A Structural Estimation on Capital Market Distortions in UK and Chinese Manufacturing Firms

Zheng Michael Song* and Guiying Laura Wu†

September 2011

Abstract

Capital market distortions which drive wedges across firms in their marginal revenue product of capital will lower the aggregate revenue total factor productivity. However, inferring such distortions from the dispersion of marginal revenue product of capital is subject to a set of identification issues: unobserved heterogeneities in production technology and market power, investment frictions with idiosyncratic shocks, and measurement errors in the data. This paper estimates the capital market distortions and accounts for alternative sources of dispersion using a structural econometric approach. We find a magnitude of distortions that is much smaller than the literature but is still significantly large, which implies a 9.3% and 37.5% aggregate revenue total factor productivity loss in UK and China, respectively. We show that missing unobserved heterogeneities in production technology and market power will cause a severe upward bias in the estimation of distortions. Using a generalized marginal revenue product of capital approach, we find that small, private-owned firms without political connection in east China face unfavorable capital market distortions.

JEL Classification: E22, D92, C15

Key Words: capital market distortions, aggregate TFPR, unobserved heterogeneities

*Booth School of Business, University of Chicargo; Email: Zheng.Song@chicagobooth.edu.
†Division of Economics, Nanyang Technological University; Email: guiying.wu@ntu.edu.sg.

1 Introduction

Understanding the large and persistent differences in output per capita across countries has been the central theme in economics for a long time. The macro-growth literature based on a homogeneous aggregate production function suggests the crucial role of aggregate capital accumulation. However, the recent micro-development literature, as surveyed in Banaerjee and Duflo (2005), has found enormous heterogeneity in the average revenue product of capital within the same country. This makes capital misallocation a potentially important explanation to the cross-country output per capita differences.

Imagine two firms A and B in a single economy. Optimal investment implies that each firm must equalize its marginal benefit of investment, that is the marginal revenue production of capital (MRPK hereafter), to its marginal cost of investment, that is the generalized user cost of capital taking into account potential investment constraints. If capital market distortions cause a lower generalized user cost of capital for firm A than for firm B, the MRPK of firm A must be lower than that of firm B. For given aggregate capital stock, simply reallocating capital from firm A to firm B will increase aggregate output or the measured aggregate TFP for this economy. Thus even for countries with same level of aggregate capital stock, different capital allocation could lead to totally different output per capita. Restuccia and Rogerson (2008) formulate this idea in a growth model with idiosyncratic distortions and heterogeneous firms.

The significance of capital misallocation motives three related research questions in this paper. First, to what extent is capital market distorted within countries at different income levels? Second, what are the quantitative impacts of such distortions on the aggregate economy? And finally, what policies and institutions are related to such distortions?

Hsieh and Klenow (2009) provide a seminal empirical procedure to address these questions. In a standard model of monopolistic competition with heterogeneous firms, the authors show that the negative effect of capital misallocation on aggregate revenue total factor productivity (TFPR hereafter) can be summarized by the variance of log MRPK. Since the MRPK is not observable, the average revenue product of capital or the sales-to-capital ratio is taken as the natural proxy. Using this framework, Hsieh and Klenow (2009) calculate that the aggregate TFPR loss is tremendous: 86-115% in China, 100-128% in India and around 30-43% in the U.S.

This MRPK approach is undoubtedly neat and useful. However, as well-recognized by Hsieh and Klenow (2009), there are also important limitations in interpreting the variance of log sales-to-capital ratio as a measure of capital misallocation. This is because in addition to idiosyncratic distortions, many other factors will drive a dispersion in the log sales-to-capital ratio as well. First, firms may use different technologies in production. To produce the same quantity of output, all else being equal, a firm utilizing more capital will have a lower sales-to-capital ratio than a firm utilizing less capital. Second, firms could also have different market power in the product markets. To obtain the same amount of sales revenue, all else being equal, a firm with more market power will display a higher sales-to-capital ratio than a firm with less market power. Furthermore, investment might be subject to various frictions. If shocks in production or demand are idiosyncratic, a firm hit by positive shocks will accumulate more capital and result in a lower sales-to-capital ratio than a firm hit by negative shocks. Finally, even if none of these factors is relevant, dispersion in the observed log sales-to-capital ratio could also arise purely as a result of measurement errors in the sales revenue and capital stock data.

