A Structural Estimation on Capital Market Distortions in the Chinese Manufacturing*

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

This version: May 2012
VERY PRELIMINARY AND INCOMPLETE

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 identifying and estimating capital market distortions in environments where all the above elements can be present. Using a representative firm-level data in the Chinese manufacturing from 2004 to 2007, we find a magnitude of distortions that is smaller than that in the literature but still substantial. Capital market distortions imply aggregate revenue total factor productivity losses of 40% in the Chinese manufacturing. We also estimate distortions in a comparable sample of Compustat firms. Improving capital allocation efficiency in the Chinese manufacturing to the level among the Compustat firms would increase China's manufacturing output by 31%. Finally, we present empirical evidence that firm size, age, ownership and political connections can significantly affect capital allocation in the Chinese manufacturing.

JEL Classification: E22, D92, C15

Keywords: capital market distortions, aggregate TFPR, unobserved heterogeneities

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

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

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

1 Introduction

Resource allocative efficiency differs across countries. The difference has recently been found important to account 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 (referred to as the ARP approach henceforth). The validity of the calibration hinges on two conditions: (1) average revenue products are proportional to marginal revenue products; (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) only applies to those with homogenous 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 measurement errors. Our goal is to identify the unobserved heterogeneity in 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 full 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 have merely second-order effects on these moments. Not surprisingly, capital adjustment costs and measurement errors mainly manifest themselves in the within-group variations. Therefore, the identification condition on the unobserved heterogeneities, including capital market distortions, carries over to the general case.

[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.

Finally, matching moments on the investment rate and the revenue growth pin down simultaneously capital adjustment costs and idiosyncratic risks, while the within-group standard deviations and the serial correlations discipline measurement errors in the data.

We apply the estimation method to a representative firm-level data in the Chinese manufacturing. In particular, we focus on a balanced panel from 2004 to 2007 covering 107,579 firms. The estimated heterogeneity in capital goods price is significant and sizable. Capital market distortions imply aggregate revenue total factor productivity losses of 40% in the Chinese manufacturing. The magnitude is smaller than that in the literature but still substantial [references to be written]. We also estimate distortions in an comparable sample of Compustat firms. Improving capital allocation efficiency in the Chinese manufacturing to the level among the Compustat firms would increase China's output by 31%.

Despite its potential biases, the ARP approach has a virtue of simplicity. This motivates us to propose a generalized ARP approach, which, on the one hand, takes care of the important unobserved heterogeneities as suggested by the structural estimation and, on the other hand, maintains the tractability. To this end, we calibrate the unobserved heterogeneities by solving a nonlinear equation system which matches a set of between-group moments of the revenue-capital and profit-revenue ratios in a panel. The idea is to back out the heterogeneity in 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 washed out through the time-series average of the revenue-capital ratio within each firm. Applying the generalized ARP approach to the Chinese manufacturing dataset, we find the results to be a first-order approximation of the capital goods price heterogeneity estimated from the full-fledged structural approach.

The good approximation of the generalized ARP approach highlights the importance of using the between-group variation of the revenue-capital ratio to identify capital market distortions. Following the insight, we regress the time-series mean of the capital-revenue ratio of each firm on its characteristics. The purpose is to show what are the policies or institutional arrangements lying hidden behind the veil of distortions. The four-digit industry dummies and the time-series mean of the profit-revenue ratio are added to control (imperfectly) the unobserved heterogeneities in capital output elasticity and markups. We find that in the Chinese manufacturing, small, young and non-state firms tend to face significantly higher capital goods prices than their counterparts that are large, mature, state-owned.

Among the growing literature studying the role of particular distortions, Midrigan and Xu (2009) evaluate the importance of non-convex adjustment costs, financing frictions and uninsurable investment risk. They find such frictions can account for the bulk of within-firm time-series variation in log revenue-capital ratio but at most 10% cross-section dispersion. This motives us to decompose the overall variance in log revenue-capital ratio into time-series and cross-section dimensions, and explicitly model capital market distortions in addition to investment frictions. In terms of estimation, Cooper and Haltiwanger (2006) and Bloom (2009) first adopt the method of simulated moments to recover structural parameters of capital adjustment costs. They show it is possible to distinguish the capital adjustment costs from the stochastic process using information on both investment rate and sales growth rate, which provides an important step for our identification strategy. However, we also 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 discuss the identification conditions. 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 reports empirical evidence for the sources of capital market distortions in the Chinese manufacturing and Section 7 concludes.

2 The Model

Our analysis is based on a monopolistic competition economy 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, such as Abel and Eberly (1994). In this section, we first obtain a static profit function by maximizing instantaneous profit 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. In Appendix 8.1, we show how capital market distortions will affect aggregate TFPR in this model economy.

2.1 Production and Demand

Firm i in period t uses productive capital stock $\hat{K}_{i,t}$, labor $L_{i,t}$ and intermediate input $M_{i,t}$ to produce $Q_{i,t}$ unit of product i, according to a Cobb-Douglas technology with constant returns to scale:

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

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

The product of firm i is demanded in a monopolistic product market according to a 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 the randomness in demand. $P_{i,t}$ denotes the price of product i in period t and $0 < \eta_i < 1$ 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 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 profit:

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

where $Y_{i,t} \equiv P_{i,t} Q_{i,t}$ denotes sales revenue and $\pi_{i,t}$ is variable profit.[2] The first order conditions imply constant intermediate input and labor cost shares:

$$\begin{aligned} \frac{w_{i,t} L_{i,t}}{Y_{i,t}} &= \beta_i (1 - \eta_i), \\ \frac{m_{i,t} M_{i,t}}{Y_{i,t}} &= (1 - \alpha_i - \beta_i)(1 - \eta_i). \end{aligned}$$

The factor shares would reduce to β_i and $1 - \alpha_i - \beta_i$ in the competitive environment with infinitely large demand elasticity ($\eta_i = 0$). Substituting these first-order conditions into (1) yields

$$\frac{\pi_{i,t}}{Y_{i,t}} = \alpha_i (1 - \eta_i) + \eta_i = \eta_i (1 - \alpha_i) + \alpha_i. \tag{2}$$

Equation (2) shows that the variable profit is a constant proportion of revenue, which is codetermined by α_i and η_i. In the limiting case perfect competition, the profit-to-revenue ratio would reduce to α_i. It is worth emphasizing that the labor, intermediate input and profit shares are independent of factor prices as a result of optimal choice for variable inputs.

The optimization establishes the following profit function:

$$\pi_{i,t} = Z_{i,t}^{\gamma_i} \hat{K}_{i,t}^{1-\gamma_i}, \tag{3}$$

[2] Following the investment literature (e.g., Abel and Eberly, 1999), 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 produtivity literature (e.g., Hsieh and Klenow, 2009), Y is simply the quantity of output, which is equivalent to Q in our model.

where

$$\gamma_i \equiv 1 - \frac{\alpha_i(1-\eta_i)}{\eta_i + \alpha_i(1-\eta_i)}, \tag{4}$$

and

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

A combination of equation (2), (3) and (4) leads to the following revenue function:

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

The profit and revenue function have utilized two reparameterization. First, $(1-\gamma_i)$ captures the firm-specific curvature of the functions, which increases with the capital output elasticity α_i and decreases with the inverse of demand elasticity η_i. 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 these components and their stochastic processes, it is ultimately $Z_{i,t}$ that matters in its investment decision. Therefore $Z_{i,t}$ is a summary statistics of the "profitability" (Cooper and Haltiwanger, 2007) or "business environment" (Bloom, 2009)
.

Without the loss of generality,[3] we assume $Z_{i,t}$ follows a trend stationary AR(1) process:

$$\begin{aligned} \log Z_{i,t} &= \mu t + z_{i,t}, \\ z_{i,t} &= \rho z_{i,t-1} + e_{i,t}, \end{aligned} \tag{5}$$

where $0 < \rho < 1$, $e_{i,t} \overset{i.i.d.}{\sim} N(0,\sigma^2)$, and $z_{i,0} = 0$. The standard deviation of the shocks σ is the parameter characterizing the level of uncertainty.

2.2 Capital Market Distortions

There is a long list of factors that may cause capital market distortions. In stead of studying the role of each specific channel, this paper aims to understand the overall effect of all the potential distortions. Therefore similar to Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), we use τ_i to generically refer to the effect of various capital market distortions on the purchase

[3] The stochastic process of $Z_{i,t}$ can be endogenously obtained from its definition, if we assume $A_{i,t}$, $X_{i,t}$, $w_{i,t}$ and $m_{i,t}$ follow a similar trend stationary AR(1) process. For equation (5) to hold, the key assumption is that these four random variables share a common level of persistence, ρ, and the shocks to each of these random variables are independent. In addition, μ and σ^2 could be heterogeneous across firms due to firm-specific output elasticities and markups. Since our interest is to study how shocks to $\log Z_{i,t}$ affect firm's factor demand decisions, we assume homogeneous μ and σ^2 in the benchmark estimation for simplicity. Section 4.4 shows that a relaxation of the assumption will not cause any substantial changes to our main estimation results.

price of capital that are heterogeneous across firms. This implies that the actual capital goods price faced by firm i in period t is

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

where P_t^K is the average capital goods price in the economy. For example, a positive value of τ_i corresponds to a firm with no access to finance hence facing an actual capital goods price higher than the average price; while an investment tax credit is represented by a negative value of τ_i.

2.3 Capital Adjustment Costs

Meanwhile, in a variety settings, capital adjustment costs have been adopted by the investment literature to summarize frictional elements that reduce, delay or protract investment (Khan and Thomas, 2006). Following Cooper and Haltiwanger (2006) and Bloom (2009), we consider three forms of capital adjustment costs that are homogeneous across firms:

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

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.

Since capital goods prices are allowed to differ across firms, the model is disciplined by restricting the capital adjustment cost function, G, to be the same for all firms. If G were also firm-specific, as will be shown in investment decision below, a firm facing high capital adjustment costs would manifest such costs as having a high τ_i^K.

By paying the cost of purchasing capital and adjusting capital, the new investment $I_{i,t}$ contributes to the productive capital stock, $\hat{K}_{i,t}$, immediately in period t, which depreciates at the end of that period.[4] The law of motion for capital is therefore given by

$$K_{i,t+1} = (1-\delta)\hat{K}_{i,t} = (1-\delta)\left(K_{i,t}+I_{i,t}\right), \tag{6}$$

[4] This timing assumption is adopted for three reasons. First, in the absence of capital adjustment costs, the implications of the model would be the same as those of a static economy (Hsieh and Klenow, 2009). In particular, the "efficient" allocation would feature an equalization of MRPK across firms. Under the alternative timing assumption that capital takes one period to build, idiosyncratic shocks may generate heterogeneous MRPK even in the efficient allocation (Collard-Wexler, Asker and De Loecker, 2011). Second, technically, our timing assumption allows for a closed-form solution to the investment problem in the absence of capital

where δ is the constant depreciation rate common across firms.

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. Risk-neutral investors allocate capital until the required rate of return on capital is equalized across different firms. Let the required rate of return be r, at which investors discount future dividends. The investment problem is then defined by the stochastic Bellman equation:

$$V(Z_{i,t}, K_{i,t}) = \max_{I_{i,t}} \left\{ \begin{array}{c} \pi(Z_{i,t}, K_{i,t}; I_{i,t}) - P^K_{i,t} I_{i,t} - G(K_{i,t}; I_{i,t}) \\ +\frac{1}{1+r} E_t \left[V(Z_{i,t+1}, K_{i,t+1}) \right] \end{array} \right\}, \tag{7}$$

where $Z_{i,t+1}$ and $K_{i,t+1}$ follow the law of motion (5) and (6), respectively.

