Division of Economics, EGC
School of Humanities and Social Sciences
Nanyang Technological University
14 Nanyang Drive
Singapore 637332

A Structural Estimation on Capital Market Distortions in Chinese Manufacturing

Zheng (Michael) SONG
and Guiying (Laura) WU

31 January 2013

EGC Report No: 2013/06

HSS-04-88
Tel: +65 67905689
Email: D-EGC@ntu.edu.sg
http://egc.hss.ntu.edu.sg

ECONOMIC GROWTH CENTRE
Division of Economics
School of Humanities & Social Sciences

Discussion Paper

A Structural Estimation on Capital Market Distortions in Chinese Manufacturing*

Zheng (Michael) Song† Guiying (Laura) Wu‡

This version: January 2013

Abstract

Capital market distortions lower aggregate productive efficiency by misallocating resources. The existing literature infers such distortions from the dispersion of the average revenue product of capital. However, the methodology is subject to a set of identification issues: unobserved heterogeneities in production technology and market power; capital adjustment costs with idiosyncratic shocks; and measurement errors in the data. This paper develops a structural econometric approach of estimating capital market distortions in environments where all the above factors can be present. Using representative firm-level data from Chinese manufacturing from 2004 to 2007, we find that capital market distortions imply aggregate revenue losses of 40 percent. We also estimate distortions for U.S. manufacturing firms in Compustat. Improving capital allocation efficiency to the level observed among the Compustat firms would increase China's manufacturing revenue by 31 percent. Finally, we propose a simplified approach, which addresses the identification issues in a much more tractable way.

JEL Classification: C15, D92, E22, O16, O47

Keywords: capital market distortions, Chinese manufacturing, structural estimation, unobserved heterogeneities

*We would like to thank Chang-Tai Hsieh, Kjetil Storesletten, Shang-Jin Wei, Daniel Xu, Xiaodong Zhu, seminar and conference participants at many institutions for helpful comments. We also thank Chao Duan and Jie Luo for excellent research assistance.

†University of Chicago, Booth School of Business. Email: zheng.song@chicagobooth.edu.

‡Nanyang Technological University, Division of Economics. Email: guiying.wu@ntu.edu.sg.

1 Introduction

Resource allocative efficiency differs across countries. The differences have recently been found important in accounting for the large cross-country difference in aggregate productive efficiency (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009).[1] A cornerstone of the quantitative analysis is to estimate unobserved market distortions. Hsieh and Klenow (2009) calibrate the distortions by matching the dispersions of average revenue products (henceforth referred to as the ARP approach).[2] The validity of the calibration hinges on two conditions: (1) average and marginal revenue products have the same dispersion; and (2) the dispersion of marginal revenue products, a mirror image of price heterogeneity, reflects the magnitude of market distortion. Both conditions are strict. Condition (1) applies only to environments with homogeneous output and demand elasticities. Condition (2) will not necessarily hold in a dynamic environment with adjustment costs. Violation of either of the conditions would lead to a biased estimation.

This paper develops a new method of estimating distortions in a more general environment, where none of the conditions has to hold. Specifically, our model incorporates unobserved firm heterogeneities in factor goods prices, output and demand elasticities. In addition, the model has a rich structure of capital adjustment costs and allows for measurement errors in data. Our goal is to identify the unobserved heterogeneity in the capital goods price, a generic representation of capital market distortions. To this end, we use the simulated method of moments (SMM hereafter) to estimate the model by matching a large set of empirical moments in panel dataset.

We first illustrate the identification analytically, in a simple model without capital adjustment costs and measurement errors. The key finding is that the parameters governing capital market distortions and the unobserved heterogeneities in output and demand elasticities can be just-identified by the means and between-group dispersions of the revenue-capital and profit-revenue ratios and the correlation between the two ratios. Numerical simulations show that capital adjustment costs and measurement errors merely have second-order effects on these moments. Instead, their effects are essentially manifested in the within-group variations.

[1] Hsieh and Klenow (2009), for instance, show that reducing the magnitude of market distortions in China and India to that in the U.S. would boost total factor productivity by at least 30 and 40 percent in China and India, respectively.

[2] This is also called "the indirect approach" by Restuccia and Rogerson (2013). See their paper for a review of the literature that adopts the approach to assess misallocation.

Therefore, the identification condition on the unobserved heterogeneities, including capital market distortions, carries over to the more general environment where capital adjustment costs and measurement errors are also present. Finally, matching moments on the investment rate and the revenue growth rate pin down capital adjustment costs and the stochastic process of idiosyncratic risks simultaneously, while the within-group standard deviations and the serial correlations discipline measurement errors in the data.

We apply the estimation method to representative firm-level data in Chinese manufacturing. In particular, we focus on a balanced panel from 2004 to 2007 covering 107,579 firms. The estimated heterogeneity in the capital goods price is significant and sizable. Capital market distortions imply aggregate revenue losses of 40 percent in Chinese manufacturing. The magnitude is smaller than that estimated by the ARP approach, but still substantial. We also estimate distortions in samples of U.S. manufacturing firms in Compustat. Improving capital allocation efficiency to the level among the Compustat firms would increase China's manufacturing revenue by 31 percent.

Despite its potential biases, the ARP approach has the virtue of simplicity. This motivates us to propose a generalized ARP approach, which, on the one hand, takes care of the unobserved heterogeneities and, on the other hand, maintains some tractability. To this end, we calibrate the unobserved heterogeneities by solving a nonlinear equation system that matches a set of the moments of the revenue-capital and profit-revenue ratios in a panel. The basic idea is to back out the heterogeneity in the capital goods price from the between-group variation in the revenue-capital ratio, where the effects of capital adjustment costs and measurement errors have largely been eliminated through the time-series average of the revenue-capital ratio within each firm. The generalized ARP approach provides a first-order approximation of the capital goods price heterogeneity estimated from the full-fledged structural approach in the Chinese and Compustat datasets.

It is certainly interesting to understand the policies or institutional arrangements lying hidden behind the veil of distortions. Although firm-specific capital goods prices are not directly observed, both the structural estimation and the generalized ARP approach suggest that the between-group variation of the revenue-capital ratio plays a key role in identifying and estimating capital market distortions. Motivated by this insight, we regress the time-series mean of the revenue-capital ratio of each firm on a set of firm characteristics. We find that in Chinese manufacturing, small, young and non-state firms tend to face significantly higher capital goods prices than their counterparts that are large, mature, and state-owned. Firms located in western and northeastern provinces appear to enjoy lower capital goods prices.

These results are broadly consistent with the findings from a growing literature on financial market imperfections in China (e.g., Dollar and Wei, 2007; Brandt et al., 2013).

The generalized ARP approach also allows us to deliver a rough estimate of labor market distortions without a full specification on labor adjustment costs and measurement errors. The magnitude of labor market distortions turns out to be much smaller than that of capital market distortions in the Chinese manufacturing. A complete removal of labor market distortions would increase the aggregate revenue by merely 4.3 percent. Furthermore, the generalized ARP approach allows us to incorporate a potentially important technological adoption decision that may affect the estimation of capital market distortions. Specifically, we assume that firms can choose between a labor-intensive and a capital-intensive technology at entry. Firms facing higher capital goods prices are more likely to take the labor-intensive industry. We find that the endogeneity of the capital output elasticity accounts for 15 percent of the estimated capital market distortions.

Within the growing literature studying the role of particular distortions, Midrigan and Xu (2012) evaluate the importance of a particular collateral constraint on aggregate productive efficiency. They find a quantitatively small effect through the misallocation channel. The main insight is that self-financing can undo the losses caused by the collateral constraint.[3] Notice that the capital market distortions in our paper are defined in a much broader sense. They entail not only the collateral constraint, but also all kinds of policies and institutional arrangements that may give birth to heterogeneity in the capital goods price at the firm level. Asker et al. (2011) show that much of the cross-sectional dispersion in the average product of capital in developing economies can be caused by the time-series volatility of productivity. We address this concern by exploiting the between-group dispersion.

In terms of estimation, Cooper and Haltiwanger (2006) and Bloom (2009) first adopt SMM to recover structural parameters of capital adjustment costs. They show that it is possible to distinguish the capital adjustment costs from the stochastic process using information on both investment rate and revenue growth rate, which provides an important step for our identification strategy. We contribute to the empirical investment literature by estimating unobserved heterogeneities and measurement errors using a structural approach.

The rest of the paper is organized as follows. Section 2 outlines the model economy with capital adjustment costs and unobserved heterogeneities in production technology and market power. Section 3 presents the empirical specification and discusses the identification conditions.

[3] The importance of credit market imperfections on aggregate productive efficiency is far from being settled, however. See Caselli and Gennaioli (2013), Buera and Shin (2010), Buera et al. (2011) and Moll (2012) for different results.

Section 4 describes the Chinese manufacturing data and reports the main empirical results. The generalized ARP approach is developed and applied in Section 5. Section 6 concludes.

2 The Model

Our analysis is based on a partial equilibrium model with two features. First, capital output elasticity and markups are allowed to differ across firms. Second, firms face heterogeneous capital goods price due to capital market distortions. The model is otherwise standard in the investment literature (e.g., Abel and Eberly, 1994). As usual, we first obtain a static profit function by maximizing instantaneous profits with respect to variable inputs. The intertemporal investment decision is made to maximize the discounted sum of future profits in the presence of capital adjustment costs.

2.1 Production and Demand

Firm i in period t uses productive capital stock, labor and intermediate input, denoted by $\hat{K}_{i,t}$, $L_{i,t}$ and $M_{i,t}$, respectively, to produce $Q_{i,t}$ units of product i. The production technology exhibits constant returns to scale and takes a Cobb-Douglas form:

$$Q_{i,t} = A_{i,t}\hat{K}_{i,t}^{\alpha_i} L_{i,t}^{\beta_i} M_{i,t}^{1-\alpha_i-\beta_i},$$

where $A_{i,t}$ is stochastic, representing randomness in productivity. $\alpha_i > 0$ and $\beta_i > 0$ denote firm-specific capital and labor output elasticities, respectively, with $\alpha_i + \beta_i < 1$.

The product of firm i is demanded in a monopolistic product market according to an isoelastic downward-sloping demand curve,

$$Q_{i,t} = X_{i,t} P_{i,t}^{-\frac{1}{\eta_i}},$$

where $X_{i,t}$ is stochastic, representing randomness in demand. $P_{i,t}$ denotes the price of product i in period t, and $\eta_i \in (0, 1)$ is the inverse of firm-specific demand elasticity with respect to price.

Denote $w_{i,t}$ the wage rate and $m_{i,t}$ the intermediate input price for firm i in period t. For a given productive capital stock $\hat{K}_{i,t}$, firm i chooses variable inputs, $L_{i,t}$ and $M_{i,t}$, optimally to maximize its instantaneous variable profits, denoted by $\pi_{i,t}$:

$$\pi_{i,t} = \max_{L_{i,t},\, M_{i,t}} \{Y_{i,t} - w_{i,t}L_{i,t} - m_{i,t}M_{i,t}\}, \tag{1}$$

where $Y_{i,t} \equiv P_{i,t}Q_{i,t}$ denotes sales revenue.[4] The first-order conditions imply constant intermediate input and labor cost shares:

$$\frac{w_{i,t}L_{i,t}}{Y_{i,t}} = \beta_i(1-\eta_i), \tag{2}$$

$$\frac{m_{i,t}M_{i,t}}{Y_{i,t}} = (1-\alpha_i-\beta_i)(1-\eta_i). \tag{3}$$

The factor shares would reduce to β_i and $1-\alpha_i-\beta_i$ in the competitive environment in which the demand elasticity goes to infinity (i.e., $\eta_i = 0$). Substituting these first-order conditions into (1) yields

$$\frac{\pi_{i,t}}{Y_{i,t}} = \eta_i(1-\alpha_i) + \alpha_i. \tag{4}$$

(4) shows that variable profits are a constant proportion of revenue, which is co-determined by α_i and η_i. In the limiting case of perfect competition, the profit-revenue ratio would reduce to α_i. The Cobb-Douglas production technology guarantees that the labor, intermediate input and profit shares are independent of factor prices.

The optimization also establishes a profit function:

$$\pi_{i,t} = Z_{i,t}^{\gamma_i}\hat{K}_{i,t}^{1-\gamma_i}, \tag{5}$$

where

$$\gamma_i \equiv 1 - \frac{\alpha_i(1-\eta_i)}{\eta_i + \alpha_i(1-\eta_i)}, \tag{6}$$

and

$$Z_{i,t} = \left(\frac{\eta_i}{\gamma_i}\right)^{\frac{1}{\gamma_i}} \left[(1-\eta_i)^{1-\alpha_i}\left(\frac{\beta_i}{w_{i,t}}\right)^{\beta_i}\left(\frac{1-\alpha_i-\beta_i}{m_{i,t}}\right)^{1-\alpha_i-\beta_i}\right]^{\frac{1}{\eta_i}-1} X_{i,t}A_{i,t}^{\frac{1}{\eta_i}-1}.$$

A combination of equations (4), (5) and (6) leads to a revenue function:

$$Y_{i,t} = \frac{\gamma_i}{\eta_i}Z_{i,t}^{\gamma_i}\hat{K}_{i,t}^{1-\gamma_i}. \tag{7}$$

(5) and (7) will be intensively used in the following analysis. Two parameterizations in (5) and (7) are worth mentioning. First, $1-\gamma_i$ captures the firm-specific curvature of the two functions. In the competitive environment where $\eta_i = 0$, both profits and sales become proportional to $\hat{K}_{i,t}$. Second, $Z_{i,t}$ encompasses randomness from productivity, demand and factor prices of variable inputs. Although firm i may know the realization of each of the components and their corresponding stochastic processes, it is ultimately $Z_{i,t}$ that matters for the firm's investment

[4] Following the investment literature (e.g., Abel and Eberly, 1994), we use Q to denote the quantity of output and refer to the product of the price and quantity of output as sales revenue, Y. In the productivity literature (e.g., Hsieh and Klenow, 2009), Y is simply the quantity of output, which is equivalent to Q in our model.

decision. Therefore, $Z_{i,t}$ is a summary statistics of the "profitability" (Cooper and Haltiwanger, 2007) or "business environment" (Bloom, 2009) .

Without loss of generality, we assume that $Z_{i,t}$ follows a trend stationary AR(1) process:

$$\begin{aligned} \log Z_{i,t} &= \mu t + z_{i,t}, \\ z_{i,t} &= \rho z_{i,t-1} + e_{i,t}, \end{aligned} \tag{8}$$

where $0 < \rho < 1$, $e_{i,t} \overset{i.i.d.}{\sim} N(0, \sigma^2)$, and $z_{i,0} = 0$.[5] The standard deviation of the shocks, σ, is the parameter characterizing the level of uncertainty. μ and σ^2 may also be firm-specific. Since our interest is to study how shocks to $\log Z_{i,t}$ affect firms' factor demand decisions, we assume homogeneous μ and σ^2 in the benchmark case. Section 4.5 shows that a relaxation of the assumption will not cause any substantial changes to our main results.

2.2 Capital Market Distortions

There is a long list of factors that may cause capital market distortions. Instead of studying the role of each specific channel, this paper aims to understand the overall effect of all the potential distortions. Similar to Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), we use τ_i to generically capture the effect of various capital market distortions on the purchase price of capital for firm i. The actual capital goods price faced by firm i in period t is, thus, equal to

$$P_{i,t}^K = (1 + \tau_i)\, P_t^K, \tag{9}$$

where P_t^K denotes the average capital goods price in the economy. A positive value of τ_i, for instance, may correspond to a firm that has limited access to external financing and, hence, is subject to a higher than average capital goods price. A negative value of τ_i, on the other hand, may represent an investment tax credit. Denote σ_τ the standard deviation of $\log(1 + \tau_i)$ across firms. Appendix 7.1 shows that σ_τ is a summary statistics of the magnitude of capital market distortions: The aggregate revenue total factor productivity (henceforth TFPR) gains of removing the distortions are proportional to σ_τ^2 in the model economy.[6] The primary goal of our paper is, thus, to estimate σ_τ.

[5] The stochastic process of $Z_{i,t}$ can be endogenously obtained from its definition, if we assume that each of $A_{i,t}$, $X_{i,t}$, $w_{i,t}$ and $m_{i,t}$ follow a similar trend stationary AR(1) process. For (8) to hold, a sufficient condition is that these four random variables share a common level of persistence, ρ, and the shocks to each of these random variables are independent.

[6] See Foster et al. (2008) for the distinction between physical productivity, TFPQ, and revenue productivity, TFPR.