This paper contributes to these identification issues by estimating a fully structural and parametric investment model. We use unobserved heterogeneity in the capital goods prices as a generic representation of idiosyncratic capital market distortions. We show that such heterogeneity will lead to a dispersion in log sales-to-capital ratio and lower aggregate TFPR. We then account for other sources of dispersion in log sales-to-capital ratio with three devices. First, we take into account the possibility of firm-specific production technology and market power by allowing for unobserved heterogeneities in the capital share in the production function and in the inverse of demand elasticity. Second, we model investment frictions through three forms of capital adjustment costs, the quadratic adjustment costs, irreversibility and fixed adjustment costs. Finally, we explicitly model and estimate potential measurement errors in the data.

Structural parameters are then estimated by the method of simulated moments (MSM hereafter). That is to match a large set of simulated moments with empirical moments, which characterize the distribution and dynamics of four variables: log sales-to-capital ratio, profit-to-sales ratio, investment rate and sales growth rate. We illustrate how this set of moments may identify the parameters of our interest. First, using moments on investment rate and sales growth rate simultaneously estimate the capital adjustment costs and the exogenous stochastic process. Second, the within-group standard deviation and the serial correlation of these four variables restrict the measurement errors in capital stock, sales revenue and gross profit. Finally, the bewteen-group standard deviations of log sales-to-capital ratio and profit-to-sales ratio, together with the cross correlation between these two variables jointly identify the unobserved heterogeneities in capital goods prices, capital share and market power.

We estimate this model using two panels from UK and Chinese manufacturing firms. Even after controlling for the unobserved heterogeneities in production and demand, capital adjustment costs and measurement errors, we still find significant evidence of unobserved heterogeneities in capital goods prices in both UK and China, and the magnitude in China is about 1.6 times of that in UK. Counterfactual simulations indicate that the implied aggregate TFPR loss is 9.3% in UK and 37.5% in China. Had capital in China been reallocated to the efficiency level in UK, the aggregate TFPR in China would enhance by 20%.

Specification tests find that, among the three novel features considered in this paper, a model without the unobserved heterogeneities in capital share and market power will substantially overestimate the capital market distortions. However, a model without capital adjustment costs or measurement errors does not necessarily lead to such bias. This is because both capital adjustment costs and measurement errors mainly cause dispersion in the time-series dimension of log sales-to-capital ratio. In contrast, capital market distortions, unobserved heterogeneities in capital share and inverse of demand elasticity will all cause dispersion in the cross-section dimension of log sales-to-capital ratio. Meanwhile, only unobserved heterogeneities in capital share and inverse of demand elasticity will lead to dispersion in the cross-section dimension of profit-to-sales ratio. This implies that the MRPK approach may still deliver reliable estimates on distortions, if panel data on capital stock, sales revenue and gross profit are available. We call this a "generalized MRPK approach", that is to back out capital market distortions from the between-group dispersion of log sales-to-profit ratio conditional on between-group dispersion in the profit-to-sales ratio.

This methodological contribution allows us to link the unobservable capital market distortions with observable firm characteristics to address the third research question of this paper. We find that firms that are small, private-owned, without political connection and in east China have faced significantly higher capital goods prices than their counterparts that are large, non-private-owned, with political connection and in other areas of China. Therefore aggregate TFPR and aggregate output would increase substantially, if capital was reallocated towards these unfavorable firms. Such policy implications are consistent with the explanation on China's remarkable economic growth in Song, Storesletten and Zilibotti (2011).

In addition to Hshieh and Klenow (2009), we also build on several important papers. 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 sales-to-capital ratio but at most 10% cross-section dispersion. This motives us to decompose the overall variance in log sales-to-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 investment model and derives the measure of aggregate TFPR loss. Section 3 presents empirical specification and describes the data. Section 4 discusses identification and reports the empirical results. Section 5 links capital market distortions to observable firm characteristics by a generalized MRPK approach . Section 6 discusses alternative explanations. Section 7 concludes.

2 The Model

2.1 Production and Demand

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

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

where $A_{i,t}$ represents the randomness in productivity; the firm-specific capital share β_i satisfies $0 < \beta_i < 1$.