Two remarks are in order. First, we assume that investors are risk-neutral or there is a complete market for risk-averse investors to diversify all the idiosyncratic risks. Section [to be written] will discuss the robustness of our results with respect to market incompleteness and the potential correlation between aggregate and firm-specific shocks.

Second, define J_t the Jorgensonian user cost of capital,

$$J_t \equiv P^K_t - \frac{1-\delta}{1+r} E_t \left[P^K_{t+1} \right]. \tag{8}$$

For simplicity, we assume the average capital goods price to be constant and then normalized it to unity throughout the following analysis. So, according to (8), $J_t = J$, where $J \equiv \frac{r+\delta}{1+r}$. We will check the robustness of our results to the assumption in Section 5.2.

In the presence of capital adjustment costs, there is generally no analytical solution to the optimal investment problem. However, the analytical solution in the case without adjustment costs provides an important benchmark for model properties. If $G(Z_{i,t}, K_{i,t}; I_{i,t}) = 0$, the optimal investment rate is a linear function of $Z_{i,t}$ relative to inherited capital stock $K_{i,t}$:[5]

$$\frac{I_{i,t}}{K_{i,t}} = \left[\frac{1-\gamma_i}{(1+\tau_i) J_t} \right]^{\frac{1}{\gamma_i}} \left(\frac{Z_{i,t}}{K_{i,t}} \right) - 1, \tag{9}$$

adjustment costs, which does not involve any expectation term (Bloom, 2009). This provides a convenient benchmark for the analysis of capital adjustment costs. Finally, in the data, the revenue-to-capital ratio, a key moment for identifying capital market distortions, has similar empirical distribution regardless whether the denominator is $\hat{K}_{i,t}$ or $K_{i,t}$.

[5] The first-order condition that $\hat{K}_{i,t} = \left[\frac{1-\gamma_i}{(1+\tau_i) J_t} \right]^{\frac{1}{\gamma_i}} Z_{i,t}$ establishes (9).

where Equations (9) 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 and is smaller than an otherwise identical firm but facing favorable capital market distortions ($\tau_i < 0$).

In general when $G(K_{i,t}; I_{i,t}) > 0$, the investment policy can be solved out using numerical dynamic programming method. Figures A1.1-A1.3 in the technical appendix illustrate these policies under different forms of adjustment costs. The 45^o straight line is the investment policy in equation (9) and is plotted as the benchmark without capital adjustment costs. As highlighted by these figures, first, irrespective to 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 adjustment. Finally, the optimal investment rate follows a 'barrier control' policy when $b^i > 0$ and a 'jump control' policy when $b^f > 0$. All the results are standard in the investment literature.

3 Structural Estimation

The goal of this paper is to quantify the effect of capital market distortions on aggregate TFPR loss using the above framework.[6] Since the capital market distortions τ_i are not observable directly, one has to infer τ_i from observable variables. The section illustrates why one the one hand the ARP approach provides important insight on the inference of capital market distortions using revenue-to-capital ratio, but on the other hand might deliver a contaminated inference. We then propose a structural econometric approach using the simulated method of moments.

3.1 The ARP Approach

To illustrate the ARP approach, we consider the model without capital adjustment costs. When $G(K_{i,t}; I_{i,t}) = 0$, (7) solves

$$\alpha_i\,(1-\eta_i)\,\frac{Y_{i,t}}{\hat{K}_{i,t}} = (1+\tau_i)\,J. \tag{10}$$

The left- and right-hand sides of (10) represent the marginal revenue product of capital and the firm-specific user cost of capital, respectively. Rearranging (10), we have

$$\log\left(\frac{Y_{i,t}}{\hat{K}_{i,t}}\right) = \log J + \log\left(1+\tau_i\right) - \log\left[\alpha_i\left(1-\eta_i\right)\right]. \tag{11}$$

[6] Appendix 8.1 proves that the aggregate TFPR losses in an economy with homogenous γ is equal to $\frac{1-\gamma}{2\gamma}\sigma^2_{\tau K}$.

(11) is the cornerstone of the ARP approach in the recent misallocation literature. It highlights why investment optimality allows the inference on the unobservable variance of $\log(1+\tau_i)$ from the observable variance of $\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$. However, the key challenge in this indirect inference is that, besides capital market distortions, the unobserved heterogeneities in α_i and η_i will also cause a dispersion in the average revenue product of capital. As will be shown later, in the full-fledged model with capital adjustment costs and measurement errors, the dispersion can further be increased, leading to a biased estimator on capital market distortions.

3.2 A Structural Econometric Approach

In contrast to the simple ARP approach, this paper proposes a structural econometric approach. By fitting the investment model directly to the data on profit-to-revenue ratio ($\pi_{i,t}/Y_{i,t}$), log revenue-to-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 ($\Delta\log Y_{i,t}$), a structural estimation simultaneously recovers unobserved heterogeneities in τ_i, α_i and η_i, capital adjustment costs, and measurement errors. The following empirical specification imposes the structure of the unobserved heterogeneities and measurement errors for estimation.

3.3 Empirical Specification

3.3.1 Unobserved Heterogeneities

There are three forms of unobserved heterogeneities in this model. Instead of estimating specific values of τ_i, α_i and η_i for each firm, our key interest is a consistent estimate for the variance of $\log(1+\tau_i)$. Therefore we assume 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:

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

By definition both capital output elasticity α_i and inverse of demand elasticity η_i are positive numbers between 0 and 1, therefore we assume

$$\log\alpha_i \overset{i.i.d}{\sim} TN\left(\mu_{\log\alpha},\sigma^2_{\log\alpha}\right),$$
$$\log\eta_i \overset{i.i.d}{\sim} TN\left(\mu_{\log\eta},\sigma^2_{\log\eta}\right).$$

That is each firm i has a firm-specific α_i and η_i, where $\log\alpha_i$ is drawn independently from an identical truncated normal distribution with mean $\mu_{\log\alpha}$ and standard deviation $\sigma_{\log\alpha}$; and

$\log \eta_i$ is drawn independently from an identical truncated normal distribution with mean $\mu_{\log \eta}$ and standard deviation $\sigma_{\log \eta}$.

The investment policy under different $(\tau_i, \alpha_i, \eta_i)$ is different. Hence the dynamic programming problem described in equation (7) must be solved for each firm i at each value of $(\tau_i, \alpha_i, \eta_i)$, which is infeasible even for a small sample. Therefore this paper adopts a standard approach used in the literature modelling unobserved heterogeneities in structural estimation, for example, Eckstein and Wolpin (1999), to allow for a finite type of firms. Specifically, in our benchmark specification, we assume there are $3 \times 3 \times 3$ types of firms. Each comprising a fixed proportion $1/(3 \times 3 \times 3)$ of the population, where the type set is defined as $F = \{(\tau_u, \alpha_v, \eta_x) : u = 1, 2, 3; v = 1, 2, 3; x = 1, 2, 3\}$. In section 4.5, we experiment whether the results are robust if we increase the types of firms to $5 \times 5 \times 5$ at the cost of "curse of dimensionality".

3.3.2 Measurement Errors

In addition to a rich structure of heterogeneities, another novelty of our empirical specification is to allow for potential measurement errors in key variables. Our structural estimation employs data on four variables: profit $\pi_{i,t}$, revenue $Y_{i,t}$, capital stock $K_{i,t}$ and investment expenditure $I_{i,t}$. In our benchmark specification, we assume

$$\begin{aligned} K_{i,t} &= K_{i,t}^{true} \exp(e_{i,t}^K), \quad e_{i,t}^K \overset{i.i.d}{\sim} N(0, \sigma_{meK}^2), \\ Y_{i,t} &= Y_{i,t}^{true} \exp(e_{i,t}^Y), \quad e_{i,t}^Y \overset{i.i.d}{\sim} N(0, \sigma_{meY}^2), \\ \pi_{i,t} &= \pi_{i,t}^{true}(1 + e_{i,t}^\pi), \quad e_{i,t}^\pi \overset{i.i.d}{\sim} U(0, \sigma_{me\pi}^2). \end{aligned}$$

Here variables with the "true" superscripts denote the true underlying variables which are not measured accurately in the data; variables without the superscripts denote the observed variables from the data. $e_{i,t}^K$ and $e_{i,t}^Y$ are the 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. $e_{i,t}^\pi$ is the measurement error in profit, which follows a uniform distribution with mean zero and standard deviation $\sigma_{me\pi}$.

There are two features in the specification of the measurement errors. First, the multiplicative structure and the log-normality assumption guarantee positive values of capital stock and sales revenue. Second, we only consider transitory measurement errors so as to distinguish measurement errors from unobserved heterogeneities. In section 4.5, we test whether the results are robust, if we model measurement error in investment instead of capital and allow the measured capital to accumulate the measurement error in investment according to the law of motion of capital.

3.4 Simulated Method of Moments

We apply the simulated method of moments (SMM) to estimate this fully parametric investment model.[7] The key idea of SMM is to estimate deep model parameters by matching simulated moments from the model with empirical moments from the data. To be specific, the **SMM estimator** Θ^* solves the minimal quadratic distance problem (Gouriéroux and Monfort, 1996):

$$\hat{\Theta}^* = \arg\min_{\Theta} \left(\hat{\Phi}^D - \frac{1}{S} \sum_{s=1}^{S} \hat{\Phi}_s^M (\Theta) \right)' \Omega \left(\hat{\Phi}^D - \frac{1}{S} \sum_{s=1}^{S} \hat{\Phi}_s^M (\Theta) \right), \tag{12}$$

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

Suppose the empirical dataset is a panel with N firms and T years. Given the unobserved heterogeneities across firms, the asymptotics is for fixed T and $N \to \infty$. At the efficient choice for the weighting matrix Ω^*, the SMM procedure provides a global specification test of the overidentifying restrictions of the model:

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

[Insert Table 1 here]

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 characterizing the magnitude of capital market distortions, σ_τ; mean and standard deviation of the log capital output elasticity, $\mu_{\log\alpha}$ and $\sigma_{\log\alpha}$; mean and standard deviation of the log inverse of the demand elasticity, $\mu_{\log\eta}$ and $\sigma_{\log\eta}$; three parameters measuring the magnitude of different capital adjustment costs $\left(b^q, b^i, b^f\right)$; the trend growth rate, μ; the standard deviation of shocks or the level of uncertainty, σ; and

[7] 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.

standard deviations of the measurement errors in capital, revenue and profit, σ_{meK}, σ_{meY}, and $\sigma_{me\pi}$.

The lower panel of Table 1 lists $\hat{\Phi}^D$, the set of moments to match. The choice of the moments is guided by two principles. First, $\hat{\Phi}^D$ is a comprehensive set of moments which characterize the distribution and dynamics of key variables that one would expect to match from a well-specified investment model. Second and more importantly, these moments are a priori informative about the parameters that we seek to estimate. Specifically, $\hat{\Phi}^D$ includes the means (*mean*), between-group standard deviations (*bsd*), within-group standard deviations (*wsd*), coefficients of skewness (*skew*) and serial correlation (*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 $\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$. This gives a total of 21 moments.