2.3 Capital Adjustment Costs

We introduce capital adjustment costs as a representation of frictions that reduce, delay or protract investment (Khan and Thomas, 2006). These frictions tend to bias the estimates of σ_τ in the ARP approach by affecting the average revenue product of capital (Asker et al., 2011). Following Cooper and Haltiwanger (2006) and Bloom (2009), we consider three forms of capital adjustment costs that are homogeneous across firms:

$$G(K_{i,t}; I_{i,t}) = \frac{b^q}{2}\left(\frac{I_{i,t}}{K_{i,t}}\right)^2 K_{i,t} - b^i P^K_{i,t} I_{i,t} \mathbf{1}_{[I_{i,t}<0]} + b^f \mathbf{1}_{[I_{i,t}\neq 0]} \pi_{i,t},$$

where $K_{i,t}$ denotes the capital stock of firm i at the beginning of period t; $I_{i,t}$ is the new investment of firm i in period t; and $G(K_{i,t}; I_{i,t})$ represents the function of capital adjustment costs, with $\mathbf{1}_{[I_t<0]}$ and $\mathbf{1}_{[I_t\neq 0]}$ being indicators for negative and non-zero investment. Here, b^q measures the magnitude of quadratic adjustment costs. b^i can be interpreted as the difference between the purchase price and the resale price expressed as a percentage of the purchase price of capital goods. Finally, b^f stands for the fraction of variable profit loss due to any non-zero investment.

Notice that our model is disciplined by restricting the capital adjustment cost function, G, to be the same for all firms. If G were also heterogeneous, a firm facing high capital adjustment costs, holding all else equal, would manifest such costs as a high τ_i. A caveat is that G may vary across industries. An auto production line, for instance, is clearly more irreversible than office furniture. Allowing G to differ across industries, however, would give essentially the same estimated σ_τ.[7]

By paying costs of purchasing and adjusting capital, the new investment, $I_{i,t}$, contributes to the productive capital stock, $\hat{K}_{i,t}$, immediately in period t. $\hat{K}_{i,t}$ depreciates at the end of that period. The law of motion for capital is, therefore, given by

$$K_{i,t+1} = (1-\delta)\hat{K}_{i,t} = (1-\delta)\left(K_{i,t} + I_{i,t}\right), \tag{10}$$

where δ is the constant depreciation rate common across firms.

The distinction between $K_{i,t}$ and $\hat{K}_{i,t}$ is motivated for two reasons. First, under the standard timing assumption where $\hat{K}_{i,t} = K_{i,t}$, capital takes one period to build. Therefore, even if capital adjustment costs are absent, idiosyncratic shocks still cause MRPK to differ across firms in the efficient allocation (Asker et al., 2011). In contrast, our timing assumption yields

[7] Specifically, we estimate the model using two subsamples that consist of firms in the ten least and most capital-intensive industries. The manufacturing capital intensity rank follows Song et al. (2011). The results are available upon request.

the same efficient allocation condition of equalized MRPK as the one in a static economy (Hsieh and Klenow, 2009). Such a distinction helps to isolate the effect of capital adjustment costs. Second, technically, with the absence of capital adjustment costs, our timing assumption allows for a closed-form solution to the investment problem, which does not involve any expectation term (Bloom, 2009). This provides a convenient analytical benchmark. Finally, it is reassuring that the average revenue-capital ratio, a key variable for estimating σ_τ, has similar empirical distributions regardless of whether the denominator is $\hat{K}_{i,t}$ or $K_{i,t}$.

2.4 Investment Decision

The presence of capital adjustment costs implies that investment is an intertemporal decision. At the beginning of each period t, optimal investment is chosen to maximize the discounted present value of dividends, which is the variable profit net of investment expenditure and capital adjustment costs. Investors are risk-neutral.[8] They invest such that the required rate of return to capital is equalized across different firms. Future dividends are discounted at the required rate of return, denoted by r. The investment problem is then defined by the stochastic Bellman equation:

$$V(Z_{i,t}, K_{i,t}) = \max_{I_{i,t}} \left\{ \begin{array}{c} \pi(Z_{i,t}, K_{i,t}; I_{i,t}) - P^K_{i,t} I_{i,t} - G(K_{i,t}; I_{i,t}) \\ +\frac{1}{1+r} E_t \left[V(Z_{i,t+1}, K_{i,t+1}) \right] \end{array} \right\}, \tag{11}$$

where $Z_{i,t+1}$ and $K_{i,t+1}$ follow the law of motion (8) and (10), respectively.

In the presence of capital adjustment costs, there is generally no analytical solution to the optimal investment problem. However, it is instructive to begin with the case without adjustment costs, which allows a simple closed-form solution and provides an important benchmark. When $G(Z_{i,t}, K_{i,t}; I_{i,t}) = 0$, the optimal investment rate is a linear function of $Z_{i,t}/K_{i,t}$:

$$\frac{I_{i,t}}{K_{i,t}} = \left[\frac{1-\gamma_i}{(1+\tau_i) J_t} \right]^{\frac{1}{\gamma_i}} \left(\frac{Z_{i,t}}{K_{i,t}} \right) - 1, \tag{12}$$

where J_t stands for the Jorgensonian user cost of capital,

$$J_t \equiv P^K_t - \frac{1-\delta}{1+r} E_t \left[P^K_{t+1} \right]. \tag{13}$$

(12) implies that the optimal investment rate is increasing in $Z_{i,t}$ but decreasing in $(1+\tau_i) J_t$. Intuitively, a firm facing unfavorable capital market distortions ($\tau_i > 0$) invests less than a firm that is facing favorable distortions ($\tau_i < 0$) but otherwise identical.

[8]The risk-neutrality assumption is equivalent to having a complete market without aggregate shocks in which risk-averse investors diversify all idiosyncratic risks. A relaxation of the assumption may cause r to vary across firms in a number of ways. Sections 4.6 and 5.1 will discuss some of the possibilities and how the estimation results would be affected accordingly.

When $G(K_{i,t}; I_{i,t}) > 0$, the investment policy can be solved numerically. Three standard features are worth mentioning. First, regardless of the form of adjustment costs, the optimal investment rate is always a non-decreasing function of $Z_{i,t}/K_{i,t}$. Second, when $b^q > 0$, capital accumulation is through a series of small and continuous adjustments. Finally, the optimal investment rate follows a "barrier control" policy when $b^i > 0$ and a "jump control" policy when $b^f > 0$.

For simplicity, we will assume a constant average capital goods price and normalize it to unity throughout the rest of the paper. Therefore, (13) implies that $J_t = J$, where $J \equiv \frac{r+\delta}{1+r}$. Two remarks are in order. First, a time-varying P_t^K, which may naturally fluctuate over business cycles, would affect the intertemporal investment decision. However, as will be shown below, the time-series variation of capital appears not to be important for our estimation. Moreover, Chinese manufacturing experienced fast but stable growth during the 2004-2007 period, with real GDP growth rates varying modestly between 10.1 percent and 11.9 percent. Consequently, as will be shown below, the aggregate statistics of the key variables are all stable over the sample period.

2.5 The ARP Approach

The primary goal of this paper is to estimate σ_τ. Since τ_i is not directly observable, one has to infer τ_i from observed variables. The ARP approach delivers a contaminated inference, though. The estimation bias can be illustrated analytically in the model without capital adjustment costs. When $G(K_{i,t}; I_{i,t}) = 0$, (11) solves

$$\alpha_i (1 - \eta_i) \frac{Y_{i,t}}{\hat{K}_{i,t}} = (1 + \tau_i) J. \tag{14}$$

The left- and right-hand sides of (14) represent the marginal revenue product of capital and the firm-specific user cost of capital, respectively. Rearranging (14), we have

$$\log \left(\frac{Y_{i,t}}{\hat{K}_{i,t}} \right) = \log J + \log (1 + \tau_i) - \log [\alpha_i (1 - \eta_i)]. \tag{15}$$

(15) is the cornerstone of the ARP approach in the misallocation literature. It highlights how σ_τ is inferred from the observed dispersion of the average revenue product of capital or revenue-capital ratio, $\log \left(Y_{i,t}/\hat{K}_{i,t} \right)$. However, a key challenge in the indirect inference is that, besides capital market distortions, unobserved heterogeneities in α_i and η_i also cause the revenue-capital ratio to differ across firms. As will be shown later, a full-fledged model with capital adjustment costs and measurement errors will further increase the dispersion of the revenue-capital ratio, resulting in an even more biased estimator of σ_τ.

3 Structural Estimation

We now propose a structural econometric approach, aiming to simultaneously recover unobserved heterogeneities in τ_i, α_i and η_i, capital adjustment costs and potential measurement errors in the data. This section starts with the empirical specification and proceeds by presenting the estimation method. The rest of the section will be devoted to the conditions through which the parameters are identified.

3.1 Empirical Specification

Unobserved Heterogeneities. There are three forms of unobserved heterogeneities in the model. For our key interest, we assume that each firm i has a firm-specific τ_i, where $\log(1+\tau_i)$ is drawn independently from an identical normal distribution with mean zero and standard deviation σ_τ:

$$\log(1+\tau_i) \overset{i.i.d}{\sim} N\left(0, \sigma_\tau^2\right). \tag{16}$$

We further assume

$$\log \alpha_i \overset{i.i.d}{\sim} TN\left(\mu_{\log\alpha}, \sigma_{\log\alpha}^2\right), \tag{17}$$

$$\log \eta_i \overset{i.i.d}{\sim} TN\left(\mu_{\log\eta}, \sigma_{\log\eta}^2\right). \tag{18}$$

Here, TN stands for a truncated normal distribution since both α_i and η_i are between 0 and 1. In words, each firm i has a firm-specific α_i and η_i, where $\log\alpha_i$ ($\log\eta_i$) is drawn independently from an identical truncated normal distribution with mean $\mu_{\log\alpha}$ ($\mu_{\log\eta}$) and standard deviation $\sigma_{\log\alpha}$ ($\sigma_{\log\eta}$). Notice that the above specification assumes α_i and η_i to be exogenous. As we will show below, this assumption helps to identify the parameters in (16), (17) and (18). Section 5.3 will discuss the possibility for τ_i to affect α_i. We will also provide a way of correcting the bias caused by the endogeneity of α_i.

The investment policies associated with various $(\tau_i, \alpha_i, \eta_i)$ differ from each other. Hence, we need to solve the dynamic programming (11) at each possible value of $(\tau_i, \alpha_i, \eta_i)$. To reduce the computational burden, this paper adopts a standard approach in the literature (e.g., Eckstein and Wolpin, 1999) by allowing only for a finite type of firms. Specifically, our benchmark specification assumes $3\times3\times3$ types of firms. Each consists of a fixed proportion; i.e., $1/(3\times3\times3)$, of the population. The type set is defined as $\digamma = \{(\tau_u, \alpha_v, \eta_x) : u = 1,2,3; v = 1,2,3; x = 1,2,3\}$. Section 4.5 will experiment with whether the results are robust when increasing the types of firms to $5\times5\times5$ at the cost of "curse of dimensionality."

Measurement Errors. In addition to the rich structure of unobservable heterogeneities, another feature of our empirical specification is to allow for measurement errors. The benchmark specification assumes that

$$K_{i,t} = K_{i,t}^{true}\exp(e_{i,t}^K), \quad e_{i,t}^K \overset{i.i.d}{\sim} N(0,\sigma_{meK}^2), \tag{19}$$

$$Y_{i,t} = Y_{i,t}^{true}\exp(e_{i,t}^Y), \quad e_{i,t}^Y \overset{i.i.d}{\sim} N(0,\sigma_{meY}^2), \tag{20}$$

$$\pi_{i,t} = \pi_{i,t}^{true}(1+e_{i,t}^{\pi}), \quad e_{i,t}^{\pi} \overset{i.i.d}{\sim} N(0,\sigma_{me\pi}^2). \tag{21}$$

Here, variables with and without the "true" superscript denote the true states and their observed counterparts in the data, respectively. $e_{i,t}^K$ and $e_{i,t}^Y$ are measurement errors in capital and revenue, which are drawn independently from an identical normal distribution with mean zero and standard deviation σ_{meK} and σ_{meY}, respectively. $e_{i,t}^{\pi}$ stands for measurement errors in profit, which follow a normal distribution with mean zero and standard deviation $\sigma_{me\pi}$.

Two features are worth mentioning. First, the multiplicative structure and the log-normality assumption guarantee positive values of capital stock and sales revenue. Second, we consider only transitory measurement errors so as to distinguish measurement errors from unobserved heterogeneities. However, abstracting persistent measurement errors in capital may lead to biased estimates of σ_τ. To address this concern, we will model measurement errors in investment and allow $K_{i,t}$ to accumulate the measurement errors according to the law of motion of capital. Section 4.5 below shows that the introduction of the persistent measure errors in capital has little effect on our main findings.

3.2 Simulated Method of Moments

We apply the simulated method of moments (SMM) to estimate the structural model.[9] Specifically, the SMM estimator Θ^* solves the following minimal quadratic distance problem (Gouriéroux and Monfort, 1996):

$$\hat{\Theta}^* = \arg\min_{\Theta} \left(\hat{\Phi}^D - \frac{1}{S}\sum_{s=1}^{S} \hat{\Phi}_s^M(\Theta) \right)' \Omega \left(\hat{\Phi}^D - \frac{1}{S}\sum_{s=1}^{S} \hat{\Phi}_s^M(\Theta) \right), \tag{22}$$

where Θ is the vector of parameters of interest; $\hat{\Phi}^D$ is a set of empirical moments estimated from an empirical dataset; $\hat{\Phi}^M(\Theta)$ is the same set of simulated moments estimated from a

[9] The SMM has been widely employed in the recent empirical investment literature. For example, in addition to Cooper and Haltiwanger (2006) and Bloom (2009), Cooper and Ejarque (2003) and Eberly, Rebelo and Vincent (2008) evaluate the Q-model; Bond, Söderbom and Wu (2008) study the effects of uncertainty on capital accumulation; Schündeln (2006), Henessy and Whited (2007) and Bond, Söderbom and Wu (2007) estimate the cost of financing investment, all through this structural econometric approach.

simulated dataset based on the model; S is the number of simulation paths; and Ω is a positive definite weighting matrix.

Suppose that the empirical dataset is a panel with N firms and T years. We use the asymptotics of fixed T and large N. At the efficient choice for Ω^*, the SMM procedure provides a global specification test of the overidentifying restrictions of the model:

$$
\begin{aligned}
OI &= \frac{NS}{1+S}\left(\hat{\Phi}^D - \frac{1}{S}\sum_{s=1}^{S}\hat{\Phi}_s^M(\Theta)\right)' \Omega^* \left(\hat{\Phi}^D - \frac{1}{S}\sum_{s=1}^{S}\hat{\Phi}_s^M(\Theta)\right) \\
&\sim \chi^2\left[\dim\left(\hat{\Phi}\right) - \dim(\Theta)\right].
\end{aligned}
$$

The upper panel of Table 1 lists Θ, the set of parameters to estimate. There are a total of 13 parameters, including the key parameter σ_τ; mean and standard deviation of $\log\alpha$, $\mu_{\log\alpha}$ and $\sigma_{\log\alpha}$; mean and standard deviation of $\log\eta$, $\mu_{\log\eta}$ and $\sigma_{\log\eta}$; capital adjustment costs parameters, b^q, b^i and b^f; the trend growth rate, μ; standard deviation of idiosyncratic shocks, σ; and standard deviations of measurement errors in capital, revenue and profit, σ_{meK}, σ_{meY}, and $\sigma_{me\pi}$.

[Insert Table 1]

The lower panel of Table 1 lists $\hat{\Phi}^D$, the set of moments to match. There are 21 moments. The choice of the moments is guided by two principles. First, $\hat{\Phi}^D$ is a comprehensive set of moments that characterize the distribution and dynamics of the variables that one would expect to match from a well-specified investment model. Second, and more importantly, these moments are informative about the parameters to estimate. Specifically, $\hat{\Phi}^D$ includes means ($mean$), between-group standard deviations (bsd), within-group standard deviations (wsd), coefficients of skewness ($skew$) and serial correlations ($scorr$) for $\pi_{i,t}/Y_{i,t}$, $\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$, $I_{i,t}/K_{i,t}$ and $\Delta\log Y_{i,t}$, together with the cross correlation ($bcorr$) between the between-group $\pi_{i,t}/Y_{i,t}$ and $\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$. The following section will establish the identification conditions through which Θ can be estimated by matching these moments.

3.3 Identification

For illustrative purposes, we start with a simple model with unobserved heterogeneities only. Capital adjustment costs and measurement errors will be added into the model step-by-step. The simple model allows for a closed-form solution, which helps to analytically establish the conditions for identifying σ_τ. We will show next that these conditions remain to be the core of recovering σ_τ in the full-blown model with capital adjustment costs and measurement errors restored.