Each product i is demanded in a monopolistic product market according to a stochastic isoelastic downward-sloping demand curve,

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

where $X_{i,t}$ represents the randomness in demand; $0 < \eta_i < 1$ is the inverse of firm-specific demand elasticity with respect to price.

In each year t, for a given productive capital stock $\widehat{K}_{i,t}$, the realization of intermediate input cost m_t, wage rate w_t, productivity $A_{i,t}$ and demand condition $X_{i,t}$, firm i chooses variable inputs $M_{i,t}$ and $L_{i,t}$ to maximize its instantaneous gross profit,

$$\pi_{i,t} = \max_{M_{i,t},\, L_{i,t}} \{P_{i,t}Q_{i,t} - m_t M_{i,t} - w_t L_{i,t}\}$$

Optimization yields the maximized value of gross profit,

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

where

$$1 - \gamma_i = \frac{\beta_i(1-\eta_i)}{\eta_i + \beta_i(1-\eta_i)} \tag{2}$$

and

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

Equation (1) has used two reparameterization. First, $1 - \gamma_i$ is the firm-specific capital share in the solved out profit function, which increases with capital share β_i and decreases with the inverse of demand elasticity η_i. Second, $Z_{i,t}$ is known as 'TFPR' or 'revenue total factor productivity' following Foster, Haltiwanger and Syverson (2008), which incorporates the randomness from factor prices, productivity and demand. Assume that the log of $Z_{i,t}$ follows a trend stationary AR(1) process,

$$\begin{aligned} \log Z_{i,t} &= \mu t + z_{i,t} \\ z_{i,t} &= \rho z_{i,t-1} + e_{i,t} \end{aligned} \tag{3}$$

where $0 < \rho < 1$, $e_{i,t} \overset{i.i.d.}{\sim} N(0, \sigma^2)$, and $z_{i,0} = 0$. Equation (3) implies a common level of persistence, growth rate and uncertainty in the TFPR. However, the TFPR shocks are idiosyncratic.

Denote sales as $Y_{i,t} \equiv P_{i,t}Q_{i,t}$.[1] The first order conditions of the profit maximization problem imply a constant material and labor cost share

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

so that the maximized gross profit is always a constant proportion of sales,

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

Consider if $\eta_i \to 0$, Equation (4) would imply that $\frac{\pi_{i,t}}{Y_{i,t}} \to \beta_i$, $\frac{m_{i,t}M_{i,t}}{Y_{i,t}} \to 1-\alpha_i-\beta_i$ and $\frac{w_{i,t}L_{i,t}}{Y_{i,t}} \to 1-\beta_i$. Intuitively in the special case of perfect competition, each production factor earns a share from the sales revenue according to its share in the production function.

2.2 Distortions and Frictions

Investment may be subject to both distortions and frictions. 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 price of capital that are heterogeneous across firms. Therefore the actual capital goods price faced by firm i 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.

Meanwhile, in a variety setting, 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,

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

[1] We follow the investment literature, such as Abel and Eberly (1999), in using Q to denote the quantity of output, calling the product of price and quantity of output as sales and denoting it as Y. In the produtivity literature, such as Hsieh and Klenow (2009), Y is simply the quantity of output, which is equivalent to the Q in our model.

where $\mathbf{1}_{[I_t<0]}$ and $\mathbf{1}_{[I_t\neq 0]}$ are indicators for negative and non-zero investment; b^i can be interpreted as the difference between the purchase price and the sale price expressed as a percentage of the purchase price of capital goods; b^f is interpreted as the fraction of gross profit loss due to any non-zero investment; and b^q measures the magnitude of quadratic adjustment costs.

By paying the cost of purchasing capital and adjusting capital, new investment $I_{i,t}$ contributes to productive capital stock $\widehat{K}_{i,t}$ immediately in year t, which depreciates at the end of the year.[2] The law of motion for capital is therefore given by

$$K_{i,t+1} = (1-\delta)\left(K_{i,t} + I_{i,t}\right) \equiv (1-\delta)\widehat{K}_{i,t} \tag{5}$$

where δ is the constant depreciation rate common across firms.