3.5 Identification

This section lay out the identification conditions for estimating the parameters governing the unobserved heterogeneities, capital adjustment costs and measurement errors, respectively. For illustrative purposes, we start with a model without any of these features and label it as Model A in Table 2.1.[8] In this baseline model there is virtually no variation in $\pi_{i,t}/Y_{i,t}$ and $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$, either across firms or over time. Furthermore, $I_{i,t}/K_{i,t}$ and $\Delta \log Y_{i,t}$ are highly volatile and negatively serially correlated. Through Table 2.1 to Table 2.3, unobserved heterogeneities, capital adjustment costs and measurement errors are added into the model step-by-step. The simple model with the unobserved heterogeneities only allows closed-form solution, which helps to establish analytically the conditions for identifying σ_τ, the parameter of key interest. We then show that the identification conditions in the simple model remain to be the core of recovering σ_τ in the full-blown model with capital adjustment costs and measurement errors restored.

3.5.1 Identification of Unobserved Heterogeneities

A simple model without capital adjustment costs and measurement errors delivers two key equations for identifying unobserved heterogeneities

$$\begin{aligned}
\frac{\pi_{i,t}}{Y_{i,t}} &= \alpha_i(1-\eta_i)+\eta_i = \eta_i(1-\alpha_i)+\alpha_i, \\
\log\frac{Y_{i,t}}{\hat{K}_{i,t}} &= \log J + \log\left(1+\tau_i^K\right) - \log\left[\alpha_i\left(1-\eta_i\right)\right].
\end{aligned}$$

[8] In all the simulations reported in Table 2, we impose $r = 0.15$, $\delta = 0.05$, $\mu_{\log\alpha} = \mu_{\log\eta} = -2.30$, $\rho = 0.90$, $\mu = 0.05$ and $\sigma = 0.30$, simulate a panel of 100000 firms and 24 years, and calculate moments using data in the last 4 years.

These equations imply that, first, none of the unobserved heterogeneities would have any effect on $wsd\,(\pi/Y)$ and $wsd\left(\log\left(Y/\hat{K}\right)\right)$. Only $bsd\,(\pi/Y)$ and $bsd\left(\log\left(Y/\hat{K}\right)\right)$ will vary with these heterogeneities. Second, σ_τ can easily be recovered from $bsd\left(\log\left(Y/\hat{K}\right)\right)$ if $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$ are known, while $bsd\,(\pi/Y)$ is solely determined by $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$. Finally, the additional moment $bcorr\left(\pi/Y, \log\left(Y/\hat{K}\right)\right)$ provides identification to further separate $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$:

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

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

Table 2.1 illustrates these properties by imposing $\sigma_\tau = 0.5$, $\sigma_{\log\alpha} = 0.5$ and $\sigma_{\log\eta} = 0.5$ from column (1) to (3), respectively. In column (1), only $bsd\left(\log\left(Y/\hat{K}\right)\right)$ increases with $\sigma_\tau > 0$. In column (2), $\sigma_{\log\alpha} > 0$ increases both $bsd\,(\pi/Y)$ and $bsd\left(\log\left(Y/\hat{K}\right)\right)$ and causes a negative $bcorr\left(\pi/Y, \log\left(Y/\hat{K}\right)\right)$. In column (3), $\sigma_{\log\eta} > 0$ also increases both $bsd\,(\pi/Y)$ and $bsd\left(\log\left(Y/\hat{K}\right)\right)$ but causes a positive $bcorr\left(\pi/Y, \log\left(Y/\hat{K}\right)\right)$. Finally, under the log normality assumption for α and η, for a given value of Jorgensonian user cost of capital J, the two mean parameters $\mu_{\log\alpha}$ and $\mu_{\log\eta}$ together with the heterogeneity parameters $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$ also joint determines $mean\,(\pi/Y)$ and $mean\left(\log\left(Y/\hat{K}\right)\right)$. Model B lists the moments where $\sigma_\tau = \sigma_{\log\alpha} = \sigma_{\log\eta} = 0.5$.

In summary, in the simple model without capital adjustment costs and measurement errors, the five parameters, $\mu_{\log\alpha}$, $\mu_{\log\eta}$, $\sigma_{\log\alpha}$, $\sigma_{\log\eta}$ and σ_τ are exactly identified by five moments: (1) the mean of $\pi_{i,t}/Y_{i,t}$; (2) the mean of $\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$; (3) the between-group standard deviation of $\pi_{i,t}/Y_{i,t}$; (4) the between-group standard deviation of $\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$; and (5) the cross correlation between the between-group $\pi_{i,t}/Y_{i,t}$ and $\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$.

[Insert Table 2.1 here]

3.5.2 Identification of Capital Adjustment Costs

The key challenge in identifying capital adjustment costs is to distinguish them from the stochastic process. Following Bloom (2009), our key identification strategy is to use information on both investment rate and revenue growth rate:

$$\begin{aligned} \frac{I_{i,t}}{K_{i,t}} &\simeq \Delta \log \hat{K}_{i,t} + \delta = f\left(\Delta \log Z_{i,t}; b^q, b^i, b^f\right) \\ \Delta \log Y_{i,t} &\equiv \gamma_i \Delta \log Z_{i,t} + (1-\gamma_i)\, \Delta \log \hat{K}_{i,t} \end{aligned}$$

The economic rational comes from the fact that $\Delta \log Y_{i,t}$ is a linear combination of $\Delta \log Z_{i,t}$ and $\Delta \log \hat{K}_{i,t}$ where $\Delta \log Z_{i,t}$ depends on the growth rate μ and the standard deviation of idiosyncratic shocks σ. Instead, the investment rate $I_{i,t}/K_{i,t}$, which is the sum of $\Delta \log \hat{K}_{i,t}$ and depreciation rate δ, depends on both $\Delta \log Z_{i,t}$ and the capital adjustment costs $\left(b^q, b^i, b^f\right)$. This implies capital adjustment costs have a first-order effect on $I_{i,t}/K_{i,t}$ but a second-order effect on $\Delta \log Y_{i,t}$. Finally, different investment policies illustrated in Figure A1.1-A1.3 implies the possibility to distinguish the three forms of adjustment costs themselves.

Table 2.2 starts with the Model B and illustrates the moments by imposing $b^q = 0.25$, $b^i = 0.25$ and $b^f = 0.025$ from column (1) to (3), respectively. Across these columns, overall the moments for $I_{i,t}/K_{i,t}$ are much more sensitive than those for $\Delta \log Y_{i,t}$ to changes in capital adjustment costs. This distinguishes the capital adjustment costs from the stochastic process. Comparing different columns, $b^q > 0$ and $b^i > 0$ both decrease $wsd\,(I/K)$ and increase $scorr\,(I/K)$; $b^i > 0$ and $b^f > 0$ both increase $skew\,(I/K)$; while $b^f > 0$ has little effect on $wsd\,(I/K)$ and $scorr\,(I/K)$. This distinguishes different forms of capital adjustment costs from each other.

Finally, to investigate whether the conditions for identifying σ_τ in the simple model are still valid in the presence of capital adjustment costs, we check how capital adjustment costs will affect the five moments listed in section 3.5.1. According to Table 2.2, although $mean\left(\log\left(Y/\hat{K}\right)\right)$, $bsd\left(\log\left(Y/\hat{K}\right)\right)$ and $bcorr\left(\pi/Y, \log\left(Y/\hat{K}\right)\right)$ do vary with capital adjustment costs, the magnitude of the change is very small compared with those initial values in Model B. In contrast, the effect of capital adjustment costs is largely to increase $wsd\left(\log\left(Y/\hat{K}\right)\right)$. This implies that capital adjustment costs have a first-order effect on $wsd\left(\log\left(Y/\hat{K}\right)\right)$, but only a second-order effect on $bsd\left(\log\left(Y/\hat{K}\right)\right)$, while unobserved heterogeneities, σ_τ, $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$, remain to have a first-order effect on $bsd\left(\log\left(Y/\hat{K}\right)\right)$. We then simulate a model with all three forms of capital adjustment costs and label it Model C.

[Insert Table 2.2 here]

3.5.3 Identification of Measurement Errors

Using Model C as benchmark, Table 2.3 illustrates which moments are informative about measurement errors by simulating $\sigma_{meK} = 0.25$, $\sigma_{meY} = 0.25$ and $\sigma_{me\pi} = 0.25$ from column (1) to (3), respectively. Among the three measurement errors, $\sigma_{meK} > 0$ will only affect moments on $\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$, $I_{i,t}/K_{i,t}$; $\sigma_{meY} > 0$ will only affect moments on $\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$, $\pi_{i,t}/Y_{i,t}$ and $\Delta \log Y_{i,t}$; while $\sigma_{me\pi} > 0$ will only affect moments on $\pi_{i,t}/Y_{i,t}$. This implies the possibility to distinguish three measurement errors from each other using moments on these four variables.

To investigate whether the conditions for identifying σ_τ in the simple model are still valid in the presence of measurement errors, we check how measurement errors will affect the five moments listed in section 3.5.1. The finding is that measurement errors make the profit-to-revenue and log revenue-to-capital ratios more dispersed. However, the effects are mainly on the within-group standard deviation instead of the between-group standard deviation. In other words, σ_{meK} and σ_{meY} have a first-order effect on $wsd\left(\log\left(Y/\hat{K}\right)\right)$ but only a second-order effect on $bsd\left(\log\left(Y/\hat{K}\right)\right)$; similarly, σ_{meY} and $\sigma_{me\pi}$ have a first-order effect on $wsd\left(\pi/Y\right)$ but only a second-order effect on $bsd\left(\pi/Y\right)$. In contrast, unobserved heterogeneities, σ_τ, $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$, remain to have a first-order effect on $bsd\left(\log\left(Y/\hat{K}\right)\right)$ and $bsd\left(\pi/Y\right)$.

Furthermore, although both capital adjustment costs and measurement errors have a first-order effect on $wsd\left(\log\left(Y/\hat{K}\right)\right)$, they have different effect on investment rate and revenue growth rate. In particular, $\sigma_{meK} > 0$ and $\sigma_{meY} > 0$ increase $wsd\left(I/K\right)$ and $wsd\left(\Delta\log Y\right)$ and reduce $scorr\left(I/K\right)$ and $scorr\left(\Delta\log Y\right)$, respectively, while capital adjustment costs have the opposite or no effect on these moments, as illustrated in section 3.5.2. This fact distinguishes the measurement errors from the capital adjustment costs.

[Insert Table 2.3 here]

4 Empirical Results

4.1 Data

The empirical exercises of this paper are based on the annual firm-level data from the Chinese Industry Survey (1998-2007). It includes all industrial firms that are identified as being either state-owned, or are non-state firms with sales revenue above 5 million RMB, contributing nearly 90% of the gross output in manufacturing. The survey was implemented by the National Bureau of Statistics on a yearly base since 1998 and a census was conducted in year 2004. We refer it as the NBS dataset hereafter. Appendix 8.2 provides detailed information on how we clean the data and define the variables and why we refine to a panel from year 2004 to 2007 in our main empirical exercises.

4.2 Predetermined Parameters

In addition to those 13 structural parameters listed in Table 1, the depreciation rate δ and the discount rate r also affect the investment decision through the Jorgensonian user cost of capital J. The calibration of δ is based on the law of motion of capital (6)

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

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. This allows us to calibrate δ by matching the difference between the average log investment rate, $\log(1 + I_{i,t}/K_{i,t})$, and the average revenue growth rate, $\Delta \log Y_{i,t}$, which is 0.05 in the data.

