Productivity and Liquidity Management Under Costly Financing *

Felix Zhiyu Feng† Jianyu Lu‡ Jing Wang§

September 2016

Abstract

We explore theoretically and empirically the relationship between firm productivity and liquidity management in the presence of financial frictions. We build a dynamic investment model and show that, counter to basic economic intuition, more productive firms could demand *less* capital assets and hold *more* liquid assets compared to less productive firms when financing costs are sufficiently high. We empirically test this prediction using a comprehensive dataset of Chinese manufacturers and find that more productive firms indeed hold less capital and more cash. We do not, however, observe this for U.S. manufacturers. Our study suggests a larger capital misallocation problem in markets with significant financing frictions than previously documented.

JEL Classification: G31, G32, G11, G15

Key Words: Liquidity management, productivity, financing, emerging economy

*All authors are from the University of Notre Dame. We thank Jeffrey Bergstrand, Wyatt Brooks, Tim Fuerst, Barney Hartman-Glaser, Pengjie Gao, and Joe Kaboski for helpful comments. We also thank the seminar and conference participants at Duke University, University of Notre Dame, the Financial Management Association annual meetings, the International Conference on Computing in Economics and Finance, the Royal Economic Society annual meetings, and the Southwestern Financial Association annual meetings. All errors are our own.

†916 Flanner Hall, Notre Dame, IN, 46556. Phone: (574)631-0428, Email: ffeng@nd.edu

‡846 Flanner Hall, Notre Dame, IN, 46556. Phone: (206)660-4798, Email: jlu4@nd.edu

§709 Flanner Hall, Notre Dame, IN, 46556. Phone: (574)485-8068, Email: jwang19@nd.edu

1 Introduction

One of the most fundamental decisions firms make is that of resource allocation. Broadly speaking, firms either allocate resources toward acquiring capital assets, such as land, factory and machinery, or toward holding liquid assets, such as cash, receivables, and inventory. The return on capital assets is usually risky and varies greatly across firms with different levels of productivity. In contrast, liquid assets usually generate uniform, risk-free return. How much to invest in growth versus liquidity naturally depends on the productivity of capital and the cost of financing. This decision of how to optimally allocate limited resources is especially relevant when financing is very costly, as it is for firms in emerging markets.

The objective of this paper is to investigate, theoretically and empirically, how firms with different levels of productivity allocate resources between capital assets and liquid assets in the presence of costly financing. We develop a tractable continuous time model to study the joint decision of capital investment and liquidity management. In the model, firms allocate resources between investing in capital, which has risky returns that vary for firms depending on their productivity, and holding cash, which earns risk-free return for all firms. Firms pay out dividends when their total assets are sufficiently high, and must refinance when total assets are too low. Because future dividends are discounted, firms prefer to pay out early.[1] On the other hand, because refinancing is costly, firms like to hold precautionary savings. Optimal liquidity management therefore represents a balance between paying out (or consuming) early and avoiding costly refinancing.

Solving the model leads to the surprising conclusion that it could be optimal for more productive firms to have *less* capital and *more* cash when financing frictions are sufficiently severe. The underlying mechanism is that in equilibrium, more productive firms endogenously choose to pay out at lower asset levels. Their greater productivity allows them to expedite consumption relative to less productive firms.[2] This in turn expedites the need to

[1]Every firm's objective is to maximize discounted dividends or shareholders' consumption stream.

[2]That more productive firms pay out earlier appears to contradict the conclusion drawn from traditional constrained growth models, such as Buera et al. (2011), Song et al. (2011) and Moll (2013). These models, however, generally do not have cash, and therefore firms can only alleviate financial constraint by delaying consumption. In our model, cash is an alternative solution, on which high productivity firms optimally choose to rely more so than do low productivity firms. In other words, holding liquid assets allow high productivity firms to endogenously take advantage of their productivity superiority and consume earlier.

refinance. As a result, high productivity firms are more risk averse and hence rely on cash holdings to mitigate downsizing risk. When financing becomes sufficiently costly, the risk aversion of high productivity firms dominates their relative advantage of holding capital. This theoretical prediction leads to an empirically testable hypothesis: a positive correlation between firm productivity and liquid assets/cash holdings in under-developed markets. This conclusion is in contrast to that obtained in a financially frictionless environment, where basic economic intuition suggests that more productive firms should acquire more capital.

We test our hypothesis using a large sample of Chinese manufacturing firms from 1999 to 2007. We find that, consistent with our theoretical prediction, firms in the highest productivity decile on average have a 32% higher net liquid to total assets ratio and a 21% higher cash to total assets ratio compared to those in the lowest decile. This pattern is highly robust, salient regardless of firm size, ownership structure, industry affiliation, or how productivity and liquidity are measured.[3] On the other hand, the productivity of manufacturers in the US–an operating environment where financing frictions are less severe–is negatively correlated with the holding of liquid assets, which is consistent with economic intuition.

We quantitatively analyze our model using parameters estimated from observed sample moments. Our model generates a distribution of liquid asset holdings that matches the observed pattern. Furthermore, we are able to infer the magnitude of some unobservable variables in the model. For example, the inferred magnitude of refinancing cost would suggest that financing frictions are much more severe in countries with less developed financial markets than previously documented. Finally, we find that lowering the cost of refinancing leads to a significant improvement in total productivity.

Our finding adds to the current understanding of liquidity management and external financing. Corporate finance studies on the determinants of firm liquidity abound, and precautionary savings is one of the primary theories.[4] Almeida et al. (2004), Khurana et al. (2006) and Riddick and Whited (2009) all argue that the sensitivity of corporate cash holdings (or savings) to cash flow is an indicator of costly external financing. On the one hand,

[3]In estimating productivity, in addition to using the value added measure reported in the data divided by capital, we also follow a recent technique developed by Ackerberg et al. (2015).

[4]See Bates et al. (2009) for a review of studies on firm liquidity and Gao et al. (2013) for a summary of those related to precautionary savings.

we document an observation consistent with this view: a positive correlation between liquidity and productivity (prominently) exists only in the Chinese sample. Our explanation for this observation hinges on the substantial external financing costs that Chinese firms face operating in an environment where the financial market is far from being developed. On the other hand, we complement the aforementioned studies by demonstrating the important role financial constraints play in the liquidity management and investment decisions of firms that face the *same* level of external finance costs but have *varying* levels of productivity. Moreover, we show that the severity of financing frictions matters: sufficiently severe financing frictions can incentivize high productivity firms to avoid costly external financing by maintaining higher internal liquidity.

The fact that high productivity firms may invest less proportionally in capital under significant financing frictions suggests a potentially serious capital misallocation problem. The literature on development economies, including Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Greenwood et al. (2010), Buera et al. (2011), and Song et al. (2011), has identified broad capital misallocation in emerging economies and argues that it leads to productivity losses.[5] In particular, recent work by Whited and Zhao (2016) compares the efficiency of capital structure between firms in the US and those in China based on implications of the Hsieh and Klenow (2009) model. They use the same datasets as ours (Compustat and the NBS survey in China) and find substantially less efficient allocation between debt and equity among Chinese firms. We reveal another potential layer of misallocation between fixed and liquid assets: not only do high productivity firms hold less capital than what is socially optimal, they actually trade off more capital for liquidity than do low productivity firms. Our understanding of emerging market economies would be incomplete without considering this hitherto undocumented potential for exacerbated capital misallocation. In support of this statement, we study the welfare implications of our model in the quantitative analysis section and show that a reduction of refinancing costs can indeed lead to a more efficient allocation of capital.

The model and data used in this paper contribute to both the theoretical methodology as well as the empirical scope of the research on liquidity management. On the modeling side,

[5] See Restuccia and Rogerson (2013) for a more thorough literature review of the relevant studies.

our model is close to Bolton et al. (2011) and DeMarzo et al. (2012) but differs in several ways: first, firms in our model solve a dynamic portfolio choice problem; second, capital price is endogenous in our model as we consider a general equilibrium approach, and we manage to aggregate total firm wealth to solve capital price, which is a challenging task in general equilibrium models; third, we consider firms with heterogeneous productivity rather than a single representative firm. On the empirical side, our paper joins those that examine liquidity management either for private firms or for firms outside the US, for example Dittmar et al. (2003), Brav (2009), Lins et al. (2010), Bigelli and Snchez-Vidal (2012), and Gao et al. (2013). However, these studies do not measure productivity and are therefore mute on the efficiency implications of firms' liquidity management. To our knowledge, this is the first large-scale study of private firms from an emerging economy to document a positive correlation between firms' productivity and their liquidity demand–a counterintuitive observation for which we offer one explanation.

2 Model

This section presents a continuous time dynamic model based on the standard liquidity management model of Bolton et al. (2011) but extended to allow heterogeneous productivity and a general equilibrium of capital. In the model, firms solve a portfolio choice problem between capital and cash, subject to costly refinancing. The solution offers several testable hypotheses. In particular, we find that more productive firms may hold proportionally more liquid assets compared to less productive firms if the cost of refinancing is sufficiently high.

2.1 Basic Environment

There is a continuum of firms who are risk neutral with discount rate ρ. Each firm keeps two types of assets: physical capital for production and cash for meeting the firm's liquidity needs. Capital does not depreciate, but its output is risky. A firm holding K_t units of capital can produce dY_t units of output according to

$$dY_t = K_t dA_t \ , \tag{1}$$

where $dA_t = \mu_s dt + \sigma dZ_t$ is the *idiosyncratic* productivity of each firm. $s \in \{l, h\}$ is a state variable denoting firms' productivity which for simplicity takes two values: either a low type μ_l or a high type μ_h. The fraction of high productivity firms is π. The production technology of capital is constant returns to scale. Finally, the technology shocks, represented by the standard Brownian motion term Z_t, are assumed to be i.i.d. across all firms. In other words, there is no aggregate shock. Let P_t be the price of capital, which firms take as given. The return per unit of capital is

$$dR_t = \frac{dA_t + dP_t}{P_t} \ . \tag{2}$$

In contrast, holding cash is risk free. Cash earns an interest rate of r which we assume is exogenous.[6] Let W_t denote a firm's total assets and α_t the fraction of W_t allocated to capital investment. The dynamics of W_t follows

$$dW_t = \alpha_t W_t dR_t + (1 - \alpha_t) W_t r dt \ . \tag{3}$$

We focus on the case where $\alpha_{s,t} \in [0, 1]$. That is, firms keep a positive balance in both physical capital and cash inventory. This is consistent with studies that show in the past decade firms in both emerging economies as well as in the US have become net lenders rather than borrowers.[7]

[6] The condition $r < \rho$ is necessary for a non-trivial solution: if the risk free interest rate is higher than firms' discount rate, firms can just hoard cash and postpone paying out dividends forever. Nevertheless, this condition does not necessarily mean that firms are more impatient than their creditors. It can be interpreted as the carrying cost of cash a la Bolton et al. (2011), which lowers the effective interest rate of holding cash.

[7] See Armenter (2012), Armenter and Hnatkovska (2012), Karabarbounis and Neiman (2012) and Gruber and Kamin (2015) for evidence in the US and other developed economies, and Cardarelli and Ueda (2006) and Bayoumi et al. (2012) for evidence in emerging economies. On the one hand, this assumption could be read as simply a no (net) borrowing constraint due to, for instance, contract enforceability in the credit market. In practice, lack of enforcement is prevalent in emerging economies, where financial markets are less developed. Allen et al. (2005), Lu and Tao (2009), Du et al. (2012) provide evidence of financing frictions in China in particular. On the other hand, such restriction does not necessarily imply that we rule out borrowing completely. Firms can still be allowed to hold debt, but it is common both in theory and in practice that borrowing is highly constrained in less developed markets, which can result in an overall positive liquidity balance. In all, we regard the assumption $\alpha_{s,t} \in [0, 1]$ as a reduced form of capturing the result after firms have balanced all liquid assets (cash and receivables) and all liabilities (debt and payables).

Firms maximize the present value of dividends minus the cost of financing. We use the word "dividend" in a general sense: we do not restrict ourselves to public firms, and therefore dividends simply refer to firm value that is consumed rather than saved by firm owners. Due to production risks, firms will distribute dividends only when they have accumulated sufficient wealth. Standard argument implies that lump sum dividends $W_t - \overline{W}$ are made when W_t exceeds a reflecting boundary $\overline{W}$. Meanwhile, production risks can also bring large losses of firm wealth. When W_t is too low, firms must refinance.[8] In reality, refinancing is seldom inexpensive. Whether firms refinance through debt or equity, there usually are substantial costs involved, including search costs, agency costs, underwriting fees, and other costs due to new covenants or change of control. We abstract away from the micro-market details by introducing a reduced form costly refinancing scenario: we assume refinancing occurs when W_t is below some threshold $\underline{W}$; firms pay a marginal cost $1 + \delta > 1$ for each unit of asset raised; and they raise just enough external wealth that makes $\underline{W}$ also a reflecting boundary.[9] These assumptions allow a simple yet general enough model to focus our attention on liquidity management. In the Appendix we discuss further the technique of rationalizing this reflecting refinancing boundary, including ruling out constant refinancing from firms' optimal equilibrium choice set by introducing an additional fixed refinancing cost.

