&D investment is likely to be

[17] See Appendix D for the details.

[18] In the model, the firm's cash flow includes internal profits and interest payments from holding the one-period financial asset.

Figure 6: Optimal choices of R&D and net worth accumulation



constrained by the available financial resources. In other words, firms are going to invest more in R&D investment as they receive more funds. Given that firms are financially constraint, increase the firm's net worth will have a positive impact on firm's R&D investment. Therefore, the model predicts a positive correlation between R&D investment and productivity as well as net worth. The right panel describes the optimal decision of wealth accumulation. Conditional on current net worth, firms of higher productivity tend to save more, anticipating that they are likely to be constrained in the future. In comparison, firms with lower productivity tend to spend more in current period. Conditional on productivity, poorer firms save more to increase its net worth in the future. The binding financial constraint increases the marginal benefits of saving, hence giving firms stronger incentives to accumulate its wealth and escape from the financial constraint.

These results are different from existing theoretical R&D models with financial constraint being absent. These model are silent about the relation between firm's wealth and R&D investment (for example, see Aw et al. (2011), Doraszelski and Jaumandreu (2013) and Eaton and Kortum (2007)). In the middle two columns of Table 4, I present the results for the correlation between R&D investment and state variables. The partial correlation between $\ln(a_t)$ and $\ln(1+x_t)$ stays close to the model. This suggests that firms may face certain level of financial constraint in undertaking R&D investment.[19] Conditional on the net worth, I also see a positive partial correlation between R&D investment and productivity. This confirms that more productive firms tend to undertake more R&D investment. This mechanic relation is modelled in many R&D investment models. Note that the correlation coefficient is weaker in the data, which may suggest that the costs of R&D investment may be positively

[19]Another possible explanation is the cost heterogeneity, which is abstracted in our benchmark model. I have also tried to control firm fixed effects, this positive correlation remain stable.

correlated with the productivity. Currently this is not formally modelled. In our model, productivity and wealth jointly determine the financial status of a firm, which further influences a firm's R&D decision. To some extent, the positive relation between R&D and net worth also provides indirect evidence that financial constraint matters for R&D investment.

The decision rule of future net worth accumulation is shown in the last two columns. The model predicts quite well in terms of the correlation between future net worth and current net worth. This partial correlation in the data is estimated to be 0.954, and the model predicts it to be 0.914. The correlation between future net worth and current productivity is relatively lower in the data. This may imply that there are other unobserved factors affecting the accumulation of assets.

4 Quantitative Analysis

In this section, I present the results of quantitative analysis based on the estimated model. Within the sample, I first show the static and dynamic TFP losses caused by financial constraints. To better understand the impact of financial constraint on productivity dynamics, I also simulate the model forward to evaluate the transition dynamics and the productivity loss in the steady state. In order to understand the role of R&D investment and endogenous productivity in determining the relation between financial constraints and TFP, I compare the results with that of the exogenous productivity model. Lastly, I perform several policy analyses based on the quantitative model.

4.1 Aggregate TFP losses caused by financial constraints

I set year 2008 as the initial period and treat the productivity and net worth recorded in the data as fundamentals. Using the estimated model, I simulate the model forward for 3 periods. Given the simulated state variables a_t and ϕ_t, I calculate the actual aggregate TFP based on (25) and the efficient aggregate productivity using (26) after removing the capital misallocation. A larger value of θ translates into a less tightening borrowing constraint. To investigate the impact of finance on TFP losses, I choose different values for θ and simulate the model to obtain their counterfactual $TFPQ^*$ and other interested outcomes. I then compute the static and dynamic aggregate productivity losses employing Equation (28).

The results of aggregate productivity losses are reported in Table 5. In the estimated models, the exogenous productivity model and the endogenous productivity model predict a similar size of productivity loss from capital misallocation within the length of sample. The average productivity loss from capital misallocation is 37%. But in the model with R&D investment and endogenous productivity, the additional TFP loss from under investment in R&D is around 29%. This implies that productivity loss from the R&D channel is quantitatively important. The impact of finance on aggregate TFP will almost double as I consider the endogenous response of R&D investment. The

Table 5: Dynamic and static TFP losses caused by financial constraint

Value of θ	TFP losses	Year			Average
		2009	2010	2011	
0.324	$Static_{ex}$	0.37	0.37	0.38	0.37
	$Static_{en}$	0.37	0.37	0.38	0.37
	Dynamic	0.22	0.31	0.35	0.29
	Total	0.59	0.68	0.72	0.66
0.5	$Static_{ex}$	0.36	0.35	0.36	0.36
	$Static_{en}$	0.36	0.36	0.36	0.36
	Dynamic	0.21	0.30	0.33	0.28
	Total	0.56	0.65	0.69	0.63
0.7	$Static_{ex}$	0.34	0.33	0.33	0.33
	$Static_{en}$	0.34	0.33	0.32	0.33
	Dynamic	0.19	0.27	0.30	0.25
	Total	0.52	0.60	0.62	0.58
0.9	$Static_{ex}$	0.28	0.26	0.26	0.27
	$Static_{en}$	0.29	0.27	0.26	0.27
	Dynamic	0.15	0.20	0.22	0.19
	Total	0.43	0.47	0.48	0.46

Note: $Static_{ex}$ and $Static_{en}$ represent the static productivity loss predicted by the exogenous productivity model and endogenous productivity model, respectively.

dynamic productivity loss starts at 0.22 in year 2009 and grows to be 0.35 in year 2011. This is because the dynamic TFP loss not only reflects current R&D effort but also past R&D activities. As I increase θ, both the static productivity loss and dynamic productivity loss decreases. When $\theta = 0.9$, the total TFP loss predicted by the endogenous productivity model is 46%, with the static productivity loss being 27% and dynamic loss being 19%. In the estimated model, both the dynamic loss and static productivity loss are relatively stable when the improvement of financial institution is mild. When θ becomes larger, the decrease in TFP loss is more sensitive to the increase in θ. This is because more firms with relatively high productivity get ride of financial constraints.

To further explore the aggregate implications of better financial institutions, I present a summary of other outcome variables in Table 6. To save space, I display the average values of different variables over three years. Both the static loss and dynamic loss are associated with the fraction of constrained firms. In the benchmark model, 44% of the firms are constrained. When I increase θ to be 0.9, only 28% of firms are constrained. The dynamic productivity loss is caused by the decrease in R&D investment. In the estimated model, the fraction of firms undertaking R&D investment is around 11%, while in the non-constrained model 79% of the firms choose to invest in R&D. Not only the extensive margin

Table 6: Aggregate implications of financial constraint: Other outcome variables

Value of θ	0.324	0.5	0.7	0.9	1.0
ζ	0.44	0.41	0.37	0.28	0.00
$\Pr(R\&D^{+})$	0.11	0.15	0.22	0.41	0.79
$\log(\overline{R\&D}^{+})$	0.34	0.42	0.66	1.05	2.92
$\log(\bar{y})$	4.59	4.79	5.07	5.64	15.96
$\log(\bar{c})$	0.87	1.01	1.27	1.92	10.06

Note: $\Pr(RD^{+})$ means the share of firms undertaking R&D investment. $\log(\overline{R\&D}^{+})$ is the log of average R&D investment for firms with positive R&D investment.

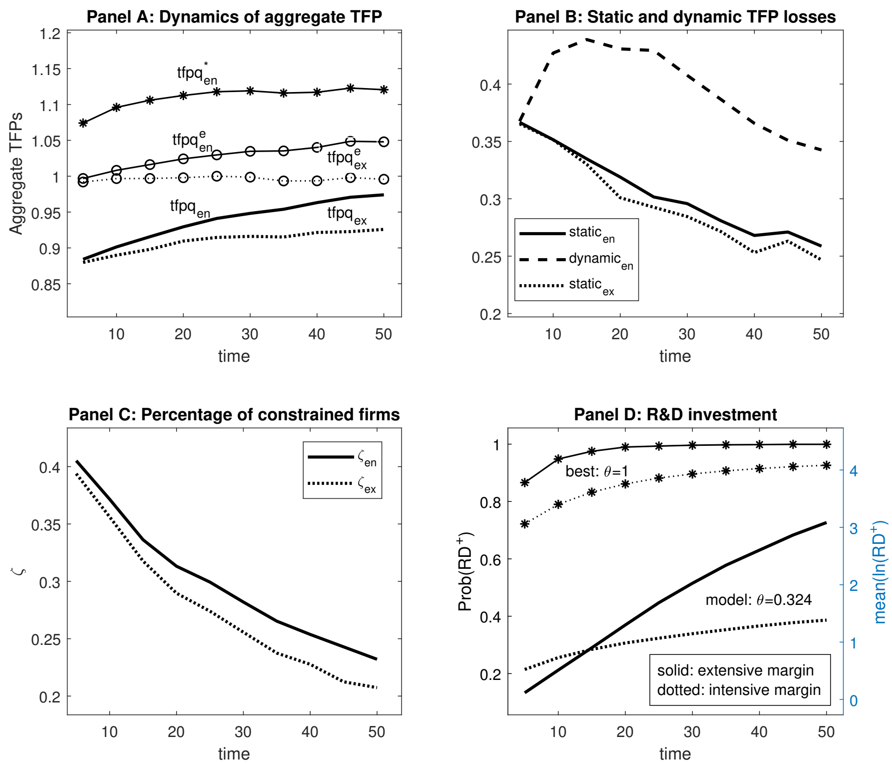
matters for the dynamic loss, the intensive margin of R&D investment also plays a role in affecting the productivity evolution. Focusing on the firms with positive R&D investment, I see the average R&D investment increases as I increase θ. In particular, the log of average R&D investment (for positive-R&D firms) jumps from 0.34 in the estimated model to 1.05 when $\theta = 0.9$. In the absence of financial frictions, this number in enlarged to be 2.92. Lastly, financial frictions also cause substantial losses in output and consumption.

4.2 Transition dynamics of TFP losses

Existing studies have emphasized the importance of technology adoption in the evaluation of the impact of financial frictions on aggregate TFP (Midrigan and Xu, 2014; Restuccia and Rogerson, 2017; Vereshchagina, 2018), but much less is known about the transition dynamics of the TFP losses from the static and dynamic channels. It is interesting to see how the self-financing can undo financial constraints along the transition. Specifically, how will the incentives to perform R&D activities will alter the efficacy of self-financing? To answers these questions, I examine the impact of financial constraints on productivity dynamics and aggregate variables. Given firms' initial states observed in 2008, I simulate the model forward for 50 years and compute the interested variables.

I present the results in Figure 7. Panel A shows the transition dynamics of aggregate TFP. For the endogenous productivity model, I am interested three levels of aggregate efficiency: actual aggregate TFP ($tfpq_{en}$), aggregate TFP with efficient capital allocation ($tfpq^{e}_{en}$), and the first-best TFP ($tfpq^{*}_{en}$). As time evolves, $tfpq_{en}$ increases steadily with a growth rate higher than $tfpq_{ex}$. This is driven by two forces: first, a reduction in capital misallocation because of wealth accumulation. Second, the investment in R&D activities drives up the fundamental productivity. In the world where the capital constraint is absent, the fundamental productivity is also increasing over time as firm's R&D efforts continuously contribute to the productivity growth. The fundamental aggregate TFP is a key feature of endogenous productivity process. As I observe in the Panel A of Figure 7, the efficient productivity $tfpq^{e}_{ex}$ is stable over time. This implies that firm's ability of earning profits and demand of external

Figure 7: Transition dynamics of the financial constraint



Note: Labels with 'en' represent the model with endogenous productivity, labels with 'ex' denote the model with exogenous productivity. The model simulation starts at year 2008. To focus on the transition path and the impact of initial conditions, I drop the first five years. I normalize all the aggregate TFP using the initial level of $tfpq^e_{en}$.

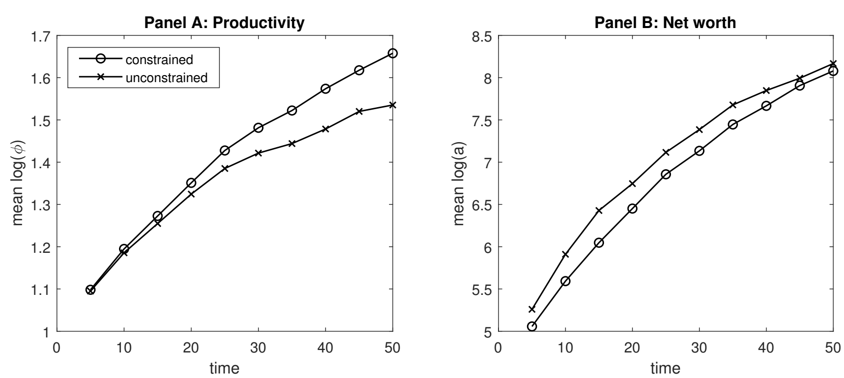
financing are unchanged over time. As firms accumulate wealth, the actual aggregate TFP ($tfpq_{ex}$) also increases because of a lower degree of capital misallocation.