All the identification conditions will be examined in a simulated panel of 100,000 firms and 24 years, where we calculate moments using data from the last 4 years. The construction is consistent with the size of a balanced panel from China's industrial survey involving about 100,000 firms over 2004-2007. All the simulations in this section assume that $r = 0.15$, $\delta = 0.05$, $\mu_{\log\alpha} = \mu_{\log\eta} = -2.30$, $\rho = 0.90$, $\mu = 0.05$ and $\sigma = 0.30$. We experiment with various parameter values, and the results reported below turn out to be robust.[10]

3.3.1 Identification of Unobserved Heterogeneities

In the simple model without capital adjustment costs and measurement errors, there are five parameters to estimate: σ_τ, $\mu_{\log\alpha}$, $\sigma_{\log\alpha}$, $\mu_{\log\eta}$ and $\sigma_{\log\eta}$. The identification conditions are established by (4) and (15). The exogeneity of τ_i, α_i and η_i implies that $\mu_{\log\alpha}$ and $\mu_{\log\eta}$ can easily be backed out by targeting the means, $mean\,(\pi/Y)$ and $mean\left(\log\left(Y/\hat{K}\right)\right)$. The identification of σ_τ, $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$ is based on two properties implied by the two equations. First, none of σ_τ, $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$ would have any effect on the within-group standard deviations, $wsd\,(\pi/Y)$ and $wsd\left(\log\left(Y/\hat{K}\right)\right)$, or on the serial correlations, $scorr\,(\pi/Y)$ and $scorr\left(\log\left(Y/\hat{K}\right)\right)$. Only the between-group standard deviations, $bsd\,(\pi/Y)$ and $bsd\left(\log\left(Y/\hat{K}\right)\right)$, vary with σ_τ, $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$. Second, since τ_i, α_i and η_i are uncorrelated, (4) and (15) imply that $bsd\left(\log\left(Y/\hat{K}\right)\right)$ and $bsd\,(\pi/Y)$ are determined by σ_τ, $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$. So, an additional moment is needed to identify σ_τ, $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$. To this end, we introduce the cross correlation, $bcorr\left(\pi/Y, \log\left(Y/\hat{K}\right)\right)$, which follows:

$$
\begin{aligned}
bcorr\left(\pi/Y, \log\left(Y/\hat{K}\right)\right) &\equiv corr\left[\frac{1}{T}\sum_{t=1}^{T}\pi_{i,t}/Y_{i,t}, \frac{1}{T}\sum_{t=1}^{T}\log\left(Y_{i,t}/\hat{K}_{i,t}\right)\right] \\
&\begin{cases} <0, \text{ if } \sigma_{\log\alpha} > 0 \text{ and } \sigma_{\log\eta} = 0 \\ >0, \text{ if } \sigma_{\log\alpha} = 0 \text{ and } \sigma_{\log\eta} > 0 \end{cases}.
\end{aligned} \tag{23}
$$

Intuitively, higher markups increase both the profit-revenue and revenue-capital ratios, while a larger capital output elasticity increases the profit-revenue ratio but decreases the revenue-capital ratio. In extreme cases, if there is no heterogeneity in η (α), the profit-revenue ratio would be negatively (positively) correlated with the revenue-capital ratio.

Table A.1 in Appendix 7.2 illustrates these properties numerically. Column (1) starts with a model with no unobserved heterogeneities. We then add positive σ_τ, $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$, respectively. Column (2) shows that only $bsd\left(\log\left(Y/\hat{K}\right)\right)$ responds to σ_τ. In Column (3), $\sigma_{\log\alpha} > 0$ increases both $bsd\,(\pi/Y)$ and $bsd\left(\log\left(Y/\hat{K}\right)\right)$ and leads to a negative

[10] For instance, the results are qualitatively the same when we impose the parameter values according to the estimates using the Chinese or U.S. firm data.

$bcorr\left(\pi/Y, \log\left(Y/\hat{K}\right)\right)$. $\sigma_{\log\eta} > 0$ leads to a positive $bcorr\left(\pi/Y, \log\left(Y/\hat{K}\right)\right)$ in Column (4), though it has the same effects on $bsd\,(\pi/Y)$ and $bsd\left(\log\left(Y/\hat{K}\right)\right)$ as $\sigma_{\log\alpha} > 0$. The last column lists the moments in the model where all the unobserved heterogeneities are present. A comparison between Columns (2) to (4) also shows that σ_τ and $\sigma_{\log\alpha}$ have a first-order effect on $bsd\left(\log\left(Y/\hat{K}\right)\right)$, while $bsd\left(\log\left(Y/\hat{K}\right)\right)$ does not vary much with $\sigma_{\log\eta}$. One can see the reason from (15). Since $\log(1-\eta_i)$ is close to but bounded at zero, increasing $\sigma_{\log\eta}$ will lead to less variation in $\log(1-\eta_i)$ than the effect of increasing $\sigma_{\log\alpha}$ on $\log\alpha_i$. This property will be used for motivating a reduced-form approach in Section 5.1.

In summary, the five parameters in this simple model are exactly identified by the five core moments. Section 3.3.4 will show that the identification conditions carry over to the full-blown model.

3.3.2 Identification of Capital Adjustment Costs

We next extend the simple model by incorporating capital adjustment costs. Following the routine in the literature (e.g., Bloom, 2009), our identification uses information on both the investment rate, $I_{i,t}/K_{i,t}$, and the revenue growth rate, $\Delta\log Y_{i,t}$, to identify b^q, b^i and b^f. As an illustrative example, we add positive b^q, b^i and b^f, respectively, to Column (1) in Table A.2, which is replicated from Column (5) in Table A.1. Column (5) lists the moments when b^q, b^i and b^f are all positive. Two results are relevant for identification. First, the moments for $I_{i,t}/K_{i,t}$ turn out to be much more sensitive than those for $\Delta\log Y_{i,t}$ in response to changes in capital adjustment costs. This difference distinguishes capital adjustment costs from the stochastic process of $\log Z_{i,t}$. Moreover, one may find that $b^q > 0$ and $b^i > 0$ decrease $wsd\,(I/K)$ and increase $scorr\,(I/K)$, and that $b^i > 0$ and $b^f > 0$ increase $skew\,(I/K)$, while $b^f > 0$ has little effect on $wsd\,(I/K)$ and $scorr\,(I/K)$. These properties distinguish different forms of capital adjustment costs from each other.

3.3.3 Identification of Measurement Errors

We now add measurement errors. The main challenge is how to identify measurement errors in an environment with capital adjustment costs. Once again, let us start with Column (1) in Table A.3, which is replicated from the last column in Table A.2. Columns (2) to (4) in Table A.3 reveal which moments are informative about measurement errors by adding positive σ_{meK}, σ_{meY} and $\sigma_{me\pi}$, respectively. Column (5) reports the moments when σ_{meK}, σ_{meY} and $\sigma_{me\pi}$ are all positive. Specifically, σ_{meK} only affects moments on $\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$ and $I_{i,t}/K_{i,t}$; σ_{meY} only affects moments on $\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$, $\pi_{i,t}/Y_{i,t}$ and $\Delta\log Y_{i,t}$; and $\sigma_{me\pi}$ only affects moments on $\pi_{i,t}/Y_{i,t}$. The three types of measurement errors can, thus, be distinguished from each other.

Furthermore, although capital adjustment costs and measurement errors have qualitatively similar effects on $wsd\left(\log\left(Y/\hat{K}\right)\right)$, their effects may differ on other moments. In particular, $\sigma_{meK} > 0$ and $\sigma_{meY} > 0$ increase $wsd\,(I/K)$ and $wsd\,(\Delta \log Y)$ and reduce $scorr\,(I/K)$ and $scorr\,(\Delta \log Y)$, respectively, while capital adjustment costs have the opposite or no effect on these moments, as shown in Section 3.3.2. These properties allow us to separate measurement errors from capital adjustment costs.

3.3.4 Identification of σ_τ in the Full-Blown Model

Finally, we investigate whether the conditions for identifying σ_τ in the simple model remain valid in the presence of capital adjustment costs and measurement errors. Since $bsd\left(\log\left(Y/\hat{K}\right)\right)$ is key to estimating σ_τ, we compute this moment from simulated economies with different parameter values for unobserved heterogeneities, capital adjustment costs and measurement errors.[11] Panels A and B of Figure 1 plot $bsd\left(\log\left(Y/\hat{K}\right)\right)$ with respect to changes in σ_τ, $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$. In line with the results in Table A.1, σ_τ and $\sigma_{\log\alpha}$ have first-order effects on $bsd\left(\log\left(Y/\hat{K}\right)\right)$, while the effect of $\sigma_{\log\eta}$ is small.

Panel C plots $bsd\left(\log\left(Y/\hat{K}\right)\right)$ with respect to b^q, b^i and b^f. Compared to the variations of $bsd\left(\log\left(Y/\hat{K}\right)\right)$ with σ_τ and $\sigma_{\log\alpha}$ in Panels A and B, the variations of $bsd\left(\log\left(Y/\hat{K}\right)\right)$ with b^q, b^i and b^f have a much smaller magnitude: The three lines in Panel C look almost flat. The reason is simple. Capital adjustment costs, together with shocks to $Z_{i,t}$, cause the revenue-capital ratio to vary over time. Therefore, on the one hand, a larger b^q, b^i or b^f may substantially increase the within-group standard deviation, $wsd\left(\log\left(Y/\hat{K}\right)\right)$; on the other hand, the between-group standard deviation, $bsd\left(\log\left(Y/\hat{K}\right)\right)$, does not vary much with b^q, b^i or b^f. Moreover, one may expect even weaker effects on $bsd\left(\log\left(Y/\hat{K}\right)\right)$ as T gets larger.[12]

That capital adjustment costs have merely second-order effects on $bsd\left(\log\left(Y/\hat{K}\right)\right)$ is an important finding. It highlights the merit of isolating the between-group dispersion in panel data. As Asker et al. (2011) show, capital adjustment costs may explain a big trunk of the overall dispersion of the revenue-capital ratio, casting doubt on the importance of capital market distortions. Our finding suggests a simple way of separating unobserved heterogeneities from capital adjustment costs by matching the between- and within-group dispersions, respectively.

[11] The parameterization in Column (5) of Table A.3 is set as the benchmark.

[12] We experiment with $T = 6, 10$ and 20 and find the slopes of $bsd\left(\log\left(Y/\hat{K}\right)\right)$ with respect to b^q, b^i or b^f to be monotonically decreasing.

Panel D plots $bsd\left(\log\left(Y/\hat{K}\right)\right)$ with respect to σ_{meK} and $\sigma_{me\pi}$.[13] Although Table A.3 shows that measurement errors increase $wsd\left(\log\left(Y/\hat{K}\right)\right)$, their effects on $bsd\left(\log\left(Y/\hat{K}\right)\right)$ have the same order of magnitude as those of capital adjustment costs. This is not surprising since measurement errors are i.i.d by construction. The longer the panel is, the weaker are the effects of measurement errors on the between-group dispersion. We will discuss in Section 4.5 how the results would change if measurement errors in capital entailed a persistent component.

In summary, none of b^q, b^i, b^f, σ_{meK}, σ_{meY} and $\sigma_{me\pi}$ is quantitatively important for $bsd\left(\log\left(Y/\hat{K}\right)\right)$, while σ_τ and $\sigma_{\log\alpha}$ continue to have first-order effects on $bsd\left(\log\left(Y/\hat{K}\right)\right)$. These properties imply that incorporating capital adjustment costs and measurement errors does not invalidate the conditions for identifying σ_τ in the simple model. Furthermore, they will serve as a cornerstone for the generalized ARP approach proposed in Section 5.

4 Empirical Results

4.1 Data

The empirical exercises of this paper are based mainly on the annual firm-level data from the Chinese Industry Survey, which the National Bureau of Statistics has conducted yearly since 1998. The dataset (henceforth, the NBS dataset) includes all industrial firms that are identified as state-owned or as non-state firms with sales revenue above 5 million RMB.[14] We will use a balanced panel from year 2004 to 2007, covering a total of 107,579 firms. Since the number of firms increased by a third in 2004, when an economic census was conducted, a panel with years earlier than 2004 will involve far fewer firms. Appendix 7.3 provides detailed information on how to clean the data and to construct some of the key variables in the model.

Our simulations in the structural estimation assume firms to be around their balanced growth paths. In particular, the estimation is to match moments of the four variables, $\pi_{i,t}/Y_{i,t}$, $\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$, $I_{i,t}/K_{i,t}$ and $\Delta\log Y_{i,t}$. So, we need to check the stationarity of the four variables from a fast-changing economy like China's. Table A.4 in the Appendix reports the annual mean value of each of the four variables in the balanced panel. It is certainly hard to give a formal test, given the fact that the panel has merely four time-series observations. Nevertheless, one can still see that except for the falling investment rate, none of the other three variables features an obvious trend.

[13] The effect of σ_{meY} is identical to that of σ_{meK}.

[14] These firms contribute about 90 percent of the gross industrial output.

4.2 Predetermined Parameters

In addition to the 13 structural parameters listed in Table 1, the depreciation rate, δ, and the discount rate, r, also affect the investment decision through J, the Jorgensonian user cost of capital. We set $\delta = 0.05$ to construct real capital stock data (see Appendix 7.3 for details). The choice of δ is based on the law of motion of capital (10), which implies that

$$\begin{aligned} \log\left(1+\frac{I_{i,t}}{K_{i,t}}\right) &= \triangle \log \hat{K}_{i,t} - \log(1-\delta) \\ &\simeq \triangle \log \hat{K}_{i,t} + \delta. \end{aligned}$$

Bloom (2000) shows that when a firm is on its balanced growth path, the gap between capital stock with and without adjustment costs is bounded. In particular, both $\triangle \log \hat{K}_{i,t}$ and $\Delta \log Y_{i,t}$ will grow at the same rate in the long run. So, the above equation implies that δ should match the difference between $\log(1 + I_{i,t}/K_{i,t})$ and $\Delta \log Y_{i,t}$, which yields $\delta = 0.05$.

Bai, Hsieh and Qian (2006) find a high and rather stable aggregate rate of return to capital in China over the period 1978-2004, ranging from 20 percent to 25 percent in most years. The rate of return is even higher for the secondary sector, which includes mining, manufacturing and construction. We impose a conservative value, $r = 0.20$, for manufacturing firms in the sample.

We set $\rho = 0.90$ in the benchmark case. Following Cooper and Haltiwanger (2006), one may calibrate ρ by applying system GMM (Blundell and Bond, 1998) to estimate a dynamic panel data model of $\log \pi_{i,t}$. The regressors include $\log \pi_{i,t-1}$, $\log \hat{K}_{i,t}$, $\log \hat{K}_{i,t-1}$ and year dummies. The estimated autoregressive coefficient is 0.41, in contrast to 0.89 in Cooper and Haltiwanger (2006). The substantially lower estimate for China may reflect the attenuation bias due to large measurement errors in the profit data, which will be confirmed by our structural estimation. We will investigate the sensitivity of our estimates to different values of δ, r and ρ in Section 4.5.

4.3 Structurally Estimated Parameters

Table 2 presents our structural estimation results.[15] The first and second columns of the left panel report the optimal estimates and the corresponding numerical standard errors. Simulated moments at the optimal estimates are listed in the right panel. We also report the corresponding empirical moments, for which the standard errors are obtained by bootstrapping.

[15] Wu (2009) reports the technical details on how to solve the minimal quadratic distance problem of (22), to draw optimal weighting matrix from the data and to calculate the numerical standard errors for the estimates.

[Insert Table 2]

σ_τ has an estimated value of 0.71, which is significantly different from zero. According to (16), $\sigma_\tau = 0.71$ implies that a firm with τ_i at the 75th percentile would face a capital goods price 1.6 times higher than the price for a firm with τ_i at the 25th percentile. Notice that the significant and sizable estimate of σ_τ is obtained after we take care of heterogeneous capital output elasticities and markups, capital adjustment costs and measurement errors. We will discuss below the extent to which the result would change by omitting any of the sources of potential estimation bias.

The estimates of $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$ are also significant and quantitatively large, supporting the presence of heterogeneities in α_i and η_i. Under the log-normality specification (17), the estimated $\mu_{\log\alpha}$ and $\sigma_{\log\alpha}$ imply that α_i has a mean of 0.086 and a standard deviation of 0.052.[16] By (18), the estimated $\mu_{\log\eta}$ and $\sigma_{\log\eta}$ imply that η_i has a mean of 0.078 and a standard deviation of 0.065. Overall, the simulated moments provide a close fit to the five core moments, which are key to identifying the unobserved heterogeneities, as discussed in Section 3.3.1.