2.3 Investment Decision

The presence of capital adjustment costs implies that investment is an intertemporal decision. At the beginning of each year t, optimal investment is chosen to maximize the discounted present value of dividends, which is the gross profit net of investment expenditure and capital adjustment costs. Investors allocate capital until the required rate of return on capital is equalized across different firms. Suppose this required rate of return is r, at which investors discount future dividends. The investment problem is then defined by the stochastic Bellman equation,

$$V(Z_{i,t}, K_{i,t}) = \max_{I_{i,t}} \left\{ \begin{array}{c} \pi(Z_{i,t}, K_{i,t}; I_{i,t}) - P^K_{i,t} I_{i,t} - G(Z_{i,t}, K_{i,t}; I_{i,t}) \\ +\frac{1}{1+r} E_t\left[V(Z_{i,t+1}, K_{i,t+1})\right] \end{array} \right\} \tag{6}$$

together with the law of motion (3) and (5).

It is known that in the presence of adjustment costs, there is generally no analytical solution to the optimal investment problem. However, the analytical solution in the frictionless case 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 TFPR relative to inherited capital stock

$$\frac{I_{i,t}}{K_{i,t}} = H_{i,t}\left(\frac{Z_{i,t}}{K_{i,t}}\right) - 1 \tag{7}$$

so as to meet the imbalance between the optimal productive capital stock and the level of TFPR

$$\widehat{K}_{i,t} = H_{i,t} Z_{i,t} \tag{8}$$

where the slope term is defined as

$$H_{i,t} = \left[\frac{1-\gamma_i}{(1+\tau_i)\, J_t}\right]^{\frac{1}{\gamma_i}}$$

[2] Compared with alternative lagged timing assumption, such as $K_{t+1} = (1-\delta)K_t + I_t$, our timing assumption does not affect the qualitative implication of the model, but allows for a closed-form solution to the investment problem in the absence of any capital adjustment costs, which does not involve any expectation term. This provides a convenient benchmark for studying the effects of capital adjustment costs.

which depends on γ_i, the combination of firm-specific production technology and demand elasticity, capital market distortions τ_i and the Jorgensonian user cost of capital

$$J_t = P_t^K - \left(\frac{1-\delta}{1+r}\right) E_t \left[P_{t+1}^K\right]$$

In general when $G(Z_{i,t}, K_{i,t}; I_{i,t}) > 0$, the investment policy can be solved out using numerical dynamic programming method. Figures 1.1-1.3 illustrate these policies under different forms of adjustment costs. The 45^o straight line is the investment policy in Equation (7) and is plotted as the frictionless benchmark. As highlighted by these figures, first, irrespective to the form of adjustment costs, the optimal investment rate is always a non-decreasing function of $\frac{Z_{i,t}}{K_{i,t}}$. 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 $b^f > 0$.

2.4 Losses in Aggregate TFPR

Both distortions and frictions will cause aggregate TFPR loss. To study their effects on the loss, consider N firms in the economy with same β and η. Suppose each firm i in year t draws TFPR $Z_{i,t}$ and accumulates capital stock $\widehat{K}_{i,t}$.

Without any distortion nor friction, according to Equation (8), the productive capital $\widehat{K}^*_{i,t}$ must be linear proportional to $Z_{i,t}$,

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

This implies that in the first-best allocation, each firm gets a share of capital proportional to the share of its TFPR,

$$\widehat{K}^*_{i,t} = \frac{Z_{i,t}}{Z_t^*} \widehat{K_t}$$

so that the first-best aggregate sales revenue is given by

$$Y_t^* = \frac{\gamma}{\eta} Z_t^{*\gamma} \widehat{K_t}^{1-\gamma}$$

where $\widehat{K}_t = \sum_{i=1}^N \widehat{K}_{i,t}$ is the existing aggregate capital stock; and $Z_t^* = \sum_{i=1}^N Z_{i,t}$ is the first-best aggregate TFPR.