Bai, Hsieh and Qian (2005) infer the aggregate real rate of return to capital in China is around 20-25% from 1978 to 2004. This rate of return is even higher for the secondary sector, which includes mining, construction and manufacturing, and tends to increase over time since 1990. Therefore we impose a conservative value $r = 0.20$ for the manufacturing firms in our sample period.

The calibration of ρ follows exactly Cooper and Haltiwanger (2006). Specifically, we estimate a dynamic panel data model of $\log \pi_{i,t}$ by system GMM (Blundell and Bond, 1998). 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 the presence of measurement errors in the profit data which will be discovered by our structural estimation. We therefore impose $\rho = 0.90$ in the benchmark case. A later section considers the sensitivity of the estimates to imposing different values for δ, r and ρ.

4.3 Structurally Estimated Parameters

Table 3 presents our structural estimation results. The first and second columns of the left panel report the optimal estimates of the structural parameters and the corresponding numerical standard errors. Simulated moments at these optimal estimates are listed in the right panel to compare with their empirical counterparts, for which we also calculate the standard errors by bootstrapping.

The estimated σ_τ is significantly different from zero, suggesting the prevalence of the capital goods price heterogeneity. The significant and quantitatively large estimates of $\sigma_{\log \alpha}$ and $\sigma_{\log \eta}$ provide evidence for the presence of the firm-specific capital output elasticity and markups in the data. Under the log-normality assumption, the estimated $\mu_{\log \alpha}$ and $\sigma_{\log \alpha}$ imply that the capital output elasticity in the three-factor production function α_i has a mean of 0.086 and standard deviation of 0.052[9]. The estimated $\mu_{\log \eta}$ and $\sigma_{\log \eta}$ imply the inverse of demand elasticity, η_i, has a mean of 0.078 and standard deviation of 0.065. Overall, the simulated moments provide a close fit to the five core moments for identifying unobserved heterogeneities as discussed in section 3.5.1.

The structural estimation finds two out of the three forms of capital adjustment costs to be quantitatively important. In particular, a combination of quadratic and fixed adjustment costs fit the data best.[10] According to the identification conditions in section 3.5.2, a positive b^q is consistent with the fact that both the investment rate and revenue growth rate are positively serial correlated, while a positive b^f is driven by the larger skewness of investment rate compared with that of revenue growth rate. Quantitatively, the estimate of b^q implies that quadratic adjustment costs increase the user cost of capital by 4.5%; the estimated b^f suggests any investment or disinvestment will cause a loss of 3.4% of the variable profit in that period.

The estimated μ is 0.08, which implies the model is able to capture the high economic growth rate in China. At this growth rate, the model simulates higher investment growth rate but lower revenue growth rate compared with those in the data. σ is estimated to be 0.42, which implies firms in our sample do face idiosyncratic shocks. At this level of uncertainty, the model generates slightly higher within-group standard deviations of investment rate and revenue growth rate, but slightly lower between-group standard deviations of these two variables.

Two out of three measurement errors we consider turn out to be significantly different

[9] Both the average and dispersion values of α are close to those in the literature that estimates 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 found 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. Using such an aggregate model, they found a capital share of 0.385 for the U.S. economy.

[10] Similar to Cooper and Haltiwanger (2006) and Bloom (2009), we also find 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 the Compustat. Consistent with the fact that 90% firms in our sample are reported to be single-plant enterprises, we find a combination of $b^q > 0$ and $b^f > 0$ fits the data best as Cooper and Haltiwanger (2006) do.

from zero. Consistent with the usual concern that capital stock is poorly measured in firm-level data, σ_{meK} turns out to be at the similar magnitude of σ, which implies a noise-to-signal ratio around 1. In contrast, the model estimates a virtually zero σ_{meY}, which means sales revenue is much better measured in the data. Since variable profit is defined as the difference between sales revenue and costs of goods sold, while costs of goods sold include many sub-items, it is not surprising that the model finds a large value for $\sigma_{me\pi}$ as well. Consistent with the identification conditions in section 3.5.3, with the presence of these measurement errors, the model is able to closely match the within-group standard deviations of the four variables.

[Insert Table 3 here]

4.4 Specification Tests

There are three new features in this paper compared with the existing distortion literature: unobserved heterogeneities in capital output elasticity and inverse of demand elasticity, capital adjustment costs and measurement errors. Table 4 reports specification tests for three restricted models, in order to understand the effect of missing each feature on the estimates of the capital market distortion. The preferred full model is listed in column (1) as benchmark.

Column (2) shows the results of imposing no unobserved heterogeneities in capital output elasticity and inverse of demand elasticity, that is $\sigma_{\log\alpha} = \sigma_{\log\eta} = 0$. As a sharp contrast to the benchmark, the estimated σ_τ increases significantly from 0.706 to 0.924 in this restricted model. The model also severely overestimates capital adjustment costs, measurement errors in revenue and profit, and underestimates the growth rate and level of uncertainty. Although the model can still fit the general features of log revenue-to-capital ratio, it fails to match the pattern in the dispersion and persistence of profit-to-revenue ratio. Neither can it match the negative correlation between these two variables. As a result, the overall fit of this restricted model degenerates enormously.

Column (3) reports the results of imposing no capital adjustment costs, that is $b^q = b^i = b^f = 0$. This model substantially underestimates the level of uncertainty and overestimates the measurement errors in revenue. However, the estimate for σ_τ is just 7% lower than that of the benchmark result. This is because we have used information on variables both in levels, namely profit-to-revenue ratio and log revenue-to-profit ratio and in growth rates, namely investment rate and revenue growth rate. A model missing capital adjustment costs fails to match moments on variables in growth rates, but is still able to fit moments on variables in levels, which are mainly determined by the five parameters governing production, demand and user cost of capital.

Column (4) reports the results of imposing no measurement errors, that is $\sigma_{meK} = \sigma_{meY} = \sigma_{me\pi} = 0$. This model generates too little within-group standard deviations but too much serial correlations for profit-to-revenue ratio and log revenue-to-capital ratio. Although it tends to generate different patterns for capital adjustment costs and the stochastic process, the estimate for σ_τ only slightly deviates from the benchmark result. This is because we have separated the within-group standard deviations from the between-group standard deviations. A model without measurement errors fails to match within-group standard deviations, but is still able to fit the between-group standard deviations, which are mainly determined by unobserved heterogeneities.

[Insert Table 4 here]

4.5 Robustness Tests

Table 5 presents results for a set of robustness checks. Recall the benchmark model in column (1) has imposed depreciation rate $\delta = 0.05$, discount rate $r = 0.20$, and serial correlation $\rho = 0.90$. Columns (2) and (3) show the results for the same model but imposing $\delta = 0.03$ and 0.07, respectively. In order to match the average investment rate and revenue growth rate, the estimates for the growth rate μ decreases when the depreciation rate increases, as one may expect. Nevertheless, the key parameter of interest σ_τ is robust to the choice of depreciation rate. Columns (4) and (5) test the sensitivity of imposing $r = 0.15$ and 0.25, respectively. The estimated σ_τ tends to increase with the discount rate. However a 50% change in r only causes a less than 10% change in σ_τ. Columns (6) and (7) report the results by imposing $\rho = 0.85$ and 0.95, respectively. Although there is a modest variation in the estimates for capital adjustment costs and measurement errors across different values of ρ, the estimated σ_τ is not sensitive to the choice of serial correlation.

[Insert Table 5 here]

Column (8) and (9) investigate whether the empirical results would change if the number of type in each heterogeneity is increased from 3 to 5, and if. the number of path for simulation is increased from 5 to 10. These two changes will increase the time for estimation by around 2 and 1.5 times, respectively. But they cause virtually no change in any of the estimates. Column (10) and (11) ask how the introduction of unobserved heterogeneities in the growth rate and level of uncertainty will affect our empirical findings. We therefore assume the growth rate follows a uniform distribution with mean μ and standard deviation σ_μ; and the level of uncertainty follows a uniform distribution with mean σ and standard deviation σ_σ. Introducing an additional dimension of heterogeneity implies the numerical dynamic programming has an additional state variable, therefore the time for estimation increases by around 2.5 times. As reported in Table 5, although a model with heterogeneous growth rate and level of uncertainty fits the data better, our key parameter of interest σ_τ is robust to these additional heterogeneities. Finally, column (12) studies what would happen to the estimates if we model measurement error in investment, and if such measurement error contaminates the measured capital stock over time. The estimated capital adjustment costs are much higher under this specification; however the variation in σ_τ is less than 5%. Overall, a model with measurement error in investment fits the data much worse compared with a model with measurement error in capital stock, which has been adopted in our benchmark specification.

[Insert Table 5-continue here]

4.6 Counterfactual Simulations

The estimated structural model provides a useful framework to quantify the effects of distortions on aggregate TFPR. Table 6.1 simulates such effects according to equation (14). Since there are heterogeneities in both capital output elasticity and inverse of demand elasticity, these effects are simulated for different type of firms and the average effects are reported in the last row.

Evaluating at the optimal estimates listed in Table 3, our investment model predicts that the actual aggregate TFPR in China is 39.7% lower than the efficient benchmark, due to presence of capital market distortions and capital adjustment costs. A model with capital market distortions only simulates a 38.2% loss in aggregate TFPR, while a model with capital adjustment costs only generates a 2.1% loss. This suggests that although both distortions and frictions will cause aggregate TFPR loss, quantitatively, the vast majority of the loss is due to the capital market distortions. Had we not controlled for potential unobserved heterogeneities in capital output elasticity and inverse of demand elasticity, the estimated σ_τ would have implied a magnitude two-thirds larger. All else being equal, the losses in aggregate TFPR increase monotonically with $1-\gamma$, the capital elasticity in the profit or revenue function. Intuitively, an economy made of firms with larger capital share in production function and less market power in product market demands more capital stock hence suffer more from capital market distortions.

[Insert Table 6.1 here]

Our finding that capital adjustment costs cause an aggregate TFPR loss around 1-3% is very similar to that estimated in Midrigan and Xu (2009). In contrast, our estimation on the effect of capital market distortions is much smaller than what is calibrated in Hsieh and Klenow (2009). Of course, one importance difference lies in that, Hsieh and Klenow (2009) claim to estimate the aggregate TFPR loss due to distortions in both product market and capital market, while this paper only claims to estimate the effect of capital market distortions. Had we interpreted any difference in firm's market power as the result of product market distortions, the heterogeneity in η_i in our specification would be isomorphic to the heterogeneity in τ_i^Y – the measure of product market distortions – in Hsieh and Klenow (2009). Then, our model would predict a 53.2% aggregate TFPR loss due to distortions in both capital and product market distortions.

5 The Generalized ARP Approach

Our structural estimation finds a statistically significant heterogeneity in the capital goods price. Eliminating the estimated heterogeneity, which we interpret as capital market distortions, would increase the aggregate TFPR by 38.2%. The sizable efficiency gain from capital reallocation naturally raise the following question: What have caused the capital market distortions? This section develops a generalized ARP approach, which allows us to address a set of important issues regarding question.