The structural estimation finds evidence for quadratic and fixed adjustment costs.[17] According to the identification conditions in Section 3.3.2, a positive b^q reflects a positive serial correlation between the investment rate and revenue growth rate, while a positive b^f shows a larger skewness of the investment rate than that of the revenue growth rate. The point estimates imply that the presence of quadratic adjustment costs would increase the user cost of capital by 4.5 percent and any investment or disinvestment would cause a loss of 3.4 percent of variable profits in that period.

The estimated μ is 0.08, in line with Brandt et al.'s (2012) estimate of the TFP growth in Chinese manufacturing over 1998-2007. The result is a compromise between a higher investment growth rate and a lower revenue growth rate in the simulated economy, compared with

[16] Both the mean and standard deviation values of α_i are close to those in the literature that estimates the capital output elasticity in a three-factor model. For example, Jorgenson, Gollop and Fraumeni (1987) estimate capital output elasticities in 28 U.S. manufacturing industries by production function regression over intermediate input, labor input and capital input. They find that the capital share estimate varies from 0.0486 (apparel and other fabricated textile products) to 0.333 (tobacco), with a mean at 0.098 (electric machinery and equipment supplies). Such estimates of α should be distinguished from those in an aggregate value-added production function with capital and labor inputs only. In the same paper, using a two-factor production function, Jorgenson et al. find an aggregate capital share of 0.385 for the U.S. economy.

[17] Similar to Cooper and Haltiwanger (2006) and Bloom (2009), we also find that only one form of the non-convex adjustment costs is necessary to fit the data. To be specific, Cooper and Haltiwanger (2006) find $b^q > 0$ and $b^f > 0$ for plants in the Longitudinal Research Database; Bloom (2009) finds $b^q > 0$ and $b^i > 0$ for large firms in Compustat. Consistent with the fact that 90 percent firms in our sample are reported to be single-plant enterprises, we find that a combination of $b^q > 0$ and $b^f > 0$ fits the data best, as Cooper and Haltiwanger (2006) do.

those in the data. σ has an estimated value of 0.42, suggesting a high level of uncertainty for Chinese firms.

Two of the three measurement errors we consider turn out to be statistically significant. Consistent with the usual concern about the accuracy of capital and profit data at the firm level, both σ_{meK} and $\sigma_{me\pi}$ are significant and quantitatively large. In contrast, the model finds σ_{meY} to be virtually zero, implying a much better measurement of revenue in the NBS data. As we will show below, measurement errors in capital and profits substantially improve the match of the within-group standard deviations.

4.4 Specification Tests

Our structural approach has an edge over the ARP approach by taking into account unobserved heterogeneities in α_i and η_i, capital adjustment costs and measurement errors. Omitting any of the three components may bias the estimate of σ_τ. To evaluate the importance of each of the three components, Table 3 reports specification tests for three restricted models. The full-blown model is taken as the benchmark, with estimation results listed in Column (1).

[Insert Table 3]

Column (2) reports the results with homogeneous α_i and η_i – i.e., $\sigma_{\log\alpha} = \sigma_{\log\eta} = 0$. The estimated σ_τ increases from 0.706 to 0.924, implying a bias of 31 percent by omitting the unobserved heterogeneities. Moreover, the model fails to match the data in a number of aspects, including all moments of the profit-revenue ratio (except for the mean) and the correlation between the revenue-capital and profit-revenue ratio. As a result, the overall fitness of the restricted model degenerates substantially.

Column (3) reports the results with no capital adjustment costs – i.e., $b^q = b^i = b^f = 0$. The estimate for σ_τ is just 7 percent lower than the benchmark result. For reasons discussed above, the unobserved heterogeneities can essentially be identified by the five core moments, on which capital adjustment costs have little impact. Nevertheless, without capital adjustment costs, the model cannot match the positive serial correlation in the investment rate and revenue growth.

Column (4) reports the results with no measurement errors – i.e., $\sigma_{meK} = \sigma_{meY} = \sigma_{me\pi} = 0$. The estimate for σ_τ is very close to the benchmark result, with a difference of 3 percent. Like capital adjustment costs, measurement errors have only second-order effects on the between-group standard deviations. Consequently, the estimation of the unobserved heterogeneities is largely unaffected by measurement errors. Regarding the fitness, the restricted model generates too small within-group standard deviations of the profit-revenue and revenue-capital ratios and too much serial correlation in these two ratios.

4.5 Robustness Tests

Table 4 presents results for a set of robustness checks. Column (1) corresponds to the benchmark model, where $\delta = 0.05$, $r = 0.20$ and $\rho = 0.90$. Columns (2) and (3) show the results with $\delta = 0.03$ and 0.07, respectively. Columns (4) and (5) test the sensitivity to the discount factor by imposing $r = 0.15$ and 0.25, respectively. Columns (6) and (7) report the results with $\rho = 0.85$ and 0.95, respectively. Overall, we see only some modest variations in the estimates. In particular, the estimated σ_τ, ranging from 0.67 to 0.75, appears to be robust to the alternative choices of δ, r and ρ.

[Insert Table 4 here]

Column (8) increases the number of type in each dimension of heterogeneity from 3 to 5. The alternative simulation specification triples the estimation time but causes virtually no change in any of the estimates.

Columns (9) and (10) allow the long-run growth rate of $Z_{i,t}$, μ, and the level of uncertainty, σ, to be firm-specific. Specifically, in Column (9), μ_i follows a normal distribution with mean μ and standard deviation σ_μ, while σ_i follows a normal distribution with mean σ and standard deviation σ_σ in Column (10). Introducing an additional dimension of heterogeneity involves an additional state variable. The estimation time increases by 2.5 times accordingly. Not surprisingly, allowing more heterogeneities improves the overall fitness. The estimate of σ_τ, however, is almost unaffected.

Column (11) replaces measurement errors in capital with measurement errors in investment. The alternative specification implies a persistent effect on the measurement of capital through capital accumulation. We find much larger capital adjustment costs. The estimated σ_τ, however, increases little, by less than 3 percent.

4.6 Heterogeneities in J and the Compustat Benchmark

One caveat with our empirical strategy is that any heterogeneity in J, the Jorgensonian user cost of capital, would show up as capital market distortions. Since $J = \frac{r+\delta}{1+r}$, any intrinsic heterogeneity in δ or r, if it exists, may bias the estimate of σ_τ upwards.

Heterogeneous δ appears if firms have different combinations of plant and equipment that depreciate at different rates. r would also be firm-specific if we relax the assumption of risk-neutrality. Through a standard market beta channel, for instance, risk-averse investors would assign higher discount rates to firms with $Z_{i,t}$ that are more correlated to aggregate shocks. It is obviously hard to tease out all the potentially important heterogeneities in J. Appendix 7.4 provides some back-of-the-envelope calculations, suggesting a limited role of market beta in explaining the estimated σ_τ.[18]

As an alternative approach, we estimate σ_τ for U.S. manufacturing firms in Compustat and then use the estimate to control the effect of heterogeneous J. The logic is as follows. These U.S. Compustat firms are arguably least likely to be affected by capital market distortions. As long as the heterogeneity in J has similar magnitudes across Chinese and U.S. firms, the difference between the estimated σ_τ in China and the U.S. would isolate the magnitude of capital market distortions in Chinese manufacturing.

We then estimate the same structural model by two Compustat samples over 2002-2005.[19] The first one is a sample of U.S. manufacturing firms with sales revenue above 0.6 million U.S. dollars in 2004 prices, referred to as Compustat I henceforth. Compustat I is broadly comparable to the NBS sample, in which most firms have sales above 5 million RMB. The second sample follows Bloom (2009) by including U.S. manufacturing firms with sales revenue above 10 million U.S. dollars in 2000 prices and more than 500 employees, referred to as Compustat II henceforth. This is, therefore, a more homogeneous sample composed of large firms only.

We set $r = 0.10$ since the required rate of returns to capital tends to be much lower for Compustat firms.[20] The values of δ and ρ remain unchanged. Table 5 reports the results. As expected, the estimated σ_τ in Compustat I is 0.30, much smaller than its counterpart of 0.71 in China. The estimate becomes even smaller and turns statistically insignificant in Compustat II. Three remarks are in order. First, the estimates of σ_τ are consistent with the impressionistic view that China has a more distorted capital market. The small estimated σ_τ by Compustat II is also in line with the findings that larger firms are less likely to be affected by financial market imperfections. Second, one may want to take the estimated σ_τ in Compustat I as a proxy for the upper bound of the magnitude of heterogeneity in J. The difference between the estimated

[18] Market incompleteness may also generate heterogeneous r across firms with different levels of uncertainty (e.g., Angeletos and Panousi, 2011). This can be considered one type of financial market imperfection. A back-of-the-envelope calculation in Appendix 7.4 suggests that market incompleteness accounts for a fourth of the estimated capital market distortions in Chinese manufacturing.

[19] We construct capital stock and deflate the data strictly following Bloom (2009). To be specific, capital stocks for firm i in industry m in year t are constructed by the perpetual inventory method: $K_{i,t} = (1-\delta) K_{i,t-1} (P_{m,t}/P_{m,t-1}) + I_{i,t}$, initialized using the net book value of capital, where $I_{i,t}$ is net capital expenditures on plant, property, and equipment, and $P_{m,t}$ is the industry-level capital goods deflator from Bartelsman et al. (2000). Sales revenue and cost of goods sold are deflated by the CPI. We use a sample from 2002 to 2005 since $P_{m,t}$ is not available after 2005.

[20] This is consistent with the fact that Compustat samples have much lower revenue-capital ratios than the NBS sample (see Table 5 for details).

σ_τ in the NBS data and Compustat I can, thus, be interpreted as the lower-bound estimate of capital market distortions in Chinese manufacturing. Third, the statistically insignificant σ_τ in Compustat II suggests that neither capital market distortions nor heterogeneity in J appears to be important for large U.S. manufacturing firms.

[Insert Table 5]

It is worth mentioning that Table 5 finds a much larger quadratic capital adjustment cost for Compustat firms, which is driven by the considerably smaller dispersion, less skewness and more persistence in the investment rate in Compustat firms (see the empirical moments in Tables 2 and 5). Not surprisingly, Compustat firms are, on average, much larger. For instance, firms in Compustat I have mean (median) employees of 10.4 (0.8) thousand; in contrast, the corresponding numbers are only 0.33 (0.13) for Chinese firms in the NBS sample. Therefore, one explanation for the larger quadratic capital adjustment costs in Compustat firms is that their investment has been consolidated across several plants within the firm.[21]

4.7 Counterfactual Simulations

The structural model provides a useful framework to quantify the effect of the estimated σ_τ on the aggregate TFPR. Specifically, (30) in Appendix 7.1 allows us to calculate such effect for each type of firms with different α_i and η_i. The aggregate result is thus obtained by averaging the effects across different types of firms.

Consider the model with parameter values set by the estimates in Table 2. Removing capital market distortions by imposing zero σ_τ would increase the aggregate TFPR by 38.2 percent.[22] Three remarks are in order. First, a similar exercise shows the importance of modeling heterogeneous α_i and η_i. Using the estimates in Column (2) of Table 3, we find that if one omits the heterogeneities in α_i and η_i, the gain from removing capital market distortions would be increased by two thirds.[23] Second, Hsieh and Klenow (2009) estimate aggregate TFPR losses caused by distortions in both product and capital markets, while the present paper aims to estimate the effect of capital market distortions alone. One may want to interpret the heterogeneity in markups as product market distortions since the heterogeneity in η_i is isomorphic to the heterogeneity in τ_i^Y – the measure of product market distortions in

[21] See Bloom (2009) on the reason why quadratic adjustment costs and aggregation smoothing are two alternative mechanisms in generating condensed, symmetric and persistent investment rate series.

[22] In contrast, removing capital adjustment costs by imposing zero b^q, b^i and b^f would increase the aggregate TFPR by merely 2.1 percent. This is in the same order of magnitude as the effect of capital adjustment costs estimated by Midrigan and Xu (2009).

[23] Recall that according to (31) in Appendix 7.1, the aggregate TFPR losses are proportional to the variance (rather than the standard deviation) of τ_i.

Hsieh and Klenow (2009). Then, our model would predict an aggregate TFPR gain of 53.2 percent by removing both product and capital market distortions. Finally, if we reduce σ_τ to the estimate in the first Compustat sample, which can be taken as a proxy for the magnitude of heterogeneity in J, our model would still generate an aggregate TFPR gain of 31.4 percent.

5 The Generalized ARP Approach

Our structural approach provides a less biased estimator of capital market distortions by teasing out the effects of some unobserved heterogeneities, capital adjustment costs and measurement errors. Nevertheless, the structural estimation is a lot more complicated and involves high computational costs. This section proposes a generalized ARP approach, which extends the ARP approach along two dimensions. First, while the ARP approach uses cross-sectional data only, the generalized ARP approach will explore between-group dispersions and correlations in panel data. This helps to eliminate much of the potential bias caused by capital adjustment costs and measurement errors, for reasons discussed in Sections 3.3.2 and 3.3.3. Second, the ARP approach estimates σ_τ by matching the dispersion of the revenue-capital ratio. The generalized ARP approach will, instead, explore first and second moments of both the revenue-capital and profit-revenue ratios. The goal is to control the unobserved heterogeneities, as illustrated in Section 3.3.1.

Specifically, the generalized ARP approach uses the five core moments to back out the five parameters governing the unobserved heterogeneities: σ_τ, $\mu_{\log\alpha}$, $\mu_{\log\eta}$, $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$. In summary, the estimation of σ_τ solves a nonlinear equation system:

$$\Theta = F\left(\mu_{\log\alpha}, \mu_{\log\eta}, \sigma_{\log\alpha}, \sigma_{\log\eta}, \sigma_\tau\right),$$

where $\Theta = [mean\,(\pi/Y)\,, mean\left(\log\left(Y/\hat{K}\right)\right), bsd\,(\pi/Y)\,, bsd\left(\log\left(Y/\hat{K}\right)\right), bcorr\left(\pi/Y,\ \log\left(Y/\hat{K}\right)\right)]'$, and F denotes a function based on (4) and (15) that returns a vector value. Notice that F is built upon the simple model in Section 3.3.1. Omitting capital adjustment costs and measurement errors makes the generalized ARP approach much more tractable. Most importantly, it will not affect the estimation results much since capital adjustment costs and measurement errors have only second-order effects on the five core moments.

Table 6 compares the estimates from the structural estimation and the generalized ARP approach in all the three samples. Notably, the two approaches generate very close estimates of σ_τ in both the NBS data and Compustat I, where the structural estimation finds statistically significant σ_τ. The estimate of σ_τ is essentially equal to zero when applying the generalized ARP approach for Compustat II. This is also consistent with the statistically insignificant estimate found by the structural approach. Moreover, most of the estimates of the other four parameters by the generalized ARP approach are also close to those estimated by the structural estimation. The only two exceptions are $\mu_{\log\eta}$ and $\sigma_{\log\eta}$ for Compustat II. In terms of computational costs, the benchmark structural estimation typically takes hours to converge, while the generalized ARP approach delivers results within a few seconds.

[Insert Table 6]

5.1 Regressions on Firm Characteristics

The applicability of the generalized ARP approach motivates a reduced-form regression that helps us understand how τ_i is correlated with firm characteristics. Notice that $bsd\left(\log\left(Y/\hat{K}\right)\right)$ is the only moment informative for σ_τ. All the other four moments in the generalized ARP approach are deployed to tease out the effects of heterogeneities in α_i and η_i on $bsd\left(\log\left(Y/\hat{K}\right)\right)$. In other words, after controlling for heterogeneities in α_i and η_i, we may interpret the average revenue-capital ratio, $\frac{1}{T}\sum_{t=1}^{T}\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$, as a proxy for the firm-specific capital goods price or, equivalently, τ_i. This suggests a simple regression of $\frac{1}{T}\sum_{t=1}^{T}\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$ on firm characteristics. We add four-digit industry dummies and the average profit-revenue ratio, $\frac{1}{T}\sum_{t=1}^{T}\log\left(\pi_{i,t}/Y_{i,t}\right)$, to the control variables. The idea is that industry dummies control heterogeneities in α_i and η_i across industries, and the average profit-revenue ratio takes care of the within-industry heterogeneities.[24] Specifically, we run the following regression:

$$\frac{1}{T}\sum_{t=1}^{T}\log\left(Y_{i,t}/\hat{K}_{i,t}\right) = b_0 + b_1\cdot\frac{1}{T}\sum_{t=1}^{T}\log\left(\pi_{i,t}/Y_{i,t}\right) + b_2\cdot D_i + b_3\cdot X_i + \xi_i, \tag{24}$$

where D_i represents a vector of industry dummies and X_i is a vector of firm characteristics. The parameter of interest is b_3, which captures the relationship between the revenue-capital ratio and firm characteristics through the channel of capital market distortions. Such a relationship is hard to reveal with structural models.