Panel B shows the transition dynamics of TFP losses. In the exogenous productivity model, the only TFP loss is the static loss caused by capital misallocation. Over time, some firms can overcome the financial constraint through self-financing. This results in a reallocation of capital towards more productive firms and an improvement in the aggregate TFP. This mechanism is also important in explaining the pattern observed in the endogenous productivity model in which I also see a decreasing trend of TFP loss from capital misallocation. The interesting finding is that the transitional speed is slower in the endogenous model. The endogenous growth of fundamental productivity counteracts the potency of self-financing in undoing financial constraints. Two competing forces are at work.

As firms become more productive, they earn more profits. This may relax the firm's financial constraints. On the other hand, more productive firms need a larger amount of external financing, which may exacerbate the financial constraint. However, the model with exogenous productivity fails to provide a framework for analyzing the dynamic interaction between financial constraints and productivity changes. The estimated model shows the second force dominates. This is supported by Panel C which shows that firms in the exogenous productivity model has a faster speed of escaping from financial constraint.

More importantly, I find a non-monotone trend of the dynamic TFP loss as time evolves. The dynamic productivity loss increases first and declines afterwards. This is mainly caused by the gap in R&D investment between the estimated model and the counterfactual case where $\theta = 1$. This is depicted in Panel D. At the beginning, the gap in R&D investment between these two scenarios is large both for the extensive margin and intensive margins. Because the difference in current productivity reflects all of the past R&D activities, the dynamic TFP loss increases as the differences in R&D investment pile up. As firms accumulate wealth, they catch up by investing more in R&D. This is especially effective at the extensive margin. As more firms are undertaking R&D investment, the loss in R&D investment shrinks. On the other hand, R&D activities in more distant history tend to be less important for current productivity due to a discounting factor. At the beginning, the dynamic TFP loss actually dominates the evolution of the total productivity loss along the transition path. As a result, the total TFP loss also increases first and then declines. This indicates that endogenous R&D investment matters for the understanding of transition dynamics of TFP losses. Because endogenous productivity growth, the ability of self-financing in easing firms' financial constraints is weaker.

Figure 8: Characteristics of constrained firms vs. unconstrained



As a supporting evidence for our discussion above, lastly I show the characteristics of constrained firms and unconstrained firms of the endogenous productivity model in Figure 8. The constrained (unconstrained) firms refer to those firms whose capital constraints are (not) binding. Panel A shows the mean of logged productivity for the constrained and unconstrained firms. Over time, relatively more productive firms are constrained, indicating that more productive firms are more difficult to get rid of the financial constraint through self-financing. On the other hand, relatively poorer firms are constrained, but the gap between the unconstrained firms and constrained firms are narrowed over time. This is mainly because firms who are more financially constrained have stronger incentives to save in order to grow out of the financial constraint. With the endogenous R&D investment, a firm compares the benefits of improving its future productivity with increasing future net worth. The trade-off between R&D investment and net worth accumulation makes it harder for the productive firms to escape from the financial constraint.

4.3 TFP losses in the steady state

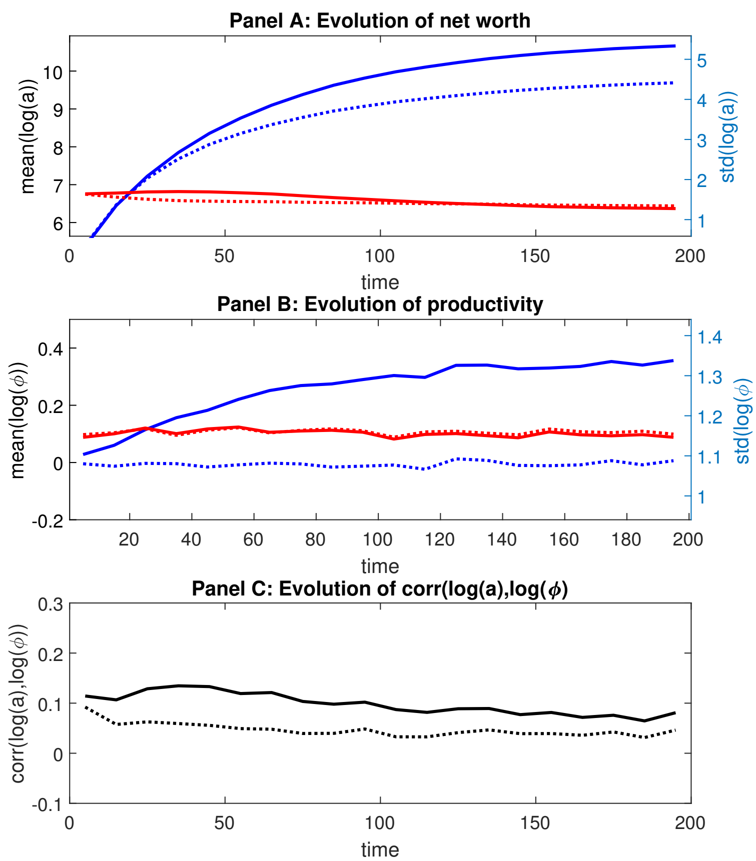
After examining the implication of R&D investment for the transition dynamics of aggregate TFP, I now investigate how R&D investment determines aggregate outcomes in the steady state. In Figure 9, I show that the joint distribution of (a, ϕ) converges to a stationary distribution. With R&D investment, firms are able to improve its productivity, produce more goods, and make more profits. This in turn allows firms to accumulate more wealth. Despite the model are calibrated using the same data, the endogenous productivity model features a richer and more productive economy in the stationary equilibrium.

In Table 7, I show these interested indicators.[20] The first indicator I am interested is the fraction of constrained firms, denoted by ζ. Recall that in our model with static capital investment decision, ζ is also the average cash flow-investment sensitivity. With the option of R&D investment, relatively more firms are constrained. Accordingly, I observe a larger static aggregate TFP loss (18.4%) in the endogenous productivity model than in the model with exogenous productivity (16.7%). In contrast, the aggregate TFP is much higher in the model with endogenous R&D investment. This implies that the main difference in the aggregate efficiency is driven by a difference in the fundamental productivity distribution. Most importantly, in the steady state I still see a substantial dynamic TFP loss, which accounts for 20% of the efficient aggregate TFP. In the last two columns, I present the absolute changes in aggregate TFP loss comparing to the within sample analysis. In the exogenous productivity model, the improvement in aggregate production efficiency is solely driven by a more efficient allocation of capital. With R&D investment, there is also a substantial decrease in the static TFP loss. But the absolute change is smaller than that observed in exogenous productivity model. This is because R&D investment drives up productivity and leads to relatively more constrained firms eventually in the steady state.

I also observe a decline in dynamic TFP loss (around 9%) in the steady state compared to the initial periods of the sample. The reason is that the accumulation of net worth allows some firms to overcome the capital constraint and undertake more R&D investment. However, because of a

[20]These variables are the averaged outcome of the last ten periods of the simulation.

Figure 9: Transition of state variables to the steady state



Note: In all panels, solid (dotted) lines represent the endogenous (exogenous) productivity model. For Panel A and Panel B, blue (red) lines represent the mean (standard deviation) of the interested variable.Starting for the initial period of the sample, I simulate the model for 200 periods.

persistent effect of R&D on a firm's productivity, the decrease in dynamic TFP loss is only .09, which is less than the half of the reduction in static TFP loss. In summary, the endogenous R&D investment affects the productivity dynamics in two ways. First, the endogenous productivity growth makes the static productivity loss more persistent over time. Second, the enduring impact of R&D investment on productivity greatly weakens the efficacy of self-financing in reducing dynamic productivity loss.

I conclude by connecting our findings to an insightful study by Moll (2014) on the the role of self-financing in undoing financial constraints. Using a tractable dynamic general equilibrium model

Table 7: Characteristics of the steady state

Indicators	ζ	$tfpq$	*TFP losses*		*ΔTFP losses*	
			Static	*Dynamic*	*Static*	*Dynamic*
Exogenous productivity	0.13	3.86	0.167	0.00	-0.21	0
Endogenous productivity	0.15	4.17	0.184	0.20	-0.19	-0.09

in which heterogeneity entrepreneurs face collateral constraints, Moll shows that the persistence of idiosyncratic productivity shocks determines both the size of steady-state TFP losses and the speed of transitions. When shocks are persistent, steady-state TFP losses are small but transition to the steady state takes a long time. The mechanism is producers are more able to accumulate wealth through saving when the productivity shocks are more correlated over time. This provides a good benchmark in understand our results. Since R&D investment enters into the productivity process, the productivity process in our model features an endogenous persistence of productivity. The auto-correlation of productivity depends on the state variables. In particular, the persistence of productivity is $corr(\ln(\phi_{t+1}), \ln(\phi_t)) = \rho + \gamma corr(\ln(\phi_t), \ln(x_t(a_t, \phi_t))$. The endogenous component of the productivity persistence is determined by the correlation between current productivity and R&D investment. When the correlation is positive, model with endogenous R&D investment has a larger productivity persistence. Using the argument by Moll (2014), I immediately know that this will prolong the transition and reduce the static TFP loss in the steady state. In this case, how come I find a larger static TFP loss in the endogenous productivity model?

The analysis above has ignored the level effect of R&D investment. By diverting some economic resources to innovation investment, producers also enhance their fundamental productivity over time. This weakens the efficacy of self-financing along the transition as productive firms require more external financing. Even though firms can partly undo the TFP loss through self-financing along the transition, the steady-state TFP loss is larger in the endogenous productivity model because the fundamental productivity is higher. Introducing endogenous R&D investment triggers a race between the accumulation of assets and productivity growth. The TFP loss along transition and in the steady-state is determined by relative speed of wealth accumulation and productivity enhancing. Our empirical models shows that the productivity-enhancing channel wins the race and causes a large TFP loss during the transition and in the steady state.

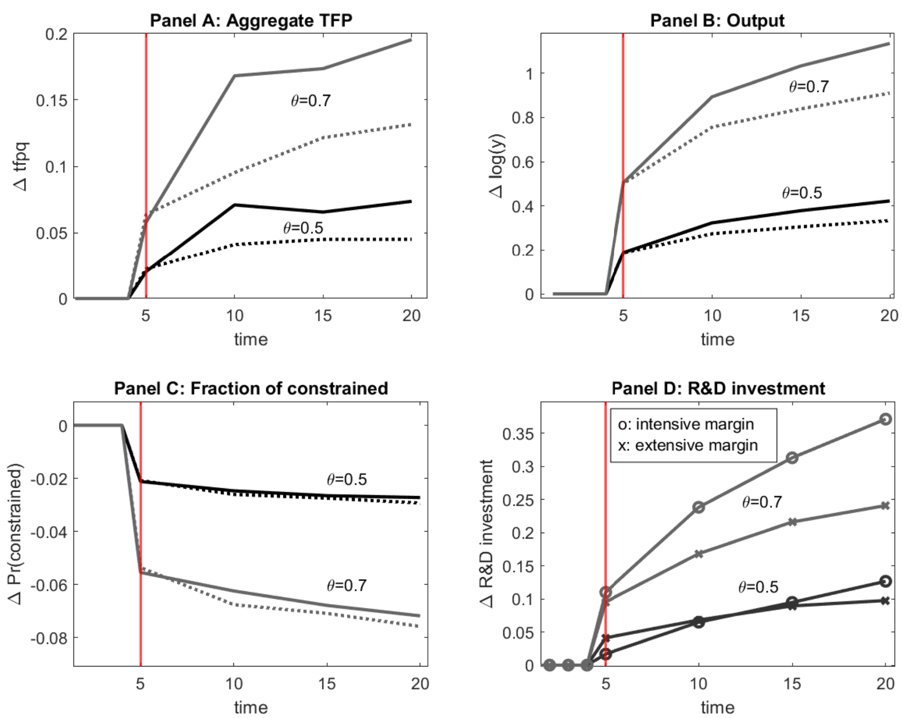
4.4 Policy analysis

4.4.1 Financial policy

I conduct two counterfactual experiments to analyze two kinds of financial policy. First, ask how incorporating R&D investment affects the policy implication of a financial reform. I then employ the estimated model to analyze consequences of a credit crunch for both the endogenous productivity model and the exogenous productivity model.

Financial reform A financial reform is an action that improves the efficiency of financial system permanently. I study this financial reform by enlarging θ perpetually at some time and keep track of the evolution of interested variables.[21] To understand the role of R&D investment, I perform the same experiment for both of the endogenous productivity and exogenous productivity models. I simulate the model starting from year 2008 and introduce the financial reform at the fifth year.[22] I entertain with two values of θ: 0.5 and 0.7, indicating different levels of financial deepening.