Finally, to properly define an equilibrium, we need to introduce a market for capital. We assume a fixed capital (for example, land) supply of one unit.[10] We then define the equilibrium of this economy:

Definition 1. *An equilibrium consists of firms' optimal payout and refinancing decisions plus their liquidity management choice α_t and the dynamics of firm assets (3). The price of capital is determined by a market clearing condition*

$$\pi \int \alpha_{h,t} W_{h,t} dF(W_{h,t}) + (1 - \pi) \int \alpha_{l,t} W_{l,t} dF_l(W_{l,t}) = P_t \ , \tag{4}$$

[8]Firms can also choose to liquidate. We assume liquidation yields a low enough return such that firms always prefer refinancing over liquidation.

[9]It does not matter qualitatively whether δ is the same for both high and low productivity firms. As later analysis shows, our theoretical predictions rely mainly on the overall level of δ, rather than its distribution.

[10]One can equivalently assume that capital can be created but with an adjustment cost so that firms do not accumulate infinite capital.

where $F(W_{s,t})$ is the distribution of assets among firms with productivity s. The market clearing condition stems from the fact that aggregate investment in capital by all firms must equal the total value of capital, which is one times the price of capital.

Furthermore, we define a *steady state* equilibrium in which the distribution of firm value $F(W_{s,t})$ does not vary with time (so the subscript t can be dropped). We show that such steady state equilibrium exists and, due to its tractability, focus our analysis on its properties.

We next proceed to establish formally the firm's optimization problem. Then, we solve for the optimal choice of capital holdings $\alpha_{s,t}$ (or cash holdings $1-\alpha_{s,t}$) and demonstrate its variation across productivity levels.

2.2 Model Solution

We first solve the model in the simplest way by treating capital price P_t as a constant. Then, we show that P_t is indeed a constant in the steady state equilibrium. Under a constant P_t, we can drop the time subscript and write $dR_t = \frac{dA_t}{P}$. That is, the return on capital is the return from the output only. We also observe that the solution method for high and low productivity firms is identical except for the different value μ represents, allowing us to drop the productivity index s as well. Substituting dR_t back into dW_t and using Ito's lemma, the value of the firm solves the HJB equation:

$$\rho V(W_t) = \max_{\alpha_t \in [0,1]} \left[\left(\frac{\mu}{P} - r \right) \alpha_t W_t + r W_t \right] V'(W_t) + \frac{1}{2} \frac{\alpha_t^2}{P^2} \sigma^2 W_t^2 V''(W_t) \; . \tag{5}$$

Firms pay out dividends when W exceeds the payout boundary $\overline{W}$ and refinance when W falls below the refinancing boundary $\underline{W}$. These boundaries are determined by the following conditions:

$$V'(\overline{W}) = 1 \; , \tag{6}$$

$$V'(\underline{W}) = 1 + \delta \; , \tag{7}$$

The above characterization of the firm's optimization problem is intuitive: optimal firm value V_t is a function of the firm's assets W_t. Higher weight on capital α_t means firms can produce more but at the same time bear more risk. At both the payout and refinancing boundaries, the marginal value of assets inside the firm equals the marginal value of dividends or the value of new external equity.

Differentiating equation (5) with respect to α_t, the usual first order condition implies

$$\alpha_t = -\frac{\mu P - rP^2}{\sigma^2}\frac{V'(W_t)}{W_t V''(W_t)} \ . \tag{8}$$

This solution of α_t takes the typical Merton form: net return of the risky asset less the risk free interest rate divided by the variance of return and multiplied by the inverse of the coefficient of relative risk aversion.

The HJB equation (5) can potentially have multiple solutions. We focus on one particular solution: constant α_t.[11] This means we conjecture and later verify that $V(W_t)$ satisfies $\frac{V'(W_t)}{V''(W_t)} = -\frac{1}{\gamma}W_t$ for some constant γ. The HJB equation is then simplified to

$$\rho V = \left(\frac{\phi^2}{2\gamma} + r\right) W_t V'(W_t) \ . \tag{9}$$

where

$$\phi = \frac{\mu - rP}{\sigma} \tag{10}$$

We refer to ϕ as the *risk-adjusted return* of a firm's capital. ϕ depends on each firm's productivity μ. As we assume r, P and σ are uniform for all firms, a high productivity firm also has a high level of risk-adjusted return on capital.

The firm's HJB has the following closed-form solution:

$$V(W) = \theta W^{1-\gamma} \ , \tag{11}$$

where θ is a constant coefficient to be determined by matching the boundary conditions, and γ measures the firm's (endogenously implied) degree of risk aversion. A higher γ implies a more risk averse firm. While firms in this model are risk-neutral by nature, they are effectively risk averse due to the concern of avoiding costly refinancing, a standard result of dynamic investment models. We relegate to the Appendix the algebraic details on the

[11]This is the solution in a standard Merton problem if there is sufficient adjustment cost for α_t

derivation of parameters for $V(W)$.

Combining (11) with the formula for α, we arrive at the following lemma:

Lemma 1. *The fraction of capital held by a firm of productivity $s \in \{l, h\}$, α_s is given by*

$$\alpha_s = \frac{\phi_s P}{\gamma_s \sigma} \,. \tag{12}$$

where ϕ_s is the risk-adjusted return of capital defined in (10) and γ_s is the firm's (endogenously implied) degree of risk aversion

Lemma 1 states that a firm's liquidity choice, $1-\alpha$, is driven by two forces. A firm holds more capital and thus less cash if it has a higher return from capital and a lower degree of risk aversion. While more productive firms naturally have higher return from capital, the following result shows that they are also endogenously *more* risk averse than low productivity firms:

Lemma 2. *The (endogenously implied) degree of risk aversion of a firm of productivity $s \in \{l, h\}$, γ_s satisfies $0 < \gamma_l < \gamma_h < 1$.*

While this might seem somewhat surprising, it is a natural result of firms' optimal liquidity management. To see why, we explicitly solve for the payout boundary $\overline{W}$ from (6), which yields[12]

$$\overline{W} = (1+\delta)^{\frac{1}{\gamma}} \, \underline{W} \,. \tag{13}$$

Therefore, Lemma 2 implies $\overline{W}_h < \overline{W}_l$; that is, high productivity firms pay out dividends earlier. On the one hand, firms discount the future, so early dividends are more valuable. One the other hand, earlier dividends mean that W stops growing earlier and there is higher chance of hitting the refinancing boundary. Therefore, high productivity firms that take advantage of their ability to pay out earlier have less slack between the payout boundary and the refinancing boundary. As such, these firms have stronger incentive to maintain higher asset levels and are more averse to losses than low productivity firms.[13]

[12]See Appendix for the algebraic details.

[13]Examining the differences between our model and models of dynamic risk management such as Rampini and Viswanathan (2010) and Moll (2013) reveal another justification for the early payout of high productivity

It should be noted that Lemma 2 and equation (13) do not necessarily imply that we expect high productivity firms to be *on average* smaller than low productivity firms in the data. Equation (13) is about the payment boundary, whereas average firm size also depends on the *distribution* of firm size. It is possible that in equilibrium more productive firms are clustered near the payment boundary resulting in bigger average size. We also assume uniform refinancing cost for simplicity. In practice it is very likely that refinancing cost is correlated with productivity, for example due to differentially skilled management teams. Despite these possibilities, our main result regarding productivity and liquidity does not rely on how productivity and average firm size are correlated. In the quantitative analysis section, we use the observed moments of size in the data to parameterize our model and therefore do not make specific predictions regarding average size alone.[14]

The final step is to replace the endogenous price of capital P in equation (12) in Lemma 1 by demonstrating that there is a steady state equilibrium in which P is indeed a constant. Substituting (8), the first order condition of α_t, into (3), the dynamics of W_t, implies that in between $\underline{W}$ and $\overline{W}$, W_t follows

$$\frac{dW_t}{W} = \left[\frac{(\mu - rP)^2}{2\gamma\sigma^2} + r\right] dt + \frac{\mu - rP}{\gamma\sigma} dZ_t \ . \tag{14}$$

Equation (14) suggests that W_t is a geometric Brownian motion with two reflecting barriers or RGBM for short. This type of processes has closed form stationary distributions with easy to compute moments. Consider a generic RGBM W_t, where $\frac{dW_t}{W_t} = \mu_W dt + \sigma_W dZ_t$ between an upper reflecting barrier $\overline{W}$ and a lower reflecting barrier $\overline{W}$. Let $\eta \equiv \frac{2\mu_W}{\sigma_W^2}$. Zhang

firms. In those models, firms face borrowing constraints but can save internal net worth until they are large enough that the constraints no longer bind. Firms with higher productivity grow faster and become unconstrained earlier. In our model, the reflecting boundary means that firms cannot grow infinitely large. The Brownian motion implies that no matter how large a firm gets, there is always the possibility of large enough losses such that refinancing is necessary. Consequently, firms never become unconstrained and must all optimally balance between expediting payout and maintaining sufficient liquidity.

[14] Since the majority of our sample is private firms with no information on payout, we are not able to directly test the relationship between productivity and the payment boundary other than through how it affects firms' liquidity policy.

and Du (2010) shows that the stationary distribution of W_t is given by the density function

$$f(W) = \frac{\eta - 1}{\overline{W}^{\eta-1} - \underline{W}^{\eta-1}} W^{\eta-2} \,. \tag{15}$$

This density function is a power function, which is consistent with the cross-sectional distribution of firm size found by Ai et al. (2013). The power distribution can also be easily integrated to obtain its expectation:

$$E(W) = \int_{\underline{W}}^{\overline{W}} \frac{\eta - 1}{\overline{W}^{\eta-1} - \underline{W}^{\eta-1}} W^{\eta-1} dW = \frac{(\eta - 1)}{\eta \left(\overline{W}^{\eta-1} - \underline{W}^{\eta-1}\right)} \left(\overline{W}^{\eta} - \underline{W}^{\eta}\right) \,. \tag{16}$$

Putting $E(W)$ back into the market clearing condition (4) allows us to solve for the equilibrium capital price P. Combining it with the formula for liquidity management choice α, we have the following conclusion regarding α as a function of productivity:

Proposition 1. *In the steady state equilibrium, a firm's optimal holdings of capital α_s are a function of the firm's productivity on capital μ_s, the volatility of return on capital σ, and the refinancing cost δ. In particular, $\alpha_l < \alpha_h$ if δ is low, and $\alpha_l > \alpha_h$ if δ is sufficiently high.*

Proposition 1 highlights the importance of refinancing costs in the relationship between firms' liquidity management and productivity. Intuitively, a higher refinancing cost makes both low and high productivity firms more risk averse, but high productivity firms are more sensitive to refinancing as they accumulate less wealth before dividend payout. Recall Lemma 1, which shows that a firm's optimal capital holdings are positively correlated with its return on capital but inversely correlated with its implied degree of risk aversion, and Lemma 2, which shows that more productive firms are effectively more risk averse. When refinancing cost is low, the effect of the return on capital dominates, and more productive firms hold more capital and less cash. In contrast, when refinancing cost is high, the effect of risk aversion dominates, and more productive firms may hold less capital and more cash.

Proposition 1 implies the following testable hypothesis:

Hypothesis 1. *More productive firms hold more cash and/or liquid assets than less productive firms when there is a substantial cost of external refinancing. The correlation between* *productivity and cash/liquid asset holdings is zero or negative when refinancing cost is low.*

We test this hypothesis in the next section for two economies with significantly different levels of financial development: an emerging economy (China) versus a developed economy (the U.S.). There is abundant empirical evidence that external financing is indeed costly in emerging economies. La Porta et al. (1997, 1998, 2000) show that countries with less formal creditor rights protection and weaker legal enforcement systems have smaller domestic loans, narrower debt markets, and greater difficulty in raising external financing. Allen et al. (2005) and Ayyagari et al. (2010) find costly informal financing in a country with less developed formal financial markets. We also provide supporting evidence of such financing frictions when we quantitatively solve for the refinancing cost in China in section 4.

3 Empirical Results

In this section we empirically test our hypothesis. Our tests focus on data from China, where substantial frictions are known to exist in its financial markets, and compare our results to those obtained using US data. We define two liquidity measures: net liquid assets to total assets ratio and net cash holdings to total assets ratio. We find that there is a robust positive correlation between both liquidity measures and productivity in the Chinese sample, a pattern distinctively different from what is observed in the US sample.

3.1 Data and Summary Statistics

The primary dataset we use is the Annual Survey of Industrial Production from 1999 to 2007 conducted by the National Bureau of Statistics of China (NBS). It has been used in development economic studies such as Cai and Liu (2009), Hsieh and Klenow (2009) and Khandelwal et al. (2013). The dataset includes all state-owned firms as well as non-state-owned firms with sales exceeding five million yuan (about $700,000) per year and spans 37 two-digit manufacturing industries and 31 provinces or province-equivalent municipal cities. It contains over 200,000 firms a year on average, significantly larger and more representative than datasets that rely on publicly available information such as firm disclosure.