Table 7 presents the regression results. The baseline model uses firm age and size as X_i. All else being equal, it predicts that the capital goods price of a firm is 3 percent lower if a firm is one year older, and 4 percent lower if a firm has 1000 more employees. This is consistent with the typical finding in the large literature on capital market imperfections.[25] Quantitatively,

[24] Since the profit-revenue ratio is co-determined by α_i and η_i, it cannot control both of the within-industry heterogeneities in α_i and η_i. However, since $bsd\left(\log\left(Y/\hat{K}\right)\right)$ is not sensitive to $\sigma_{\log\eta}$, one can safely omit the heterogeneity in η_i in the regression aiming to illustrate the correlation between τ_i and firm characteristics.

[25] For instance, among many others, Fazzari, Hubbard and Peterson (1988) show that younger and smaller firms tend to face a higher user cost of capital due to financial constraints. See, also, Hadlock and Pierce (2010) for more recent evidence.

the point estimates suggest that age and employment can account for 24 percent and 6 percent of the observed $bsd\left(\log\left(Y/\hat{K}\right)\right)$, respectively.[26]

[Insert Table 7]

The regional disparity of capital returns documented by Brandt et al. (2013) indicates a role of firm location. NBS classifies all 31 provinces into four regions: eastern, central, western and northeastern. We take the eastern region as the reference group. Dummy variables for the central (CENTRAL), western (WESTERN) and northeastern (NORTHEASTERN) regions are added to X_i in the second regression. Consistent with Brandt et al. (2013), we find significant and large coefficients for the western and northeastern provinces, suggesting that the state government has been heavily subsidizing firms in these regions. The capital goods price for firms in northeastern provinces, for instance, appears to be a quarter lower than that for firms in the eastern provinces. Overall, the point estimates suggest that the geographic location can account for 9 percent of the observed $bsd\left(\log\left(Y/\hat{K}\right)\right)$.[27]

There is a growing literature on heterogeneous financing costs across firms with different ownership in China. State firms often have much better access to external financing than non-state firms (e.g., Dollar and Wei, 2007; Song et al., 2011). To check the role of ownership, we take state-owned enterprises (SOE) as the reference group. Dummy variables for collective-owned enterprises (COE), domestic private enterprises (DPE), foreign-invested enterprises (FIE) and other ownership types (OTHERS) are included in the third regression.[28] All the ownership dummies turn out to be positive and highly significant, suggesting lower capital goods prices for state firms. Specifically, the capital goods prices of COE, DPE, FIE and "others" are 50 percent, 45 percent, 10 percent and 26 percent higher, respectively, than that of SOE. The point estimates also indicate that the ownership heterogeneity can account for 18 percent of the observed $bsd\left(\log\left(Y/\hat{K}\right)\right)$.[29]

5.2 Labor Market Distortions

The generalized ARP approach also motivates a simple way of estimating the magnitude of labor market distortions, which we have not addressed. Thanks to the Cobb-Douglas production function, the estimation of σ_τ is independent of labor market distortions. Nevertheless,

[26] The standard deviation of age and employment is equal to 6.92 and 1.15, respectively.

[27] 73.6, 12.4, 8.7 and 5.4 percent of firms are located in eastern, central, western and northeastern regions, respectively.

[28] The category of "OTHERS" includes cooperative units, joint ownership units, limited liability corporations and share-holding corporations Ltd.

[29] SOE, COE, DPE, FIE and "others" account for 4.6, 7.3, 41.1, 23.8 and 23.3 percent of all firms in the NBS data, respectively.

it is interesting to see which of two factor markets is more distorted in China. The result below suggests that China has a much more distorted capital market and, therefore, provides a justification for our paper to focus on the capital market.

Analogous to (9), we assume that the actual wage rate by firm i in period t is

$$w_{i,t} = \left(1 + \tau_i^L\right) w_t, \tag{25}$$

where w_t is the average wage rate and τ_i^L is a firm-specific component caused by labor market distortions. $\log\left(1 + \tau_i^L\right)$ follows a normal distribution with mean zero and standard deviation σ_{τ^L}. Substituting (25) back into (2) yields

$$\log \frac{Y_{i,t}}{L_{i,t}} = \log w_t + \log\left(1 + \tau_i^L\right) - \log\left[\beta_i \left(1 - \eta_i\right)\right]. \tag{26}$$

(26) is akin to (15). By (2), $bsd\left(\log\left(\beta\left(1 - \eta\right)\right)\right)$ can be obtained by computing the between-group standard deviation of the labor income share.[30] Therefore, a combination of (26) and (2) implies a simple way of backing out σ_{τ^L}. For the following two reasons, we refer to this method as the generalized ARP approach for the estimation of labor market distortions. First, it explores the between-group dispersion of the revenue-labor ratio in panel data, which might mitigate the effects of potential labor adjustment costs and measurement errors. Second, like the estimation of capital market distortions, the estimation of labor market distortions is affected by unobservable heterogeneities other than distorted factor prices. When it comes to labor market distortions, the estimation is less involved simply because we can directly use the labor income share to proxy $\beta_i \left(1 - \eta_i\right)$.[31]

An important concern is that labor quality differs across firms. We then construct efficency units of labor as follows:

$$L_{i,t}^e = \sum_h \exp\left(b \cdot s\left(h\right)\right) L_{i,t}\left(h\right),$$

where $L_{i,t}^e$ stands for efficiency units of labor in firm i at period t; b is the Mincerian rate of return to education; $s\left(h\right)$ denotes years of schooling for education group h; and $L_{i,t}\left(h\right)$ is the number of workers in education group h. We set $b = 10$ *percent* and let $s\left(h\right)$ be 6, 9, 12 and 16 for primary-school-and-below, middle-school, high-school and college-and-above educated

[30] The ratio of labor compensations (including benefits) to sales is unusually low (around 0.08) in the NBS data. One important reason is the severely under-reported labor compensations (Qian and Zhu, 2011). To address this concern, we use the difference between costs of goods sold and material costs to back out the actual labor compensations. This leads to a labor share of 0.125, averaged over the sample period. Recall that our point estimates in Table (2) imply a capital share of 0.13. Therefore, the labor share is roughly equal to the capital share in Chinese manufacturing, which is in line with the aggregate statistics and the finding in Qian and Zhu (2011).

[31] Notice that we cannot infer $\alpha_i \left(1 - \eta_i\right)$ from one minus the sum of the intermediate input and labor shares. The latter involves profits and, thus, equals $\alpha_i \left(1 - \eta_i\right) + \eta_i$. See also (4).

employees, respectively.[32] Since the firm-level data on educational composition of employees are available only in the 2004 NBS sample, we assume a constant educational composition in each firm over time; i.e., $L_{i,t}(h)/L_{i,t} = L_{i,2004}(h)/L_{i,2004}$ for $t > 2004$.

Using the raw labor data yields an estimated σ_{τ^L} of 0.289, while using efficiency units of labor reduces the estimated σ_{τ^L} to 0.230. The difference implies a non-trivial heterogeneity in the educational composition across firms in the Chinese manufacturing. Notice that both estimated values of σ_{τ^L} are substantially smaller than the estimates of capital market distortions. Accordingly, a removal of labor market distortions by reducing σ_{τ^L} from 0.23 to zero would increase the aggregate TFPR by merely 4.3 percent.[33]

5.3 Endogenous Capital Output Elasticity

Our empirical specification assumes α_i to be uncorrelated with τ_i. This assumption does not seem innocuous if firms can choose α_i. Song et al. (2011), for instance, show that in a two-sector model, financially constrained firms would enter the labor-intensive industry, while financially integrated firms choose to stay in the capital-intensive industry. They also find evidence from Chinese manufacturing that seems consistent with the theory. In the context of our model economy, the above findings imply a negative correlation between α_i and τ_i, which may bias the estimate of σ_τ. To address the issue, this section first develops a model that incorporates a technological choice on α_i. Then, we show a simple way of estimating σ_τ by adapting the generalized ARP approach to the new environment.

Assume that, before entering the market, each firm must make an irreversible choice on $\alpha_i \in \{\alpha_l, \alpha_h\}$, with $\alpha^l < \alpha^h$. Formally,

$$\alpha_i^* = \arg \max_{\alpha_i \in \{\alpha_l, \alpha_h\}} E\left[V\left(\alpha_i\right)\right], \tag{27}$$

where $E\left[V\left(\alpha_i\right)\right]$ stands for the ex-ante expected firm value conditional on the technological choice of α_i. In the simple model without capital adjustment costs, $E\left[V\left(\alpha_i\right)\right]$ solves a static optimization problem according to (11). Dropping the irrelevant time subscript t and using (5), (12) and the facts that $I_i = \delta \hat{K}_i$, we have

$$\begin{aligned} E\left[V\left(\alpha_i\right)\right] &= \frac{1+r}{r} E\left[Z_i\left(\alpha_i\right)\gamma_i \hat{K}_i^{1-\gamma_i(\alpha_i)} - P_i^K I_i\right] \\ &= \frac{1+r}{r}\left[1 - \frac{\delta\left(1-\gamma_i\left(\alpha_i\right)\right)}{J}\right]\left[\frac{1-\gamma_i\left(\alpha_i\right)}{\left(1+\tau_i\right)J}\right]^{\frac{1}{\gamma_i(\alpha_i)}-1} E\left[Z_i\left(\alpha_i\right)\right], \end{aligned} \tag{28}$$

[32] Zhang et al. (2005) estimated returns to education in urban areas of six provinces from 1988 to 2001. The returns had an averaged 10.3 percent in 2001.

[33] See Appendix 7.1 for the formula of calculating the aggregate TFPR losses caused by labor market distortions.

where

$$\gamma_i(\alpha_i) = 1 - \frac{\alpha_i(1-\eta_i)}{\eta_i + \alpha_i(1-\eta_i)},$$

$$E[Z_i(\alpha_i)] = \left[\frac{\eta_i}{\gamma_i(\alpha_i)}\right]^{\frac{1}{\gamma_i(\alpha_i)}} \left[(1-\eta_i)^{1-\alpha_i}\left(\frac{1-\alpha_i-\beta_i}{m_i}\right)^{1-\alpha_i-\beta_i}\right]^{\frac{1}{\eta_i}-1} Z,$$

and Z, independent of α_i, is an unimportant constant.[34] Since $\gamma_i(\alpha_l) > \gamma_i(\alpha_h)$, substituting (28) back into (27) yields

$$\alpha_i^* = \begin{cases} \alpha_l & \text{if } \tau_i > \Pi(\alpha_l, \alpha_h) \\ \\ \alpha_h & \text{if } \tau_i < \Pi(\alpha_l, \alpha_h) \end{cases}, \tag{29}$$

where

$$\Pi(\alpha_l, \alpha_h) = \frac{1}{J}\left(\frac{J-\delta(1-\gamma_i(\alpha_h))}{J-\delta(1-\gamma_i(\alpha_l))} \frac{[1-\gamma_i(\alpha_h)]^{\frac{1}{\gamma_i(\alpha_h)}-1}}{[1-\gamma_i(\alpha_l)]^{\frac{1}{\gamma_i(\alpha_l)}-1}} \frac{E[Z_i(\alpha_h)]}{E[Z_i(\alpha_l)]}\right)^{\frac{\gamma_i(\alpha_h)\gamma_i(\alpha_l)}{\gamma_i(\alpha_l)-\gamma_i(\alpha_h)}} - 1.$$

When $\tau_i = \Pi(\alpha_l, \alpha_h)$ – i.e., firms are indifferent between α_l and α_h – we assume that α_i will be chosen in a random way. (29) shows that a firm would optimally choose the more capital-intensive technology if τ_i is sufficiently low.

5.3.1 Estimation

We maintain the assumption that $\log\eta_i \overset{i.i.d}{\sim} TN\left(\mu_{\log\eta}, \sigma^2_{\log\eta}\right)$ and $\log(1+\tau_i) \overset{i.i.d}{\sim} N\left(0, \sigma^2_\tau\right)$. $r = 0.20$ and $\delta = 0.05$ are imposed for the same reasons explained in Section 4.2. For simplicity, we assume no heterogeneity in β_i and m_i such that $\beta_i = \beta$ and $m_i = m$. So, there are seven parameters, α_l, α_h, β, m, $\mu_{\log\eta}, \sigma_{\log\eta}$ and σ_τ, left to be estimated.

For any given β and m, the generalized ARP approach can be directly applied to estimate α_l, α_h, $\mu_{\log\eta}, \sigma_{\log\eta}$ and σ_τ by matching the five core moments. The only difference is that $\mu_{\log\alpha}$ and $\sigma_{\log\alpha}$ are now replaced with α_l and α_h. We calibrate β to match the aggregate labor income share in the NBS data. To pin down m, we assume that half of the firms would choose α_h in the absence of capital market distortions (i.e., $\tau_i = 0$). As will be shown below, the estimation turns out to be insensitive to the distribution of α_i in the distortionless environment.

The benchmark results are reported in the first row of Table 8. Compared with those in Table 6, one may find that the technological choice model yields a lower σ_τ (0.68 vs. 0.58). In other words, the endogeneity of the capital output elasticity accounts for 15 percent of

[34] Here, we let β_i be independent of α_i. Alternatively, under the assumption of constant returns to scale, we may let $1-\alpha_i-\beta_i$ be indepdendent of α_i. Then, $\left((1-\alpha_i-\beta_i)/m_i\right)^{1-\alpha_i-\beta_i}$ in $E[Z_i(\alpha_i)]$ should be replaced with $(\beta_i/w_i)^{\beta_i}$. Accordingly, the following estimation would involve $1-\alpha_i-\beta_i$ and w_i, rather than β_i and m_i. However, such alternation has no effect on the estimation of other parameters including σ_τ.

the estimated capital market distortions. The second and third rows report the estimates when m is calibrated to 20 percent and 80 percent of firms choosing α_h in the distortionless environment, respectively. Our key estimate, σ_τ, remains largely unchanged.

[Insert Table 8]

6 Conclusion

Misallocation has been viewed as a promising candidate for explaining the large TFP differences across countries. To evaluate the importance of misallocation, we need to estimate distortions at disaggregate levels. The popular ARP approach relies on the dispersion of revenue productivity to back out distortions. However, the dispersion may reflect factors other than distortions. This paper contributes to the literature in two aspects. The first is methodological. To address the identification issue, we present a model that incorporates (i) unobserved heterogeneities in the capital output elasticity and markups; (ii) capital adjustment costs; and (iii) measurement errors. All three factors contribute to the dispersion of the revenue-capital ratio and, hence, bias the estimate upwards. We then develop identification conditions that isolate capital market distortions in a structural way. Secondly, when applying the method to a firm-level panel dataset from Chinese manufacturing, we find quantitatively large capital market distortions. In contrast, for large Compustat firms, the magnitude of the distortions turns out to be statistically insignificant.

To be sure, there are potentially other factors that may bias the estimate upwards. Our estimate can, thus, be taken as a more reasonable upper bound than those from the ARP approach. It is also worth emphasizing that our model does not entail any selection effects, which tend to amplify the effect of distortions. In a recent paper, Midrigan and Xu (2012) find a large effect of the credit constraint on misallocation via entry. A future research direction would be to incorporate entry and exit in the structural estimation. Moreover, like most others in the literature, our paper accommodates static misallocation only.[35] It would be interesting in exploring dynamic welfare implications of the distortions.

References

[1] Abel, Andrew B. and Janice C. Eberly (1994), "A Unified Model of Investment under Uncertainty," *American Economic Review*, 84, 1369-1384.

[35] One exception is Michael Peters (2011).

[2] Angeletos, George-Marios and Vasia Panousi (2011), "Financial integration, entrepreneurial risk, and global dynamics," *Journal of Economic Theory*, 146, 863-896.

[3] Asker John, Allan Collard-Wexler and Jan De Loecker (2011), "Productivity Volatility and the Misallocation of Resources in Developing Economies," NBER Working Paper 17175.

[4] Bai, Chong-En, Chang-Tai Hsieh and Yingyi Qian (2006), "The Returns to Capital in China," *Brookings Papers on Economic Activity*, 37, 61-102.

[5] Banerjee, Abhijit V. and Esther Duflo (2005), "Growth Theory through the Lens of Development Economics," in Philippe Aghion & Steven Durlauf (ed.), *Handbook of Economic Growth*, edition 1, volume 1, chapter 7, 473-552.