Section 3.5 has illustrated that in principle, among the three novel features considered in this paper, capital adjustment costs are crucial for matching moments on variables in growth rate; measurement errors are critical in matching moments on the time-series dimension; while unobserved heterogeneities are essential in matching moments on the cross-section dimension. The specification tests in section 4.4 further establish that in our empirical exercise, a model without the unobserved heterogeneities in capital output elasticity and inverse of demand elasticity will seriously overestimate the unobserved heterogeneity in capital goods prices. In contrast, a model without capital adjustment costs or measurement errors does not necessarily lead to such bias, if we separate variables in levels from those in growth rate, and if we separate between-group standard deviations from within-group standard deviations.

This implies the exact identification conditions for the five parameters established in section 3.5.1 using the five moments are the core of recovering σ_τ, even if there are capital adjustment costs and measurement errors. In other words, one would pin down the five parameters by solving the following five simultaneous equations

$$\begin{aligned}
mean\,(\pi/Y) &= F_1\left(\mu_{\log\alpha},\mu_{\log\eta},\sigma_{\log\alpha},\sigma_{\log\eta},\sigma_\tau\right), \\
mean\left(\log\left(Y/\hat{K}\right)\right) &= F_2\left(\mu_{\log\alpha},\mu_{\log\eta},\sigma_{\log\alpha},\sigma_{\log\eta},\sigma_\tau\right), \\
bsd\,(\pi/Y) &= F_3\left(\mu_{\log\alpha},\mu_{\log\eta},\sigma_{\log\alpha},\sigma_{\log\eta},\sigma_\tau\right), \\
bsd\left(\log\left(Y/\hat{K}\right)\right) &= F_4\left(\mu_{\log\alpha},\mu_{\log\eta},\sigma_{\log\alpha},\sigma_{\log\eta},\sigma_\tau\right), \\
bcorr\left(\pi/Y,\ \log\left(Y/\hat{K}\right)\right) &= F_5\left(\mu_{\log\alpha},\mu_{\log\eta},\sigma_{\log\alpha},\sigma_{\log\eta},\sigma_\tau\right).
\end{aligned}$$

We name this set of identification conditions as the generalized ARP approach, recognizing its important inheritance from the conventional ARP approach. To check the validity of the generalized ARP approach, Table 7.1 compares the estimates of these five parameters from full structural estimation and the generalized ARP approach. These two approaches generate very similar estimates for $\mu_{\log\alpha}$and $\mu_{\log\eta}$. Using the full structural estimates as benchmark, the generalized ARP approach slightly underestimates σ_τ, $\sigma_{\log\alpha}$ and overestimates $\sigma_{\log\eta}$. However, the bias for all these three parameters is only around 5%. In this sense, we claim the generalized ARP approach provides the first-order approximation to the five parameters of our key interest.

[Insert Table 7.1 here]

There are several advantages of this approach, which make it a useful tool to study capital market distortions. First, it takes into account the heterogeneities in production technology and market power, which have been demonstrated to cause large bias in σ_τ in the conventional ARP approach. Second, by using the between-group standard deviation of profit-to-revenue ratio and log revenue-to-capital ratio, it filters the effect of capital adjustment costs and measurement errors. Therefore one could still get a first-order approximation for σ_τ without doing the full structure estimation. Third, it only requires panel information on profit, revenue and capital stock, which are widely available in most firm-level dataset.

5.1 Capital Market Distortions in Different Sectors

The generalized ARP approach allows us to explore the capital market distortions in different sectors, without going to the full structural estimation. Table 7.2. reports the estimates of the five parameters for six sectors at the 4-digit level. The first important finding is that the generalized ARP approach generates large and positive values for $\sigma_{\log\alpha}$ and $\sigma_{\log\eta}$ for each sector. This implies even within the 4-digit level, there are still substantial heterogeneities in capital output elasticity and inverse of demand elasticity. It therefore highlights the importance of modelling such heterogeneities as firm-specific instead of sector-specific.

Second, the variation in the values of $\mu_{\log\alpha}$ and $\mu_{\log\eta}$ across sectors are consistent with our conventional observation. For example, compared with the average values for all the sectors, both $\mu_{\log\alpha}$ and $\mu_{\log\eta}$ are higher in automobile parts, wine and tobacco, which means firms in these sectors are more capital intensive and have more market power. The paper sector has higher $\mu_{\log\alpha}$ but lower $\mu_{\log\eta}$, which implies this is a relatively capital intensive sector but faces a more competitive product market. The garment sector has lower $\mu_{\log\alpha}$, consistent with the fact that it is usually a labor and material intensive sector.

Finally, the key parameter of interest σ_τ varies from 0.61 to 0.69 across five out of six sectors, which is about the same magnitude for the full sample. However, the generalized ARP finds a virtually zero σ_τ in the tobacco sector, a indication of almost no capital market distortion. As discussed in section 6, we think this is probably driven by the fact that 98% firms in tobacco sector are state-owned enterprises.

[Insert Table 7.2 here]

5.2 The Evolution of Capital Market Distortions

The firm-specific τ_i implies the capital goods price heterogeneity to be time-invariant. Capital market efficiency may change over time. For instance, Hsieh and Klenow (2009) have found the 2005 revenue-to-capital ratio in Chinese manufacturing firms to be less dispersed than the 1998 ratio, an indication of improving capital market efficiency. We now apply the generalized ARP approach to study how the capital market distortions have been evolving from 1998 to 2007.

Since this approach is only applicable to panel data, we split the NBS dataset into three periods, each of which is made of four years:1998-2001, 2001-2004, and 2004-2007. To maintain compatibility across different periods, we clean the data in 1998-2001 and 2001-2004 with the same criteria as we do with 2004-2007, the benchmark sample of our empirical exercise.

The lower panel of Table 7.3 lists the five moments in each period. Although the between-group dispersion of log revenue-to-capital ratio indeed decreases over time, so does the between-group dispersion of the profit-to-revenue ratio. Another salient feature of the moments is that the average log revenue-to-capital ratio has increased substantially over time. If the production technology and market condition are broadly constant over time, this is a sign of higher required rate of return to capital. Indeed according to Bai, Hsieh and Qian (2005), the aggregate required rate of return to capital has increased by 10% from 1998 to 2004 for the secondary sector. Therefore we impose $r = 0.10$ and 0.15 for 1998-2001 and 2001-2004, respectively. The generalized ARP approach predicts $\sigma_\tau = 0.954$ and 0.7608 for these two earlier periods. Taking $\sigma_\tau = 0.684$ in the most recent period as benchmark, we conclude an improvement in the capital market efficiency in China since later 1990.[11]

[Insert Table 7.3 here]

5.3 Heterogeneities in J and the Compustat Benchmark

One caveat in our empirical strategy is to attribute all the unobserved heterogeneity in the user cost of capital to firm-specific capital goods prices $(1+\tau_i)$, and assume a common Jorgensonian user cost of capital J in the key identification condition equation (11). Since $J = \frac{r+\delta}{1+r}$, if there was any intrinsic heterogeneity in either δ or r even in the absence of any policy and institution distortion, our estimated σ_τ would overestimate the magnitude of capital market distortions of our true interest.

Such concern could be relevant for δ. For example, different firms may have different capital stock combinations of plant and equipment, which naturally depreciate at different rates. Such concern could also be relevant for r. If we relax the assumption of risk-neutrality and if there are aggregate shocks, different firms may have heterogeneous r induced by firm-specific beta. Theories that take into account asymmetric information in the capital market would also endogenously predict different r across firms with different characteristics, for example, age and size.

Instead of netting out these possibilities from our estimated σ_τ directly, our fundamental interest is to understand those non-intrinsic policies and institutions that have caused the capital market distortions. Since publicly-traded firms in the Compustat are usually taken as a benchmark with least distortions, Table 7.4. therefore applies the generalized ARP approach to a panel of Compustat firms from 2002 to 2005.[12] We consider three different samples. First, a full sample without trimming any firms. Second, a sample of firms with sales revenue more than 1 million US dollars in 2004 price. It is therefore broadly comparable with the NBS sample, which only includes firms with sales revenue more than 5 million RMB in 1998 price. Finally, following Bloom (2009), a sample with sales revenue more than 10 million US dollars

[11] Had we not allowed r to increase over time, the GMRPK approach would predict even higher σ_τ for earlier periods, which indicates even more capital market efficiency gain over time.

[12] 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: $K_{i,t} = (1-\delta) K_{i,t} (P_{m,t}/P_{m,t-1}) + I_{i,t}$, initialized using the net book value of capital, where $I_{i,t}$ is net capital expenditure on plant, property, and equipment, and $P_{m,t}$ is the industry-level capital goods deflator from Bartelsman, Becker and Grey (2000). Sales revenue and cost of goods sold figures come from accounts after deflation using the CPI. We consider a sample from 2002 to 2005 instead of from 2004 to 2007 as in China, because the $P_{m,t}$ is not available after 2005.

in 2000 price and number of employees more than 500. It is therefore a homogeneous sample only made of large firms.

[Insert Table 7.4 here]

All the three samples of Compustat have higher profit-to-revenue ratio and lower log revenue-to-capital ratio, compared with the NBS dataset. Not surprisingly, this implies on average firms in Compustat are more profitable and investors have a lower required rate of return on capital. Across different samples, when more small firms are trimmed out, the parameters characterizing unobserved heterogeneities get smaller. In particular, σ_τ decreases from 0.461 in the full sample, to 0.311 in the NBS comparable sample and to almost zero in a sample with homogeneous firms.

Given that the second Compustat sample is most comparable with the NBS dataset, we take it as our benchmark to do a back-of-the envelope calculation. Recall that the generalized ARP approach has predicted $\sigma_\tau = 0.684$ in China. Under the assumptions that, first, the intrinsic heterogeneities in δ and r are similar for firms in comparable samples; second, firms in Compustat face no policy or institution distortion; and finally, the the intrinsic heterogeneities in δ and r are uncorrelated with policy and institution distortions, the proportion of the heterogeneity in the user cost of capital driven by China-specific distortions is

$$\frac{\sqrt{0.684^2 - 0.311^2}}{0.684} \times 100\% = 89\%.$$

With the Compustat benchmark, we can also investigate hypothetical questions by controlled experiment. For example, what would happen if these Chinese firms had been operating in an environment such as those in the Compustat. Table 6.2. simulates the aggregate TFPR loss in China by reducing σ_τ to 0.311. We find that all else being equal, averaging across different type of firms, the aggregate TFPR losses in China would decline to 7.7%. This implies that without any additional investment, the GDP of China would increase by 30.5% if the existing aggregate capital stock in China could be reallocated across existing Chinese firms to equalize their user cost of capital similar to the level in the U.S..

6 Regressions on Firm Characteristics

The above exercises indicate that, first, the majority of the estimated capital goods price heterogeneity in China are associated with its policy and institution distortions in the capital market. Second, the efficiency gain would be substantial if such distortions could be eliminated. Since the distortions we model are generic and unobservable, it motives us to link such distortions with some observable firm characteristics for policy implications.