The data report firm production information as well as balance sheet information. Cai and Liu (2009) describes the collection procedure and reliability of the data set as follows:

NBS collects the data to compute the Gross Domestic Product. For that purpose, every industrial firm in the dataset is required to file an annual report of production activities and accounting and financial information with the NBS. The information reported to the NBS should be quite reliable, because the NBS has implemented standard procedures in calculating the national income account since 1995 and has strict double checking procedures for above-scale firms. Moreover, firms do not have clear incentives to misreport their information to the NBS, because such information cannot be used against them by other government agencies such as the tax authorities. Misreporting of statistical data was commonly suspected for some time in China, the most notorious was local GDP data provided by local governments. However, the national income accounting of above-scale firms is done by the central NBS, hence is much less subject to manipulation by local governments. The NBS designates every firm in the dataset a legal identification number and specifies its ownership type. Firms are classified into one of the following six primary categories: state-owned enterprises (SOEs), collective firms, private firms, mixed-ownership firms (mainly joint stock companies), foreign firms, and Hong Kong, Macao and Taiwan firms. The NBS does not treat publicly listed companies in China separately, which are all grouped under the mixed-ownership category. By the end of 2005, there were about 1,300 publicly listed companies in Chinas two stock exchanges; only slightly over 700 of them were manufacturing firms, a fraction of our sample.

Our variables of interest are firms' liquidity management and productivity. The value of gross output is measured by the total value of goods produced in each period, which differs from sales due to changes in inventories and is also more closely linked to the intermediate inputs. Intermediate inputs mainly measure the cost of materials used in production. The tax on value added in China is 17% for most firms (with the exception of very small firms, which are rarely included in our dataset). The dataset provides information on value added, which is defined as the value of gross output net of intermediate inputs plus tax. A measure of productive used is the valued added per capital, which is obtained from dividing the value added by the value of fixed assets.

Measuring productivity with value added per capital may raise several concerns. First, production also requires labor and intermediate goods. A firm that employs more workers may produce more than another firm with the same amount of capital. Secondly, capital and labor are endogenously chosen given each firm's underlying technology. Finally, if for some reason capital is mis-measured, then value added per capital may be subject to the same bias. To overcome these issues, we construct an alternative productivity measure: total factor productivity (TFP). Commonly used in development economics, TFP is the residual from regressing output on capital, labor, and intermediate inputs, and is arguably the most unbiased measure of productivity available. The specific estimation procedure has undergone constant improvement. We adopt the latest technique in this field following Ackerberg et al. (2015). This technique has been used in the most recent empirical studies on productivity such as Greenstone et al. (2010), Brandt et al. (2012), and De Loecker and Warzynski (2012). Details of the Ackerberg et al. (2015) are laid out in the Appendix. The calculated TFP is strongly positively correlated with the value added per capital measure reported in our dataset, suggesting that there is unlikely to be any systemic bias or measurement error in capital or value added in our sample. [15]

Using TFP as the productivity measure also eliminates the concern that value added per capital is driven by decreasing returns to scale in capital, and thus its correlation with liquidity demand is merely a reflection of the difference in liquidity management between large and small firms. The procedure used in Ackerberg et al. (2015) explicitly removes the potential confounding factor of size on productivity, as TFP is defined as the residual term in a regression with capital and labor as explanatory variables. In our empirical analysis below, we also control for firm size in all regressions to further reduce such concern.

We define two measures that reflect firms' liquidity management decisions. The first is firm's net liquid assets to total assets ratio. Net liquid assets is the difference between liquid assets and liquid debt. Our second liquidity measure is cash holdings to total assets ratio, which is commonly used in corporate finance studies of liquidity management. However, examining only the level of cash holdings can be biased, as a cash-rich firm may also hold a large amount of debt and is a net borrower, whereas another firm may have less cash but also

[15] We also construct value added per worker as an alternative productivity measure. Using value added per worker produces almost the exact results as using value added per capital which further confirms that our results are not driven by unobservable measurement errors in capital.

no debt. Therefore we use a "net" cash measure, which equals cash plus accounts receivables minus accounts payables[16]. Throughout this paper, the term"cash holdings" always refers to the"net" cash measure.

Table 1 Panel A summarizes the main firm variables.[17] We then compare our sample of Chinese manufacturing firms to their US counterpart from Compustat North America. We find that the firms in our sample are much smaller, in terms of total employment, profit and size. Most balance sheet variables are within a comparable range except for the composition of debt. Chinese firms use more liquid debt and significantly less long-term debt. This reflects the difficulty of financing for most Chinese firms, as long-term debt usually comes with stronger collateral requirement, better credit-history, and more covenant restrictions that are typically difficult to satisfy in a market with significant financing frictions.

[Insert Table 1 here]

As a result of significant economic transitions in China, the firms in our sample have particularly complicated ownership structures. There are two general ways of defining firm ownership. One way is based on paid-up capital, which is capital funded by investors. Each firm has six types of investors: state, collective, legal-person, individual, HMT (Hong Kong, Macau, Taiwan), and foreign. We define ownership based on the type of investor that supplies the largest share of capital. As such, we have six ownership types: SOE (State-owned enterprise), COE (Collective-owned enterprise)[18], DPE (Domestic private-owned enterprise),

[16]It is also common in the literature to add short-term investments and subtract the current portion of long-term debt from a cash holdings measure. However, in our sample the median value for both terms is 0 and we therefore exclude them from our cash holdings measure. Adding them back does not alter our findings qualitatively

[17]The number of observations for TFP is smaller because TFP calculation requires a panel of firms that survive throughout the sample period.

[18] Tian (2000) in "Property Rights and the Nature of Chinese Collective Enterprises" describes the COE as follows: *Collective-owned is a special type of ownership seen in China. It is like Township and Village Enterprises (TVEs). It has at least five features. First, the TVE has no well-defined owners, although a TVE is conceptually owned by the people of a community. There exists no well-defined relation-to-person shares. Therefore, by the standard definition of property rights, the TVE has no well-defined owners. Second, in the majority of TVEs, funds are drawn mainly from the assets of the people of the community, although some are drawn from government loans. Only a small portion of TVEs, the so-called red-hatted collective TVEs, are individual, partnership, or cooperative stock-sharing enterprises, which, in order to obtain low interest rate loans or for ideological reasons, are registered as collective enterprises. Third, the people of the community cannot share directly in the profits of the local TVEs. Only employees are compensated in the form of wages. Fourth, most TVEs are controlled by their communitys administration, particularly the chief leaders. The*

HMT (Hong Kong, Macau, Taiwan-owned enterprise), FE (Foreign-owned enterprise) and LPE (Legal person-owned enterprise). Legal person could be an individual, company, or other entity that has legal rights and obligations.

Panel B of Table 1 summarizes the distribution of ownership in our sample based on paid-up capital. The majority of the firms in our sample are private. One drawback of this classification is that, within the "legal person" category, which accounts for 22% of the sample, we cannot further distinguish investor identity among individual, company, or other entity. To circumvent this issue, we also construct an alternative measure of ownership as the fraction of state-own capital. This by definition is a continuous variable and thus can capture the incremental effect state ownership has on firms' liquidity policy when used as an explanatory variable in a regression. The average state ownership is 10% in our sample while the median is 0%. That is, most firms are 100% privately owned.

3.2 Observations

A. Graphic illustrations

To uncover the relationship between firm productivity and liquidity management, we sort firms into deciles based on their productivity. Within each decile, we compute the average of our two liquidity measures. Our first productivity measure is value added per capital, which is directly computed from the data. The distribution of our liquidity measures conditional on value added per capital is shown in Figure 1.

[Insert Figure 1 here]

Our second productivity measure is TFP, which is calculated according to the latest technique as proposed by Ackerberg et al. (2015). The distribution of liquidity conditional on TFP is shown in Figure 2.

[Insert Figure 2 here]

chief leader can appoint the manager of a TVE and may participate in production decision making. Fifth, TVE capital cannot be transferred or sold. When an individual leaves a community, he automatically loses common ownership of that TVE.

It is evident from these two distributions shown above that more productive firms hold *more* liquid assets and cash. The pattern is robust to the two productivity measures as well the two liquidity measures, albeit the dispersion of net liquid assets is larger than that of cash holdings, and the dispersion across productivity deciles is larger using value added per capital than using TFP. Moreover, these figures reveal that firms in our sample are on average net savers, which justify our model assumption $\alpha \in [0, 1]$.

We conduct a range of robustness checks to confirm that our finding is not driven by any other firm characteristics. We control for firm ownership, industry fixed effects as well as firm size. Figure 3 shows the result for SOE firms only and DPE firms only. Ownership is defined by paid-up capital. The concern that firm ownership matters in China is well grounded. Major banks in China are state-owned, and these SOEs usually have a much lower hurdle for external financing. We find that despite SOEs' generally keeping a lower fraction of assets for liquidity, there is no systematic difference between SOE and DPE in terms of the correlation between liquid asset holdings and productivity. This finding is consistent with Bayoumi et al. (2012). We also define ownership according to firms' registration type, and the same pattern still holds.

[Insert Figure 3 here]

It should be noted that while studies have argued that SOE faces lower financing cost than DPE, the fact that liquidity asset holdings is higher for more productive firms in both groups of firms is not inconsistent with our theoretical prediction. Proposition 1 states that liquidity and productivity should be positively correlated when financing is costly enough. It is possible that in an overall under-developed financial market, financing costs are high enough for both SOE and DPE such that Proposition 1 holds. Later in this section we provide contrasting liquidity policy using a sample of US public firms, which, regardless of their ownership structures, should have lower financing costs compared to most Chinese firms.

We also check the robustness of our finding against potential industry fixed effects. It is possible that the distribution of productivity differs dramatically from one industry to another. We thus remove industry fixed effects using two different approaches. In one approach, we subtract the average industry productivity from each firm's productivity, thus taking into account the difference in average productivity across industries. With the second approach, we first construct productivity deciles within each industry and then aggregate firms that belong to the same decile across all industries to create overall deciles. Results are shown in Figure 4. Again, we find no systematic difference from our baseline analysis. We also plot the same graphs of our liquidity measures against productivity deciles for each industry individually. Results are not reported in the interest of space, but we find a strong positive correlation between liquidity and productivity in all industries except for the tobacco industry which is heavily regulated and should not be treated the same as other manufacturing firms.

[Insert Figure 4 here]

B. Regression Results

We provide regression results to formally support the aforementioned findings. We regress our liquidity measures, net liquid assets and cash ratio, against our two productivity measures, TFP and valued added per capital. For control variables, we follow Gao et al. (2013), which also studies liquidity policy from a sample of private firms, in constructing these control variables. Due to data limitation, the inclusion of certain control variables reduces our sample size significantly.[19] We therefore report regression results for two sets of controls: in the first set, we include basic firm characteristics that as firm size, age, ownership and well as variables that are available throughout the sample: sales growth, leverage, capital investment, foreign sales, etc. The second, more extensive set additionally includes cash flow and R&D expenditure, which are only available in our data starting in 2005.[20] Industry and year fixed effects are also included, and standard errors are clustered in all regressions.

Table 2 reports three OLS regressions of firms' net liquid assets to total assets ratio against our productivity measures. The coefficients on productivity in all regressions are positive

[19] Also, not all variables in Gao et al. (2013), such as dividend, are available in our data. However, our sample includes a much larger number of firms.

[20] Moreover, controlling for cash flow and sales rules out the possibility that high productivity firms hold more liquid assets or cash simply because they have more unspent cash income at the time the survey is conducted. However, since cash flow is only available from 2005-2007, we are then unable to include cash flow volatility as a control due to the data limitation.

and significant, both statistically and economically. For example, a one standard deviation increase in TFP on average leads to a 2.8% increase in net liquid asset holdings, which is fairly substantial given that the average net liquid assets ratio is around 4.6% in the sample. Economic significance is also stronger using value added per capital as the productivity measure: a one standard deviation increase in value added per capital on average implies a 1.9% increase in net liquid assets.

[Insert Table 2 here]

Table 3 reports similar regressions of cash holdings. Results are consistent with those from regressing net liquid assets: all coefficients on productivity are positive and significant. On average, a one standard deviation increase in TFP and value added per capital implies a 1.7% and 1.2% increase in cash holdings, respectively.

[Insert Table Table 3 here]

We check the robustness of our regression results in a number of alternative specifications. For example, defining ownership using the firm's registered type; using a subsample of 100% privately-owned firms; including more control variables such as industry-median adjusted net liquid assets or cash (that is, the ratio of a firm's liquid assets/cash to the industry median). Results hold in all the robustness checks and are thus omitted here.

C. Comparison to Compustat US firms

To further test the prediction of Hypothesis 1, we investigate whether the correlation between firm productivity and liquidity found in the Chinese sample holds for firms in the US, where financial markets are known to be far more developed. Our sample of US firms comes from Compustat. In a recent study, İmrohoroğlu and Tüzel (2014) calculate TFP for Compustat firms using a modern technique similar to ours. We thus construct productivity deciles for Compustat firms using their TFP data directly. We also compute the net liquid assets to total assets ratio as well as the cash holdings to total assets ratio for the Compustat firms.

The results are shown in Figure 5. Unlike for Chinese firms, the relationship with productivity varies across different liquidity measures. The net liquid assets ratio displays a *negative* correlation with productivity except for the highest decile, whereas cash holdings exhibit a less obvious pattern. Furthermore, the range over which the net liquid assets ratio and cash holdings vary across productivity deciles is smaller for Compustat firms. As a further robustness check, we reverse engineer İmrohoroğlu and Tüzel (2014) to compute the value added per capital measure for Compustat firms as well. We do not find any clear patterns of net liquid assets or cash holdings for Compustat firms ranked by value added per capital. We omit the graph in the interest of space.