Figure 10: Effects of financial reform



Note: In Panels A, B, and C, dotted lines denote the exogenous productivity model while solid lines represent the endogenous productivity model with R&D investment.

The results of this counterfactual exercise are presented in Figure 10. Panel A shows the response of aggregate TFP upon the initiation of financial reform. The aggregate TFP increases immediately

[21] A larger θ may reflect the improvement in monitoring technology which increases the cost of defaulting.

[22] One can also introduce the policy shock at the steady state. In our case, I find qualitatively similar results. To save space, I only present the results of introducing the financial reform at non steady-state equilibrium.

because of the reduction in capital misallocation as more financial resources are allocated to more productive firms. In the year of starting the financial reform, the increase of $tfpq$ is almost the same for the exogenous productivity model and endogenous productivity model. No matter whether the productivity is exogenous, a deeper financial reform leads to a larger increase in the aggregate TFP. However, one year after the financial reform, with the endogenous R&D investment the aggregate TFP grows more than that of the model with exogenous productivity. This pattern persists afterwards. The change in output display a similar trend, showing that the aggregate output is co-moving with the aggregate TFP in the same direction.

The financial reform also decreases the level of financial constraint. Panel C shows the the change in ζ. Improving the financial system to be $\theta = 0.5$, the percentage of constrained firms drops about 2% (in levels) in the first year. If I change θ to be 0.7, ζ turns to be around 6% (in levels) lower in the first period. As time moves forward, the speed of decline in the exogenous productivity model is slightly slower than the endogenous productivity model with R&D investment. In the model, ζ can also be interpreted as the investment-cash flow sensitivity. Recall that the degree of capital misallocation is positively related to ζ. I can also infer that the capital misallocation decreases less in the endogenous productivity model. In Panel D, I show the change in R&D investment at both the extensive margin and intensive margins in response to the financial reform. At both margins, I see an increase in R&D investment. The increase in R&D investment shift the fundamental productivity distribution to the right and hence pushes up the aggregate TFP. This explains the gains in aggregate TFP is larger in the endogenous productivity model even though the reduction in misallocation is relatively less.

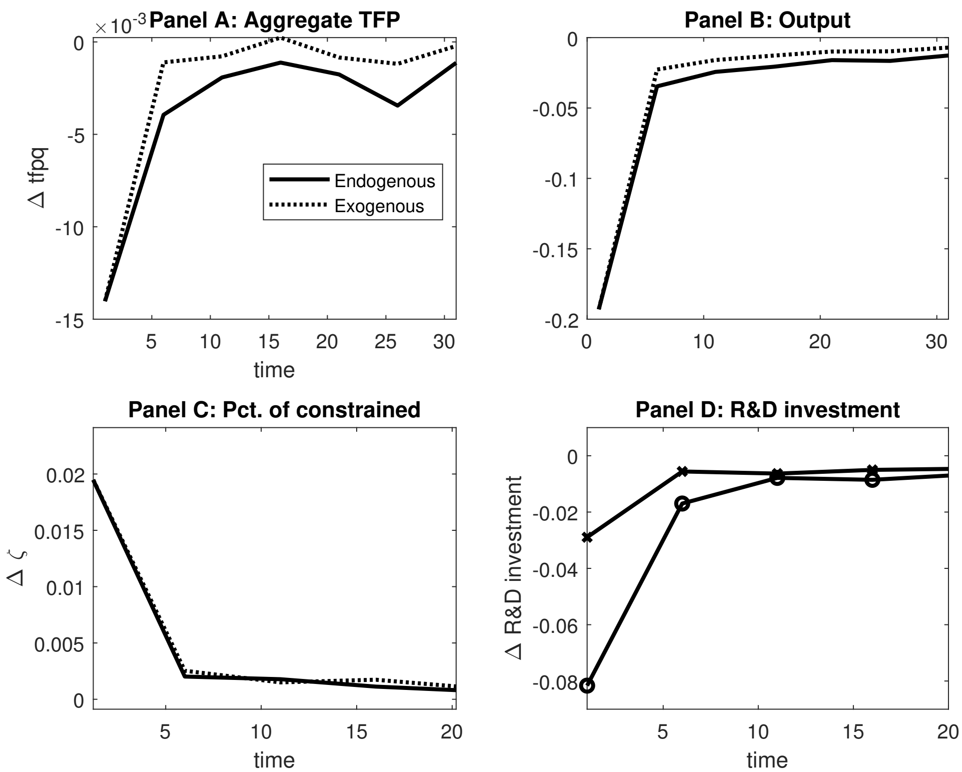
The innovation incentive amplifies the impact of the financial reform on aggregate TFP and aggregate output. Because the growth of fundamental productivity, firms need more external finance for physical capital investment. This in turn may exacerbate the firm's financial constraint. Empirically, this could be tested investigated by comparing the response of productivity and financial constraint to the financial reform for countries (or industries) with different levels of R&D intensity. For countries (or industries) that are more innovative, I expect that the productivity growth is higher while the reduction in financial constraint is lower. Due to data constraint, I do not provide formal empirical tests here. I think this is an interesting empirical question to be studied in the future.

Credit crunch A credit crunch is a tightening of credit supply, which is reflected by a decrease in θ. When a credit crunch hits the economy, I see a decrease in aggregate TFP and output.[23] In the same time, relatively more firms are being constrained. In the model with R&D investment, the recover from a credit crunch is slower and the credit crunch has a long-lasting negative impact on the economy (see Panel A and Panel B in Figure 11). This is mainly because of the decrease of R&D investment which reduces the firm's individual productivity.[24] The effect of the plummeted

[23]Experiments at the steady state generate a similar result, except that the transition back to the steady state is faster.

[24]As supporting evidence, global R&D has experienced strong decline during the financial crisis between 2008 and 2009 (OECD, 2009)

Figure 11: Recover from a credit crunch



Note: the shock of credit crunch is introduced in period 1.

R&D investment continuously lower the aggregate productivity and output. Therefore, the R&D investment not only amplifies the gains from financial reform, but also magnifies the losses from a credit crunch. Considering the response of R&D investment in the evaluation of such policies is important for evaluating the consequence of financial crisis.

4.4.2 Pledgeability of intangible assets

Lastly, I investigate the impact of polices that promote the market of intellectually property rights. To separate the effect of financial reform (a larger θ) from these polices, I fix θ and focus on the counterfactual analysis of a change in η.

Values of η and R&D investment In the model, R&D investment also contributes to the amount of pledged assets which helps reduce the financial constraints. It is interesting to see how the level of pledgeability of intangibles affects the firm's choice of R&D investment. A larger η implies that the pledgeable intangible assets increase dis-proportionally more for relatively more productivity firms, which, on average, makes intangible collateral more important in financing firms.[25] In the parameterized model, the ratio of intangibles to tangibles in the collateral constraint is 2.54%.[26] This number is relatively small compared to US. Loumioti (2012) finds that using intangible as collateral increases loan size by approximately 18%. This implies the value of η to be 1.579. I show the relation between $MRPK$ and productivity in Figure 12. In the left panel, I show the case when firms have low level of net worth. In this case, the cut-off productivity above which firms are constrained are relatively low. For these constrained firms, increasing η slightly decreases the amount of pledgeable intangible assets. This increases the shadow prices of capital, rendering these firms more likely to be financially constrained. However, as firms get more productive, a larger η means more intangible collateral is available to the firms when obtaining external financing. For richer firms that have high net worth, the cut-off productivity above which firms face financial constraint is high. As a result, a larger η leads to a lower $MRPK$ for the constrained firms. Because firms with high productivity and low net worth are easier to be financially constrained, a larger η will relieve their financial constraints relatively more.

Figure 12: Values of η and MRPK

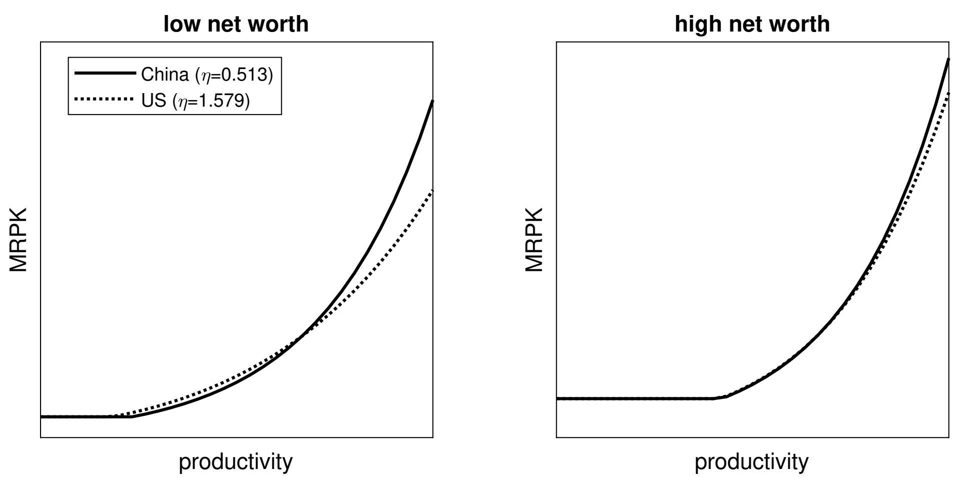


Table 8 shows a comparison between China of US in terms of TFP losses and R&D participation. When we improve the collateralization of intangibles in China to be as US, both the static TFP loss and the dynamic TFP loss decrease. However, the dynamic TFP loss declines (from 29.1% to 22.0%) more than the static TFP loss (from 37.3% to 36.9%). This is because the endogenous response of R&D

[25] See Appendix B for the micro-foundation for the relation between η and the potential of using intangible assets as the collateral.

[26] This ratio is calculated as the average of $\theta\phi_{it}^{\eta}/a_{it}$ by pooling all firms together.

investment when increasing the pledgeability of intangibles. At the extensive margin, the percentage of firms undertaking R&D investment increases from 10.6% to 19.7%. As more firms undertake R&D investment, the dynamic TFP loss decreases. The relative mild decrease in static TFP losses implies that policies aimed at increasing the usage of intangible assets as collateral may be more effective in increasing TFP through stimulating R&D investment.

Table 8: Relation between η and R&D investment

		TFP losses:		R&D investment:	
Country	η	*static*	*dynamic*	$Pr(R\&D^{+})$	$\log(\overline{R\&D}^{+})$
China	0.512	0.373	0.291	0.106	0.341
US	1.579	0.369	0.220	0.197	0.311

In what follows I provide reduce-form evidence supporting the counterfactual outcome. I first introduce the policy background of the intellectual property mortgage financing in China, then I present the related data and empirical results.

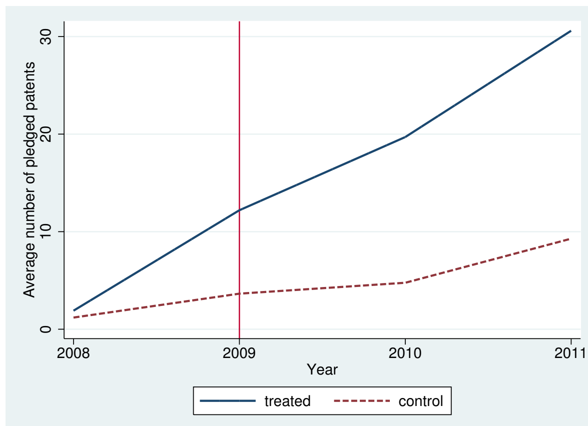
Policy background Intellectual property mortgage financing refers to an enterprise or individual using the legally owned property right of patent right, trademark right and copyright to apply for financing from banks. Using intellectual property rights as collateral is common in developed countries like United States and many European countries. For example, in 2013, 38% of US patenting firms had once pledged patents as collateral to obtain external financing (Mann, 2018). In China, partly due to the lack of protection on intellectual property rights, intellectual property mortgage financing in only at the beginning stage. In 2006, Chinese government has chosen three pilot regions (Pudong District of Shanghai, Beijing, and Wuhan) to launch the intellectual property pledge financing. After three years of trial, the State Intellectual Property Office of China (SIPO)[27] decided to promote and deepen this practice across the country. In 2009, SIPO launched two groups of pilot units for supporting intellectual property mortgage financing. The pilot units include intellectual property offices of 12 cities and/or regions. Their main task is to reduce the costs for firms using intellectual property financing. In employ this policy shock happened in 2009 to investigate the impact of pledgeability of intellectual property rights on R&D investment.