[6] Bartelsman Eric J., Randy A. Becker, and Wayne B. Gray (2000), "NBER Productivity Database," available at http://www.nber.org/nberces.

[7] Bartelsman, Eric J., John C. Haltiwanger and Stefano Scarpetta (2009), "Cross-Country Differences in Productivity: The Role of Allocation and Selection," NBER Working Paper 15490 (forthcoming, *American Economic Review*)

[8] Bloom, Nick (2000), "The Dynamic Effects of Real Options and Irreversibility on Investment and Labour Demand," Institute for Fiscal Studies Working Paper No.W00/15.

[9] Bloom, Nick (2009), "The Impact of Uncertainty Shocks," *Econometrica*, 77, 623-685.

[10] Bloom, Nick, Stephen R. Bond and John Van Reenen (2007), "Uncertainty and Investment Dynamics," *Review of Economic Studies*, 74, 391-415.

[11] Blundell, Richard and Stephen R. Bond (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," Journal of Econometrics, 87, 115-143.

[12] Bond, Stephen R., Måns Söderbom and Guiying Laura Wu (2007), "Investment and Financial Constraints: Empirical Evidence for Firms in Brazil and China," mimeo, University of Oxford.

[13] Bond, Stephen R., Måns Söderbom and Guiying Laura Wu (2008), "A Structural Estimation for the Effects of Uncertainty on Capital Accumulation with Heterogeneous Firms," mimeo, University of Oxford.

[14] Brandt, Loren, Johannes Van Biesebroeck and Yifan Zhang (2012), "Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing," *Journal of Development Economics*, 97(2), 339-351.

[15] Brandt, Loren, Trevor Tombe and Xiaodong Zhu (2013), "Factor Market Distortions Across Time, Space and Sectors in China," *Review of Economic Dynamics*, 16, 39-58.

[16] Buera, Francisco J. and Yongseok Shin (2010), "Financial Frictions and the Persistence of History," mimeo.

[17] Buera, Francisco J., Joseph P. Kaboski and Yongseok Shin (2011), "Finance and Development: A Tale of Two Sectors," American Economic Review, 101, 1964-2002.

[18] Caselli, Francesco (2005), "Accounting for Cross-Country Differences," in Philippe Aghion & Steven Durlauf (ed.), *Handbook of Economic Growth*, edition 1, volume 1, chapter 9, 679-741.

[19] Caselli, Francesco and James Feyrer (2007), "The Marginal Product of Capital," *Quarterly Journal of Economics*, 122, 535-568.

[20] Caselli, Francesco and Nicola Gennaioli (2013), "Dynastic Management," *Economic Inquiry*, 51, 971-996.

[21] Cooper, Russell W. and Joao Ejarque (2003), "Financial Frictions and Investment: A Requiem in Q," *Review of Economic Dynamics*, 6, 710-728.

[22] Cooper, Russell W. and John Haltiwanger (2006), "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*, 73, 611-634.

[23] Dollar, David and Shang-jin Wei (2007), "Das (Wasted) Kapital: Firm Ownership and Investment Efficiency in China," NBER Working Paper, No. 13103.

[24] Eberly, Janice C., Sergio Rebelo and Nicolas Vincent (2008), "Investment and Value: A Neoclassical Benchmark," NBER Working Paper, No. 13866.

[25] Eckstein, Zvi and Kenneth I. Wolpin (1999), "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities," *Econometrica*, 67, 1295-1339.

[26] Fazzari, S.M., R. G. Hubbard, and B. C. Petersen (1988), "Financing Constraints and Corporate Investment," Brookings Papers on Economic Activity, 1, 141-195.

[27] Foster, Lucia, John Haltiwanger and Chad Syverson (2008), "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98, 394-425.

[28] Hsieh, Chang-Tai and Peter J. Klenow (2009), "Misallocation and Manufacturing TFP in China and India," *Quarter Journal of Economics*, 124, 1403-1448.

[29] Jorgenson, Dale W., Frank M. Gollop and Barbara M. Fraumeni (1987), *Productivity and U.S Economic Growth*, toExcel Press.

[30] Gouriéroux, Christian and Alain Monfort (1996), *Simulation-Based Econometric Methods*, Oxford University Press.

[31] HENNESSY, Christopher A. and Toni M. Whited (2007), "How Costly Is External Financing? Evidence from a Structural Estimation," *Journal of Finance*, 62, 1705-1745.

[32] Khan, Aubhik and Julia K. Thomas (2006), "Adjustment Costs," *The New Palgrave Dictionary of Economics*, Palgrave Macmillan.

[33] Mankiw, N. Gregory and Matthew D. Shapiro (1986), "Risk and Return: Consumption Beta Versus Market Beta," *Review of Economics and Statistics*, 68, 452-459.

[34] Midrigan, Virgiliu and Daniel Yi Xu (2012), "Finance and Misallocation: Evidence from Plant-Level Data," mimeo.

[35] Moll, Benjamin (2012), "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation," mimeo.

[36] Perkins, Dwight H. and Thomas G. Rawski (2008), "Appendix to Forecasting China's Economic Growth to 2025," mimeo.

[37] Peters, Michael (2011), "Heterogeneous Mark-Ups and Endogenous Misallocation," mimeo.

[38] Qian, Zhenjie and Xiaodong Zhu (2012), "Why Is Labor Income Share So Low in China?" mimeo.

[39] Schündeln, Matthias (2006), "Modeling Firm Dynamics to Identify the Cost of Financing Constraints in Ghanaian Manufacturing," mimeo.

[40] Song, Zheng, Kjetil Storesletten and Fabrizio Zilibotti (2011), "Growing Like China," *American Economic Review*, 101, 202-241.

[41] RESTUCCIA, Diego, and Richard ROGERSON (2008), "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 11, 707-720.

[42] RESTUCCIA, Diego, and Richard ROGERSON (2013), "Misallocation and Productivity," *Review of Economic Dynamics*, 16, 1-10.

[43] WU, Guiying Laura (2009), *Uncertainty, Investment and Capital Accumulation: A Structural Econometric Approach*, D.Phil. Thesis, University of Oxford.

[44] ZHANG, Junsen, Yaohui ZHAO, Albert PARK and Xiaoqing SONG (2005), "Economic Returns to Schooling in Urban China, 1988 to 2001," *Journal of Comparative Economics*, 33, 730-752.

7 Appendix

7.1 Aggregate TFPR Losses

Both distortions and frictions may cause aggregate productive efficiency loss. To quantify these effects, consider N firms in the economy with the same α, η, thus γ. Define the "efficient" benchmark as the capital allocation in an economy without capital adjustment costs or capital market distortions. Denote $\hat{K}^*_{i,t}$ firm i's productive capital stock in the efficient allocation. $\hat{K}^*_{i,t}$ must be linear proportional to $Z_{i,t}$,

$$\hat{K}^*_{i,t} = \left(\frac{1-\gamma}{J}\right)^{\frac{1}{\gamma}} Z_{i,t}.$$

This implies that in the efficient allocation, each firm gets a share of capital proportional to the share of its $Z_{i,t}$,

$$\hat{K}^*_{i,t} = \frac{Z_{i,t}}{Z^*_t}\hat{K}_t,$$

where $\hat{K}_t = \sum_{i=1}^N \hat{K}_{i,t}$ is the existing aggregate productive capital stock; and $Z^*_t = \sum_{i=1}^N Z_{i,t}$ is an aggregation of $Z_{i,t}$.

The efficient allocation has the following aggregate sales revenue:

$$Y^*_t \equiv \frac{\gamma}{\eta}\sum_{i=1}^N (Z_{i,t})^\gamma \left(\hat{K}^*_{i,t}\right)^{1-\gamma} = \frac{\gamma}{\eta} Z^{*\gamma}_t \hat{K}^{1-\gamma}_t.$$

In contrast, the actual aggregate sales revenue with capital allocation of $\left\{\hat{K}_{i,t}\right\}_{i=1}^N$ is

$$Y_t \equiv \frac{\gamma}{\eta}\sum_{i=1}^N \left(Z^\gamma_{i,t}\hat{K}^{1-\gamma}_{i,t}\right) = \frac{\gamma}{\eta} Z^\gamma_t \hat{K}^{1-\gamma}_t,$$

where

$$Z_t \equiv \left[\sum_{i=1}^N \left(\frac{\theta_{i,t}}{Z^*_t}\right)^{1-\gamma} Z_{i,t}\right]^{\frac{1}{\gamma}},$$

and

$$\theta_{i,t} \equiv \frac{\hat{K}_{i,t}}{\hat{K}^*_{i,t}} = \frac{Z^*_t \hat{K}_{i,t}}{Z_{i,t}\hat{K}_t}.$$

Here, $\theta_{i,t}$ denotes the wedge between the actual and efficient capital. Z^γ_t is referred to as the aggregate revenue total factor productivity (TFPR). Note that in the efficient allocation, $\theta_{i,t} = 1$, and the aggregate TFPR, Z^γ_t, is identical to $(Z^*_t)^\gamma$. The aggregate TFPR losses due to the misallocated $\left\{\hat{K}_{i,t}\right\}_{i=1}^N$ can, thus, be represented by the difference between Y_t and Y^*_t:

$$\begin{aligned}\log Y_t - \log Y^*_t &= \log Z^\gamma_t - \log (Z^*_t)^\gamma \\ &= \log\left(\sum_{i=1}^N \left(Z^\gamma_{i,t}\hat{K}^{1-\gamma}_{i,t}\right)\right) - (1-\gamma)\log\left(\sum_{i=1}^N \hat{K}_{i,t}\right) - \gamma\log\left(\sum_{i=1}^N Z_{i,t}\right) \quad (30)\end{aligned}$$

To highlight how capital market distortions will lower the aggregate TFPR, consider the special case without adjustment cost and with common $Z_{i,t}$ across firms. With a large number of firms ($N \to \infty$), there is a closed-form solution to the aggregate TFPR losses:

$$\begin{aligned} \Delta \log TFPR_t &= -\frac{1}{2}\frac{1-\gamma}{\gamma} Var\left[\log\left(1+\tau_i\right)\right] \\ &= -\frac{1}{2}\frac{\alpha\left(1-\eta\right)}{\eta} Var\left[\log\left(1+\tau_i\right)\right]. \end{aligned} \tag{31}$$

In other words, the negative effect of capital market distortions on aggregate TFPR can be summarized by the variance of $\log\left(1+\tau_i\right)$, and the magnitude of the effect increases with $1-\gamma$, the capital share in the profit or revenue function.

Finally, we can derive the aggregate TFPR losses caused by labor market distortions specified in Section 5.2 in a similar fashion.

$$\Delta \log TFPR_t = -\frac{1}{2}\frac{\beta\left(1-\eta\right)}{\eta} Var\left[\log\left(1+\tau_i^L\right)\right]. \tag{32}$$

7.2 Identification

[Insert Table A.1 to A.3]

7.3 Data

Brandt et al. (2012) provide an excellent description of the dataset and implement a series of consistency checks. Therefore, we strictly follow them in constructing a panel and cleaning the data.

In our application, $Y_{i,t}$ is defined as sales revenue of products plus changes in the inventory of finished products. Variable profit $\pi_{i,t}$ is constructed by subtracting the cost of goods sold from the sales revenue. Ideally, variable profit should be the difference between sales revenue and the cost of labor and intermediate inputs. However, the cost of intermediate inputs is not available in the dataset; and cost of labor is known to be poorly measured in the Chinese context. Instead, costs of goods sold are reported by all the firms in the dataset. By accounting definition, cost of goods sold (COGS) refers to the inventory costs of those goods a firm has sold during a particular period. The key components of cost generally include: parts, raw materials and supplies used; labor, including associated costs such as payroll taxes and benefits; and overhead of the business allocable to production. Therefore, we think that the difference between sales revenue and cost of goods sold provides a good proxy to the variable profit in our investment model.

We deflate the revenue and profit using the GDP deflator for the secondary industry from the *China Statistical Yearbook*. The survey does not contain information on investment expenditures. However, firms report the book value of their fixed capital stock at original purchase prices. Since these book values are the sum of nominal values for different years, they should not be used directly. Thus, we construct our capital stock series using the following formula:

$$K_{i,t} = (1-\delta)K_{i,t-1} + \left(BK_{i,t} - BK_{i,t-1}\right)/P_t,$$

where $BK_{i,t}$ is the book value of capital stock for firm i in year t; P_t is the price index of investment in fixed assets in year t constructed by Perkins and Rawski (2008). The initial book value of capital stock is taken directly from the dataset for firms founded later than 1998. For firms founded before 1998, we predict it to be

$$BK_{i,t_0} = BK_{i,t_1}/\left(1+g_i\right)^{t_1-t_0}$$

where BK_{i,t_0} is the projected initial book value of capital stock when firm i was born in year t_0; BK_{i,t_1} is the book value of capital stock when firm i first appears in our dataset in year t_1; and g_i is the average capital stock growth rate of firm i for the period we observe in the data since year t_1.

Investment expenditure $I_{i,t}$ is then recovered according to equation (10). We experiment with a depreciation rate δ from 3 percent to 10 percent and pin it down to be 5 *percent*, which is the average difference between the constructed investment rate and the revenue growth rate.

Four key variables for estimation are then constructed by definition: profit-revenue ratio $(\pi_{i,t}/Y_{i,t})$, log revenue-capital ratio $\left(\log\left(Y_{i,t}/\hat{K}_{i,t}\right)\right)$, investment rate $(I_{i,t}/K_{i,t})$ and revenue growth rate $(\Delta \log Y_{i,t})$. We exclude outliers by trimming the top and bottom 5 percent of observations for each variable in each year. The model assumes that firms are on the balanced-growth path. In the presence of capital adjustment costs, however, it may take years for firms to reach their balanced-growth path. Therefore, we exclude firms that are less than 5 years old when they first enter our dataset. Furthermore, our investment model does not consider entry and exit, which means that the model's implications are valid only for existing and ongoing firms. Finally, many existing non-state-owned firms with sales revenue beyond RMB 5 millions were missing from the survey in the early years but have appeared in the NBS data since 2004 thanks to the industrial census conducted in that year. For these reasons, our empirical exercise utilizes a sample of firms surviving from 2004 to 2007 and being at least 5 years old in 2004. This gives us a balanced panel consisting of 107,579 firms and spanning 4 years. The annual mean values of each of the four variables in the balanced panel are reported in the following table.

[Insert Table A.4]

7.4 Back-of-the-Envelope Calculations for the Effects of Market Beta and Market Incompleteness

Recall that our empirical exercise has estimated that

$$\log(1+\tau_i) \overset{i.i.d.}{\sim} N\left(0, 0.7143^2\right).$$

If all the heterogeneity comes from J_i, $\log J_i$ would follow a log-normal distribution with a standard deviation of 0.7143. Since $\log J_i = \log \frac{r_i+\delta}{1+r_i} \simeq \log(r_i+\delta)$, it implies that

$$\log(r_i+\delta) \overset{i.i.d.}{\sim} N\left(\log(0.20+0.05), 0.7143^2\right).$$

We can, thus, back out $std(r_i) = 0.2632$.

To see the capacity of market beta in generating $std(r_i)$, consider a typical CAPM,

$$r_i = r_f + (r_m - r_f)\cdot beta_i,$$

where r_f is the interest rate on riskless assets, r_m is the expected market return and $r_m - r_f$ is the expected market risk premium.

If all the heterogeneity in J_i is driven solely by heterogeneous market beta, $std(beta_i)$ has to be 5.3 to match $std(r_i) = 0.2632$ with a five-percent risk premium. Since most firms in the Chinese dataset are not listed, it is not feasible to calculate the dispersion of market beta. Picking up the number from Mankiw and Shapiro (1986), $std(beta_i)$ is only 0.38 for 464 U.S. stocks over 92 quarters. Even if we double the number for $std(beta_i)$, $std(r_i)$ would be 0.038, merely 14 percent of the 0.2632 that is needed to explain the estiamted σ_τ.

To see the importance of market incompleteness in generating heterogeneity in r_i, following the notations in equation (18) in Angeletos and Panousi (2011), at the steady state, we have

$$r_i = r_f + \sqrt{\frac{2\theta\gamma(d-r_f)}{\theta+1}}\sigma_i,$$

where $d > 0$ is the discount factor, $\gamma > 0$ is the coefficient of relative risk aversion, and $\theta > 0$ is the elasticity of intertemporal substitution. Now, all we need is a measure of σ_i and its dispersion in the data.

In our model, sales revenue is a linear combination of TFPR and capital stock (see equation 7), where the log of $Z_{i,t}$ follows (8). Idiosyncratic risks can be introduced through shocks $e_{i,t}$. In particular, we assume

$$e_{i,t} \overset{i.i.d.}{\sim} N(0, \sigma_i^2),$$

where σ_i parameterizes the level of the risk in firm i.