[Insert Figure 5 here]

We complement our comparison of Chinese manufacturers and their US Compustat counterpart with regression results. Table 4 shows the results from regressing the two liquidity measures against TFP, obtained from İmrohoroğlu and Tüzel (2014). We use the sample period of 1999-2007 and include all the controls used in Table 2 and 3, consistent with regressions run for the Chinese sample. We find a negative coefficient on productivity for net liquid assets, and an insignificant coefficient for cash holdings. This is contrary to the regression results for the Chinese sample but is consistent with economic intuition. The economic magnitude of the coefficient is also smaller. With a one standard deviation increase in TFP, the net liquid assets ratio decreases by 0.59% while the sample average is 28%.

[Insert Table 4 here]

From both graphical illustrations and regression results, we find no evidence of a positive correlation between productivity and liquidity demand for the sample of US manufacturers in Compustat. It should be noted that whereas Compustat contains large public firms, our Chinese sample consists of mostly small private ones. Nevertheless, we show that our findings for the Chinese sample are robust to ownership and firm size. Furthermore, this difference is consistent with Hypothesis 1, since the difference in financing costs should be the greatest between large US public firms and small Chinese private firms.

In sum, we empirically verify that corporate savings, measured by either net liquid assets or cash holdings, has a positive correlation with productivity in a developing economy (China), a pattern that is not observed in a developed economy (US). This is consistent with our theoretical prediction that financing frictions lead more productive firms to maintain more liquidity due to a stronger precautionary savings motive.

4 Quantitative Analysis

The purpose of this section is two-fold: first, while the empirically documented liquidity management behavior is consistent with that predicted by our model, the lack of identification means we fall short of proving that the liquidity policies we observe are *caused* by costly financing. To further validate our theory, we calibrate the model-implied distribution of cash holdings using parameters according to comparable studies as well as moments calculated from our data. We show that the model implied distribution matches the observed distribution well, evidence that our model is likely a reasonable approximation of liquidity management in practice.

The second purpose of this section is to quantitatively infer the level of refinancing cost, which is not directly observable from the data. Furthermore, the calibration exercise allows us to examine the welfare gain in terms of the change in aggregate productivity if refinancing costs were lowered. Overall, we find refinancing costs in China to be almost twice that previously documented for the U.S., and that reducing them can lead to significant welfare improvements.

4.1 Parameter Choices

To begin, we set $r = 3\%$, following the average one-year interest rate on deposits in China in our sample period. We set $\rho = 6.5\%$, following the one-year interest rate on commercial loans. In the model, $\rho - r$ can be interpreted as the wedge between firm owners' required rate of return and the risk-free savings rate, which is commonly proxied by the loan-deposit spread. In China, the loan-deposit spread is set by the People's Bank of China (the central bank) and therefore does not move from day to day. Historically it has been stable, around 3.5% from 1998-2008. The actual loan rate could vary, especially across firms with different ownership structures; private firms usually bear higher rates than state-owned firms. Given that SOEs account for only 5% of our sample, we decide on 6.5% as the uniform rate at which firms discount future dividends. Alternatively, the $\rho - r$ spread can be modeled as the "carrying cost" of cash due to, for example, agency costs of cash. Bolton et al. (2011) sets a carrying cost of 1%. We argue that it is likely higher for the firms in our sample, which are mostly small private firms and are much more likely to be plagued with severe agency problems as found by La Porta et al. (1997, 1998, 2000).

We estimate the average expected return on capital μ to be 15% by utilizing the value added measure in our data combined with information on labor and capital. However, a precise mapping of a productivity measure into the mean and volatility of risk-adjusted productivity shocks μ and σ proves to be challenging. The procedure we follow is broadly consistent with Bai et al. (2006), Hsieh and Klenow (2009), Asker et al. (2012), as well as the standard literature on industrial organization. There is no uniform method for determining the range of μ; nor is there consensus on the average value for various developing markets. We assign $\mu = 15\%$ for firms in the 5th productivity decile of our sample and estimate μ_s for the rest of the deciles based the ratio of each group's TFP estimate over the TFP of the 5th decile. As such, $\mu = 10\%$ for firms in the 1st decile and $\mu = 40\%$ for firms in the 10th decile, a range that is comparable to the variation in TFP.[21]

In the model, the volatility of capital return σ is fixed for simplicity. In practice, σ could vary across firms and be correlated with firms' productivity. We calculate σ to be 0.4 for most firms except those in the largest productivity deciles. We have to manually assign σ to a number of firms due to the brevity of our sample period. The Appendix provides more details on the procedure used to calculate average μ and average σ.

For variables not directly observable from data–payout boundary $\overline{W}$, refinancing boundary $\underline{W}$, price of capital P, risk-adjusted return on capital ϕ_s, firm's degree of risk aversion γ_s–we utilize their mathematical relationship with other model parameters. In particular, we take advantage of the fact that our steady-state distribution of firm assets (15) satisfies the power law and measure the coefficient of the power distribution η directly from the data. In the model, η_s is a function of risk-adjusted return on capital ϕ_s and firm's degree of risk aversion γ_s. Using the mathematical expression for η derived in the Appendix and replacing

[21]We do not claim that our parameter choices for μ are flawless, and the values we use do affect the performance of our quantitative analysis later. However, our choice for μ is not ungrounded. It is informed by our sample; moreover, we are not aware of a clearly better method of estimation.

γ_s with ϕ_s from (21), we can calculate ϕ_s from (38) and infer γ_s. All parameters and their values are summarized in Table 5.

[Insert Table 5 here]

4.2 Quantitative Results

We first test the validity of our model by computing the model-implied cash holdings ratio $1 - \alpha_s$ with the observed cash holdings for each productivity decile. Results are shown in Panel C of Table 5. Compared to the distribution of actual cash holdings, our model generates very similar distributions in terms of correlation with productivity, with the only exception being the 10th decile. This is mainly due to the fact that our estimate for μ relies on the observed dispersion of productivity in our sample, which is most extreme in the right tail.[22] In general, our model predicts well the distribution of cash holdings conditional on productivity and, in particular, how it correlates with increasing productivity.

Next, we use equation 13 to calculate the refinancing cost δ. Our calibration implies that δ equals to 14%. This is quite substantial compared to what is found for large public firms in the US, as estimated by Altinkilic and Hansen (2000) for example. But it is broadly in line with studies on Chinese firms, such as Zou and Qian (2005), Shen (2007), and Park and Shen (2008). There are two main reasons why refinancing cost is higher in our study than that found in Altinkilic and Hansen (2000): one, it is well-known that financial markets in China are much less developed than in the US. Two, most previous studies only focus on public (listed) firms, while our sample comprises of mostly small private firms with more limited access to financing. Overall, the implied cost of refinancing is consistent with the common knowledge that financial markets in China are far less developed than in the U.S. and with the prediction of Proposition 1 that severe financing frictions are an important driver of the positive correlation between cash holdings and productivity.

[22] In fact, if we remove the 5% observations from each tail, and recalculate the productivity dispersion, we would have $\mu = 29.0\%$ for the last decile, which implies a cash holdings ratio of 35.2%–much closer to the observation from data.

4.3 Capital Allocation and Welfare

The fact that high productivity firms hold less capital suggests a potentially large capital misallocation problem. Misallocation due to financial frictions has been extensively discussed in the literature. Buera et al. (2011), Hsieh and Klenow (2009), Moll (2013) argue that asset misallocation plays a significant role in explaining cross-sectional productivity discrepancies across different industries and countries. Our study contributes to the debate. We argue that financial frictions, such as costly refinancing, could result in high productivity firms simultaneously deploying an insufficient level of capital and maintaining an excessive level of liquidity.

We conduct a quantitative experiment to examine how much the distribution of capital, cash, and the overall productivity of capital in the economy would change, if financial frictions were reduced. More specifically, we hypothesize a reduction of δ from the value we calculated above, and solve the payout boundary $\overline{W}$ and firm's portfolio weight on capital α_s for each productivity decile accordingly. Since α_s is a constant independent of W, we define total productivity μ^* as the total output from capital divided by the total amount of capital. That is,

$$\mu^* = \frac{\sum_{s=1}^{10}\left(\int_{\underline{W}}^{\overline{W}} \mu_s \alpha_s W_s dF(W_s)\right)}{\sum_{s=1}^{10}\left(\int_{\underline{W}}^{\overline{W}} \alpha_s W_s dF(W_s)\right)} \tag{17}$$

We calculate that if refinancing cost δ_s were reduced by 20% from its current level, total productivity μ^* would increase by 12%; if δ_s were reduced by 50%, total productivity would increase by 26%. To understand the mechanism behind such productivity gain, note that payout boundaries for all firms are lower under a smaller refinancing cost, because firms have a weaker precautionary savings motive and are able to pay out dividends sooner. All firms hold more capital and less cash. However, with a large reduction in refinancing cost, high productivity firms may hold more capital than low productivity firms. Overall, reducing the cost of refinancing has a significant positive effect on total productivity, thanks to the reallocation of capital from low to high productivity firms.

5 Conclusion

In recent years, firms worldwide have been holding an increasing amount of cash and liquid assets, prompting many to ask why corporations save so much. This paper is a step towards answering that question by seeking to understand how a firm's liquidity management relates to its productivity and the welfare consequences of such management. We show theoretically and empirically that there exists an equilibrium where high productivity firms actually hold more liquid assets and less capital than low productivity firms, due to an overwhelmingly strong precautionary savings motive associated with significant financing frictions. Improvements to the financial market, such as reduction of financing costs, could redistribute capital more efficiently across firms, leading to an overall increase in aggregate productivity.

We capture financial frictions in the model through a simple refinancing cost to highlight the basic mechanism and yet maintain the model's tractability. In practice, financial frictions may come from various sources. We do not distinguish in detail how each specific friction affects capital allocation and corporate savings behaviors, which is an area of growing interest for both theoretical and empirical studies. Our model also applies to asset allocation from a mergers and acquisitions perspective. It can be modified, following Maksimovic and Phillips (2002), Yang (2008), and Warusawitharana (2008), to study M&As in the presence of strong liquidity demand and how M&As impact the overall efficiency of the economy.

We restrict our study to cross-sectional analysis of asset distribution and liquidity management. Many equally interesting questions remain regarding the time-series of these corporate decisions. Eisfeldt and Rampini (2006) shows that the redistribution of firm wealth is strongly pro-cyclical, whereas the socially optimal reallocation should be counter-cyclical. While we resort to a steady-state analysis due to our short-sample period, our model is equipped to incorporate aggregate shocks and transitory dynamic analysis. Furthermore, our sample ends in 2007, right before the global financial crisis. Research such as Campello et al. (2011) shows that firms with various levels of financing constraints manage liquidity differently during crisis times. It remains a question whether the same is true of firms with different levels of productivity. Our model is flexible in capturing the implications of firm characteristics on liquidity management and asset allocation from a time-series perspective.

Appendix A: Tables and Figures

Table 1 – Summary statistics

This table describes the main variables in the firm level dataset from 1999 to 2007. Total assets is the book value of total assets in billions of RMB. Total factor productivity (TFP) is calculated following the methodology in Ackerberg et al. (2015). Value added per capital is computed as the ratio of total value added per 1,000 RMB fixed assets. MNC is an indicator variable taking the value of one if the fraction of foreign sales to total sales exceeds 20%, and zero otherwise. State capital participation is the percentage of paidup capital under state-owned enterprise ownership. Net liquid asset is defined as liquid asset - liquid liability and scaled by total assets. Cash is defined as cash holdings + account receivables - account payable and scaled by total assets. Capital investment is the change of fixed assets scaled by total assets. All variables are winsorized at 0.5‰. Variables with a * indicates variables scaled by total assets.