City-level data To verify the effects of this policy, I have manually collected the contracts of pledged patents between 2008 and 2011 from the website of SIPO.[28] The records include a identifier for each patent, the date, the name of the pledger, the name of the pledgee, and the period of validity. I have identified the city of each patent holder for all of the pledged patents included in the database. This

[27]Now renamed as China National Intellectual Property Administration (CNIPA).
[28]The website address is http://www.sipo.gov.cn/tjxx/zlqzyhtdjxgxx/.

gives us information on the total number of pledged patents in different regions. In the graph below, I show the trends of average number of pledged patents for the treated cities and control cities. I can see that the average number of pledged patents for the treated cities had experienced rapid growth since 2008, climbing from close to zero to be over 30 in 2011. In contrast, the growth of the number of pledged patents in the control city is much slower, remaining to be under 10 in 2011. This graphic evidence support that these pilot regions did encourage firms to use patents as the collateral to access external financing from banks.

Figure 13: Number of pledged patents for treated and control groups



To investigate the impact of the increased pledgeability of intellectual properties on the R&D investment. I have constructed a dataset of city-year level R&D investment. The sample period is between 2006 and 2011 so that I am able to control for the pre-trend of R&D investment before the introduction of the policy. The R&D data are from two sources. For years of 2006 and 2007 I have acquired the R&D data from China Industrial Survey. For years between 2008 and 2011, R&D data are from China Innovation Survey. Note that our dataset only only contains the R&D expenditure by firms. But given that firms are the main undertaker of R&D investment, I expect that this will not overturn our empirical results.

Empirical strategy and results To test whether introducing intellectual property mortgage financing had stimulated R&D investment, I adopt the Difference-In-Difference (DID) empirical strategy and specify following econometric model.

$$RD_{ct} = \beta_0 + \beta_1 time_t \times treat_c + \boldsymbol{\gamma}\mathbf{Z}'_{ct} + u_t + u_c + \varepsilon_{ct} \tag{39}$$

where RD_{ct} is the level of R&D investment measured in billions of RMB, $time_t$ is a dummy equal to one for years after 2009 (including 2009), $treat_c$ is a dummy indicating the treated cities. I exclude

Beijing, Shanghai, Wuhan for their early exposure to a similar policy. I define cities belonging to the same province with the pilot cities as the treated cities. This allows us to capture the potential spillover and competition effects that will also increase R&D investment by firms in neighboring cities. By this classification, there are 98 cities out of 344 cities are treated cities in the full sample. $\mathbf{Z}'_{ct}$ include other city-year level control variables that may also affect the R&D performance of the city. In particular, I control for the financial development, trade openness, level of foreign direct investment, GDP, and GDP per capita. u_t and u_c represent the time and city fixed effects, respectively. ε_{ct} is the error term with mean zero.

Table 9: Impact of intellectual property mortgage financing on R&D investment

	Dependent variable: $R\&D$ expenditures			
	(1)	(2)	(3)	(4)
$time_t \times treat_c$	2.254***	0.480***	4.792***	0.642*
	(0.563)	(0.166)	(1.247)	(0.326)
$\mathbf{Z}'_{ct}$	No	Yes	No	Yes
N	1978	1127	596	371

Note: A full set of city and time fixed effects are controlled in all of the regressions. Heterosckedaticity-robust standard errors are in parentheses; *** $p < 0.01$ * $p < 0.1$.

The estimation results are reported in Table 9. In columns (1) and (2), I use the full sample of cities. The coefficient estimate shows that after I control for appropriate city-level factors, on average the treated cities had experienced 0.48 billion RMB additional increase in R&D investment. The result is significant at 1% significance level. I have also tried to use the sample of cities which have at least one pledged patent recorded in the database of pledged patents to form a more reasonable control group. In this case, I end up with 42 treated cities out of 100 cities. The regression results are reported in the last two columns of the table. As shown in column (4), I see that the average increase in R&D investment for the treated cities is 0.64 billion RMB and significant at 10% significance level. These results show that enhancing the pledgeability of intellectual property had stimulated the R&D investment by Chinese firms. This is consistent with the implication of the model.

5 Extensions and Robustness

In this section, I discuss several extensions of our model. These extensions aims to provide robustness checks to our benchmark results. The details on computation are delegated to the appendix.

5.1 Unobserved heterogeneity in R&D costs

In R&D investment model where the financial market is frictionless, the R&D activities are explained by the unobserved cost heterogeneity.[29] In these models, the impact of financial constraint is loaded to the R&D costs and cannot be separated from other factors. By modeling the financial constraint directly, our model explicitly provide a quantification of financial factors in affecting the firm's R&D investment. In the benchmark model, the cost function of undertaking R&D investment is common to all firms. In this case, the cost heterogeneity that affects R&D investment is attributed to the firm's financial condition. In other words, I rely on the heterogeneity in net worth and productivity to explain the R&D investment. This may lead to a bias in appraising the impact of financial constraint on R&D investment. Think of a firm which has no R&D investment. It could either be high costs of undertaking R&D or severe financial constraints that prevent them from investing in R&D.

Table 10: Heterogeneity in R&D costs

μ_f	σ_f	μ_d	σ_d
5.597	1.793	-2.859	3.642
(0.256)	(0.240)	(0.547)	(0.069)

Note: Standard errors are in the parenthesis.

To account for the cost heterogeneity in R&D investment, I introduce randomness to the cost function parameters. In particular, I assume f and d follow independent joint log-normal distributions: $\ln f \sim \mathbf{N}(\mu_f, \sigma_f^2)$, $\ln d \sim \mathbf{N}(\mu_d, \sigma_d^2)$.[30] To reduce the computation burden, I discretize the normal distributions and estimate a model with 9 combinations of different values for (f, d). I then apply MCMC algorithm to obtain the estimates for R&D costs. Computational details are delegated to Appendix F. Table 10 shows the estimation results. The estimated R&D costs display certain degree of dispersion. Especially for the variable cost parameter d, the standard deviation is larger than the absolute value of its mean, implying a normal distribution with heavy tails. In Table 11, I display the results on productivity loss with R&D cost heterogeneity. Both of the dynamic loss and static loss are close to the benchmark model. In particular, I find a slightly larger dynamic productivity loss from distortions to R&D investment when I introduce the R&D cost heterogeneity to the model. This shows that the cost heterogeneity is unlikely to undo the productivity loss from financial constraint. On the contrary, it actually amplified the productivity loss from distortions in R&D investment decision.[31]

[29] For example, see Aw et al. (2011), Peters et al. (2017), and Chen (2019).

[30] I also considered including a correlation coefficient. The estimation results shows the correlation is close to zero and not statistically significant.

[31] Separating the impact of heterogeneity in the fundamental R&D costs and financial constraint will be important for the future work on quantifying the impact of financial constraint on innovation and TFP.

Table 11: TFP losses in model extensions

Models		*static TFP loss*	*dynamic TFP loss*
Benchmark		0.37	0.29
Heterogeneity in R&D costs		0.39	0.36
Industrial heterogeneity	high-tech	0.36	0.30
	low-tech	0.38	0.33
Endogenous uncertainty		0.37	0.30

5.2 Sectoral heterogeneity in innovation technology

In the baseline estimation, I impose a common cost-benefits structure of the R&D investment for all the industries. However, the participation in R&D activities differ across industries. It is documented that different industries may have different costs and benefits of innovation (Peters, Roberts and Vuong, 2016). To examine the impact of heterogeneity, I classify the industries into high-tech and low-tech industries based on the *2002 Catalogue of High-tech Industry Statistics* developed by the China National Bureau of Statistics.[32] According to this classification, in the sample 56 out of 639 four-digit industries are high-tech. High-tech firms perform better in participating R&D investment. In the dataset, 43.4% of high-tech firms engage in R&D activities while 16.8% of non-high-tech firms undertake R&D investment. The R&D intensity (R&D-to-sales ratio) also differs between the high-tech industry and the low-tech industry. The low-tech industry has an R&D intensity of 0.5% while the high-tech industry 1.8%. I can expect that the participation in R&D activities is an outcome of differences in costs and benefits of R&D investment.

5.2.1 Estimation by high-tech and low-tech industries

To understand the importance of industrial heterogeneity in affecting the results of our quantitative analysis, I estimate the empirical model by high-tech and low-tech industries. First, I estimated the productivity evolution equation separately for high-tech and low-tech industries. The estimation results are reported in Table E.2. All of the coefficients estimates are significant at 1% significance level. The productivity do differ across this two groups of industries. The productivity is more persistent in the high-tech industry while the productivity-R&D elasticity is slightly lower in the high-tech industry. Moreover, the dispersion of the productivity shocks is larger in the high-tech industry. I continue to undertake the structural estimation based on the estimated productivity process. The

[32] The catalogue mainly refers to the methods adopted by the Organization for International Economic Cooperation and Development (OECD), and adopts relatively high standards according to the R&D intensity of manufacturing industry and the actual status of development of China's industry. In the 2002 Catalogue, high-tech industries are divided into 8 major sectors, covering 59 manufacturing industries and 2 software services.The eight major areas are: nuclear fuel processing, information chemical manufacturing, pharmaceutical manufacturing, aerospace manufacturing, electronics and communication equipment manufacturing, electronic computer and office equipment manufacturing, medical equipment and instrumentation manufacturing, and public software services.

structural estimation process is the same except that now I separately estimate the R&D costs for high-tech and low-tech industries. The estimation results are shown in Table 12. I find smaller fixed costs and marginal costs for R&D investment in the high-tech industries than in low-tech industries. This suggests that innovative ideas are easier to find and implement in the high-tech industries compared to the low-tech industries. In the high-tech industry, the financial constraint causes a dynamic productivity loss around 0.30 and a static productivity loss around 0.36, which is very close to the benchmark results. In the low-tech industry, the productivity losses are slightly higher. This result is mainly driven by the relatively higher productivity-R&D elasticity in the low-tech sectors. Again, these results confirms the robustness of our benchmark results.

Table 12: R&D costs parameters for high- and low-tech sectors

Sectors	Productivity evolution			R&D costs	
	ρ	γ	σ_ϵ	f	d
High-Tech	0.384	0.0513	1.289	70.14	0.031
Low-Tech	0.332	0.0564	1.273	113.49	0.135

I am also interested in the long-run dynamics of productivity loss. To this end, I also simulate the model to see the transition dynamics of the share of dynamic productivity loss in total productivity loss. Both for high-tech industries and low-tech industries, I see a declining trend for the importance of dynamic productivity loss caused by under investment in R&D. This reflects that the static loss caused by capital misallocation is declines slower as the wealth accumulates. This is also consistent with our benchmark quantitative results. This shows that our analysis is robust to considering industrial heterogeneity in R&D costs and benefits.

5.3 Innovation with endogenous uncertainty

Now I examine the robustness of our results when subject to a different innovation technology. In our benchmark specification, the future productivity is only subject to exogenous productivity shocks. But innovation is full of uncertainties and risks. To capture the impact of uncertainties underlying in the innovation process, I extend the formulation of the productivity process specified by Warusawitharana (2015). The productivity improvement is assumed to be step-by-step. Let $\kappa_{it} \in \{0, 1\}$ be the random variable representing the innovation outcome, with $\kappa_{it} = 1$ meaning the innovation outcome is successful. κ_{it} follows a binomial distribution with probability of success equal to

$$\Pr\left(\kappa_{it} = 1 | x_{it}\right) = 1 - \exp\left(-\psi x_{it}^{\vartheta}\right) \tag{40}$$

where ψ is the parameter governing the overall efficiency of R&D investment in increasing the probability of success, ϑ capturing the curvature of this innovation function. The productivity evolution equation is specified as

$$\ln \phi_{it+1} = \rho \ln \phi_{it} + h\kappa_{it} + \mu_{jt} + \xi_{it+1} \tag{41}$$

where $\xi_{it+1} \sim \mathbf{N}\left(0,\, \sigma_\xi^2\right)$ and μ_{jt} is the industry-year fixed effects. In addition to the exogenous productivity shocks, the uncertainty in the innovation outcome is captured by the randomness in κ_{it}. I employ the probability density function of ϕ_{it+1} and apply the Maximum Likelihood Estimator (MLE) to obtain estimates for the associated parameters. Employing these the estimated productivity process, I then conduct the structural estimation to obtain the R&D costs parameters using the same structure of R&D costs. Notice that under this productivity process, I do not require a large fixed cost to capture the extensive margin of R&D investment. Even the coefficient of the variable R&D costs is much smaller. This is because the productivity process with endogenous uncertainty generates lower returns to R&D investment. As I show in Appendix E, this cost-benefits structure for R&D investment entails a negative relation between R&D investment and productivity, which is not consistent with the data.