In contrast, with distortions or frictions, the actual aggregate sales revenue is

$$Y_t = \frac{\gamma}{\eta} \sum_{i=1}^N \left(Z_{i,t}^\gamma \widehat{K}_{i,t}^{1-\gamma}\right) = \frac{\gamma}{\eta} Z_t^\gamma \widehat{K}_t^{1-\gamma}$$

in which the actual aggregate TFPR can be backed out as

$$Z_t = \left[\frac{\left(\sum_{i=1}^N \frac{Z_{i,t}}{k_{i,t}^{1-\gamma}}\right)}{\left(\sum_{i=1}^N \frac{Z_{i,t}}{k_{i,t}}\right)^{1-\gamma}}\right]^{\frac{1}{\gamma}}$$

where the wedge is defined as

$$k_{i,t} \equiv \frac{Z_{i,t}}{\widehat{K}_{i,t}}$$

The aggregate TFPR loss is the difference between actual and first-best aggregate TFPR,

$$\begin{aligned}
\Delta \log TFPR_t &= \log Z_t^{\gamma} - \log Z_t^{*\gamma} \\
&= \log\left(\sum_{i=1}^{N} \frac{Z_{i,t}}{k_{i,t}^{1-\gamma}}\right) - (1-\gamma)\log\left(\sum_{i=1}^{N} \frac{Z_{i,t}}{k_{i,t}}\right) - \gamma \log\left(\sum_{i=1}^{N} Z_{i,t}\right) \qquad (9) \\
&= \log\left(\sum_{i=1}^{N}\left(Z_{i,t}^{\gamma}\widehat{K}_{i,t}^{1-\gamma}\right)\right) - (1-\gamma)\log\left(\sum_{i=1}^{N}\widehat{K}_{i,t}\right) - \gamma\log\left(\sum_{i=1}^{N} Z_{i,t}\right)
\end{aligned}$$

To highlight how capital market distortions will lower the aggregate TFPR, consider the special case when there is no adjustment cost $G(Z_{i,t}, K_{i,t}; I_{i,t}) = 0$, common TFPR shocks $Z_{i,t} = Z_{j,t}$ and large number of firms $N \to \infty$, so that there is a closed-form solution

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

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

3 Empirical Specification

3.1 The MRPK Approach

The goal of this paper is to quantify the effect of capital market distortions on aggregate TFPR loss using the above framework. Since the capital market distortions τ_i are not observable directly, one has to infer τ_i from observable variables.

In the special case of $b^q > 0$ and $b^i = b^f = 0$, there is a closed form investment Euler equation,

$$\begin{aligned}
(1-\gamma_i)\left(\frac{Z_{i,t}}{\widehat{K}_{i,t}}\right)^{\gamma_i} &= \left(P_t^K - \left(\frac{1-\delta}{1+r}\right)E_t\left[P_{t+1}^K\right]\right)(1+\tau_i) \\
&\quad + b^q\frac{I_{i,t}}{K_{i,t}} - b^q\left(\frac{1-\delta}{1+r}\right)E_t\left[\frac{I_{i,t+1}}{K_{i,t+1}}\right] - \frac{b^q}{2}\left(\frac{1-\delta}{1+r}\right)E_t\left[\left(\frac{I_{i,t+1}}{K_{i,t+1}}\right)^2\right]
\end{aligned}$$

If the average capital goods price is constant and is normalized to unity, the Jorgensonian user cost of capital is then simplified as

$$J \equiv \frac{r+\delta}{1+r} \qquad (10)$$

By omitting the higher order term $\frac{b^q}{2}\left(\frac{1-\delta}{1+r}\right) E_t\left[\left(\frac{I_{i,t+1}}{K_{i,t+1}}\right)^2\right]$ and assuming $E_t\left[\frac{I_{i,t+1}}{K_{i,t+1}}\right] = \frac{I_{i,t}}{K_{i,t}}$, the investment Euler equation can be approximated as

$$(1-\gamma_i)\left(\frac{Z_{i,t}}{\widehat{K}_{i,t}}\right)^{\gamma_i} = J\left(1+\tau_i\right)\left(1+b^q\frac{I_{i,t}}{K_{i,t}}\right) \tag{11}$$

Notice that the left hand side of Equation (11) is simply the marginal revenue product of capital or MRPK, which is proportional to the sales-to-capital ratio, due to the linear homogeneity of sales with respect to TFPR and productive capital stock,

$$MRPK_{i,t} \equiv (1-\gamma_i)\left(\frac{Z_{i,t}}{\widehat{K}_{i,t}}\right)^{\gamma_i} = \beta_i\left(1-\eta_i\right)\frac{Y_{i,t}}{\widehat{K}_{i,t}} \tag{12}$$