Bloom (2000) shows that even if capital adjustment costs are present, both revenue and capital will still grow at the same rate of μ in the long run. This implies

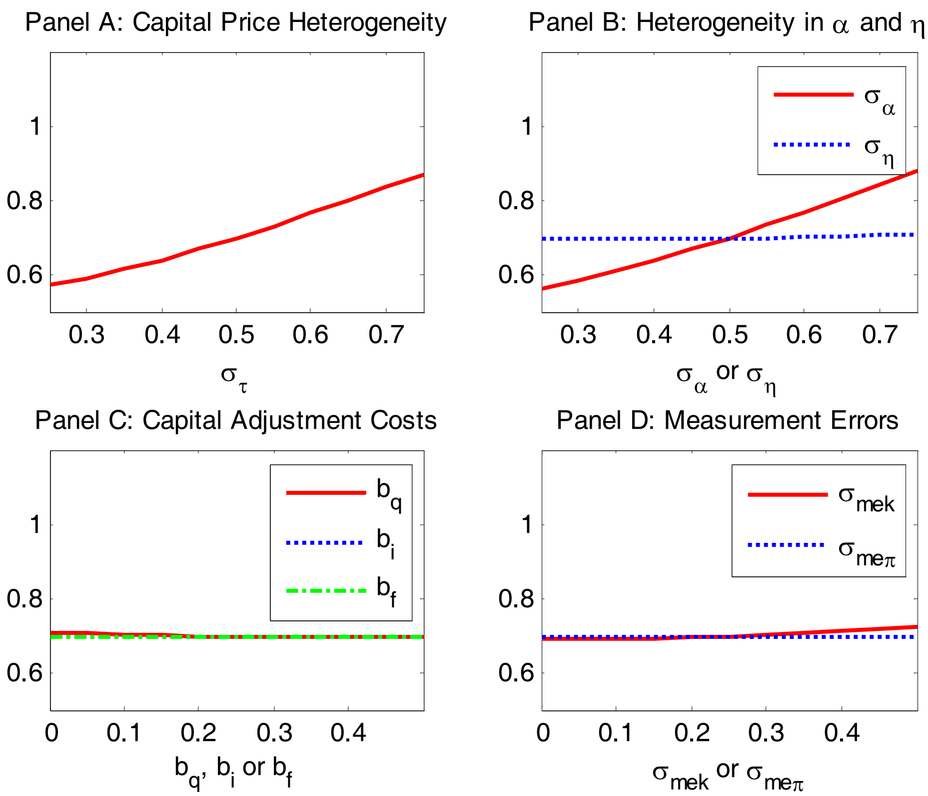
$$\begin{aligned}
\Delta \log Y_{i,t} &= \Delta \log Z_{i,t} \\
&= (\mu t + z_{i,t}) - (\mu(t-1) + z_{i,t-1}) \\
&= \mu + z_{i,t} - z_{i,t-1} \\
&\simeq \mu + e_{i,t} \quad (\text{if } \rho \rightarrow 1).
\end{aligned}$$

Therefore, when panel data are available, we can calculate the standard deviation of the sales growth rate for each firm to proxy for that firm's the level of risk:

$$\begin{aligned}
std_t (\Delta \log Y_{i,t}) &= std_t (\mu + e_{i,t}) \\
&= \sigma_i.
\end{aligned}$$

In our Chinese dataset, the standard deviation of $std_t (\Delta \log Y_{i,t})$ has an estimate of 0.1416. If all the heterogeneity in J_i is driven solely by idiosyncratic risks, $\sqrt{\frac{2\theta\gamma(d-R)}{\theta+1}}$should be as large as 1.8591. If $d - r_f = 0.05$ and $\theta = 1$, the coefficient of relative risk aversion γ has to be as large as 70 to match $std\,(r_i) = 0.2632$. Alternatively, if $\theta = 1$ and $\gamma = 5$, the standard deviation of r_i would be equal to 0.07, which accounts for about 27 percent of the estimated σ_τ.

Figure 1: Identification in the Full-Blown Model



This figure plots the sensitivity of $bsd\left(log(Y/\widehat{K})\right)$ with respect to σ_τ (Panel A), $\sigma_{log\alpha}$ and $\sigma_{log\eta}$ (Panel B), b^q, b^i and b^f (Panel C) and σ_{meK} and $\sigma_{me\pi}$ (Panel D), respectively. The benchmark parameterization is set equal to that in Column (5) of Table A.3.

Table 1: Parameters and Moments

Parameters	Definition
σ_{τ}	standard deviation of heterogeneities in capital goods price
$\mu_{log\alpha}$	mean of log capital output elasticity in production function
$\sigma_{log\alpha}$	standard deviation of log capital output elasticity
$\mu_{log\eta}$	mean of log inverse of demand elasticity
$\sigma_{log\eta}$	standard deviation of log inverse of demand elasticity
b^{q}	quadratic adjustment costs
b^{i}	partial irreversibility
b^{f}	fixed adjustment costs
μ	mean of growth rate in $Z_{i,t}$
σ	standard deviation of shocks to $Z_{i,t}$
σ_{meK}	standard deviation of measurement errors in capital stock
σ_{meY}	standard deviation of measurement errors in sales revenue
$\sigma_{me\pi}$	standard deviation of measurement errors in variable profit

Moments	Definition
mean(π/Y)	mean of profit-to-revenue ratio
mean(log(Y/Khat))	mean of log revenue-to-capital ratio
mean(I/K)	mean of investment rate
mean(ΔlogY)	mean of revenue growth rate
bsd(π/Y)	between-group standard deviation of profit-to-revenue ratio
wsd(π/Y)	within-group standard deviation of profit-to-revenue ratio
bsd(log(Y/Khat))	between-group standard deviation of log revenue-to-capital ratio
wsd(log(Y/Khat))	within-group standard deviation of log revenue-to-capital ratio
bsd(I/K)	between-group standard deviation of investment rate
wsd(I/K)	within-group standard deviation of investment rate
bsd(ΔlogY)	between-group standard deviation of revenue growth rate
wsd(ΔlogY)	within-group standard deviation of revenue growth rate
skew(π/Y)	skewness of profit-to-revenue ratio
skew(log(Y/Khat))	skewness of log revenue-to-capital ratio
skew(I/K)	skewness of investment rate
skew(dlogY)	skewness of revenue growth rate
scorr(π/Y)	serial correlation of profit-to-revenue ratio
scorr(log(Y/Khat))	serial correlation of log revenue-to-capital ratio
scorr(I/K)	serial correlation of investment rate
scorr(ΔlogY)	serial correlation of revenue growth rate
bcorr(π/Y, log(Y/Khat))	cross correlation between between-group profit-to-revenue ratio and log revenue-to-capital ratio

Table 2: Benchmark Results

Parameters	estimate	s.e.
$\sigma_{\tau K}$	0.7143	0.0033
$\mu_{log\alpha}$	-2.6058	0.0019
$\sigma_{log\alpha}$	0.5568	0.0043
$\mu_{log\eta}$	-2.8084	0.0051
$\sigma_{log\eta}$	0.7253	0.0061
b^q	0.2777	0.0038
b^i	0.0001	0.0395
b^f	0.0335	0.0006
μ	0.0802	0.0004
σ_ε	0.4253	0.0016
σ_{mK}	0.4010	0.0013
σ_{mY}	0.0007	0.1255
$\sigma_{m\pi}$	0.5777	0.0020
No. of firms	107579	
span of years	2004-2007	
mean(sales)	97	
med(sales)	20	
mean(employees)	0.33	
med(employees)	0.13	

Moments	empirical	s.e.	simulated
mean(π/Y)	0.1578	0.0002	0.1542
mean(log(Y/Khat))	1.1377	0.0025	1.1456
mean(I/K)	0.1640	0.0005	0.1729
mean(ΔlogY)	0.0963	0.0005	0.0803
bsd(π/Y)	0.0763	0.0001	0.0745
wsd(π/Y)	0.0506	0.0001	0.0488
bsd(log(Y/Khat))	0.8666	0.0011	0.8781
wsd(log(Y/Khat))	0.3470	0.0009	0.3321
bsd(I/K)	0.1991	0.0006	0.1642
wsd(I/K)	0.2027	0.0006	0.2149
bsd(ΔlogY)	0.1876	0.0004	0.1632
wsd(ΔlogY)	0.2042	0.0004	0.2187
skew(π/Y)	0.7760	0.0039	0.8539
skew(log(Y/Khat))	0.0570	0.0038	0.0037
skew(I/K)	2.2307	0.0075	2.2510
skew(dlogY)	0.1567	0.0037	0.1760
scorr(π/Y)	0.5703	0.0021	0.5993
scorr(log(Y/Khat))	0.8403	0.0009	0.8377
scorr(I/K)	0.1188	0.0030	0.2430
scorr(ΔlogY)	0.0685	0.0028	0.0526
bcorr(π/Y, log(Y/Khat))	-0.2422	0.0034	-0.2707
OI/100		183	

Note: Unit of sales is millions of RMB in 2004 prices. Imposed parameters are δ= 0.05, r= 0.20, ρ= 0.90.

Table 3: Specification Tests

	col (1) benchmark	col (2) $\sigma_{log\alpha}=\sigma_{log\eta}=0$	col (3) $b^q=b^i=b^f=0$	col (4) $\sigma_{meK}=\sigma_{meY}=\sigma_{me\pi}=0$
Parameters				
σ_τ	0.714	0.924	0.665	0.734
$\mu_{log\alpha}$	-2.606	-2.351	-2.645	-2.742
$\sigma_{log\alpha}$	0.557	**0.000**	0.587	0.500
$\mu_{log\eta}$	-2.808	-2.494	-2.716	-2.998
$\sigma_{log\eta}$	0.725	**0.000**	0.660	0.885
b^q	0.278	0.443	**0.000**	0.163
b^i	0.000	0.000	**0.000**	0.476
b^f	0.034	0.082	**0.000**	0.041
μ	0.080	0.078	0.100	0.054
σ	0.425	0.354	0.205	0.443
σ_{meK}	0.401	0.380	0.420	**0.000**
σ_{meY}	0.001	0.123	0.110	**0.000**
$\sigma_{me\pi}$	0.578	0.816	0.541	**0.000**
Moments				
mean(π/Y)	0.154	0.171	0.155	0.141
mean(log(Y/Khat))	1.146	1.011	1.151	1.218
mean(I/K)	0.173	0.168	0.206	0.127
mean(ΔlogY)	0.080	0.078	0.100	0.053
bsd(π/Y)	0.075	0.042	0.073	0.071
wsd(π/Y)	0.049	0.073	0.047	0.000
bsd(log(Y/Khat))	0.878	0.848	0.872	0.851
wsd(log(Y/Khat))	0.332	0.328	0.343	0.137
bsd(I/K)	0.164	0.146	0.145	0.136
wsd(I/K)	0.215	0.218	0.274	0.177
bsd(ΔlogY)	0.163	0.153	0.123	0.160
wsd(ΔlogY)	0.219	0.254	0.227	0.221
skew(π/Y)	0.854	0.184	0.887	0.391
skew(log(Y/Khat))	0.004	0.008	0.011	0.013
skew(I/K)	2.251	2.220	1.586	2.450
skew(dlogY)	0.176	0.213	0.002	0.370
scorr(π/Y)	0.599	-0.001	0.604	1.000
scorr(log(Y/Khat))	0.838	0.830	0.822	0.977
scorr(I/K)	0.243	0.126	-0.047	0.242
scorr(ΔlogY)	0.053	-0.149	-0.223	0.027
bcorr(π/Y, log(Y/Khat))	-0.271	-0.019	-0.304	-0.208
OI/100	183	1510	653	3127

Table 4: Robustness Tests

Parameters	(1) benchmark	(2) δ=0.03	(3) δ=0.07	(4) r=0.15	(5) r=0.25	(6) ρ=0.85	(7) ρ=0.95
σ_τ	0.714	0.705	0.730	0.670	0.746	0.712	0.729
$\mu_{log\alpha}$	-2.606	-2.675	-2.539	-2.727	-2.496	-2.595	-2.602
$\sigma_{log\alpha}$	0.557	0.566	0.543	0.606	0.524	0.559	0.549
$\mu_{log\eta}$	-2.808	-2.752	-2.909	-2.672	-2.973	-2.826	-2.812
$\sigma_{log\eta}$	0.725	0.708	0.770	0.666	0.794	0.729	0.730
b^q	0.278	0.308	0.256	0.387	0.273	0.258	0.346
b^i	0.000	0.000	0.001	0.000	0.005	0.000	0.014
b^f	0.034	0.040	0.025	0.060	0.025	0.024	0.029
μ	0.080	0.094	0.066	0.080	0.083	0.081	0.085
σ	0.425	0.426	0.430	0.412	0.447	0.427	0.416
σ_{meK}	0.401	0.387	0.409	0.379	0.410	0.405	0.411
σ_{meY}	0.001	0.001	0.000	0.000	0.001	0.000	0.004
$\sigma_{me\pi}$	0.578	0.581	0.572	0.575	0.579	0.581	0.574
Moments							
mean(π/Y)	0.154	0.152	0.155	0.153	0.155	0.154	0.154
mean(log(Y/Khat))	1.146	1.132	1.146	1.127	1.165	1.143	1.159
mean(I/K)	0.173	0.162	0.184	0.170	0.177	0.174	0.179
mean(ΔlogY)	0.080	0.094	0.066	0.080	0.083	0.081	0.084
bsd(π/Y)	0.075	0.075	0.074	0.075	0.074	0.074	0.075
wsd(π/Y)	0.049	0.049	0.049	0.049	0.049	0.049	0.049
bsd(log(Y/Khat))	0.878	0.877	0.880	0.874	0.882	0.879	0.884
wsd(log(Y/Khat))	0.332	0.326	0.333	0.319	0.337	0.333	0.337
bsd(I/K)	0.164	0.157	0.170	0.153	0.170	0.155	0.178
wsd(I/K)	0.215	0.204	0.221	0.209	0.215	0.217	0.214
bsd(ΔlogY)	0.163	0.164	0.163	0.160	0.165	0.158	0.168
wsd(ΔlogY)	0.219	0.222	0.215	0.222	0.215	0.224	0.211
skew(π/Y)	0.854	0.873	0.846	0.856	0.846	0.855	0.857
skew(log(Y/Khat))	0.004	0.005	0.004	0.007	0.006	0.002	-0.002
skew(I/K)	2.251	2.303	2.158	2.320	2.206	2.168	2.181
skew(dlogY)	0.176	0.180	0.151	0.208	0.146	0.165	0.169
scorr(π/Y)	0.599	0.598	0.598	0.608	0.590	0.596	0.600
scorr(log(Y/Khat))	0.838	0.844	0.837	0.849	0.834	0.837	0.835
scorr(I/K)	0.243	0.246	0.254	0.200	0.274	0.202	0.297
scorr(ΔlogY)	0.053	0.040	0.068	0.026	0.073	0.015	0.099
bcorr(π/Y, log(Y/Khat))	-0.271	-0.257	-0.278	-0.280	-0.275	-0.270	-0.259
OI/100	183	182	213	208	179	215	157