Panel A: Summary of Variables

Variable	Mean	Median	Std. Dev.	N
Total assets (billions of RMB)	0.0792	0.0132	0.5794	1,818,926
Value added per (1k RMB of) capital	0.0061	0.0013	0.0534	1,793,251
TFP	0.1440	0.0919	0.1722	1,046,990
Firm age	11.071	7.000	11.523	1,826,670
MNC	0.2191	0.0000	0.4136	1,834,472
State capital participation	0.1010	0.0000	0.2888	1,805,148
Fixed asset*	0.3304	0.2970	0.2172	1,805,980
Liquid asset*	0.5755	0.5882	0.2350	1,818,926
Short-term investment*	0.0030	0.0000	0.0288	1,060,147
Account receivables*	0.1832	0.1304	0.1790	1,818,926
Inventory*	0.1884	0.1449	0.1679	1,818,926
Net liquid assets*	0.0463	0.0572	0.3484	1,818,926
Cash*	0.2383	0.2269	0.2782	1,060,147
Long-term investment*	0.0107	0.0000	0.0482	1,818,926
Total debt*	0.5969	0.6000	0.3373	1,818,926
Long-term debt*	0.0562	0.0000	0.1408	1,818,926
Liquid debt*	0.5293	0.5242	0.3247	1,818,926
Account payables*	0.1610	0.0946	0.1916	1,060,147
Capital Investment*	0.0040	0.000	0.3939	1,250,485
Operating Cash Flow (Net) *	0.0504	0.0000	0.3317	812,119
R&D*	0.0018	0.0000	0.0130	961,063

Panel B: Ownership Structure

	SOE	COE	DPE	HMT	FE	LPE
N of firms	183,718	196,490	742,603	127,559	126,605	408,618
% of firms	10.01%	10.71%	40.48%	6.95%	6.90%	22.27%

Table 2 – Productivity and Net Liquid Asset to Total Asset Ratio (Chinese sample)

This table reports the OLS regression results of firms' net liquid asset to total assets ratio (NLQAT) on productivity, over the 1999-2007 sample period. Net liquid asset is defined as liquid asset - liquid liability and liquid assets is defined as short-term investment + account receivables + inventory + cash holdings. Productivity is defined as TFP in columns (1)-(2), and value added per 1,000 RMB of capital in columns (3)-(4). Basic controls include*total assets*, *sale growth*, *leverage*, *capital investment*, *firm age*, *MNC* (whether the fraction of foreign sales to total sales exceeds 20%), and *state ownership* (the fraction of state-own capital). Additional controls include *cash flow* and *R&D expenditure*. Industry and year fixed-effect is included in each regression. Standard errors are clustered at the firm level and are reported in parentheses. All variables are winsorized at 0.5‰. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Net Liquid Asset to Total Asset Ratio			
	(1)	(2)	(3)	(4)
TFP	0.1603***	0.1298***		
	(0.0021)	(0.0023)		
Value Added per capital			0.3655***	0.3849***
			(0.0127)	(0.0170)
Adjusted R^2	0.534	0.501	0.542	0.502
Observations	1,033,719	568,907	1,225,713	652,508
Basic Controls	YES	YES	YES	YES
Additional Controls	NO	YES	NO	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Cluster SE	YES	YES	YES	YES

Table 3 – Productivity and Cash Holdings (Chinese sample)

This table reports the OLS regression results of firms' cash holdings to total assets ratio (CASH) on productivity. Cash holdings is defined as cash + account receivable - account payable. Productivity is defined as TFP in columns (1)-(2), and value added per 1,000 RMB of capital in columns (3)-(4). Basic controls include*total assets*, *sale growth*, *leverage*, *capital investment*, *firm age*, *MNC* (whether the fraction of foreign sales to total sales exceeds 20%) and *state ownership* (the fraction of state-own capital). Additional controls include *cash flow* and *R&D expenditure*. Industry and year fixed-effect is included in each regression. Standard errors are clustered at the firm level and are reported in parentheses. All variables are winsorized at 0.5‰. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Cash to Total Asset Ratio			
	(1)	(2)	(3)	(4)
TFP	0.0963***	0.0920***		
	(0.0026)	(0.0026)		
Value Added per capital			0.2259***	0.2242***
			(0.0143)	(0.0153)
Adjusted R^2	0.076	0.075	0.072	0.070
Observations	661,484	568,907	773,276	652,508
Basic Controls	YES	YES	YES	YES
Additional Controls	NO	YES	NO	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Cluster SE	YES	YES	YES	YES

Table 4 – Productivity and Liquidity Management (US sample)

This table reports the OLS regression results of firms' net liquid asset to total assets ratio and cash holdings to total assets ratio on measures of productivity, using the US Compustat data from 1999 to 2007. Controls include*total assets*, *sale growth*, *leverage*, *capital investment*, *firm age*, *Tobin's q*, *MNC*, *cash flow* and *R&D expenditure*. Industry and year fixed-effect is included in each regression. Standard errors are clustered at the firm level and are reported in parentheses. All variables are winsorized at 0.5%. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Net Liquid Asset	Cash
TFP	-0.0114*	0.0011
	(0.0065)	(0.0119)
Adjusted R^2	0.335	0.414
Observations	8,197	4,194
Controls	YES	YES
Industry FE	YES	YES
Year FE	YES	YES
Cluster SE	YES	YES

Table 5 – Quantitative Analysis

Panel A and B of this table reports the parameter values used for the quantitative analysis. The lower-case W represents log value of total assets. Panel C shows the results on cash holding. The first row shows the average cash holding to total assets ratio for firms within each productivity decile obtained from our sample. The second row reports the cash-holding to total assets ratio and the implied refinancing costs

Panel A: List of Parameters Measured from Data

Variable	Definition in the Model
r	Risk free savings rate
ρ	Firm discount rate
μ	Expected return from capital
σ	Volatility of returns from capital
$\overline{w}$	Payout boundary (log)
η	Power coefficient of the distribution of firm size

Panel B: Parameter Value

Decile	1	2	3	4	5	6	7	8	9	10
r	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%
ρ	6.5%	6.5%	6.5%	6.5%	6.5%	6.5%	6.5%	6.5%	6.5%	6.5%
μ	10.0%	11.0%	12.2%	13.4%	15.0%	17.6%	18.2%	21.6%	24.0%	40%
σ	0.400	0.400	0.400	0.400	0.400	0.400	0.410	0.420	0.450	0.550
$\overline{w}$	13.33	12.98	12.75	12.54	12.32	12.18	11.95	11.78	11.58	11.41
η	1.08	1.09	1.13	1.14	1.15	1.17	1.20	1.23	1.26	1.29

Panel C: Model implied distribution of cash holdings

Decile	1	2	3	4	5	6	7	8	9	10
Data	13.8%	17.2%	19.6%	21.8%	23.7%	25.4%	27.0%	29.0%	30.9%	34.3%
Model	9.2%	12.4%	16.2%	19.6%	21.3%	24.2%	26.7%	29.8%	30.2%	16.9%

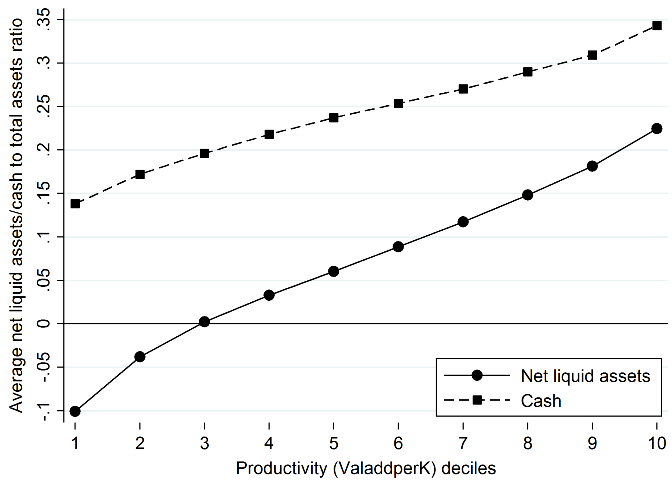


Figure 1 – Liquidity Management and Productivity (Value Added per Capital)

This graph plots the average ratio of Net Liquid Assets (liquid assets - liquid liabilities; solid line) to total assets and the average ratio of Cash holdings (cash and short - term investments + receivables - payables; dashed line) to total assets in each productivity decile based on firms value added per capital. Value added and total assets are directly reported in the dataset. Using value added per worker produces a very similar result.

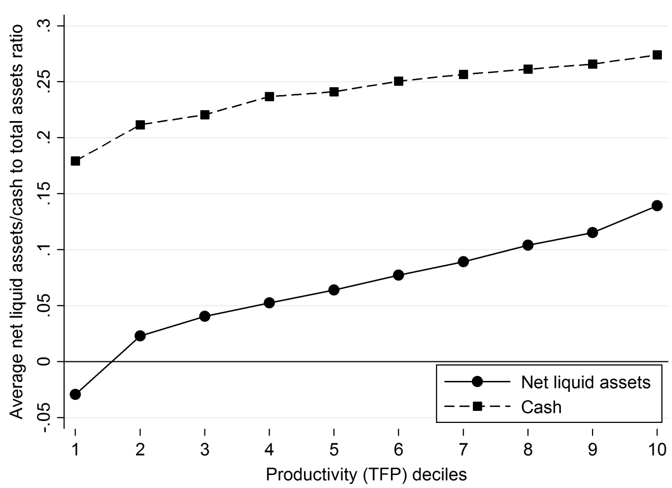
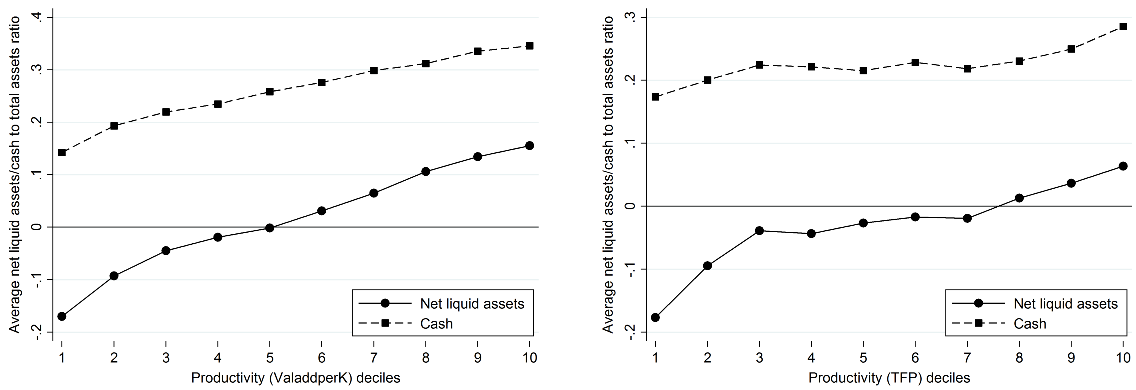
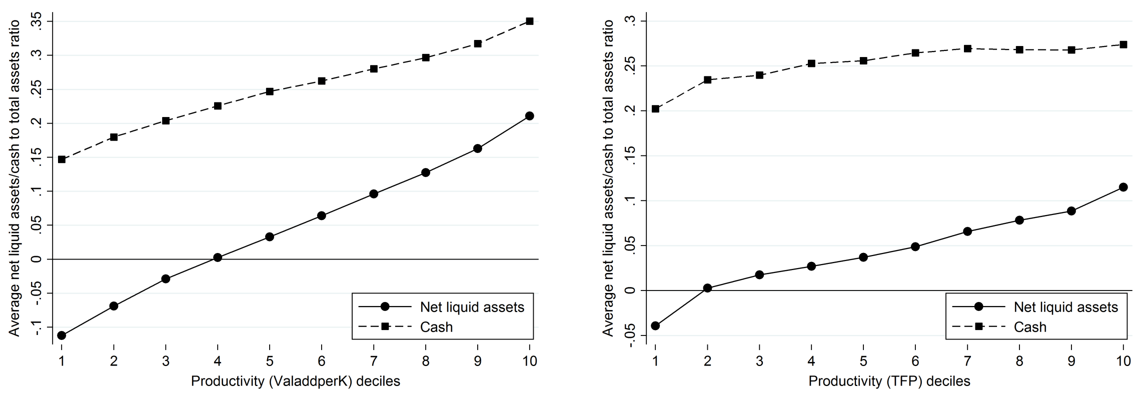


Figure 2 – Liquidity Management and Productivity (TFP)

This graph plots the average ratio of Net Liquid Assets (liquid assets - liquid liabilities; solid line) to total assets and the average ratio of Cash holdings (cash and short - term investments + receivables - payables; dashed line) to total assets in each productivity decile based on firms TFP. TFP is calculated following Ackerberg et al. (2015). Details of the calculation procedure can be found in the Appendix C



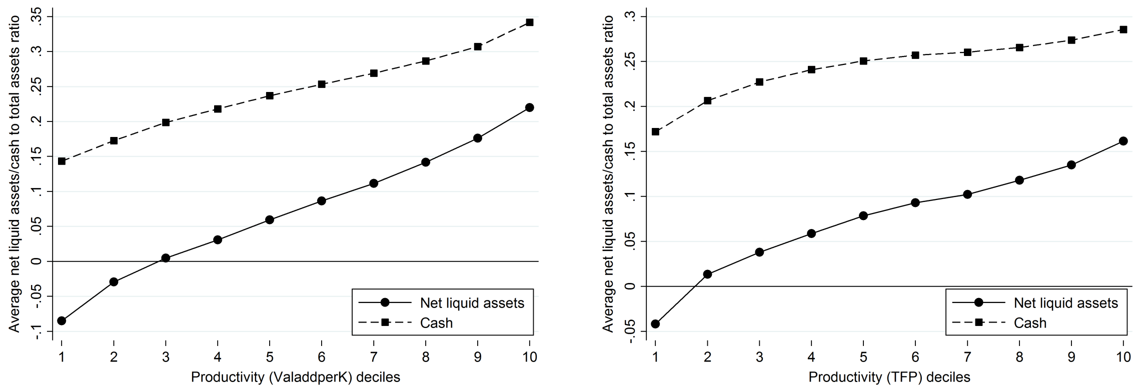
State-owned firms



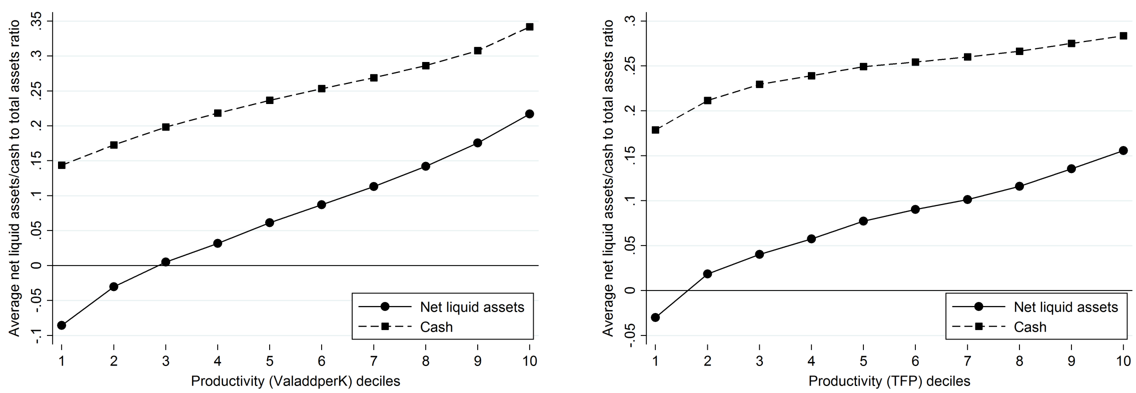
Domestic private-owned firms

Figure 3 – Liquidity Management and Productivity, Controlling for Ownership

This figure shows the average of our two liquidity measures in each productivity decile for certain firm ownerships. Solid line represents the net liquid asset to total assets ratio. Dashed line represents cash to total assets ratio. Ownership is determined by the amount of paidup capital. Results of two ownership type: SOE (state-owned enterprise) and DPE (domestic private-owned enterprises) are shown. For each ownership, the left graph uses value added per capital as the productivity measure, and right graph uses TFP as the productivity measure.