Consider the following reduced-form regression transformed from the generalized ARP approach,

$$\frac{1}{T}\sum_{t=1}^{T}\log\left(Y_{i,t}/\hat{K}_{i,t}\right) = b_0 + b_1 \cdot X_i + b_2 \cdot \frac{1}{T}\sum_{t=1}^{T}\log\left(\pi_{i,t}/Y_{i,t}\right) + b_3 \cdot D_i + \xi_i, \tag{13}$$

where X_i is a vector of firm characteristics and D_i represents a vector of industry and province dummies; $\frac{1}{T}\sum_{t=1}^{T}\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$ and $\sum_{t=1}^{T}\log\left(\pi_{i,t}/Y_{i,t}\right)/T$ are firm i's times-series means of the log revenue-to-capital and profit-to-revenue ratios. We have known from equation (11) that $\frac{1}{T}\sum_{t=1}^{T}\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$ entails not only $\log(1+\tau_i)$, but also α_i, η_i and the effect of capital adjustment costs, and measurement errors. Regression (13) adds $\frac{1}{T}\sum_{t=1}^{T}\log\left(\pi_{i,t}/Y_{i,t}\right)$ and the industry and province dummies to control the heterogeneities in α_i and η_i. In addition, (13) ignores capital adjustment costs and measurement errors, which have been found to have second-order effects on the between-group variation of the log revenue-to-capital ratio. The above two procedures therefore can be considered an approximation of $\log(1+\tau_i)$ by $\frac{1}{T}\sum_{t=1}^{T}\log\left(Y_{i,t}/\hat{K}_{i,t}\right)$. Following the logic, $b_1 \cdot X_i$ can then proxy the effect of X_i on firm i's the user cost of capital.

[Insert Table 8 here]

Table 8 presents the regression results. In the baseline model which only controls for industry and province dummies, $\frac{1}{T}\sum_{t=1}^{T}\log\left(\pi_{i,t}/Y_{i,t}\right)$ has the correct negative sign. The second regression considers the effect of firm age and size. All else being equal, it predicts the capital good price of a firm is 3% lower if a firm is one year older, and 4% lower if a firm has 1000 more employees. This is consistent with the typical findings in the large literature on capital market imperfections, for example, Fazzari, Hubbard and Peterson (1988), that younger and smaller firms tend to face higher user cost of capital due to financial constraints.

The third regression tests whether a firm with a higher beta tends to have a higher user cost of capital. Without information on firm value, it is not possible to get an exact beta for firms in our dataset. However, we construct a quasi-beta using sales revenue as illustrated in the Appendix 8.3. The empirical results indicate that all else being equal, a firm with a larger beta does have a higher user cost of capital.

Dummy variables for state-owned firms (SOE), collective-owned firms (COE), domestic private-owned firms (DPE), Hong Kong, Macau and Tai Wan-owned firms (HMT), and foreign-owned firms (FIE) are included in the fourth regression to study the effect of ownership.

Compared with the default category, which includes all other type of firms, such as share holding firms, the user cost of capital for SOE, HMT and FIE are 25%, 17% and 15% lower, while COE and DPE are paying a user cost of capital 23% and 18% higher. A large body of literature, such as Dollar and Wei, 2006, Song et al., 2011, has pointed out that in the developing economy of China, SOE have a much better access to external financing than DPE due to capital market distortions. Our empirical findings provide further information on COE, HMT and FIE and the relative user cost of capital facing each type of firms.

The fifth regression entertains the hypothesis that a firm with a political connection with the communist party faces a favorable distortions therefore has a lower user cost of capital. We use whether there is a labor union in the firm as a proxy for such political connection. Instead of being a worker association that bargains with employers over wages, benefits and working conditions, the labor union in the context of China has a different function.[13] It serves an important channel through which the communist party influences the firm. Therefore having a labor union in the firm can be an indication of its political connection. Our regression result indicates that all else being equal, a firm with a labor union has a 16% lower user cost of capital than otherwise. Using party membership as an alternative measure, Li, Meng, Wang and Zhou (2008) also find political connection to be relevant in China.

The last regression includes additional interaction terms of ownership and labor union, which investigates to what extent the different user cost of capital across ownership is driven by their political connection. The default category is therefore other type of firms without a labor union. By normalizing the user cost of capital for this category as one, Figure 1 plots the predicted user cost of capital across different type of firms, with and without a labor union, using the estimated coefficients from this regression. Interestingly, although on average firms with a labor union do have a lower user cost of capital, the effect of having a labor union is very heterogenous across different ownerships. To be specific, conditional on ownership, having a labor union reduces the user cost of capital by only 4% for HMT, 11% for FIE, but 16% for COE and DPE, and as large as 43% for SOE. Our preferred interpretation is that at least for domestic firms (SOE, COE and DPE), an important channel of capital market distortions is the political connection with the communist party.

[Insert Figure 1 here]

[13] According to the Labor Union Law of China modified in 2001, a labor union is an association made of workers at a voluntary base and led by the communist party of China; it is an important bridge that connects the party and the workers; it represents and protects the right of the workers; and it stablizes and harmonizes the relationship between the employees and employers.

7 Conclusion

[to be written]

8 Appendix

8.1 Aggregate TFPR Losses

Both distortions and frictions may cause aggregate productive efficiency loss. To quantify these effects, consider N firms in the economy with same α, η thus γ. Define the "efficient" benchmark as the capital allocation in an economy without capital adjustment costs nor 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:

$$\begin{aligned}\log Y^*_t - \log Y_t &= \log Z^\gamma_t - \log (Z^*_t)^\gamma \\ &= \log\left(\sum_{i=1}^N \left(Z^\gamma_{i,t}\hat{K}^{1-\gamma}_{i,t}\right)\right) - (1-\gamma)\log\left(\sum_{i=1}^N \hat{K}_{i,t}\right) - \gamma\log\left(\sum_{i=1}^N Z_{i,t}\right) \quad (14)\end{aligned}$$

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:

$$\Delta \log TFPR_t = -\frac{1}{2}\frac{1-\gamma}{\gamma} Var\left[\log\left(1+\tau_i\right)\right] \tag{15}$$

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

8.2 Data

Brandt, Van Biesebroeck and Zhang (2012) provide an excellent description on the dataset and implement a series of consistency checks. We therefore 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 cost of labor and intermediate inputs. However, 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 is reported by all the firms in the dataset. By accounting definition, cost of goods sold (COGS) refer 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 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 Statistic Yearbook*. The survey does not contain information on investment expenditure. 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. We therefore 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 from the data since year t_1.

Investment expenditure $I_{i,t}$ is then recovered according to equation (6). We experimented with a depreciation rate δ from 3% to 10% and pin it down to be 5%, which is the average difference between the constructed investment rate and revenue growth rate.

Four key variables for estimation, namely profit-to-revenue ratio $(\pi_{i,t}/Y_{i,t})$, log revenue-to-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 $(\Delta \log Y_{i,t})$, are then constructed by definition. We exclude outliers by trimming the top and bottom 5% observations for each variable in each year. Table A.1. reports the sample average for these variables from 1998 to 2007.

[Insert Table A here]

The implications for efficient capital allocation are for firms on the balanced-growth-path. In the presence of capital adjustment costs, it may cost several 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. The corresponding data description is provided in Table A.2. Furthermore, our investment model does not consider entry and exit, which means the model implications are only valid for existing and ongoing firms. Table A.3 and A.4 therefore report the four variables for firms that are at least 5 years old upon enter our dataset and survive 10 years and at least 4 years, correspondingly. A striking fact is that even for firms surviving the entire 10 years, their average log revenue-to-capital ratio–the key variable for our identification–is increasing over time. However, this ratio begins to stabilize since year 2004. We think it is probably driven by two facts. First, there has been massive privatization in China since the economic reform and many firms began to export after China's entry to the WTO in 2001. This implies substantial structural changes have been going on until early 2000. Second, many existing non-state-owned firms with sales revenue beyond 5 millions were missing from the survey in early years but were included in our dataset since 2004 thanks to the industrial census conduct in that year.

For these reasons, in our main empirical exercises, we only utilize a sample for firms surviving 2004 to 2007 and at least 5 years old in 2004. This gives us a balanced panel made of 107579 firms and spanning 4 years.

8.3 Construction of a Quasi-beta

Consider the following specification for $Z_{i,t}$:

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

where e_t is an aggregate shock common to all the firms and is independent of idiosyncratic shocks $e_{i,t}$; and λ_i is a firm-specific loading draw from a distribution with first moment μ_λ and second moment σ_λ^2.

When $\rho \to 1$, the revenue growth rate can be approximated as

$$\Delta \log Y_{i,t} = \Delta \log Z_{i,t} = \mu + \lambda_i e_t + e_{i,t}.$$

And the average revenue growth rate is therefore

$$\Delta \log Y_t = \frac{1}{N} \sum_{i=1}^{N} \Delta \log Y_{i,t} = \mu + e_t \frac{1}{N} \sum_{i=1}^{N} \lambda_i = \mu + \mu_\lambda e_t$$

This implies that one potential feasible strategy is to proxy $beta_i$ as

$$\begin{aligned} beta_i &= \frac{cov\left[\Delta \log Y_{i,t}, \Delta \log Y_t\right]}{var\left[\Delta \log Y_t\right]} \\ &= \frac{cov\left[\lambda_i e_t, \mu_\lambda e_t\right]}{var\left[\mu_\lambda e_t\right]} \\ &= \frac{\lambda_i \mu_\lambda var\left[e_t\right]}{\mu_\lambda^2 var\left[e_t\right]} \\ &= \frac{\lambda_i}{\mu_\lambda} \end{aligned}$$

The rationale is that a firm that has a pro-cyclical revenue ($\lambda_i > 0$) is a risky firm, and tends to have a positive beta. Thus investors should demand a higher rate of return from investments whose performance is strongly tied to the performance of the economy.

References

[1] Abel, Andrew B. and Eberly, Janice C. (1999), "The Effects of Irreversibility and Uncertainty on Capital Accumulation", *Journal of Monetary Economics*, 44, 339-377.

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

[3] Banerjee, Abhijit V. and Duflo, Esther (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.

[4] 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.

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

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

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

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

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

[10] Brandt, Loren, Tombe, Trevor and Zhu, Xiaodong (2010), "Factor Market Distortions Across Time, Space and Sectors in China", Working Papers, University of Toronto.

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

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

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

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

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

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

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

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

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

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

[21] Li, Hongbin, Meng, Lingsheng, Wang, Qian, Zhou, Li-An (2008), "Political connections, financing and firm performance: Evidence from Chinese private firms", *Journal of Development Economics*, 87, 283-299.

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

[23] Midrigan, Virgiliu and Xu, Daniel Yi (2009), "Accounting for Plant-Level Misallocation", Working Paper, New York University.

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

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

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

[27] Zhang, Junsen, Zhao, Yaohui, Park, Albert, and Song, Xiaoqing (2005), "Economic Returns to Schooling in Urban China, 1988 to 2001", *Journal of Comparative Economics*, 33, 730-752.