Table 13: Parameters for innovation with endogenous uncertainty

Productivity process					R&D costs	
ρ	σ_ξ	ψ	ϑ	h	f_1	d_1
0.335	1.262	0.076	1.080	0.573	$1.9e^{-8}$	$1.7e^{-5}$

Nevertheless, the quantitative results show a similar result for the importance of dynamic productivity losses. In the last row of Table 11, I see that the productivity loss is quite similar to that in the benchmark model. This shows that the results is robust to alternative modelling of R&D investment.

6 Conclusion

This paper studies a dynamic R&D investment model with financial frictions to help understand the role of innovation in shaping the link between finance and aggregate TFP dynamics. A parameterized model consistent with important aspects of firm-level decisions on R&D investment and net worth accumulation shows a sizeable TFP loss caused by distortions of R&D investment and capital misallocation. As time evolves, self-financing does enable some firms to grow out of their financial constraints. However, productivity-enhancing innovation investment undermines the efficacy of self-financing in reducing TFP loss. Dynamic TFP loss is more persistent than the static TFP loss. More interestingly, the dynamic TFP loss amplifies in the beginning as the effect of R&D investment on productivity persists over time. In the long run, the TFP loss generated by distortions of R&D investment is mainly caused by the intensive margin of R&D investment itself. These results show that R&D investment not only affects the size of TFP losses due to financial constraints, but it also affects the transition dynamics of aggregate TFP.

Further counterfactual analysis then shows that innovation investment amplifies the TFP gains from financial reform and causes a more longer-lasting consequence in terms of a credit crunch. This finding implies that the benefits of financial reform may be severely under-evaluated if we ignore R&D investment. The counterfactual with respect to the pledgeability of intangible assets in obtaining loans shows that improving the tangibility of collateral may be an effective measure in reducing dynamic TFP loss but has limited impact on static TFP loss; developing countries can increase TFP by establishing a better market of intellectually property rights.

Because I do not have data on intangible assets, I choose a parsimonious way of modeling the relationship between pledgeable intangible assets and productivity. The accumulation of intangibles, though, is an important channel through which R&D investment helps firms to relieve their financial constraints. Unveiling the function that intangible assets play in reducing firms' financial constraints would be an interesting avenue of future study.

References

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu, "R&D investment, exporting, and productivity dynamics," *American Economic Review*, 2011, *101* (4), 1312–44.

Bento, Pedro and Diego Restuccia, "Misallocation, establishment size, and productivity," *American Economic Journal: Macroeconomics*, 2017, *9* (3), 267–303.

Bhattacharya, Dhritiman, Nezih Guner, and Gustavo Ventura, "Distortions, endogenous managerial skills and productivity differences," *Review of Economic Dynamics*, 2013, *16* (1), 11–25.

Buera, Francisco J and Yongseok Shin, "Financial frictions and the persistence of history: A quantitative exploration," *Journal of Political Economy*, 2013, *121* (2), 221–272.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin, "Finance and development: A tale of two sectors," *American Economic Review*, August 2011, *101* (5), 1964–2002.

Buera, Francisco J, Joseph P Kaboski, and Yongseok Shin, "Entrepreneurship and financial frictions: A macro-development perspective," *Annual Review of Economics*, 2015, *7*, 409–36.

Caggese, Andrea, "Financing constraints, radical versus incremental innovation, and aggregate productivity," *American Economic Journal: Macroeconomics*, April 2019, *11* (2), 275–309.

Chava, Sudheer, Alexander Oettl, Ajay Subramanian, and Krishnamurthy V. Subramanian, "Banking deregulation and innovation," *Journal of Financial Economics*, 2013, *109* (3), 759–774.

Chen, Zhiyuan, "A cost-benefit analysis of R&D and patents: Firm-level evidence from China," *mimeo*, 2019.

Chernozhukov, Victor and Han Hong, "An MCMC approach to classical estimation," *Journal of Econometrics*, 2003, *115* (2), 293–346.

Cornaggia, Jess, Yifei Mao, Xuan Tian, and Brian Wolfe, "Does banking competition affect innovation?," *Journal of Financial Economics*, 2015, *115* (1), 189–209.

Da-Rocha, José-María, Marina Mendes Tavares, and Diego Restuccia, "Policy distortions and aggregate productivity with endogenous establishment-level productivity," Technical Report, National Bureau of Economic Research 2017.

David, Joel M. and Venky Venkateswaran, "The sources of capital misallocation," *American Economic Review*, July 2019, *109* (7), 2531–67.

Doraszelski, Ulrich and Jordi Jaumandreu, "R&D and productivity: Estimating endogenous productivity," *The Review of Economic Studies*, 2013, *80* (4), 1338–1383.

Eaton, Jonathan and Samuel Kortum, "International technology diffusion: Theory and measurement," *International Economic Review*, 1999, *40* (3), 537–570.

_ **and** _ , "Chapter 3 Patents and Information Diffusion," in "Intellectual Property, Growth and Trade," Emerald Group Publishing Limited, 2007, pp. 87–121.

Ek, Chanbora and Guiying Laura Wu, "Investment-cash flow sensitivities and capital misallocation," *Journal of Development Economics*, 2018, *133*, 220–230.

Fazzari, Steven M, R Glenn Hubbard, Bruce C Petersen, Alan S Blinder, and James M Poterba, "Financing constraints and corporate investment," *Brookings Papers on Economic Activity*, 1988, *1988* (1), 141–206.

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez, "Capital allocation and productivity in South Europe," *The Quarterly Journal of Economics*, 2017, *132* (4), 1915–1967.

Gorodnichenko, Yuriy and Monika Schnitzer, "Financial constraints and innovation: Why poor countries don't catch up," *Journal of the European Economic Association*, 2013, *11* (5), 1115–1152.

Hall, Bronwyn H. and Josh Lerner, "Chapter 14 - The financing of R&D and innovation," in Bronwyn H. Hall and Nathan Rosenberg, eds., *Handbook of The Economics of Innovation, Vol. 1*, Vol. 1 of *Handbook of the Economics of Innovation*, North-Holland, 2010, pp. 609 – 639.

Head, Keith and Thierry Mayer, "Gravity equations: Workhorse, toolkit, and cookbook," in "Handbook of International Economics," Vol. 4, Elsevier, 2014, pp. 131–195.

Hochberg, Yael V, Carlos J Serrano, and Rosemarie H Ziedonis, "Patent collateral, investor commitment, and the market for venture lending," *Journal of Financial Economics*, 2018, *130* (1), 74–94.

Hopenhayn, Hugo A, "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica: Journal of the Econometric Society*, 1992, pp. 1127–1150.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and manufacturing TFP in China and India," *The Quarterly Journal of Economics*, November 2009, *124* (4), 1403–1448.

Huggett, Mark, "The risk-free rate in heterogeneous-agent incomplete-insurance economies," *Journal of Economic Dynamics and Control*, 1993, *17* (5-6), 953–969.

Jeong, Hyeok and Robert M. Townsend, "Sources of TFP growth: occupational choice and financial deepening," *Economic Theory*, Jul 2007, *32* (1), 179–221.

Kerr, William R. and Ramana Nanda, "Financing innovation," *Annual Review of Financial Economics*, 2015, *7* (1), 445–462.

Klette, Tor Jakob and Samuel Kortum, "Innovating firms and aggregate innovation," *Journal of Political Economy*, 2004, *112* (5), 986–1018.

Loumioti, Maria, "The use of intangible assets as loan collateral," *Available at SSRN 1748675*, 2012.

Mann, William, "Creditor rights and innovation: Evidence from patent collateral," *Journal of Financial Economics*, 2018, *130* (1), 25–47.

Mestieri, Martí, Johanna Schauer, and Robert M Townsend, "Human capital acquisition and occupational choice: Implications for economic development," *Review of Economic Dynamics*, 2017, *25*, 151–186.

Midrigan, Virgiliu and Daniel Yi Xu, "Finance and misallocation: Evidence from plant-level data," *American Economic Review*, 2014, *104* (2), 422–58.

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

Nanda, Ramana and Tom Nicholas, "Did bank distress stifle innovation during the Great Depression?," *Journal of Financial Economics*, 2014, *114* (2), 273–292.

OECD, "Policy responses to the economic crisis: Investing in innovation for long-term growth," 2009.

Peters, Bettina, Mark J Roberts, and Van Anh Vuong, "Dynamic R&D choice and the impact of the firm's financial strength," *Economics of Innovation and New Technology*, 2016, pp. 1–16.

_ , _ , _ , and Helmut Fryges, "Estimating dynamic R&D choice: an analysis of costs and long-run benefits," *The RAND Journal of Economics*, 2017, *48* (2), 409–437.

Rajan, Raghuram G. and Luigi Zingales, "Financial dependence and growth," *The American Economic Review*, 1998, *88* (3), 559–586.

Restuccia, Diego and Richard Rogerson, "The causes and costs of misallocation," *Journal of Economic Perspectives*, 2017, *31* (3), 151–74.

Robb, Alicia M. and David T. Robinson, "The capital structure decisions of new firms," *Review of Financial Studies*, 2012, p. hhs072.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Growing like China," *American Economic Review*, 2011, *101* (1), 196–233.

Varela, Liliana, "Reallocation, competition, and productivity: Evidence from a financial liberalization episode," *The Review of Economic Studies*, 2018, *85* (2), 1279–1313.

Vereshchagina, Galina, "Financial constraints and economic development: the role of innovative investment," 2018 Meeting Papers 1107, Society for Economic Dynamics 2018.

Warusawitharana, Missaka, "Research and development, profits, and firm value: A structural estimation," *Quantitative Economics*, 2015, *6* (2), 531–565.

Appendices

A Math Appendix

A.1 Marginal cost of capital

The firm's optimization problem is

$$\max_{l_{it},k_{it}} \{y_{it} - \omega l_{it} - r k_{it}\} \tag{A.1}$$

$$s.t.\ k_{it} \leq \frac{a_{it}}{1-\theta} + \frac{\theta}{1-\theta}\phi_{it}^{\eta} \tag{A.2}$$

where $y_{it} = \left(\phi_{it} k_{it}^{\alpha} l_{it}^{1-\alpha}\right)^{\frac{1}{\bar{m}}}$ and $\bar{m} = \frac{\sigma}{\sigma-1}$ is the mark up. Let λ_{it} be the Lagrangian multiplier for the capital constraint, the first-order conditions are:

$$\frac{1-\alpha}{\bar{m}} \left(\phi_{it} k_{it}^{\alpha}\right)^{\frac{1}{\bar{m}}} l_{it}^{\frac{1-\alpha}{\bar{m}}-1} = \omega \qquad \Rightarrow l_{it} = \frac{(1-\alpha)\, y_{it}}{\bar{m}\omega} \tag{A.3}$$

$$\frac{\alpha}{\bar{m}} \left(\phi_{it} l_{it}^{1-\alpha}\right)^{\frac{1}{\bar{m}}} k_{it}^{\frac{\alpha}{\bar{m}}-1} = r + \delta + \lambda_{it} \qquad \Rightarrow k_{it} = \frac{\alpha y_{it}}{\bar{m}\left(\lambda_{it} + r + \delta\right)} \tag{A.4}$$

$$k_{it} = \frac{a_{it}}{1-\theta} + \frac{\theta}{1-\theta}\phi_{it}^{\eta} \text{ iff } \lambda_{it} > 0 \tag{A.5}$$

where the last condition is the complementary slackness condition. Using (A.3) and (A.4) I solve for the optimal choice of labor and capital:

$$k_{it} = \frac{1}{\bar{m}^{\sigma}} \left(\frac{1-\alpha}{\omega}\right)^{(1-\alpha)(\sigma-1)} \left(\frac{\alpha}{\lambda_{it} + r + \delta}\right)^{1+\alpha(\sigma-1)} \phi_{it}^{\sigma-1} \tag{A.6}$$

$$l_{it} = \frac{1}{\bar{m}^{\sigma}} \left(\frac{1-\alpha}{\omega}\right)^{\alpha+\sigma(1-\alpha)} \left(\frac{\alpha}{\lambda_{it} + r + \delta}\right)^{\alpha(\sigma-1)} \phi_{it}^{\sigma-1} \tag{A.7}$$

When $\lambda_{it} > 0$, I can use the binding capital constraint and (A.6) to solve for the capital prices:

$$R\left(a_{it},\, \phi_{it}\right) = D \left[\phi_{it} \left(\frac{a_{it}}{1-\theta} + \frac{\theta\phi_{it}^{\eta}}{1-\theta}\right)^{1-\bar{m}}\right]^{\frac{1}{\bar{m}+\alpha-1}} \tag{A.8}$$

where

$$D \equiv \frac{\alpha}{\bar{m}} \left(\frac{1-\alpha}{\bar{m}\omega}\right)^{\frac{1-\alpha}{\bar{m}+\alpha-1}}$$

When $\lambda_{it} = 0$, the capital price is $R\left(a_{it},\, \phi_{it}\right) = r + \delta$.