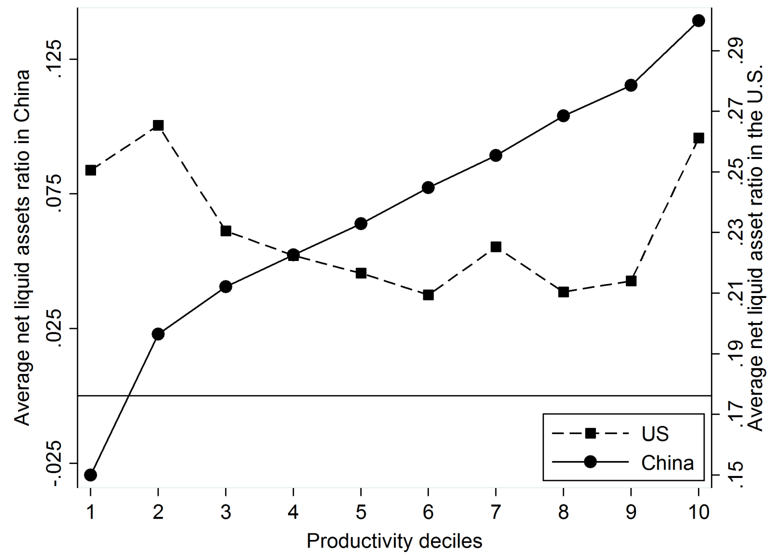
Productivity decile with industry fixed effect removed



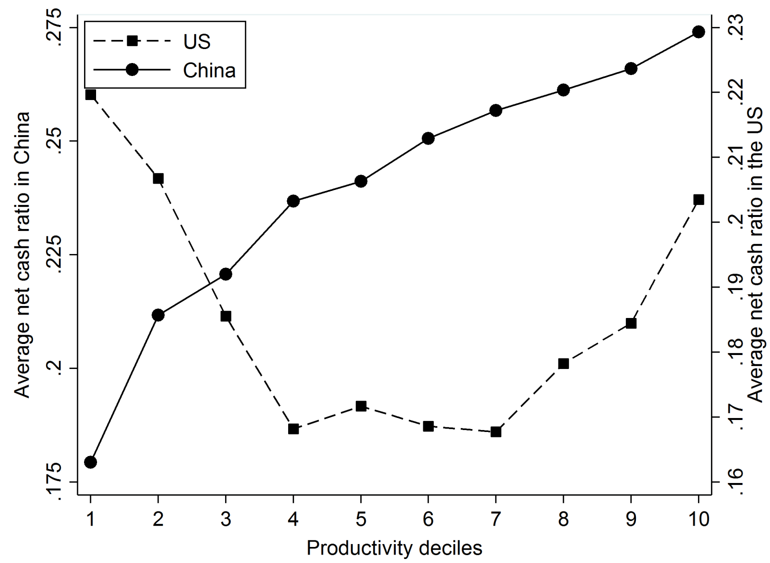
Productivity decile as the sum of 10 productivity groups for each industry

Figure 4 – Liquidity Management and Productivity, Controlling for Industry FE

This figure shows the average of our two liquidity measures in each productivity decile, controlling for industry fixed effect in calculating productivity. Solid line represents the net liquid asset to total assets ratio. Dashed line represents cash to total assets ratio. In the top two graphs, we subtract the average industry value added per capital from each firms value added per capital and then group them into deciles by the difference. In the bottom two graphs, the cross-industry productivity deciles is generated by dividing firms into deciles within each industry and then aggregate the firms in the same decile across all industries. In both rows, the left graph uses value added per capital as the productivity measure, and right graph uses TFP as the productivity measure.



Net liquid assets ratio



Cash holding

Figure 5 – Liquidity Management and Productivity Between China and U.S.

This figure compares the average of our two liquidity measures in each productivity decile between the Chinese sample and the United States Compustat sample. Solid line represents measures for the Chinese sample. Dashed line represents measures for the US Compustat sample. Productivity deciles are constructed using TFP, which, for US Compustat firms, are taken from İmrohoroğlu and Tüzel (2014). The top graph compares the average net liquid asset to total assets ratio. The bottom graph compares the average cash holdings to total assets ratio.

Appendix B: Proofs

Proof of Lemma 1: In this equilibrium, firms choose optimal portfolio according to

$$\alpha = \frac{\mu P - rP^2}{\gamma \sigma^2} \, . \tag{18}$$

Define $\phi = \frac{\mu - rP}{\sigma}$ as the Sharp ratio of capital investment, obviously ϕ must be larger than 0 in equilibrium, otherwise firms will hold cash only. Substituting ϕ into the dynamics of firm assets yields

$$\frac{dW_t}{W_t} = \left[\frac{\phi^2}{\gamma} + r\right] dt + \frac{\phi}{\gamma} dZ_t \, . \tag{19}$$

Combine (18) and (19) into firm's value function yields

$$\rho V = \left(\frac{\phi^2}{2\gamma} + r\right) W_t V'(W_t) \, . \tag{20}$$

The solution is $V(W_t) = \theta W^{1-\gamma}$, where

$$\rho = \left(\frac{\phi^2}{2\gamma} + r\right)(1-\gamma) \, . \tag{21}$$

The coefficient θ and payout boundary can be pinned down by matching the refinancing boundary condition: $V'(\underline{W}) = (1-\gamma)\theta \underline{W}^{-\gamma} = 1 + \delta$ and the payout boundary condition $V'(\overline{W}) = (1-\gamma)\theta \overline{W}^{-\gamma} = 1$, we obtain

$$\theta = \frac{1+\delta}{1-\gamma} \underline{W}^{\gamma} \, , \tag{22}$$

$$\overline{W} = (1+\delta)^{\frac{1}{\gamma}} \, \underline{W} \, . \tag{23}$$

Combining (18) and the definition of ϕ and γ implies

$$\alpha_s = \frac{\phi_s P}{\gamma_s \sigma} \tag{24}$$

□

Proof of Lemma 2:

Equation (21) implies that γ_s is the solution to an equation $\Gamma(\gamma_s) = 0$, where $\Gamma_s(x)$ takes a quadratic form:

$$\Gamma_s(x) \equiv 2rx^2 + (\phi_s^2 + 2\rho - 2r)x - \phi_s^2 \, . \tag{25}$$

Notice that $\Gamma_s(0) < 0$ and $\Gamma_s(1) > 0$, which implies $\Gamma_s(x) = 0$ always has one negative root and one positive root that is between 0 and 1. We keep positive root and therefore

$$\gamma_s = \frac{\sqrt{(\phi_s^2 + 2\rho - 2r)^2 + 8r\phi_s^2} - (\phi_s^2 + 2\rho - 2r)}{4r} \, . \tag{26}$$

Taking the derivative of the numerator with respect to ϕ_s yields

$$\frac{2\phi_s(\phi_s^2 + 2\rho - 2r)^2 + 8r\phi_s}{\sqrt{(\phi_s^2 + 2\rho - 2r)^2 + 8r\phi_s^2}} - 2\phi_s > 0 \;, \tag{27}$$

where the last inequality comes from

$$(\phi_s^2 + 2\rho - 2r) + 4r > \sqrt{(\phi_s^2 + 2\rho - 2r)^2 + 8r\phi_s^2} \;. \tag{28}$$

That is, the numerator is increasing in ϕ_s. Therefore $\phi_h > \phi_l$ implies $0 < \gamma_l < \gamma_h < 1$. □

Proof of Proposition 1:

We first proof the following useful result:

Lemma A.1: *Suppose capital price P is exogenously given, then $\alpha_l > \alpha_h$ ($\alpha_l < \alpha_h$) if P is sufficiently high (low).*

Proof: Equation (18) implies

$$\frac{\alpha_l}{\alpha_h} = \frac{\mu_l P - rP^2}{\mu_h P - rP^2} \cdot \frac{\gamma_h \sigma^2}{\gamma_l \sigma^2} = \frac{\phi_l \gamma_h}{\phi_h \gamma_l} \;. \tag{29}$$

Substituting in (26), we have

$$\frac{\alpha_l}{\alpha_h} = \frac{\sqrt{\left(\phi_h + \frac{2(\rho-r)}{\phi_h}\right)^2 + 8r} - \left(\phi_h + \frac{2(\rho-r)}{\phi_h}\right)}{\sqrt{\left(\phi_l + \frac{2(\rho-r)}{\phi_l}\right)^2 + 8r} - \left(\phi_l + \frac{2(\rho-r)}{\phi_l}\right)} \;. \tag{30}$$

Define function $H(y) \equiv \sqrt{y^2 + 8r} - y$, then $\frac{\alpha_l}{\alpha_h} = \frac{H(y_h)}{H(y_l)}$ where $y_s \equiv \phi_s + \frac{2(\rho-r)}{\phi_s}$. Note that $H'(y) < 0$ as long as $r > 0$, therefore $\frac{\alpha_l}{\alpha_h} > 1$ as long as $y_h < y_l$, or

$$\phi_h + \frac{2(\rho - r)}{\phi_h} < \phi_l + \frac{2(\rho - r)}{\phi_l} \;. \tag{31}$$

Since $\phi_s = \frac{\mu_s - rP}{\sigma}$, given any level of μ_h and μ_l, the above quation is equivalent to

$$r^2P^2 - r(\mu_h + \mu_l) + 2(\rho - r)^2\sigma^2 < 0 \;, \tag{32}$$

The left hand side of the above unequally can be analyzed as a quadratic function of P. Set the left hand side to 0, it has two roots P_1, P_2 where $0 < P_1 < P_2$, provided that

$$r(\mu_h + \mu_l) < 2\sigma\sqrt{\rho - r} \;. \tag{33}$$

However $P_2 > \frac{\mu_h + \mu_l}{2r} > \frac{\mu_l}{r}$ which would imply that $\phi_l < 0$. Therefore we only focus on the smaller root which implies that $\alpha_l > \alpha_h$ if

$$P > \frac{r(\mu_h + \mu_l) - \sqrt{r^2(\mu_h + \mu_l)^2 - 4(\rho - r)^2\sigma^2}}{2r} \;. \tag{34}$$

We now prove the proposition under endogenous P using the market clearing condition for capital. A geometric Brownian motion $\frac{dW_t}{W_t} = \mu_W dt + \sigma_W dZ_t$ with two reflecting barriers $\overline{W}$ and $\underline{W}$ has a stationary distribution. Define $\eta = \frac{2\mu_W}{\sigma_W}$, the stationary distribution is characterized by the density function:

$$f(W) = \frac{\eta - 1}{\overline{W}^{\eta-1} - \underline{W}^{\eta-1}} W^{\eta-2} \tag{35}$$

The price of capital P is the solution to the market clearing condition (4). Since α_s is constant, the market clearing condition in the steady-state equilibrium can be written as

$$P = \pi E\left[\alpha_h E(W_h)\right] + (1-\pi) E\left[\alpha_l E(W_l)\right] \ , \tag{36}$$

where

$$E(W_s) = \frac{(\eta_s - 1)}{\eta_s \left(\overline{W}_s^{\eta_s - 1} - \underline{W}^{\eta_s - 1}\right)} \left(\overline{W}_s^{\eta_s} - \underline{W}^{\eta_s}\right) \ , \tag{37}$$

$\overline{W_s}$ is given by (23), and η_s is given by

$$\eta_s = \frac{2\left(\frac{\phi_s^2}{\gamma_s} + r\right)}{\left(\frac{\phi_s}{\gamma_s}\right)^2} = 2\left(\gamma_s + \frac{r\gamma_s^2}{\phi_s^2}\right) \tag{38}$$

Finally, combing equations (4), (18), (38) and manipulating the algebra implies

$$1 + \delta = \pi_s \left(\frac{\eta_s \phi_s}{\underline{W}(\eta_s - 1)\gamma_s \sigma^2}\right)^{\gamma_s} > \pi_s \left(\frac{\alpha_s}{\underline{W}P}\right)^{\gamma_s} \tag{39}$$

Let Ω stands for the right hand side of (34), then $P > \Omega$ if

$$1 + \delta > \frac{\pi_s}{\underline{W}} > \pi_s \left(\frac{1}{\underline{W}}\right)^{\gamma_s} > \pi_s \left(\frac{\alpha_s}{\underline{W}\Omega}\right)^{\gamma_s} > \pi_s \left(\frac{\alpha_s}{\underline{W}P}\right)^{\gamma_s} \tag{40}$$

for both $s \in \{l, h\}$. It should be noted that this is a sufficient condition for theoretical argument only. In our quantitative analysis, parameters do not have to follow such condition in order for the results to hold. $\square$

Appendix C: Measuring Total Factor Productivity (TFP)

Estimating production technology and measuring productivity heterogeneity using micro-level data is the critical first step in understanding firm decisions. One of the most common measures of productivity is value added per worker or value added per unit of fixed assets. While it is simple and usually directly observable from data, this measure could be biased for several reasons.