The term on the right hand side of Equation (11) is known as the generalized user cost of capital, which depends on the Jorgensonian user cost of capital J, investment frictions $b^q\frac{I_{i,t}}{K_{i,t}}$, and the capital market distortions τ_i,

$$U_{i,t} \equiv J\left(1+\tau_i\right)\left(1+b^q\frac{I_{i,t}}{K_{i,t}}\right) \tag{13}$$

Equality between $MRPK_{i,t}$ and $U_{i,t}$ then provides an analytical approximation to the log sales-to-capital ratio,

$$\log\left(\frac{Y_{i,t}}{\widehat{K}_{i,t}}\right) = \log J + \log\left(1+\tau_i\right) - \log\left[\beta_i\left(1-\eta_i\right)\right] + \log\left(1+b^q\frac{I_{i,t}}{K_{i,t}}\right) \tag{14}$$

3.2 Identification Strategy

Equation (14) indicates that one may infer the variance of $\log\left(1+\tau_i\right)$, which is unobservable, from the variance of $\log\left(\frac{Y_{i,t}}{\widehat{K}_{i,t}}\right)$, which is unobservable. The key challenge in this indirect inference is that, not only the distortions τ_i themselves, but also will another two factors cause a dispersion in $\log\left(\frac{Y_{i,t}}{\widehat{K}_{i,t}}\right)$. One is the heterogeneities in capital share β_i and inverse of demand elasticity η_i, and the other is the capital adjustment costs $\left(b^q, b^i, b^f\right)$ when TFPR shocks are idiosyncratic. Finally, when one infers the distortions from observed $\log\left(\frac{Y_{i,t}}{\widehat{K}_{i,t}}\right)$, potential measurement errors in the data of sales revenue and capital stock may contaminate the inference as well.

To control for alternative sources of dispersion in $\log\left(\frac{Y_{i,t}}{\widehat{K}_{i,t}}\right)$, we employ an identification strategy as follows. First, following Cooper and Haltiwanger (2006) and Bloom (2009), we use information on both investment rate $\frac{I_{i,t}}{K_{i,t}}$ and log sales growth rate $\Delta\log Y_{i,t}$ to jointly estimate the capital adjustment costs and the stochastic process. Second, we use information on both log sales-to-capital ratio $\log\left(\frac{Y_{i,t}}{\widehat{K}_{i,t}}\right)$ and profit-to-sales ratio $\frac{\pi_{i,t}}{Y_{i,t}}$ to jointly estimate unobserved heterogeneities in capital goods prices, capital share and inverse of demand elasticity. Finally, we explicitly model and estimate measurement errors in key variables. The last two points require further empirical specification.

3.3 Unobserved Heterogeneities

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

Assumption 1 ***Heterogeneity in the capital goods prices:***

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

That is each firm i has a firm-specific component in the capital goods prices τ_i, where τ_i is drawn independently from an identical normal distribution with mean zero and standard deviation σ_τ.

Assumption 2 ***Heterogeneity in the capital share in production function:***

$$\log \beta_i \overset{i.i.d}{\sim} N\left(\mu_{\log\beta}, \sigma_{\log\beta}^2\right)$$

That is each firm i has a firm-specific capital share β_i, where $\log\beta_i$ is drawn independently from an identical normal distribution with mean $\mu_{\log\beta}$ and standard deviation $\sigma_{\log\beta}$.

Assumption 3 ***Heterogeneity in the inverse of demand elasticity:***

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

That is each firm i has a firm-specific inverse of demand elasticity η_i, where $\log\eta_i$ is drawn independently from an identical normal distribution with mean $\mu_{\log\eta}$ and standard deviation $\sigma_{\log\eta}$.

The log-normality assumptions on $(1+\tau_i)$, β_i and η_i are based on the restrictions that $\tau_i > -1$, $0 < \beta_i < 1$ and $0 < \eta_i < 1$. Furthermore, Assumption 1 also implies that the key parameter of our interest σ_τ, is the direct indicator for the negative effect of distortions on aggregate TFPR.

The investment policy under different $(\tau_i, \beta_i, \eta_i)$ is different. Hence the dynamic programming problem described in Equation (6) must be solved for each firm i at each value of $(\tau_i, \beta_i, \eta_i)$, 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.