Table 4: Robustness Tests – Continued

Parameters	(1) benchmark	(8) type-5	(9) $\sigma_{\upsilon} > 0$	(10) $\sigma_{\sigma} > 0$	(11) $\sigma_{meI} > 0$
σ_{τ}	0.714	0.690	0.721	0.712	0.745
$\mu_{log\alpha}$	-2.606	-2.620	-2.604	-2.592	-2.654
$\sigma_{log\alpha}$	0.557	0.557	0.551	0.556	0.577
$\mu_{log\eta}$	-2.808	-2.851	-2.805	-2.805	-2.776
$\sigma_{log\eta}$	0.725	0.692	0.728	0.719	0.716
b^q	0.278	0.284	0.325	0.308	0.405
b^i	0.000	0.001	0.000	0.000	0.479
b^f	0.034	0.034	0.039	0.031	0.059
μ	0.080	0.082	0.083	0.080	0.061
σ	0.425	0.424	0.411	0.403	0.465
σ_{meK}	0.401	0.402	0.404	0.390	..
σ_{meY}	0.001	0.000	0.002	0.001	0.001
$\sigma_{me\pi}$	0.578	0.597	0.576	0.575	0.561
σ_{υ}	..	..	0.080	..	..
σ_{σ}	..	..	..	0.151	..
σ_{meI}	..	..	..	..	0.114
Moments					
mean(π/Y)	0.154	0.148	0.154	0.155	0.153
mean(log(Y/Khat))	1.146	1.155	1.154	1.147	1.104
mean(I/K)	0.173	0.175	0.177	0.171	0.135
mean(ΔlogY)	0.080	0.082	0.083	0.080	0.060
bsd(π/Y)	0.075	0.071	0.074	0.074	0.075
wsd(π/Y)	0.049	0.048	0.048	0.049	0.047
bsd(log(Y/Khat))	0.878	0.880	0.883	0.875	0.870
wsd(log(Y/Khat))	0.332	0.331	0.334	0.324	0.144
bsd(I/K)	0.164	0.165	0.180	0.161	0.135
wsd(I/K)	0.215	0.214	0.213	0.213	0.169
bsd(ΔlogY)	0.163	0.163	0.168	0.162	0.165
wsd(ΔlogY)	0.219	0.217	0.210	0.218	0.230
skew(π/Y)	0.854	1.010	0.852	0.846	0.846
skew(log(Y/Khat))	0.004	0.013	0.004	0.006	0.029
skew(I/K)	2.251	2.193	2.225	2.295	2.412
skew(dlogY)	0.176	0.176	0.195	0.157	0.268
scorr(π/Y)	0.599	0.581	0.604	0.598	0.620
scorr(log(Y/Khat))	0.838	0.839	0.838	0.844	0.976
scorr(I/K)	0.243	0.250	0.292	0.237	0.268
scorr(ΔlogY)	0.053	0.058	0.103	0.051	0.021
bcorr(π/Y, log(Y/Khat))	-0.271	-0.319	-0.263	-0.274	-0.299
OI/100	183	229	148	179	741

Table 5: Compustat Results

Sample	Compustat I		Compustat II	
Parameters	estimate	s.e.	estimate	s.e.
σ_τ	0.3020	0.0676	0.1113	0.1822
$\mu_{log\alpha}$	-1.8745	0.0193	-2.1797	0.0225
$\sigma_{log\alpha}$	0.6271	0.0272	0.5338	0.0308
$\mu_{log\eta}$	-1.5831	0.0305	-1.6313	0.0270
$\sigma_{log\eta}$	0.5949	0.0146	0.5235	0.0186
b^q	1.1355	0.1489	0.9992	0.0741
b^i	0.0073	0.1836	0.0008	0.5095
b^f	0.0017	0.0017	0.0003	0.0136
μ	0.0263	0.0017	0.0524	0.0018
σ	0.2636	0.0096	0.2009	0.0060
σ_{meK}	0.2014	0.0117	0.1875	0.0055
σ_{meY}	0.0023	0.2542	0.0010	0.2222
$\sigma_{me\pi}$	0.2286	0.0112	0.1678	0.0089
No. of firms	1431		970	
span of years	2002-2005		1981-2000	
mean(sales)	2066		3132	
med(sales)	121		461	
mean(employees)	10.4		20.8	
med(employees)	0.8		4.3	

Sample	Compustat I		Compustat II	
Moments	empirical	simulated	empirical	simulated
mean(π/Y)	0.3929	0.3839	0.3458	0.3206
mean(log(Y/Khat))	0.5659	0.5600	0.8560	0.8194
mean(I/K)	0.1338	0.1463	0.1779	0.1752
mean(ΔlogY)	0.0900	0.0274	0.0696	0.0523
bsd(π/Y)	0.1591	0.1624	0.1251	0.1300
wsd(π/Y)	0.0450	0.0481	0.0456	0.0327
bsd(log(Y/Khat))	0.7288	0.7148	0.5297	0.5544
wsd(log(Y/Khat))	0.2301	0.1829	0.2662	0.1978
bsd(I/K)	0.0729	0.0614	0.0755	0.0217
wsd(I/K)	0.0522	0.0582	0.0667	0.0696
bsd(ΔlogY)	0.1483	0.1020	0.1040	0.0239
wsd(ΔlogY)	0.1297	0.1497	0.1182	0.1433
skew(π/Y)	0.4315	0.3018	0.3549	0.3385
skew(log(Y/Khat))	-0.2917	0.0636	-0.0427	0.0555
skew(I/K)	0.9393	0.7311	1.0737	0.6744
skew(dlogY)	0.3168	0.0435	0.6425	0.0052
scorr(π/Y)	0.9183	0.8926	0.9500	0.9374
scorr(log(Y/Khat))	0.9117	0.9245	0.9265	0.9078
scorr(I/K)	0.5630	0.4926	0.6206	0.4558
scorr(ΔlogY)	0.2057	-0.0253	0.3217	-0.0202
bcorr(π/Y, log(Y/Khat))	-0.0738	-0.1149	-0.0298	-0.0398
OI/100	10		26	

Note: See the text for the definition of Compustat I and II. Unit of sales is millions of USD in 2004 prices. Unit of employees is thousand. The imposed parameters are δ= 0.10, r= 0.10, ρ= 0.90.

Table 6: Comparison between Full Structural and GARP Approach

	NBS		Compustat I		Compustat II	
Parameters	**full structural**	**GARP**	**full structural**	**GARP**	**full structural**	**GARP**
σ_{τ}	0.7143	0.6843	0.3020	0.3095	0.1113	0.0091
$\mu_{\log\alpha}$	-2.6058	-2.6189	-1.8745	-1.9320	-2.1797	-2.2519
$\sigma_{\log\alpha}$	0.5568	0.5539	0.6271	0.6259	0.5338	0.5125
$\mu_{\log\eta}$	-2.8084	-2.8086	-1.5831	-1.4527	-1.6313	-1.4067
$\sigma_{\log\eta}$	0.7253	0.7887	0.5949	0.5246	0.5235	0.3896
Moments						
mean(π/Y)	0.1578		0.3929		0.3514	
mean(log(Y/Khat))	1.1377		0.5659		0.5542	
bsd(π/Y)	0.0763		0.1591		0.1448	
bsd(log(Y/Khat))	0.8666		0.7288		0.6064	
bcorr(π/Y, log(Y/Khat))	-0.2422		-0.0738		-0.0879	

Note: r = 0.20, δ = 0.05 for NBS and r = 0.10, δ = 0.10 for Compustat.

Table 7: Regressions on Firm Characteristics

	(1) age and size		**(2) region**		**(3) ownership**	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
log(π/Y)	**-0.3929**	0.0027	**-0.3921**	0.0027	**-0.3590**	0.0026
age	**-0.0304**	0.0002	**-0.0303**	0.0002	**-0.0291**	0.0002
emp	**-0.0399**	0.0032	**-0.0394**	0.0032	**-0.0244**	0.0021
CENTRAL			**0.0610**	0.0040	**0.0319**	0.0039
WESTERN			**-0.1673**	0.0045	**-0.1721**	0.0044
NORTHEASTERN			**-0.2495**	0.0056	**-0.2551**	0.0055
COE					**0.4973**	0.0075
DPE					**0.4481**	0.0064
FIE					**0.0986**	0.0067
OTHERs					**0.2614**	0.0063
Obs.	107579		107579		107579	
R-sq	0.2420		0.2493		0.2806	

Note: 1. 4-digital industry dummies are included in all regressions.
2. Robust standard errors are reported in the second column of each regression.
3. Age is the difference between 2004 and the year of firm foundation.
4. Emp is the number of total employees normalized by 1000.
5. Baseline group is EASTERN--dummy = 1 if province is Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, or Hainan.
CENTRAL--dummy = 1 if province is Shanxi, Anhui, Jiangxi, Henan, Hubei or Hunan.
WESTERN--dummy =1 if province is Inner Mongolia, Guangxi, Chongqin, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia or Xinjiang.
NORTHWESTERN-- dummy = 1 if province is Liaoning, Jilin or Heilongjiang.
6. Baseline group is SOE--dummy = 1 if state-owned; defined as registration type = 110, 141 and 151.
COE--dummy = 1 if collective owned; defined as registration type = 120 and 142.
DPE--dummy = 1 if domestic private-owned; defined as registration type from 170 to 174.
FIE--dummy = 1 if foreign-owned; defined as registration type from 200 to 240 or from 300 to 340.
OTHERS--dummy = 1 if other types, including cooperative units, joint ownership units, limited liability corporations and share-holding corporation Ltd.

Table 8: Estimation with Endogenous Capital Output Elasticity

	σ_τ	$\mu_{log\eta}$	$\sigma_{log\eta}$	α_l	α_h	β	m
(1)	**0.5784**	**-2.7460**	**0.7841**	**0.0551**	**0.1090**	**0.1369**	**1.2024**
(2)	0.5800	-2.7456	0.7838	0.0555	0.1097	0.1369	1.1769
(3)	0.5887	-2.7434	0.7817	0.0576	0.1142	0.1369	1.0396

Note: m and β denote the material cost relative to the unit price of capital and labor output elasticity, respectively. (1) is the benchmark case where 50% of firms would choose α_h if τ_i=0. (2) and (3) refer to the cases where 20% and 80% of firms would choose α_h if τ_i=0, respectively.

Table A.1 Illustration for Identification of Unobserved Heterogeneities

Parameters	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
	$\boldsymbol{\sigma_\tau = 0.0}$	$\sigma_\tau = 0.5$	$\sigma_\tau = 0.0$	$\sigma_\tau = 0.0$	$\sigma_\tau = 0.5$
	$\boldsymbol{\sigma_{log\alpha} = 0.0}$	$\sigma_{log\alpha} = 0.0$	$\sigma_{log\alpha} = 0.5$	$\sigma_{log\alpha} = 0.0$	$\sigma_{log\alpha} = 0.5$
	$\boldsymbol{\sigma_{log\eta} = 0.0}$	$\sigma_{log\eta} = 0.0$	$\sigma_{log\eta} = 0.0$	$\sigma_{log\eta} = 0.5$	$\sigma_{log\eta} = 0.5$
Set of Moments					
mean(π/Y)	0.157	0.157	0.167	0.167	0.170
mean(log(Y/Khat))	0.834	0.834	0.834	0.847	0.840
mean(I/K)	0.159	0.159	0.159	0.159	0.159
mean(ΔlogY)	0.050	0.050	0.050	0.050	0.050
bsd(π/Y)	0.000	0.000	**0.044**	**0.044**	0.061
wsd(π/Y)	0.000	0.000	0.000	0.000	0.000
bsd(log(Y/Khat))	0.000	**0.496**	**0.495**	**0.054**	0.682
wsd(log(Y/Khat))	0.000	0.000	0.000	0.000	0.000
bsd(I/K)	0.164	0.165	0.165	0.165	0.164
wsd(I/K)	0.321	0.321	0.321	0.321	0.321
bsd(ΔlogY)	0.162	0.162	0.162	0.162	0.162
wsd(ΔlogY)	0.252	0.252	0.252	0.252	0.253
skew(π/Y)	0.000	0.000	**1.190**	**1.190**	0.176
skew(log(Y/Khat))	0.000	0.000	0.000	1.336	0.000
skew(I/K)	1.015	1.015	1.015	1.015	1.015
skew(dlogY)	0.005	0.005	0.005	0.005	0.005
scorr(π/Y)	N.A.	N.A.	**1.000**	**1.000**	1.000
scorr(log(Y/Khat))	N.A.	**1.000**	**1.000**	**0.985**	1.000
scorr(I/K)	-0.062	-0.062	-0.062	-0.062	-0.062
scorr(ΔlogY)	-0.067	-0.067	-0.067	-0.067	-0.067
bcorr(π/Y, log(Y/Khat))	N.A.	N.A.	**-0.954**	**0.991**	-0.378

Table A.2: Illustration for Identification of Capital Adjustment Costs

Parameters	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
	$\boldsymbol{b^q = 0.0}$	$b^q = 0.25$	$b^q = 0.0$	$b^q = 0.0$	$b^q = 0.25$
	$\boldsymbol{b^i = 0.0}$	$b^i = 0.0$	$b^i = 0.25$	$b^i = 0.0$	$b^i = 0.25$
	$\boldsymbol{b^f = 0.0}$	$b^f = 0.0$	$b^f = 0.0$	$b^f = 0.025$	$b^f = 0.025$
Set of Moments					
mean(π/Y)	0.170	0.170	0.170	0.170	0.170
mean(log(Y/Khat))	0.840	0.846	0.808	0.850	0.875
mean(I/K)	0.159	0.116	0.122	0.153	0.116
mean(ΔlogY)	0.050	0.050	0.050	0.050	0.050
bsd(π/Y)	0.061	0.061	0.061	0.061	0.061
wsd(π/Y)	0.000	0.000	0.000	0.000	0.000
bsd(log(Y/Khat))	0.682	0.679	0.684	0.684	0.678
wsd(log(Y/Khat))	0.000	**0.075**	**0.083**	**0.069**	0.087
bsd(I/K)	0.164	0.094	0.116	0.171	0.095
wsd(I/K)	0.321	**0.101**	**0.167**	0.328	0.123
bsd(ΔlogY)	0.162	0.120	0.124	0.152	0.117
wsd(ΔlogY)	0.253	0.161	0.172	0.223	0.161
skew(π/Y)	0.176	0.176	0.176	0.176	0.176
skew(log(Y/Khat))	0.000	0.004	0.020	0.018	0.025
skew(I/K)	1.015	0.465	**2.519**	**2.105**	1.310
skew(dlogY)	0.005	0.001	0.636	0.475	0.304
scorr(π/Y)	1.000	1.000	1.000	1.000	1.000
scorr(log(Y/Khat))	1.000	0.988	0.986	0.988	0.985
scorr(I/K)	-0.062	**0.449**	**0.168**	-0.056	0.255
scorr(ΔlogY)	-0.067	0.059	0.028	-0.015	0.034
bcorr(π/Y, log(Y/Khat))	-0.378	-0.382	-0.381	-0.375	-0.382

Table A.3: Illustration for Identification of Measurement Errors

Parameters	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
	$\boldsymbol{\sigma_{meK} = 0.0}$	$\sigma_{meK} = 0.25$	$\sigma_{meK} = 0.0$	$\sigma_{meK} = 0.0$	$\sigma_{meK} = 0.25$
	$\boldsymbol{\sigma_{meY} = 0.0}$	$\sigma_{meY} = 0.0$	$\sigma_{meY} = 0.25$	$\sigma_{meY} = 0.0$	$\sigma_{meY} = 0.25$
	$\boldsymbol{\sigma_{me\pi} = 0.0}$	$\sigma_{me\pi} = 0.0$	$\sigma_{me\pi} = 0.0$	$\sigma_{me\pi} = 0.25$	$\sigma_{me\pi} = 0.25$
Set of Moments					
mean(π/Y)	0.170	0.170	0.175	0.170	0.175
mean(log(Y/Khat))	0.875	0.872	0.875	0.875	0.873
mean(I/K)	0.116	0.120	0.116	0.116	0.120
mean(ΔlogY)	0.050	0.050	0.050	0.050	0.050
bsd(π/Y)	0.061	0.061	0.067	0.062	0.069
wsd(π/Y)	0.000	0.000	**0.041**	**0.023**	0.047
bsd(log(Y/Khat))	0.678	0.687	0.690	0.678	0.699
wsd(log(Y/Khat))	0.087	**0.216**	**0.233**	0.087	0.306
bsd(I/K)	0.095	0.101	0.095	0.095	0.101
wsd(I/K)	0.123	**0.134**	0.123	0.123	0.134
bsd(ΔlogY)	0.117	0.117	0.167	0.117	0.167
wsd(ΔlogY)	0.161	0.161	**0.370**	0.161	0.370
skew(π/Y)	0.176	0.176	0.851	0.415	0.996
skew(log(Y/Khat))	0.025	0.013	0.018	0.025	0.008
skew(I/K)	1.310	1.651	1.310	1.310	1.651
skew(dlogY)	0.304	0.304	0.035	0.304	0.035
scorr(π/Y)	1.000	1.000	**0.639**	**0.844**	0.568
scorr(log(Y/Khat))	0.985	**0.885**	**0.869**	0.985	0.790
scorr(I/K)	0.255	**0.231**	0.255	0.255	0.231
scorr(ΔlogY)	0.034	0.034	**-0.370**	0.034	-0.370
bcorr(π/Y, log(Y/Khat))	-0.382	-0.376	-0.413	-0.374	-0.399

Table A.4: A Balanced-Panel for 2004-2007

Year	2004	2005	2006	2007
No. of firms	107579	107579	107579	107579
mean(π/Y)	0.155	0.159	0.157	0.160
mean(log(Y/Khat))	1.143	1.145	1.129	1.134
mean(I/K)		0.187	0.161	0.144
mean(ΔlogY)		0.109	0.083	0.097

Note: Top and Bottom 5 percent of observations are trimmed year-by-year for each variable.