First, the premise of using value added as a proxy for productivity is that output is a linear function of labor or capital. In actuality, however, firm output usually depends on a combination of various inputs. Value added per worker or fixed assets is likely to be a biased measure of productivity. For instance, Firm A and Firm B could have the same number of employees, but firm A produces more output because it has better machinery. Using value-added per worker to estimate productivity would lead to the erroneous conclusion that Firm A's workforce is more productive. The solution is usually to estimate a concave production function, for example a Cobb-Douglas function or a translog function.

The second issue is endogeneity. To illustrate, assume a firm has a Cobb-Douglas production function, whose natural log form is given by $y_i = \beta_0 + \beta_k k_i + \beta_\ell \ell_i$. The firm specific productivity is defined as $a_i \equiv \ln(A_i) = \beta_0$, which is usually obtained as the residual term from estimating the production function. In this case, firm's profit maximization will create an upward endogeneity bias. Assuming that at least labor is partially adjustable, the first order condition for the firm's profit maximization implies that

$$\frac{\partial \pi_i}{\partial \ell_i} = 0 \Leftrightarrow p_i A_i \beta_\ell K_i^{\beta_k} L_i^{\beta_\ell - 1} = W \tag{41}$$

Take log,

$$\ln p_i + \ln A_i + \ln \beta_\ell + \beta_k \ln K_i + (\beta_\ell - 1) \ln L_i = \ln W \tag{42}$$

$$\ln L_i = \frac{1}{1 - \beta_\ell} \left[\ln \beta_\ell + \ln p_i - \ln W_i + \beta_k k_i + a_i \right] \tag{43}$$

Since productivity is not exogenous, firms will adjust their labor demand according to the perception of their productivity. If higher productivity leads to higher demand for labor, this estimation would introduce an upward bias in β_ℓ, which in turn implies a downward-biased estimate of TFP. Also, measurement errors in the inputs, typically more severe for capital $k_i = k_i^* + u_i$, would cause a downward bias in β_k.

Besides ones described above, there are other important problems when estimating the production function: serial correlation or unobserved heterogeneity in productivity, and sample selection or the exit of unproductive firms.

An innovative approach developed by Olley and Pakes (1996) is considered a panacea for all aforementioned issues. They propose a dynamic model in which each firm makes two decisions: first, to remain in operation or to shut down, which depends on whether productivity is above a certain threshold monotonically increasing in capital stock; second, if a firm decides to stay in operation, it must then decide how much to invest, which is a function of both productivity and capital stock, $i_{it} = i_{it}(\omega_{it}, k_{it})$. To estimate the model, four key assumptions are made. First, the production function shock has two components: $a_{it} = \omega_{it} + e_{it}$, in which the unobserved productivity ω_{it} is a FOMP (First-Order Markov

Process, $P\left(\omega_{it+1} \mid \{\omega_{i\tau}\}_{\tau=0}^{t}\right) = P\left(\omega_{it+1} \mid \omega_{it}\right)$, which is more general than AR(1)). Second, labor is a flexible input chosen after observing ω_{it}, which means productivity ω_{it} is exogenous to the choice of labor. Third, capital is quasi-fixed with the time-to-build feature ($K_{it} = (1-\delta) K_{it-1} + I_{it-1}$). Finally, the investment function is strictly monotonic in ω_{it} conditional on k_{it}. The crucial assumption that productivity is a function of capital and investment allows the correction of simultaneity in the following way:

$$\begin{aligned} y_{it} &= \beta_\ell \ell_{it} + \beta_k k_{it} + \underbrace{h_t\left(k_{it}, i_{it}\right)}_{\omega_{it}} + e_{it} && (44) \\ &= \beta_\ell \ell_{it} + \phi\left(k_{it}, i_{it}\right) + e_{it} && (45) \end{aligned}$$

The equation above allows consistent estimates of β_ℓ and $\phi\left(k_{it}, i_{it}\right)$ by regressing output y_{it} on labor ℓ_{it} and a polynomial function of capital k_{it} and investment i_{it}. Given a consistent β_ℓ, we can solve the correlated effects and sample selection issues together as follows:

$$\begin{aligned} E\left[y_{it} - \beta_\ell \ell_{it} \mid \chi_{it} = 1\right] &= \beta_k k_{it} + E\left[\omega_{it} \mid \chi_{it} = 1\right] && (46) \\ &= \beta_k k_{it} + E\left[\omega_{it} \mid \omega_{it} \geq \bar{\omega}\left(k_{it}\right)\right] && (47) \\ &= \beta_k k_{it} + g\left(\omega_{it-1}, \bar{\omega}\left(k_{it}\right)\right) && (48) \end{aligned}$$

$g\left(\omega_{it-1}, \bar{\omega}\left(k_{it}\right)\right)$ is a function of the unobserved firm productivity (lagged) and the survival probability ($P_{it} = P\left(\chi_{it} = 1 \mid \omega_{it-1}, \bar{\omega}\left(k_{it}\right)\right)$), which must be estimated. Olley and Pakes (1996) suggest using $h_{it-1} = \phi_{it-1} - \beta_k k_{it-1}$ as a proxy for the first, and the predicted probability of survival from a probit or semi-parametric estimate of exit as a proxy for the latter. Then,

$$\hat{\phi}_{it} = \beta_k k_{it} + g\left(\hat{\phi}_{it-1} - \beta_k k_{it-1}, \hat{P}_{it}\right) + \xi_{it} \tag{49}$$

Since the functional form of $g\left(\cdot\right)$ is unknown, Olley and Pakes (1996) suggest modeling it as a polynomial in $h\left(\cdot\right)$ and P_{it} or as a kernel regression. In practice, as long as one maintains an unbalanced panel, the P_{it} typically does not change results too much. Finally, we can search for more control functions that rely on better data. For instance, Levinsohn and Petrin (2003) propose using intermediate inputs instead of investment in the above estimation, since there are a large number of observations with zero investment in firm-level datasets from developing countries while intermediate inputs are usually well reported in micro-level datasets.

The technique proposed by Ackerberg et al. (2015) further improves upon Olley and Pakes (1996) and is currently considered the state-of-the-art approach to estimating productivity. It has been adopted in a number of the most recent empirical studies such as Greenstone et al. (2010), Brandt et al. (2012), and De Loecker and Warzynski (2012). Ackerberg et al. (2015) argue that the first stage estimation in both Olley and Pakes (1996) and Levinsohn and Petrin (2003) does not identify the labor coefficient due to functional dependence. They suggest an alternative estimation procedure that avoids this problem. Specifically, they invert investment or the intermediate demand function which are both conditional on labor input. As a result, their moment conditions produce consistent estimates even if labor is chosen prior to other inputs, there are unobserved and serially correlated shocks to the price of labor, and/or there are firm-specific adjustment costs to labor.

Our procedure to compute the alternative TFP closely follows Ackerberg et al. (2015). We consider the production function

$$\begin{aligned} y_{it} =& \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \\ & \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{lk} l_{it} k_{it} + \beta_{lm} l_{it} m_{it} + \beta_{km} k_{it} m_{it} + \omega_{it} + \epsilon_{it} \end{aligned} \tag{50}$$

where y_{it} is the gross output of firm i in year t, l_{it} is the amount of labor employed, k_{it} is the book value of fixed capital after depreciation, m_{it} is the value of intermediate inputs, and ω_{it} is firm productivity.

To proxy for firm's productivity, we follow Levinsohn and Petrin (2003) by inverting the material demand, $m_{it} = m_t(k_{it}, \omega_{it}, z_{it})$. Assuming that m_t is a monotonic function, we can rely on $\omega_{it} = h_t(m_{it}, k_{it}, z_{it})$ to proxy for productivity in the production function estimation.

Our first stage estimation is

$$y_{it} = \phi(l_{it}, k_{it}, m_{it}, z_{it}) + \epsilon_{it} \ , \tag{51}$$

from this we recover the estimate for expected output ($\hat{\phi}_{it}$) and for the residual (ϵ_{it}). The expected output is given by

$$\begin{aligned} \hat{\phi}_{it} =& \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \\ & \beta_{mm} m_{it}^2 + \beta_{lk} l_{it} k_{it} + \beta_{lm} l_{it} m_{it} + \beta_{km} k_{it} m_{it} + h_t(m_{it}, k_{it}, z_{it}) \ . \end{aligned} \tag{52}$$

In the second stage, we rely on the law of motion for productivity to estimate all the coefficients in the production function:

$$\omega_{it} = g_{it}(\omega_{it-1}) + \xi_{it} \ , \tag{53}$$

where ξ_{it} is an idiosyncratic shock. From the first stage, we can compute productivity for any value of β, where $\beta = (\beta_l, \beta_k, \beta_m, \beta_{ll}, \beta_{kk}, \beta_{mm}, \beta_{lk}, \beta_{lm}, \beta_{km})$, using

$$\begin{aligned} \omega_{it}(\beta) =& \hat{\phi}_{it} - (\beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \\ & \beta_{mm} m_{it}^2 + \beta_{lk} l_{it} k_{it} + \beta_{lm} l_{it} m_{it} + \beta_{km} k_{it} m_{it}) \ . \end{aligned} \tag{54}$$

Given β and $\xi_{it}(\beta)$, the idiosyncratic shock to productivity can be obtained by nonparametrically regressing $\omega_{it}(\beta)$ on its lag $\omega_{it-1}(\beta)$. To obtain the estimates of the production function, we estimate the following moment conditions

$$E\left(\xi_{it}(\beta) Y_{it}'\right) = 0 \tag{55}$$

by using standard GMM technique and relying on block bootstrapping for the standard errors, where $Y_{it} = \left\{l_{it-1}, l_{it-1}^2, m_{it-1}, m_{it-1}^2, k_{it}, k_{it}^2, l_{it-1} m_{it-1}, l_{it-1} k_{it}, m_{it-1} k_{it}\right\}$. Eventually, we use the GMM estimates to recover the TFP for each firm.

Table A1 reports estimates of the average output elasticities for each input, as well as the return to scale. Almost all industries display decreasing return to scale (except Tobacco which is heavily regulated), and most industries are labor intensive, consistent with previous research.

Table A1. Average Input Elasticities of Output

This table reports estimates of the average output elasticities and returns to scale. For each industry, the coefficients are based on a value-added specification estimated using a complete polynomial of degree two with the method from Ackerberg et al. (2015). These coefficients are used to identify the production function and to compute TFP for each firm. Standard errors (in parentheses) are estimated using the bootstrap with 200 replications and are reported in the parentheses below the estimates of the coefficients.