A.2 Investment-cashflow sensitivity

Notice in the model the capital investment-cash flow sensitivity is one when a firm is financially constrained, hence ζ also represents the average investment-cash flow sensitivity, which was introduced by Fazzari et al. (1988) and has been used extensively in the empirical literature in financial economics. See appendix for the discussion on the cash flow-investment sensitivity. When the capital constraint is binding, the capital investment is independent of the cash flow. When the firm is constrained, the capital investment is given by

$$\begin{aligned} k_{it} &= \frac{1}{\bar{m}^{\sigma}} \left(\frac{1-\alpha}{\omega} \right)^{(1-\alpha)(\sigma-1)} \left[\frac{\alpha}{R\left(a_{it},\, \phi_{it}\right)} \right]^{1+\alpha(\sigma-1)} \phi_{it}^{\sigma-1} \\ &\propto \lambda\left(a_{it},\, \phi_{it+1}\right) \end{aligned}$$

It follows that $\epsilon_{it} = \frac{\partial \ln(k_{it})}{\partial \ln(\lambda(a_{it},\, \phi_{it}))} = 1$. Let P_c be the fraction of constrained firms, the average cashflow sensitivity is

$$\zeta_t = \frac{1}{N} \sum_i \epsilon_i \times \mathbb{I}(constrained) = P_c \tag{A.9}$$

A.3 Static aggregate TFP loss

A.3.1 Aggregate TFP

I first show that given the distribution of fundamentals $(a_t,\, \phi_t)$, the aggregate TFP can be written as

$$TFPQ_t = \frac{\left[\int_{i \in N} R_{it}^{\alpha(1-\sigma)} \phi_{it}^{\sigma-1} di \right]^{\frac{1}{\sigma-1}+\alpha}}{\left[\int_{i \in N} R_{it}^{\alpha(1-\sigma)-1} \phi_{it}^{\sigma-1} di \right]^{\alpha}} \tag{A.10}$$

By the definition of aggregate TFP, we know

$$TFPQ_t = \frac{Q_t}{K_t^{\alpha} L_t^{1-\alpha}} = \frac{Q_t}{\left(\int_i k_{it} d_i \right)^{\alpha} \left(\int_i l_{it} di \right)^{1-\alpha}} \tag{A.11}$$

Let's first look at the aggregate inputs. Using the optimal decisions of capital and labor, I immediately have

$$K_t = \int_i k_{it} di = \frac{P_t^{\sigma} Q_t}{\bar{m}^{\sigma}} \left(\frac{1-\alpha}{\omega} \right)^{(1-\alpha)(\sigma-1)} \alpha^{1+\alpha(\sigma-1)} \int_i R_{it}^{\alpha(1-\sigma)-1} \phi_{it}^{\sigma-1} di \tag{A.12}$$

$$L_t = \int_i l_{it} di = \frac{P_t^{\sigma} Q_t}{\bar{m}^{\sigma}} \left(\frac{1-\alpha}{\omega} \right)^{\alpha+\sigma(1-\alpha)} \alpha^{\alpha(\sigma-1)} \int_i R_{it}^{\alpha(1-\sigma)} \phi_{it}^{\sigma-1} di \tag{A.13}$$

Moreover, the aggregate industrial output is

$$
\begin{aligned}
Q_t &= \left(\int q_{it}^{\frac{\sigma-1}{\sigma}} di\right)^{\frac{\sigma}{\sigma-1}} \qquad \text{(A.14)}\\
&= \frac{P_t^{\sigma} Q_t}{\bar{m}^{\sigma}} \left[\left(\frac{1-\alpha}{\omega}\right)^{(1-\alpha)(\sigma-1)} \alpha^{1+\alpha(\sigma-1)}\right]^{\alpha} \\
&\quad \left[\left(\frac{1-\alpha}{\omega}\right)^{\alpha+\sigma(1-\alpha)} \alpha^{\alpha(\sigma-1)}\right]^{1-\alpha} \times \left[\int R_{it}^{\alpha(1-\sigma)} \phi_{it}^{\sigma-1} di\right]^{\frac{\sigma}{\sigma-1}}
\end{aligned}
$$

Plugging the expressions of K_t and L_t into (A.11), canceling the aggregate variables $P_t^{\sigma} Q_t$ and other constants, I obtain (A.10).

Now I focus on a specific joint distribution of $(a_{it},\, \phi_{it})$ and obtain the estimator for aggregate TFP loss. In particular, we assume that the net worth a_{it} and productivity ϕ_{it} follow a joint log normal distribution

$$
\begin{bmatrix} \log(a_{it}) \\ \log(\phi_{it}) \end{bmatrix} \sim \mathbf{N}\left(\begin{bmatrix} \mu_a \\ \mu_\phi \end{bmatrix}, \begin{bmatrix} \sigma_a^2 & \tilde{\rho}\sigma_a\sigma_\phi \\ \tilde{\rho}\sigma_a\sigma_\phi & \sigma_\phi^2 \end{bmatrix}\right)
$$

Let $G(a,\, \phi)$ be the CDF and $g(a,\, \phi)$ be the PDF, respectively. Define $u \equiv \log(a)$ and $v \equiv \log(\phi)$, then the density function for $(u,\, v)$ is

$$
\begin{aligned}
h(u,\, v) &= \frac{1}{2\pi\sigma_a\sigma_\phi\sqrt{1-\tilde{\rho}^2}} \exp\left[-\frac{z(u,v)}{2(1-\tilde{\rho}^2)}\right] \qquad \text{(A.15)}\\
\text{where } z &= \frac{(u-\mu_a)^2}{\sigma_a^2} - \frac{2\tilde{\rho}(u-\mu_a)(v-\mu_\phi)}{\sigma_a\sigma_\phi} + \frac{(v-\mu_\phi)^2}{\sigma_\phi^2}
\end{aligned}
$$

Applying the change of variables theorem for this bi-variate case, I have

$$
g(a,\, \phi) = \frac{1}{a\phi} h(\log(a),\, \log(\phi)) \qquad \text{(A.16)}
$$

In addition, I normalize that $P_t^{\sigma} Q_t = 1$ to simplify our analysis to be on the partial equilibrium. I define a cut-off function $\bar{\phi}(a)$ such that $\left(a,\, \bar{\phi}(a)\right)$ solves

$$
R\left(a,\, \bar{\phi}(a)\right) = r + \delta \qquad \text{(A.17)}
$$

Correspondingly, I define a function $\bar{v}(u) \equiv \log\left(\bar{\phi}(e^u)\right)$ that characterizes the cut-off function on the space of $(u,\, v)$. Now I omit the time subscripts for the sake of brevity.

Theorem 1. When the fraction of constrained firms are small, the aggregate TFP can be approximated as

$$TFPQ = N^{\frac{1}{\sigma-1}} \Upsilon_0 \left(\Theta\right)^{\frac{1}{\sigma-1}} e^{\mu_\phi + \frac{(\sigma-1)\sigma_\phi^2}{2}}$$

Proof. Applying the Law of Large Numbers, I can express the numerator and denominator of $TFPQ$ as

$$\begin{aligned}
&\left[\int R_{it}^{\alpha(1-\sigma)} \phi_{it}^{\sigma-1} di\right]^{\frac{1}{\sigma-1}+\alpha} \\
&= N^{\frac{1}{\sigma-1}+\alpha} \left[\mathbf{E}\left(R_{it}^{\alpha(1-\sigma)} \phi_{it}^{\sigma-1}\right)\right]^{\frac{1}{\sigma-1}+\alpha} \\
&= N^{\frac{1}{\sigma-1}+\alpha} \left[\int_0^\infty \int_0^\infty R\left(a,\, \phi\right)^{\alpha(1-\sigma)} \phi^{\sigma-1} g\left(a,\, \phi\right) d\phi da\right]^{\frac{1}{\sigma-1}+\alpha} \\
&\propto \left[\left(r+\delta\right)^{\alpha(1-\sigma)} \int \int_{-\infty}^{\bar{\phi}(a)} \phi^{\sigma-1} dG\left(a,\, \phi\right) + \int \int_{\bar{\phi}(a)}^\infty \phi^{\sigma-1} R\left(a,\, \phi\right)^{\alpha(1-\sigma)} dG(a, \phi)\right]^{\frac{1}{\sigma-1}+\alpha} \\
&\left[\int_{i\in N} R_{it}^{\alpha(1-\sigma)-1} \phi_{it}^{\sigma-1} di\right]^{\alpha} \\
&\propto \left[\left(r+\delta\right)^{\alpha(1-\sigma)-1} \int \int_{-\infty}^{\bar{\phi}(a)} \phi^{\sigma-1} dG\left(a,\, \phi\right) + \int \int_{\bar{\phi}(a)}^\infty \phi^{\sigma-1} R\left(a,\, \phi\right)^{\alpha(1-\sigma)-1} dG(a, \phi)\right]^{\alpha}
\end{aligned}$$

Changing the variables to be u and v , the first double integral in the bracket is

$$\begin{aligned}
\int_{-\infty}^\infty \int_{-\infty}^{\bar{v}(u)} e^{(\sigma-1)v} g\left(e^u, e^v\right) de^v de^u &= \int_{-\infty}^\infty \int_{-\infty}^{\bar{v}(u)} e^{(\sigma-1)v} h\left(u,\, v\right) dvdu \\
&= \int_{-\infty}^\infty \int_{-\infty}^{\bar{v}(u)} \frac{1}{2\pi\sigma_a\sigma_\phi\sqrt{1-\tilde{\rho}^2}} \exp\left(\left(\sigma-1\right)v - \frac{z\left(u, v\right)}{2\left(1-\tilde{\rho}^2\right)}\right) dvdu
\end{aligned}$$

Integrating from the integral inside, I have

$$\int_{-\infty}^{\bar{v}(u)} \frac{1}{2\pi\sigma_a\sigma_\phi\sqrt{1-\tilde{\rho}^2}} \exp\left((\sigma-1)v - \frac{z(u,v)}{2(1-\tilde{\rho}^2)}\right) dv$$

$$= \frac{1}{\sqrt{2\pi(1-\tilde{\rho}^2)}\sigma_a} e^{-\frac{(u-\mu_a)^2}{2(1-\tilde{\rho}^2)\sigma_a^2} - \frac{\tilde{\rho}\mu_\phi(u-\mu_a)}{(1-\tilde{\rho}^2)\sigma_a\sigma_\phi} + \frac{\left[\mu_\phi+\tilde{\rho}\sigma_\phi(u-\mu_a)/\sigma_a\right]^2-\mu_\phi^2}{2(1-\tilde{\rho}^2)\sigma_\phi^2}}$$

$$\times \int_{-\infty}^{\bar{v}(u)} \frac{1}{\sqrt{2\pi}\sigma_\phi} e^{-\frac{\left[v-\left(\mu_\phi+\tilde{\rho}\sigma_\phi(u-\mu_a)/\sigma_a\right)\right]^2-2(\sigma-1)\left(1-\tilde{\rho}^2\right)\sigma_\phi^2 v}{2\left(1-\tilde{\rho}^2\right)\sigma_\phi^2}} dv$$

$$= \frac{1}{\sqrt{2\pi}\sigma_a} e^{-\frac{(u-\mu_a)^2}{2\sigma_a^2}+(\sigma-1)\left[\frac{\tilde{\rho}\sigma_\phi(u-\mu_a)}{\sigma_a}+\mu_\phi\right]+\frac{\left(1-\tilde{\rho}^2\right)(\sigma-1)^2\sigma_\phi^2}{2}} \times$$

$$\int_{-\infty}^{\bar{v}(u)} \frac{1}{\sqrt{2\pi(1-\tilde{\rho}^2)}\sigma_\phi} e^{-\frac{\left[v-\left((\sigma-1)\left(1-\tilde{\rho}^2\right)\sigma_\phi^2+\tilde{\rho}(u-\mu_a)\frac{\sigma_\phi}{\sigma_a}+\mu_\phi\right)\right]^2}{2\left(1-\tilde{\rho}^2\right)\sigma_\phi^2}} dv$$

$$= \frac{1}{\sqrt{2\pi}\sigma_a} e^{-\frac{(u-\mu_a)^2}{2\sigma_a^2}+(\sigma-1)\left[\frac{\tilde{\rho}\sigma_\phi(u-\mu_a)}{\sigma_a}+\mu_\phi\right]+\frac{\left(1-\tilde{\rho}^2\right)(\sigma-1)^2\sigma_\phi^2}{2}} \Phi\left(\frac{\bar{v}(u)-\tilde{v}_0(u)}{\sqrt{1-\tilde{\rho}^2}\sigma_\phi}\right)$$

where $\Phi(\cdot)$ is the CDF of standard normal distribution and $\tilde{v}_0(u) \equiv (\sigma-1)\left(1-\tilde{\rho}^2\right)\sigma_\phi^2 + \tilde{\rho}(u-\mu_a)\frac{\sigma_\phi}{\sigma_a} + \mu_\phi$. This implies that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\bar{v}(u)} e^{(\sigma-1)v} h(u,\, v)\, dvdu$$

$$= e^{\frac{\left(1-\tilde{\rho}^2\right)(\sigma-1)^2\sigma_\phi^2}{2}+(\sigma-1)\mu_\phi} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_a} e^{-\frac{(u-\mu_a)^2}{2\sigma_a^2}+(\sigma-1)\frac{\tilde{\rho}\sigma_\phi(u-\mu_a)}{\sigma_a}} \Phi\left(\frac{\bar{v}(u)-\tilde{v}_0(u)}{\sqrt{1-\tilde{\rho}^2}\sigma_\phi}\right) du$$

$$= e^{\frac{(\sigma-1)^2\sigma_\phi^2}{2}+(\sigma-1)\mu_\phi} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \Phi\left(\frac{\bar{v}(u_0(y))-\tilde{v}_0(u_0(y))}{\sqrt{1-\tilde{\rho}^2}\sigma_\phi}\right) dy$$