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 2.1. Illustration for Identification of Unobserved Heterogeneities

Parameters	**Model A**	**col (1)**	**col (2)**	**col (3)**	**Model B**
	$\boldsymbol{\sigma_{\tau} = 0.0}$	$\sigma_{\tau} = 0.5$	$\sigma_{\tau} = 0.0$	$\sigma_{\tau} = 0.0$	$\sigma_{\tau} = 0.5$
	$\boldsymbol{\sigma_{\log \alpha} = 0.0}$	$\sigma_{\log \alpha} = 0.0$	$\sigma_{\log \alpha} = 0.5$	$\sigma_{\log \alpha} = 0.0$	$\sigma_{\log \alpha} = 0.5$
	$\boldsymbol{\sigma_{\log \eta} = 0.0}$	$\sigma_{\log \eta} = 0.0$	$\sigma_{\log \eta} = 0.0$	$\sigma_{\log\eta} = 0.5$	$\sigma_{\log\eta} = 0.5$
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 2.2. Illustration for Identification of Capital Adjustment Costs

Parameters	Model B	col (1)	col (2)	col (3)	Model C
	$\boldsymbol{b^q = 0.0}$	$b^q = 0.25$	$b^q = 0.0$	$b^q = 0.0$	$b^q = 0.25$
	$\boldsymbol{b^i = 0.0}$	$b^i = 0.0$	$b^i = 0.25$	$b^i = 0.0$	$b^i = 0.25$
	$\boldsymbol{b^f = 0.0}$	$b^f = 0.0$	$b^f = 0.0$	$b^f = 0.025$	$b^f = 0.025$
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 2.3. Illustration for Identification of Measurement Errors

Parameters	**Model C**	**col (1)**	**col (2)**	**col (3)**	**Model D**
	$\boldsymbol{\sigma_{meK} = 0.0}$	$\sigma_{meK} = 0.25$	$\sigma_{meK} = 0.0$	$\sigma_{meK} = 0.0$	$\sigma_{meK} = 0.25$
	$\boldsymbol{\sigma_{meY} = 0.0}$	$\sigma_{meY} = 0.0$	$\sigma_{meY} = 0.25$	$\sigma_{meY} = 0.0$	$\sigma_{meY} = 0.25$
	$\boldsymbol{\sigma_{me\,\pi} = 0.0}$	$\sigma_{me\,\pi} = 0.0$	$\sigma_{me\,\pi} = 0.0$	$\sigma_{me\,\pi} = 0.25$	$\sigma_{me\,\pi} = 0.25$
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 3. Benchmark Results

Parameters	estimate	s.e.
σ_{τ}	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	0.4253	0.0016
σ_{meK}	0.4010	0.0013
σ_{meY}	0.0007	0.1255
$\sigma_{me\,\pi}$	0.5777	0.0020

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: 1.Empirical moments are based on 107579 firms from year 2004-2007.

2. Imposed parameters are $\delta = 0.05$, $r = 0.20$, $\rho = 0.90$.

3. Simulation path is set to be S =5.

Table 4. Specification Tests

	col (1)	col (2)	col (3)	col (4)
	benchmark	$\sigma_{log\alpha} = \sigma_{log\eta} = 0$	$b^q = b^i = b^f = 0$	$\sigma_{meK} = \sigma_{meY} = \sigma_{me\pi} = 0$
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 5. Robustness Tests

Parameters	col (1) benchmark	col (2) δ =0.03	col (3) δ =0.07	col (4) r =0.15	col (5) r =0.25	col (6) ρ =0.85	col (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 5. Robustness Tests--continued

Parameters	col (1) benchmark	col (8) type-5	col (9) S =10	col (10) $\sigma_{\upsilon} > 0$	col (11) $\sigma_{\sigma} > 0$	col (12) $\sigma_{meI} > 0$
σ_{τ}	0.714	0.690	0.716	0.721	0.712	0.745
$\mu_{log\,\alpha}$	-2.606	-2.620	-2.606	-2.604	-2.592	-2.654
$\sigma_{log\,\alpha}$	0.557	0.557	0.557	0.551	0.556	0.577
$\mu_{log\eta}$	-2.808	-2.851	-2.808	-2.805	-2.805	-2.776
$\sigma_{log\eta}$	0.725	0.692	0.727	0.728	0.719	0.716
b^{q}	0.278	0.284	0.266	0.325	0.308	0.405
b^{i}	0.000	0.001	0.000	0.000	0.000	0.479
b^{f}	0.034	0.034	0.028	0.039	0.031	0.059
μ	0.080	0.082	0.083	0.083	0.080	0.061
σ	0.425	0.424	0.422	0.411	0.403	0.465
σ_{meK}	0.401	0.402	0.405	0.404	0.390	..
σ_{meY}	0.001	0.000	0.002	0.002	0.001	0.001
$\sigma_{me\pi}$	0.578	0.597	0.577	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	0.154	0.155	0.153
mean(log(Y/Khat))	1.146	1.155	1.142	1.154	1.147	1.104
mean(I/K)	0.173	0.175	0.176	0.177	0.171	0.135
mean(ΔlogY)	0.080	0.082	0.083	0.083	0.080	0.060
bsd(π/Y)	0.075	0.071	0.074	0.074	0.074	0.075
wsd(π/Y)	0.049	0.048	0.049	0.048	0.049	0.047
bsd(log(Y/Khat))	0.878	0.880	0.878	0.883	0.875	0.870
wsd(log(Y/Khat))	0.332	0.331	0.334	0.334	0.324	0.144
bsd(I/K)	0.164	0.165	0.165	0.180	0.161	0.135
wsd(I/K)	0.215	0.214	0.215	0.213	0.213	0.169
bsd(ΔlogY)	0.163	0.163	0.163	0.168	0.162	0.165
wsd(ΔlogY)	0.219	0.217	0.218	0.210	0.218	0.230
skew(π/Y)	0.854	1.010	0.853	0.852	0.846	0.846
skew(log(Y/Khat))	0.004	0.013	0.007	0.004	0.006	0.029
skew(I/K)	2.251	2.193	2.192	2.225	2.295	2.412
skew(dlogY)	0.176	0.176	0.155	0.195	0.157	0.268
scorr(π/Y)	0.599	0.581	0.601	0.604	0.598	0.620
scorr(log(Y/Khat))	0.838	0.839	0.836	0.838	0.844	0.976
scorr(I/K)	0.243	0.250	0.249	0.292	0.237	0.268
scorr(ΔlogY)	0.053	0.058	0.055	0.103	0.051	0.021
bcorr(π/Y, log(Y/Khat))	-0.271	-0.319	-0.266	-0.263	-0.274	-0.299
OI/100	183	229	199	148	179	741

Table 6.1. Aggregate TFPR Loss in China
Efficient Benchmark

type	α	η	1-γ	ΔlogTFPR overall	ΔlogTFPR distortions	ΔlogTFPR frictions
1	0.040	0.027	0.589	-0.383	-0.363	-0.020
2	0.040	0.060	0.385	-0.169	-0.160	-0.009
3	0.040	0.133	0.208	-0.071	-0.067	-0.004
4	0.074	0.027	0.724	-0.681	-0.649	-0.035
5	0.074	0.060	0.535	-0.309	-0.292	-0.017
6	0.074	0.133	0.325	-0.130	-0.123	-0.007
7	0.136	0.027	0.828	-1.034	-1.030	-0.056
8	0.136	0.060	0.679	-0.558	-0.530	-0.029
9	0.136	0.133	0.469	-0.238	-0.225	-0.013
average	0.083	0.074	0.527	-0.397	-0.382	-0.021

Table 6.2. Aggregate TFPR Loss in China
Compustat Benchmark

type	α	η	1-γ	ΔlogTFPR distortions
1	0.040	0.027	0.589	-0.069
2	0.040	0.060	0.385	-0.030
3	0.040	0.133	0.208	-0.013
4	0.074	0.027	0.724	-0.127
5	0.074	0.060	0.535	-0.056
6	0.074	0.133	0.325	-0.023
7	0.136	0.027	0.828	-0.231
8	0.136	0.060	0.679	-0.102
9	0.136	0.133	0.469	-0.043
average	0.083	0.074	0.527	-0.077

Table 7.1. Generalized ARP v.s. Full Structural

	Full Structural	Generalized ARP
Parameters		
σ_{τ}	0.7143	0.6845
$\mu_{log\,\alpha}$	-2.6058	-2.6199
$\sigma_{log\,\alpha}$	0.5568	0.5254
$\mu_{log\eta}$	-2.8084	-2.8085
$\sigma_{log\eta}$	0.7253	0.7644
Moments		
mean(π/Y)	0.1578	
mean(log(Y/Khat))	1.1377	
bsd(π/Y)	0.0763	
bsd(log(Y/Khat))	0.8666	
bcorr(π/Y, log(Y/Khat))	-0.2422	

Note: r = 0.20 and δ = 0.05 in both columns.

Table 7.2. Sub-Sector Results

sector	**all**	**garment**	**paper**	**auto parts**	**electronics**	**wine**	**tobacco**
cic code	**..**	**1810**	**2231**	**3725**	**4061**	**1521**	**1620**
Parameters							
σ_{τ}	0.6845	0.6182	0.6752	0.6281	0.6845	0.6762	0.0034
$\mu_{\log \alpha}$	-2.6199	-2.8147	-2.5671	-2.4901	-2.5424	-2.3212	-2.1724
$\sigma_{\log \alpha}$	0.5254	0.5217	0.3315	0.4755	0.5253	0.5373	0.4200
$\mu_{\log\eta}$	-2.8085	-2.8493	-3.3204	-2.6193	-2.8233	-2.1687	-2.1957
$\sigma_{\log\eta}$	0.7644	0.7550	0.9484	0.6377	0.7691	0.5791	0.3832
Moments							
mean(π/Y)	0.1578	0.1407	0.1347	0.1740	0.1632	0.2329	0.2391
mean(log(Y/Khat))	1.1377	1.3259	1.0619	1.0176	1.0597	0.9023	0.7338
bsd(π/Y)	0.0763	0.0700	0.0624	0.0709	0.0778	0.0921	0.0940
bsd(log(Y/Khat))	0.8666	0.8090	0.7579	0.7925	0.8671	0.8707	0.4186
bcorr(π/Y, log(Y/Khat))	-0.2422	-0.2195	-0.0885	-0.2671	-0.2643	-0.2580	-0.6346

Note: r = 0.20 and δ = 0.05 in all columns.

Table 7.3. Evolution of Capital Market Distortions

period	2004-2007	2001-2004	1998-2001
r	**0.2**	**0.15**	**0.1**
Parameters			
σ_{τ}	0.6845	0.7608	0.9537
$\mu_{log\ \alpha}$	-2.6199	-2.4153	-2.1661
$\sigma_{log\ \alpha}$	0.5254	0.5771	0.5260
$\mu_{log \eta}$	-2.8085	-3.1199	-3.6370
$\sigma_{log \eta}$	0.7644	1.0056	1.3390
Moments			
mean(π/Y)	0.1578	0.1705	0.1877
mean(log(Y/Khat))	1.1377	0.7470	0.2462
bsd(π/Y)	0.0763	0.0859	0.0913
bsd(log(Y/Khat))	0.8666	0.9399	1.0712
bcorr(π/Y, log(Y/Khat))	-0.2422	-0.2809	-0.2481

Note: δ = 0.05 for all columns.

Table 7.4. NBS v.s. Compustat

sample	NBS	Compustat full sample	Compustat NBS comparable	Compustat Bloom (2009)
Parameters				
σ_{τ}	0.6843	0.4612	0.3112	0.0019
$\mu_{log\ \alpha}$	-2.6189	-2.0793	-2.1435	-2.2327
$\sigma_{log\ \alpha}$	0.5539	0.6569	0.6530	0.5784
$\mu_{log \eta}$	-2.8086	-1.4040	-1.3814	-1.4873
$\sigma_{log \eta}$	0.7887	1.0038	0.9556	0.6751
Moments				
mean(π/Y)	0.1578	0.3863	0.3826	0.3514
mean(log(Y/Khat))	1.1377	0.4401	0.5066	0.5542
bsd(π/Y)	0.0763	0.1600	0.1572	0.1448
bsd(log(Y/Khat))	0.8666	0.8291	0.7531	0.6064
bcorr(π/Y, log(Y/Khat))	-0.2422	-0.0779	-0.0705	-0.0879

Note: r = 0.20 for NBS and 0.10 for Compustat; δ = 0.05 for all columns.