Industry	Labor	Capital	Return to Scale
Food Processing	0.4986	0.2026	0.7012
	(0.0109)	(0.0088)	(0.0123)
Food Production	0.5494	0.2915	0.8409
	(0.0214)	(0.0148)	(0.0202)
Beverage	0.4997	0.2715	0.7713
	(0.0308)	(0.0205)	(0.0331)
Tobacco	0.7544	0.5554	1.3098
	(0.1204)	(0.0854)	(0.0815)
Textiles	0.3859	0.2135	0.5994
	(0.0088)	(0.0059)	(0.0089)
Garments	0.5245	0.1729	0.6974
	(0.0099)	(0.0066)	(0.0108)
Leather	0.4648	0.2010	0.6658
	(0.0142)	(0.0101)	(0.0159)
Timber	0.4599	0.1886	0.6485
	(0.0179)	(0.0130)	(0.0187)
Furniture	0.5839	0.1783	0.7622
	(0.0193)	(0.0161)	(0.0215)
Papermaking	0.4548	0.2436	0.6984
	(0.0156)	(0.0113)	(0.0172)
Printing	0.5256	0.3339	0.8595
	(0.0230)	(0.0178)	(0.0256)
Cultural	0.5060	0.1895	0.6955
	(0.0190)	(0.0132)	(0.0205)
Petroleum Processing	0.4708	0.3994	0.8702
	(0.0327)	(0.0247)	(0.0240)
Raw Chemical	0.3661	0.2756	0.6417
	(0.0108)	(0.0073)	(0.0111)
Medical	0.4780	0.2770	0.7549
	(0.0188)	(0.0123)	(0.0191)
Chemical Fibers	0.4424	0.3068	0.7492
	(0.0403)	(0.0315)	(0.0346)
Rubber	0.3782	0.2784	0.6566
	(0.0276)	(0.0190)	(0.0295)
Plastic	0.4193	0.2492	0.6686
	(0.0112)	(0.0086)	(0.0126)
Non-metal Products	0.3327	0.2499	0.5826
	(0.0124)	(0.0071)	(0.0138)
Ferrous Processing	0.5510	0.3186	0.8696
	(0.0174)	(0.0119)	(0.0124)
Non-ferrous Processing	0.4762	0.2634	0.7396
	(0.0224)	(0.0157)	(0.0204)
Metal Products	0.4373	0.2415	0.6787
	(0.0109)	(0.0086)	(0.0121)
General Machinery	0.3901	0.2571	0.6473
	(0.0126)	(0.0076)	(0.0132)
Special Equipment	0.4629	0.2294	0.6923
	(0.0197)	(0.0103)	(0.0197)
Transport Equipment	0.5942	0.3008	0.8949
	(0.0140)	(0.0095)	(0.0135)
Electric Machinery	0.4946	0.2634	0.7580
	(0.0115)	(0.0086)	(0.0130)
Electronics & Telecom	0.6074	0.2609	0.8683
	(0.0126)	(0.0086)	(0.0118)
Instruments	0.5151	0.1789	0.6940
	(0.0258)	(0.0166)	(0.0266)
Crafts	0.3875	0.1786	0.5661
	(0.0169)	(0.0110)	(0.0188)
Recycling	0.4446	0.2117	0.6562
	(0.0871)	(0.0586)	(0.0910)

Appendix D: Technical Discussions

Endogenizing the Refinancing Boundary

In the paper we assume a refinancing boundary $\underline{W}$ that is reflecting and exogenous. We provide some additional discussion here how to endogenously form such boundary using the same technique as in Bolton et al. (2011): in addition to the marginal refinancing cost δ, firms must also pay a fixed cost ξ each time they raise external funding. The fixed cost ensures that firms will only refinance when assets are sufficiently low and in lump-sum rather than as a flow. These assumptions imply that the boundary condition associated with the HJB at $W = \underline{W}$ are

$$V'(\underline{W}) = 1 + \delta \ , \tag{56}$$

and

$$V(\underline{W}) = \xi + (1 + \delta)\underline{W} \ . \tag{57}$$

$W = \underline{W}$ is then endogenous and can be derived using the value-matching procedure introduced above. Specifically, substituting $V(W) = \theta W^{1-\gamma}$ into the boundary conditions above yields

$$\theta \underline{W}^{1-\gamma} = \xi + (1 + \delta)\underline{W} \tag{58}$$

and

$$(1 - \gamma)\theta \underline{W}^{-\gamma} = 1 + \delta \tag{59}$$

together one can solve for $\underline{W}$

$$\underline{W} = \frac{(1 - \gamma)\xi}{\gamma(1 + \delta)} \tag{60}$$

Measuring the Return and Volatility on Capital

We estimate the marginal return to capital μ following the procedure used in Bai et al. (2006). In particular, we assume the production function is Cobb-Douglas, where $y = zk^{\lambda_k}\ell^{\lambda_l}$. The marginal return to capital is

$$\mu = \frac{p_y y - wL}{p_k k} - \Delta = \frac{\lambda_k}{p_k k / p_y y} - \Delta \tag{61}$$

λ_k is the capital share of income. While it is not directly observable, in a Cobb-Douglas function, it equals one minus λ_l, the labor share of income, which is usually reported in the data. However, our data, the Annual Surveys of Industrial Production in China, report only wage payments and do not provide information on non-wage compensation. Therefore, we resort to the aggregate labor share in manufacturing reported in the Chinese input-output tables and the national accounts to set $\lambda_k = \lambda_l = 0.5$. Following Hsieh and Klenow (2009), we define $p_k k$ as the book value of fixed capital net of depreciation and $p_y y$ as the value added. To compute the average marginal return to capital, we use the formula

$$\bar{\mu} = \frac{\lambda_k}{\left(\sum_i p_{k_i} k_i\right) / \left(\sum_i p_{y_i} y_i\right)} \tag{62}$$

which averages out noise across different types of firms. We use a similar formula to compute the average marginal return to capital for each value added per capital decile.

To estimate productivity volatility σ, we use a balanced panel with 34,035 firms from 1999 to 2007. For simplicity, we first use OLS to estimate TFP by assuming a Cobb-Douglas production function. Then we impose an AR(1) process on the TFP and compute the standard deviation of the residuals as σ. Overall, the standard deviation of the white noise for the AR(1) process is between 0.4 and 0.55, consistent with the findings in Asker et al. (2012) in the World Bank Survey.

References

Ackerberg, D. A., Caves, K., and Frazer, G. 2015. Identification properties of recent production function estimators. *Econometrica, forthcoming.*

Ai, H., Kiku, D., and Li, R. 2013. A mechanism design model of firm dynamics: The case of limited commitment. Working Paper. University of Minnesota, University of Pennsylvania, and Purdue University.

Allen, F., Qian, M., and Qian, J. 2005. Law, finance, and economic growth in China. *Journal of Financial Economics*, 77:57–116.

Almeida, H., Campello, M., and Weisbach, M. S. 2004. The cash flow sensitivity of cash. *Journal of Finance*, 59:1777–1804.

Altinkilic, O. and Hansen, R. S. 2000. Are there economies of scale in underwriting fees? evidence of rising external financing costs. *Review of Economic Studies*, 13:191?218.

Armenter, R. 2012. The rise of corporate savings. *Business Review*, Q3:1–8.

Armenter, R. and Hnatkovska, V. 2012. The macroeconomics of firms' savings. In *2012 Meeting Papers*, number 803. Society for Economic Dynamics.

Asker, J., Collard-Wexler, A., and De Loecker, J. 2012. Productivity volatility and the misallocation of resources in developing economies. Technical report, National Bureau of Economic Research.

Ayyagari, M., Demirguc-Kunt, A., and Maksimovic, V. 2010. Formal versus informal finance: Evidence from China. *Review of Financial Studies*, 23:3048–3097.

Bai, C.-E., Hsieh, C.-T., and Qian, Y. 2006. The return to capital in China. *Brookings Papers on Economic Activity*, 37(2):61–102.

Bates, T. W., Kahle, K. M., and Stulz, R. M. 2009. Why do u.s. firms hold so much more cash than they used to? *The Journal of Finance*, 64(5):1985–2021.

Bayoumi, T., Tong, H., and Wei, S.-J. 2012. The chinese corporate savings puzzle: a firm-level cross-country perspective. In *Capitalizing China*, pages 283–308. University of Chicago Press.

Bigelli, M. and Snchez-Vidal, J. 2012. Cash holdings in private firms. *Journal of Banking and Finance*, 36(1):26 – 35.

Bolton, P., Chen, H., and Wang, N. 2011. A unified theory of tobin's q, corporate investment, financing, and risk management. *Journal of Finance*, 66:1545–1578.

Brandt, L., Biesebroeck, J. V., and Zhang, Y. 2012. Creative accounting or creative destruction? firm-level productivity growth in chinese manufacturing. *Journal of Development Economics*, 97(2):339 – 351.

Brav, O. 2009. Access to capital, capital structure, and the funding of the firm. *The Journal of Finance*, 64(1):263–308.

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

Cai, H. and Liu, Q. 2009. Competition and corporate tax avoidance: Evidence from chinese industrial firms. *The Economic Journal*, 119(537):764–795.

Campello, M., Giambona, E., Graham, J., and Harvey, C. 2011. Liquidity management and corporate investment during a financial crisis. *Review of Financial Studies*, 24:1944–1979.

Cardarelli, R. and Ueda, K. 2006. Awash with cash: Why are corporate savings so high? *World Economic Outlook, April 2006*.

De Loecker, J. and Warzynski, F. 2012. Markups and firm-level export status. *American Economic Review*, 102(6):2437–71.

DeMarzo, P., Fishman, M., He, Z., and Wang, N. 2012. Dynamic agency and the q theory of investment. *Journal of Finance*, 67:2295–2340.

Dittmar, A., Mahrt-Smith, J., and Servaes, H. 2003. International corporate governance and corporate cash holdings. *The Journal of Financial and Quantitative Analysis*, 38(1):pp. 111–133.

Du, J., Lu, Y., and Tao, Z. 2012. Contracting institutions and vertical integration: Evidence from China's manufacturing firms. *Review of Financial Studies*, 40:89–107.

Eisfeldt, A. and Rampini, A. 2006. Capital reallocation and liquidity. *Journal of Monetary Economics*, 53:369–399.

Gao, H., Harford, J., and Li, K. 2013. Determinants of corporate cash policy: Insights from private firms. *Journal of Financial Economics*, 109(3):623 – 639.

Greenstone, M., Hornbeck, R., and Moretti, E. 2010. Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 118(3):pp. 536–598.

Greenwood, J., Sanchez, J. M., and Wang, C. 2010. Financing development: The role of information costs. *American Economic Review*, 100(4):1875–91.

Gruber, J. W. and Kamin, S. B. 2015. The corporate savings glut in the aftermath of the global financial crisis. Working Paper, Board of Governors of the Federal Reserve System.

Hsieh, C.-T. and Klenow, P. 2009. Misallocation and manufacturing tfp in China and India. *Quarterly Journal of Economics*, 124:1403–1448.

İmrohoroğlu, A. and Tüzel, c. 2014. Firm-level productivity, risk, and return. *Management Science*, 60:2073–2090.

Karabarbounis, L. and Neiman, B. 2012. Declining labor shares and the global rise of corporate saving. Technical report, National Bureau of Economic Research.

Khandelwal, A. K., Schott, P. K., and Wei, S.-J. 2013. Trade liberalization and embedded institutional reform: Evidence from chinese exporters. *American Economic Review*, 103(6):2169–95.

Khurana, I. K., Martin, X., and Pereira, R. 2006. Financial development and the cash flow sensitivity of cash. *Journal of Financial and Quantitative Analysis*, 103:787–807.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. 1997. Legal determinants of external finance. *Journal of Finance*, 52:1131–1150.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. 1998. Law and finance. *Journal of Political Economy*, 106:1113–1155.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. 2000. Investor protection and corporate governance. *Journal of Financial Economics*, 58:3–27.

Levinsohn, J. and Petrin, A. 2003. Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2):317–341.

Lins, K. V., Servaes, H., and Tufano, P. 2010. What drives corporate liquidity? an international survey of cash holdings and lines of credit. *Journal of Financial Economics*, 98(1):160 – 176.

Lu, Y. and Tao, Z. 2009. Contract enforcement and family control of business: Evidence from China. *Journal of Comparative Economics*, 37:597–609.

Maksimovic, V. and Phillips, G. 2002. Do conglomerate firms allocate resources inefficiently across industries? theory and evidence. *Journal of Finance*, 57:721–767.

Moll, B. 2013. Productivity losses from financial frictions: Can self-financing undo capital misallocation? Working paper, Princeton University.

Olley, S. and Pakes, A. 1996. The dynamics of productivity in the telecommunications.

Park, A. and Shen, M. 2008. Refinancing and decentralization: Evidence from China. *Journal of Economic Behavior & Organization*, 66(3):703–730.

Rampini, A. A. and Viswanathan, S. 2010. Collateral, risk management, and the distribution of debt capacity. *The Journal of Finance*, 65(6):2293–2322.

Restuccia, D. and Rogerson, R. 2008. Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, 11(4):707–720.

Restuccia, D. and Rogerson, R. 2013. Misallocation and productivity. *Review of Economic Dynamics*, 16(1):1 – 10. Special issue: Misallocation and Productivity.

Riddick, L. A. and Whited, T. M. 2009. The corporate propensity to save. *The Journal of Finance*, 64(4):1729–1766.

Shen, H. 2007. Market segmentation, cross-listing and expected capital cost. *Journal of Financial Research*, 2:13–23.

Song, Z., Storesletten, K., and Zilibotti, F. 2011. Growing like China. *The American Economic Review*, 101(1):196–233.

Tian, G. 2000. Property rights and the nature of chinese collective enterprises. *Journal of comparative Economics*, 28(2):247–268.

Warusawitharana, M. 2008. Corporate asset purchases and sales: Theory and evidence. *Journal of Financial Economics*, 87:471–497.

Whited, T. and Zhao, J. 2016. The misallocation of finance. Working Paper. University of Michigan, SUNY Stony Brook.

Yang, L. 2008. The real determinants of asset sales. *Journal of Finance*, 63:2231–2263.

Zhang, L. and Du, Z. 2010. On the reflected geometric brownian motion with two barriers. *Intelligent Information Management*, 2:295–298.

Zou, W. and Qian, X. 2005. Financing cost, rent seeking and capital allocation inside firms. *Economic Research Journal*, 5:6–22.