Define that

$$\Upsilon_0(\Theta) \equiv \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \Phi\left(\frac{\bar{v}(u_0(y))-\tilde{v}_0(u_0(y))}{\sqrt{1-\tilde{\rho}^2}\sigma_\phi}\right) dy,$$

where $\Theta \equiv (\mu_a,\, \sigma_a,\, \tilde{\rho})$ and

$$y = \frac{u-\mu_a-(\sigma-1)\sigma_a\sigma_\phi}{\sigma_a} \Leftrightarrow u_0(y) = \sigma_a y + \mu_a + (\sigma-1)\sigma_a\sigma_\phi$$

is used to change the integrand. For the numerator, the second double integral in the bracket is

$$\begin{aligned}\int\int_{\bar{\phi}(a)}^{\infty}\phi^{\sigma-1}R\left(a,\,\phi\right)^{\alpha(1-\sigma)}d\phi da &= \int\int_{\bar{v}(u)}^{\infty}e^{(\sigma-1)v}R\left(e^{u},\,e^{v}\right)^{\alpha(1-\sigma)}g\left(e^{u},e^{v}\right)de^{u}de^{v}\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\bar{v}(u)}R\left(e^{u},\,e^{v}\right)^{\alpha(1-\sigma)}e^{(\sigma-1)v}dH\left(u,v\right)\\ &= D_{t}^{\alpha(1-\sigma)}\int\int_{-\infty}^{\bar{v}(u)}\left[e^{v}\lambda\left(e^{u},\,e^{v}\right)^{1-\bar{m}}\right]^{\frac{\alpha(1-\sigma)}{\bar{m}+\alpha-1}}e^{(\sigma-1)v}dH\left(u,v\right)\\ &= D_{t}^{\alpha(1-\sigma)}\int\int_{-\infty}^{\bar{v}(u)}\lambda\left(e^{u},\,e^{v}\right)^{\frac{\alpha}{\alpha+\bar{m}-1}}e^{\frac{v}{\bar{m}+\alpha-1}}dH\left(u,v\right)\end{aligned}$$

For the denominator, the second double integral in the bracket is

$$\int\int_{\bar{\phi}(a)}^{\infty}\phi^{\sigma-1}R\left(a,\,\phi\right)^{\alpha(1-\sigma)}d\phi da = D_{t}^{\alpha(1-\sigma)-1}\int\int_{-\infty}^{\bar{v}(u)}\lambda\left(e^{u},e^{v}\right)dH\left(u,v\right) \tag{A.18}$$

Therefore I can write down the aggregate TFP as

$$TFPQ \propto \frac{\left[\Upsilon_{0}\left(\Theta\right)e^{(\sigma-1)\mu_{\phi}+\frac{(\sigma-1)^{2}\sigma_{\phi}^{2}}{2}}+\left(\frac{D_{t}}{r+\delta}\right)^{\alpha(1-\sigma)}\int\int_{-\infty}^{\bar{v}(u)}\lambda\left(e^{u},e^{v}\right)^{\frac{\alpha}{\alpha+\bar{m}-1}}e^{\frac{v}{\bar{m}+\alpha-1}}dH\left(u,v\right)\right]^{\alpha+\frac{1}{\sigma-1}}}{\left[\Upsilon_{0}\left(\Theta\right)e^{(\sigma-1)\mu_{\phi}+\frac{(\sigma-1)^{2}\sigma_{\phi}^{2}}{2}}+\left(\frac{D_{t}}{r+\delta}\right)^{\alpha(1-\sigma)-1}\int\int_{-\infty}^{\bar{v}(u)}\lambda\left(e^{u},e^{v}\right)dH\left(u,v\right)\right]^{\alpha}} \tag{A.19}$$

$$\propto \Upsilon_{0}\left(\Theta\right)^{\frac{1}{\sigma-1}}e^{\mu_{\phi}+\frac{(\sigma-1)\sigma_{\phi}^{2}}{2}}\frac{\left[1+\left(\frac{D_{t}}{r+\delta}\right)^{\alpha(1-\sigma)}\Upsilon_{1}\left(\Theta\right)\right]^{\alpha+\frac{1}{\sigma-1}}}{\left[1+\left(\frac{D_{t}}{r+\delta}\right)^{\alpha(1-\sigma)}\Upsilon_{2}\left(\Theta\right)\right]^{\alpha}} \tag{A.20}$$

where

$$\begin{aligned}\Upsilon_{1}\left(\Theta\right) &\equiv \frac{\int\int_{-\infty}^{\bar{v}(u)}\lambda\left(e^{u},e^{v}\right)^{\frac{\alpha}{\alpha+\bar{m}-1}}e^{\frac{v}{\bar{m}+\alpha-1}}dH\left(u,v\right)}{\Upsilon_{0}\left(\Theta\right)e^{(\sigma-1)\mu_{\phi}+\frac{(\sigma-1)^{2}\sigma_{\phi}^{2}}{2}}}\\ \Upsilon_{2}\left(\Theta\right) &\equiv \frac{\int\int_{-\infty}^{\bar{v}(u)}\lambda\left(e^{u},e^{v}\right)dH\left(u,v\right)}{\Upsilon_{0}\left(\Theta\right)e^{(\sigma-1)\mu_{\phi}+\frac{(\sigma-1)^{2}\sigma_{\phi}^{2}}{2}}}\end{aligned}$$

when $\Upsilon_1$and Υ_2 are close to zero, I obtain that

$$TFPQ = N^{\frac{1}{\sigma-1}}\Upsilon_{0}\left(\Theta\right)^{\frac{1}{\sigma-1}}e^{\mu_{\phi}+\frac{(\sigma-1)\sigma_{\phi}^{2}}{2}}$$

□

A.3.2 Characterization of aggregate TFP loss

The efficient aggregate TFP is given by

$$TFPQ^e = \left(\int_i \phi_{it}^{\sigma-1} di\right)^{\frac{1}{\sigma-1}}$$
$$= N^{\frac{1}{\sigma-1}} \left(\mathbf{E}\left(\phi_t^{\sigma-1}\right)\right)^{\frac{1}{\sigma-1}}$$

Because ϕ_t has a log-normal distribution with mean μ_ϕ and variance σ_ϕ, I have

$$TFPQ^e = N^{\frac{1}{\sigma-1}} e^{(\sigma-1)\mu_\phi + \frac{(\sigma-1)}{2}\sigma_\phi^2} \tag{A.21}$$

Consider the case that $\Upsilon_1(\Theta)$ and $\Upsilon_2(\Theta)$ are close to zero, I can calculate the log of TFP loss as

$$TFP\,loss = 1 - \frac{TFPQ}{TFPQ^e}$$
$$= 1 - \Upsilon_0(\Theta)^{\frac{1}{\sigma-1}}$$

Theorem 2. When $\bar{v}(u)$ is an increasing (decreasing) function of u, the aggregate TFP loss is decreasing (increasing) in μ_a. The aggregate TFP loss is and increasing in μ_ϕ.

Proof. First, note that I can write $\tilde{v}_0(u_0(y))$ as

$$\tilde{v}_0(u_0(y)) = \tilde{\rho}\sigma_a y + (\sigma-1)\left(1-\tilde{\rho}^2\right)\sigma_\phi^2 + \tilde{\rho}(\sigma-1)\sigma_\phi^2 + \mu_\phi,$$

which is independent of μ_a. The derivative of $\Upsilon_0(\Theta)$ with respect to μ_a is

$$\frac{\partial \Upsilon_0(\Theta)}{\partial \mu_a} = \int \frac{1}{\sqrt{2\pi}\sigma_a} e^{-\frac{y^2}{2}} \Phi' \frac{\bar{v}'(u_0(y))\sigma_a}{\sqrt{1-\tilde{\rho}^2}\sigma_\phi} dy \lesseqgtr 0$$
$$\Leftrightarrow \bar{v}'(u_0(y)) \lesseqgtr 0$$

Following a similar logic, I know that

$$\frac{\partial \Upsilon_0(\Theta)}{\partial \mu_\phi} = \int \frac{1}{\sqrt{2\pi}\sigma_a} e^{-\frac{y^2}{2}} \Phi' \frac{-1}{\sqrt{1-\tilde{\rho}^2}\sigma_\phi} dy < 0$$

□

B Theory Appendix

Here I provide a micro foundation for the functional form chosen in our benchmark model. Let $V^D(a,\phi)$ be the value function when the firm defaults and $V^N(a,\phi)$ be the value function when the firm does not default. When not defaulting, the value function is given by

$$V^N(a,\phi) = V(a,\phi) = \max_{b',k',x} \left\{ \frac{c^{1-\epsilon}}{1-\epsilon} + \beta \mathbf{E} V(a',\phi') \right\} \tag{B.1}$$

subject to the constraint:

$$c + k' + (1+r)b + \mathbb{I}(x)f + \frac{d}{2}x^2 = y - \omega l + (1-\delta)k + b'$$

where b is the amount of debt in the next period, and k' is the physical capital in the next period. By definition, I always have $a = k - b$. Because of limited enforcement of contracts, a firm can default on a fraction of the face value of current debt. Therefore the firm only needs to pay $(1 + r - \mu_0)b$ to the bank. The cost of defaulting is a fraction of collateral used when borrowing from the bank: $\mu_1(1-\delta)k + \mu_2\Psi(\phi)$. In other words, the bank can seize a fraction of the capital and the value of intangible assets when the firm defaults. The firms are assumed to be only defaulting for one period and have access to the financial market next period. This implies that $V^D(a,\phi)$ can be expressed as:

$$V^D(a,\phi) = \max_{b',k',x} \left\{ \frac{c^{1-\epsilon}}{1-\epsilon} + \beta \mathbf{E} V(a',\phi') \right\} \tag{B.2}$$

subject to the constraint:

$$c + k' + (1+r-\mu_0)b + \mu_1(1-\delta)k + \mu_2\Psi(\phi) + \mathbb{I}(x)f + \frac{d}{2}x^2 = y - \omega l + (1-\delta)k + b'$$

The condition for an equilibrium in which no firm defaults is

$$\begin{aligned} & V^D(a,\phi) \le V^N(a,\phi) \\ \Rightarrow\ & (1+r)b \le (1+r-\mu_0)b + \mu_1(1-\delta)k + \mu_2\Psi(\phi) \end{aligned}$$

This immediately implies that

$$b \le \frac{\mu_1}{\mu_0}k + \frac{\mu_2}{\mu_0}\Psi(\phi) \tag{B.3}$$

Lastly, I determine the functional form of the value of intangible asset as follows. I assume that the productivity can be decomposed into organization capital ($\phi^{1-\nu}$) and intangible assets (ϕ^{ν}), where $1 > \nu > 0$. The organization capital is not pledgeable while the intangible assets can be used as a collateral when the firm borrows from the financial institutions. The intangible assets can only be used by a potential manufacturer with organization capital greater than $\phi^{1-\nu}$ and generate profits of $\chi\phi^{\sigma-1}$. Furthermore, the distribution of the organization capital follows a Pareto distribution with lower bound $\underline{\phi}$ ($\phi^{1-\nu} > \underline{\phi}$) and shape parameter m. Therefore, with probability $(\phi^{1-\nu}/\underline{\phi})^{-m}$. This implies that the value of pledged intangible asset is

$$\Psi(\phi) = \chi\underline{\phi}^m \phi^{\sigma-m(1-\nu)-1} \tag{B.4}$$

Define $\eta \equiv m(\nu-1)+\sigma-1$ and $\theta \equiv \mu_1/\mu_0$. Moreover, if I impose that $\chi\underline{\phi}^m = \mu_1$, I obtain the specification used in our benchmark model. Note that in addition to financial friction, m also captures the efficiency of the market for intellectual property rights including patents and trademarks. A smaller m means that banks can sell the pledged intellectual property rights to a productive potential buyer more easily.