Table 8. Regressions on Firm Characteristics

	(1) baseline model		**(2) age and size**		**(3) cyclicity**		**(4) ownership**		**(5) political connection**		**(6) full model**	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
log(π/Y)	**-0.3934**	0.0028	**-0.3799**	0.0027	**-0.3800**	0.0027	**-0.3470**	0.0027	**-0.3407**	0.0027	**-0.3397**	0.0027
age			**-0.0328**	0.0002	**-0.0327**	0.0002	**-0.0314**	0.0002	**-0.0291**	0.0002	**-0.0287**	0.0002
emp			**-0.0428**	0.0034	**-0.0430**	0.0034	**-0.0284**	0.0024	**-0.0245**	0.0021	**-0.0233**	0.0020
beta					**0.0011**	0.0001	**0.0010**	0.0001	**0.0010**	0.0001	**0.0010**	0.0001
SOE							**-0.2521**	0.0066	**-0.2341**	0.0065	**-0.0567**	0.0189
COE							**0.2270**	0.0054	**0.2090**	0.0054	**0.1781**	0.0085
DPE							**0.1816**	0.0033	**0.1594**	0.0033	**0.1085**	0.0053
HMT							**-0.1711**	0.0049	**-0.1756**	0.0049	**-0.2924**	0.0074
FIE							**-0.1512**	0.0049	**-0.1521**	0.0049	**-0.2365**	0.0078
LU									**-0.1647**	0.0027	**-0.2409**	0.0057
SOE_LU											**-0.1907**	0.0200
COE_LU											**0.0378**	0.0109
DPE_LU											**0.0771**	0.0066
HMT_LU											**0.2018**	0.0092
FIE_LU											**0.1341**	0.0097

Note: 1. Industry and province 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.

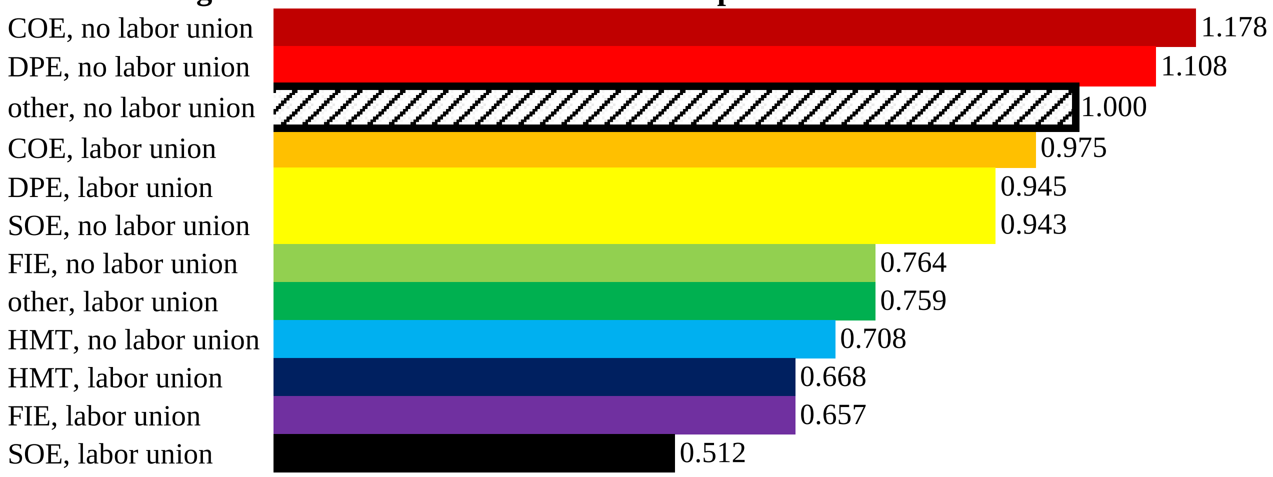
5. Beta is estimated as corr($\Delta \log Y_{i,t}$, $\Delta \log Y_t$)/var($\Delta \log Y_t$), where $\Delta \log Y_t$ is the average of $\Delta \log Y_{i,t}$.

6. SOE--dummy = 1 if state-owned; defined as registration type = 110, 141 and 151.
COE--dummy = 1 if collective owned firms; defined as registration type = 120 and 142.
DPE--dummy = 1 if domestic private-owned firms; defined as registration type from 170 to 174.
HMT--dummy = 1 if Hong Kong, Macau and Tai Wan ownerd firms; defined as registration type from 200 to 240.
FIE--dummy = 1 if foreign-owned firms; defined as registration type from 3000 to 340.

7. LU--dummy = 1 if a firm has a labor union.

8. SOE_LU--dummy = 1 if a firm is SOE and has a labor union; similar definition applies to other interaction terms.

Figure 1. The Predicted User Cost of Capital in Different Chinese Firms



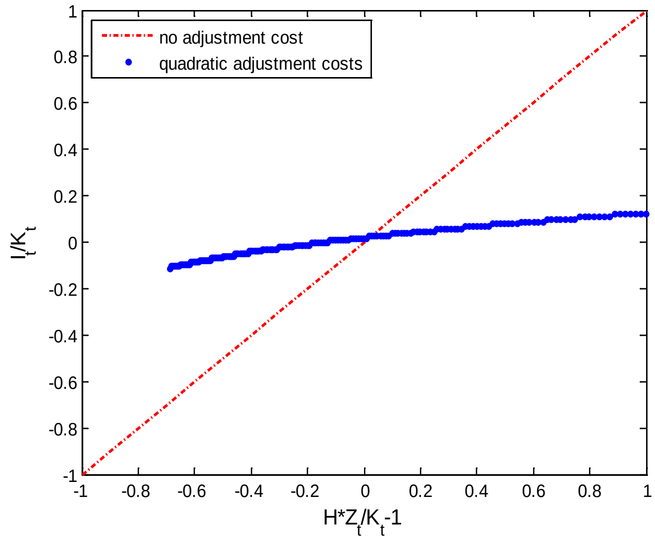


Figure A.1. Investment Policy under Quadratic Adjustment Costs

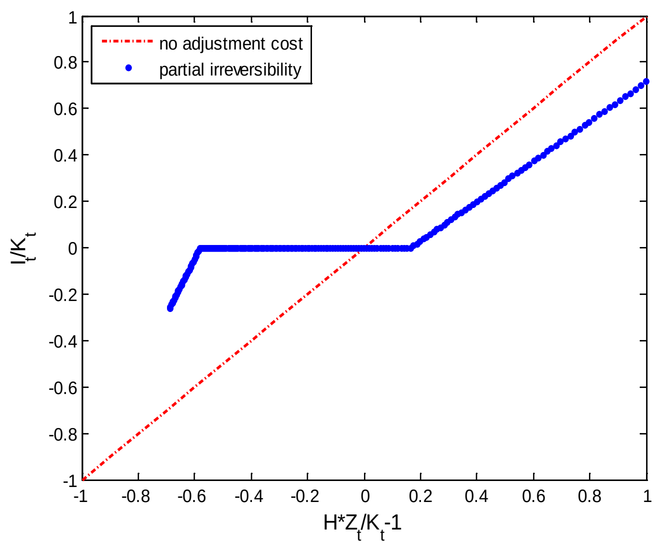


Figure A.2. Investment Policy under Irreversibility

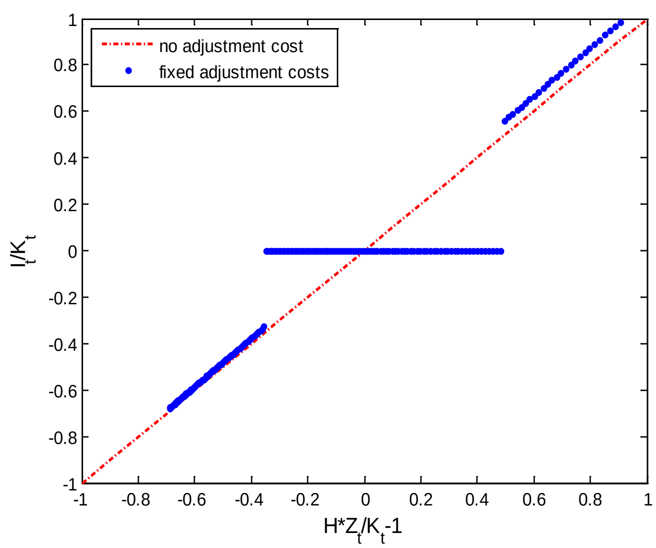


Figure A.3. Investment Policy under Fixed Adjustment Costs

Table A.1. Full Industrial Survey

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
No. of firms	149689	147112	148272	156812	166864	181179	259405	251498	276165	313041
mean(π/Y)	0.171	0.187	0.177	0.168	0.165	0.157	0.145	0.154	0.152	0.154
mean(log(Y/Khat))	0.430	0.522	0.658	0.827	0.973	1.121	1.323	1.311	1.334	1.413
mean(I/K)		0.082	0.085	0.097	0.128	0.161	0.191	0.248	0.216	0.213
mean(ΔlogY)		0.041	0.037	0.017	0.088	0.111	0.084	0.140	0.126	0.171

Table A.2. Firms at least age 5 upon Entry Dataset

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
No. of firms	106485	110233	112856	114056	118411	124708	151961	154849	174245	198882
mean(π/Y)	0.175	0.191	0.181	0.174	0.171	0.164	0.155	0.158	0.156	0.156
mean(log(Y/Khat))	0.184	0.325	0.475	0.623	0.770	0.939	1.105	1.163	1.224	1.319
mean(I/K)		0.057	0.061	0.072	0.091	0.117	0.142	0.169	0.166	0.169
mean(ΔlogY)		0.008	0.009	-0.005	0.058	0.078	0.044	0.078	0.075	0.114

Table A.3. Firms at least age 5 upon Entry Dataset and Survive 10 Years

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
No. of firms	28232	28232	28232	28232	28232	28232	28232	28232	28232	28232
mean(π/Y)	0.185	0.197	0.189	0.184	0.181	0.174	0.169	0.167	0.165	0.168
mean(log(Y/Khat))	0.346	0.418	0.477	0.493	0.561	0.621	0.613	0.644	0.650	0.681
mean(I/K)		0.089	0.088	0.085	0.085	0.096	0.093	0.103	0.096	0.093
mean(ΔlogY)		0.111	0.098	0.056	0.102	0.106	0.047	0.077	0.055	0.078

Table A.4. Firms at least age 5 upon Entry Dataset and Survive at least 4 Consecutive Years

Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
No. of firms	62422	72682	84511	99188	103486	110033	126924	120547	114876	107641
mean(π/Y)	0.180	0.195	0.184	0.177	0.173	0.165	0.156	0.160	0.158	0.160
mean(log(Y/Khat))	0.192	0.343	0.476	0.572	0.729	0.903	1.062	1.084	1.088	1.134
mean(I/K)		0.068	0.068	0.073	0.091	0.116	0.143	0.172	0.153	0.144
mean(ΔlogY)		0.060	0.045	0.004	0.068	0.088	0.050	0.089	0.069	0.097

Table A.5. 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 Buttom 5% observations are trimmed year-by-year for each variable in all the tables.