C Characterization of the Steady State

In this section, I provide a characterization of the steady state of the model. According to our aggregation of the economy, aggregate outcomes are determined by the joint distribution of the state variables (a_t, ϕ_t). Note that in the endogenous productivity model, the evolution of net worth and productivity is summarized by a non-linear first-order stochastic differential equation:

$$a_{t+1} = a'(a_t, \phi_t) \tag{C.1}$$

$$\phi_{t+1} = \rho\phi_t + \gamma \ln(x(a_t, \phi_t) + 1) + \xi_{t+1}, \tag{C.2}$$

where $a'(a_t, \phi_t)$ is the decision rule of the accumulation of future net worth, $x(a_t, \phi_t)$ is the optimal choice of R&D investment. This *non-linear vector auto-regressive model* can be approximated by following first-order linear vector auto-regressive model:

$$\begin{bmatrix} a_{t+1} \\ \phi_{t+1} \end{bmatrix} = \begin{bmatrix} a^{(0)} \\ \phi^{(0)} \end{bmatrix} + \underbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}}_{B} \begin{bmatrix} a_t \\ \phi_t \end{bmatrix} + \begin{bmatrix} 0 \\ \xi_{t+1} \end{bmatrix} \tag{C.3}$$

where

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} \frac{\partial a'(a,\phi)}{\partial a} & \frac{\partial a'(a,\phi)}{\partial \phi} \\ \gamma\frac{\partial \ln(x(a,\phi)+1)}{\partial a} & \rho + \gamma\frac{\partial \ln(x(a,\phi)+1)}{\partial \phi} \end{bmatrix}_{a=\mathbf{E}(a),\phi=\mathbf{E}(\phi)}$$

The necessary and sufficient condition ensuring the weak stationarity of this VAR(1) model is the eigenvalues of B has to be smaller than 1 in modulus.

D Data

D.1 Panel data of Chinese manufacturing firms

In this subsection, I provide the detailed information on the construction of variables used in the paper.

Net Worth a_{it}: Difference between book value of total assets and total liabilities and deflated by industry output price deflators.

(Net) Debt b_{it}: Book value of current liabilities minus cash holdings and deflated this difference with the industry output price deflators. This is a short-term measure of the debt; the model has no maturity choice.

Capital Stock k_{it}. Book value of fixed assets and deflated with the price of investment goods.[33]

Wage Bill ω_{it}. Book value of sum of salary and payments for labor insurance and other benefits including health insurance, pension insurance, and other insurance.

Labor Input l_{it}. The firm's wage bill deflated by the industry price deflator. This measurement controls for the quality difference of the labor.

Value Added $p_{it}y_{it}$. Book value of the firm's value added.

Real Output y_{it}. Nominal value added deflated by an output price deflator.

Marginal Revenue Product of Capital (MRPK) R_{it}. I follow Hsieh and Klenow (2009) to construct the marginal revenue product of capital, which is $R_{it} = \alpha\phi_{it}\left(\frac{k_{it}}{l_{it}}\right)^{\alpha-1}$.

E Supplementary Empirical Results

This section provides supplementary empirical results referred in the main text.

E.1 Productivity and leverage ratio

In this subsection, I provide the supporting evidence for the modelling of productivity as intangible assets used as collateral. I show that in the data there is a robust positive relation between productivity and leverage ratio, indicating that productive firms are more able to obtain debt financing.

E.2 Productivity processes for high- and low-tech industries

I show the OLS estimation of the productivity process when estimating separately for high-tech and low-tech industries. I do find some heterogeneity in the endogenous productivity process. Surprisingly, I find that the productivity growth effect of R&D is larger in low-tech industry. Given that

[33] In GKKV, both tangible and intangible fixed assets are included.In our data, I do not observe the firm's intangible assets. I use the fixed assets under the tangible assets.

Table E.1: Determinants of leverage ratio and R&D investment

Dependent var.:	leverage ratio	
	(1)	(2)
$\ln(\phi)$	0.562***	0.399***
	(0.092)	(0.098)
$\ln(a)$	-6.926***	-7.254***
	(0.181)	(0.145)
Controls	No	Yes
N	87329	87329
R^2	0.205	0.272

Note: control variables include firm age, ownership, province, year and 3-digit industry fixed effects. Standard errors are clustered at the city level. *** indicates significance level at 1% significance level.

low-tech firms undertake less R&D investment, it must be that the the high costs prevent them from doing that.

Table E.2: Productivity evolution for high- and low-tech industries

dependent var.: $\ln(\phi_{t+1})$		
sectors	High-Tech	Low-Tech
$\ln(\phi_t)$	0.384***	0.332***
	(0.025)	(0.006)
$\ln(x_t + 1)$	0.051***	0.056***
	(0.004)	(0.002)
industry-year FEs	Yes	Yes
$\hat{\sigma}_\xi$	1.289	1.273
N	5136	60843
R^2	0.240	0.186

Note: standard errors are heteroscedastic robust. *** $p < 0.01$

E.3 R&D with endogenous uncertainty

E.3.1 Maximum-Likelihood Estimator

In particular, the log-likelihood function is given by

$$LLF = \sum_{i,t} \log \Big\{ [1 - \exp(-\psi x_{it}^{\vartheta})] g[\frac{1}{\sigma_\xi}(\ln(\phi_{it+1}) - \rho\phi_{it} - h - \mu_{jt})] + \exp(-\psi x_{it}^{\vartheta}) g[\frac{1}{\sigma_\xi}(\ln(\phi_{it+1}) - \rho\phi_{it} - \mu_{jt})] \Big\} \tag{E.1}$$

where $g(\cdot)$ is the density function of Standard Normal distribution. The estimates for the associated parameters are obtained by maximizing the objective log-likelihood function. Table E.3 shows the estimation results. The persistence of the productivity process is estimated to be 0.335, which is close to what I obtained using the productivity process with only exogenous productivity shocks. Also, the dispersion is similar to the estimate used in the benchmark analysis. The average improvement of productivity is 0.573. The probability of successful innovation is $\Pr(\kappa_{it} = 1|x_{it}) = 1 - \exp(-0.076 x_{it}^{1.080})$. All of the coefficient estimates are significant at 1% significance level.

Table E.3: Productivity process with endogenous uncertain innovation outcomes

Parameters	ρ	ψ	ϑ	h	$\ln(\sigma_\xi)$	
Coefficients	0.335***	0.076***	1.080***	0.573***	0.233***	$N = 65979$
s.e.	(0.023)	(0.132)	(0.024)	(0.003)	(0.003)	$LLF = -109129.8$

Note: estimates are obtained using Maximum Likelihood Estimator *** $p < 0.01$

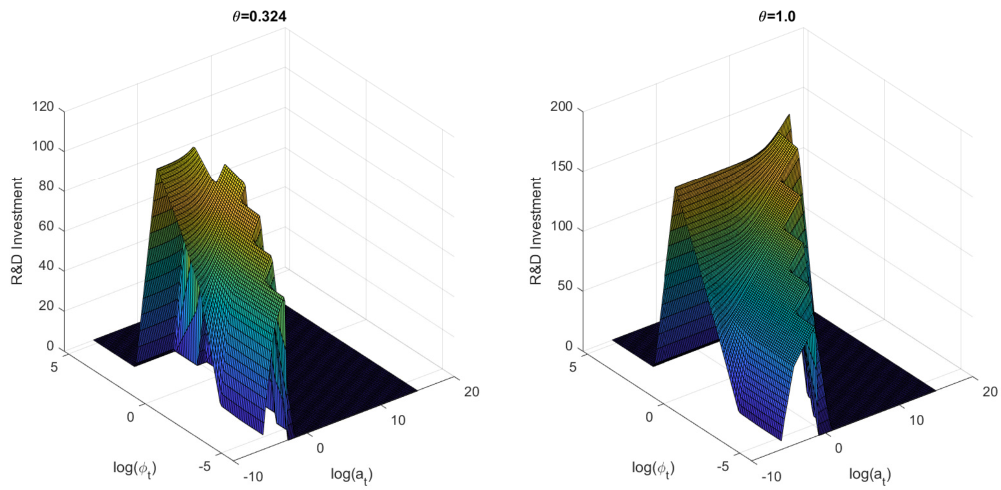
E.3.2 Optimal R&D policy under endogenous innovation uncertainty

The figure above shows the R&D policy functions when modelling the innovation process with uncertainty. It is clear that firms with relative lower productivity and smaller net worth choose to invest in R&D investment. This is inconsistent with the data which shows a positive partial correlation between R&D investment and net worth, as well as between R&D investment and productivity.

F Computation

This section describes the main computation procedures performed in this paper.

Figure E.1: Optimal R&D policy: innovation process with uncertainty



F.1 Benchmark model and simulation

Our benchmark model has two continuous state variables and two continuous dynamic choice variables. Because of the fixed cost, the choice of R&D features kinks in its decision rule. Using traditional value function iteration method requires a large amount of time to accomplish the computation. Instead, I use the collocation method to solve the value function by approximating it as a combination of known basis functions. After choosing an appropriate grid for the state variables (a, ϕ), I use the alternating search algorithm to solve the bi-variate optimization problem. I find the optimization method such as Nelder-Mead algorithm and Newton's iteration method are less efficient than the alternating search method. By fixing one variable each time, the alternating search method simplifies the optimization to be a single-variable optimization. The algorithm stops whenever the new optimal values are close to the values found in the previous round. I also employ the parallel computing in Matlab to improve the computing efficiency. To avoid the problem of local optimum, I use MCMC simulations to find the minimizer of the objective GMM estimator as suggested by Chernozhukov and Hong (2003). The estimation algorithm is as follows:

1. Given a guess of (f, d) and pooled data of net worth and productivity, I solve the value function and find the optimal decisions of future net worth and R&D investment

2. Construct the objective function using the model generated data on future net worth and R&D investment for each observation

3. Obtain the Meteropolis-Hastings MCMC chains for parameters using 2000 simulations

The estimate of (f, d) is the mean of the simulated data. After I parameterize the model, I perform simulation using the solved policy functions of R&D investment and net worth to generate the path of productivity and net worth and. Then all relevant variables are computed accordingly.

F.2 Model Extensions

F.2.1 Model with R&D costs heterogeneity

I assume that f and d follows a joint log-normal distribution with correlation ρ_{fd}:

$$\left[\begin{array}{c} \ln f_i \\ \ln d_i \end{array}\right] \sim \mathbf{N}\left(\left[\begin{array}{c} \mu_f \\ \mu_d \end{array}\right], \left[\begin{array}{cc} \sigma_f^2 & \rho_{fd}\sigma_f\sigma_d \\ \rho_{fd}\sigma_f\sigma_d & \sigma_d^2 \end{array}\right]\right)$$

I draw N realizations of f_i^0 and d_i^0 independently from the standard normal distribution $\mathbf{N}(0,1)$, putting them aside to construct the R&D costs parameters f_i and d_i.

Step 1 Using f_i's and d_i's, I construct N realizations for each of $\ln f_i^*$ and $\ln d_i^*$ as

$$\left[\begin{array}{c} f_i^* \\ d_i^* \end{array}\right] = \left[\begin{array}{cc} \sigma_f\sqrt{1-\rho_{fd}^2} & \sigma_f\rho_{fd} \\ 0 & \sigma_d \end{array}\right]\left[\begin{array}{c} f_i^0 \\ d_i^0 \end{array}\right]$$

Step 2 For any values of μ_f and μ_d, I discretize $\ln f_i^*$ and $\ln d_i^*$ according to the characteristics of normal distribution, I obtain realizations for each of $\ln f_i$ and $\ln d_i$:

$$\ln z_i = \begin{cases} \mu_z + \sigma_z & \text{if } z_i^* \geq \mu_z + \frac{\sigma_z}{2} \\ \mu_z - \sigma_z & \text{if } z_i^* < \mu_z - \frac{\sigma_z}{2} \\ \mu_z & z_i^* \in [\mu_z - \frac{\sigma_z}{2},\, \mu_z + \frac{\sigma_z}{2}) \end{cases} \quad ; z \in \{f,\, d\}$$

Since I do not find a significant estimate for ρ_{fd}, ρ_{fd} is imposed to be zero in our preferred estimation. I then choose a set of values for $(\mu_f, \sigma_f, \mu_d, \sigma_d)$ to minimize the distance between the model and the data using the same moments as in the benchmark model. To avoid the local miminum, I follow Chernozhukov and Hong (2003) to obtain the Metropolis-Hastings MCMC chains using multivariate Gaussian proposal distribution.