Assumption 4 ***Finite type of firms:*** *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 $\digamma = \{(\tau_u, \beta_v, \eta_x) : u = 1,2,3; v = 1,2,3; x = 1,2,3\}$.*

3.4 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. This is motivated by three consideration. First, measurement errors are common in firm-level data. The identification strategy discussed above employes four variables: profit-to-sales ratio $\frac{\pi_{i,t}}{Y_{i,t}}$; log sales-to-capital ratio $\log\left(\frac{Y_{i,t}}{\widehat{K}_{i,t}}\right)$; investment rate $\frac{I_{i,t}}{K_{i,t}}$ and sales growth rate $\Delta \log Y_{i,t}$. All these variables in ratio or growth rate are constructed from three variables in level: capital stock $K_{i,t}$, sales revenue $Y_{i,t}$ and gross profit $\pi_{i,t}$, which might be measured inaccurately in the data. Second and more fundamentally, measurement errors may contaminate identification. For example, neglecting measurement errors in $K_{i,t}$ or $Y_{i,t}$ may lead to an overestimation in σ_τ; in contrast, neglecting measurement errors in $\pi_{i,t}$ or $Y_{i,t}$ may overestimate $\sigma_{\log\beta}$ and $\sigma_{\log\eta}$ hence underestimate σ_τ. Last but not the least, the constant Jorgensonian user cost of capital J in Equation (10) is based on the assumption of a time-invariant average capital goods price P^K. Allowing for transitory measurement errors in capital stock is a parsimonious way to control for possible changes in the capital goods price over time.

Assumption 5 ***Measurement errors in the data:***

$$
\begin{aligned}
K_{i,t} &= K'_{i,t}\exp(e^K_{i,t}),\ \textit{where } e^K_{i,t} \overset{i.i.d}{\sim} N(0,\sigma^2_{meK}) \\
Y_{i,t} &= Y'_{i,t}\exp(e^Y_{i,t}),\ \textit{where } e^Y_{i,t} \overset{i.i.d}{\sim} N(0,\sigma^2_{meY}) \\
\pi_{i,t} &= \pi'_{i,t}\exp(e^\pi_{i,t}),\ \textit{where } e^\pi_{i,t} \overset{i.i.d}{\sim} N(0,\sigma^2_{me\pi})
\end{aligned}
$$

Here $W_{i,t}$ (where $W = K, Y$ and π respectively) denotes the observed variable, $W'_{i,t}$ denotes the true underlying variable which is not measured accurately in the data. $e^W_{i,t}$ is the measurement error in variable $W'_{i,t}$, where $e^W_{i,t}$ is drawn independently from an identical normal distribution with mean zero and standard deviation σ_{meW}.

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, sales revenue and gross profit. Second, we only allow for transitory measurement errors so as to distinguish measurement errors from unobserved heterogeneities.

3.5 Data Description

Since the estimation method we adopt is fully structural and parametric, to take into account the effect of potential model misspecification on the estimated effects, we apply the model to both UK and Chinese firms and use the results of UK as our critical benchmark.

The UK data employed in this paper has been analyzed previously by Bloom, Bond and Van Reenen (2007), for the investment dynamics under uncertainty and irreversibility. These firm-level data are taken from the consolidated accounts of manufacturing firms listed on the UK stock market and are obtained from the Datastream on-line service. The cleaned sample we use contains firm-level data for a panel of 629 firms between 1972 and 1991. On average there are 9 annual observations per firm.

The data for Chinese firms are from the Chinese Manufacturing Enterprise Survey. Questionnaires were designed and surveys were implemented by economists at the University of Michigan in collaboration with the Economics Research Institute of the Chinese Academy of Social Sciences. The survey instrument was divided into two parts. The objective part was directed to the firm's chief accountant. It is purely based on accounting information concerning the firms. The subjective part was directed to the chief managers–roughly equivalent to the CEOs–of the firms personally. It contains attitudinal and qualitative questions with multiple choices. The sample we use for estimation is made of a panel of 701 firms between 1994 to 1999. These firms were sampled almost evenly from four provinces in China (Jiangsu, Sichuan, Shanxi and Jilin) that together contribute about 20% of China's industrial output. The sample covers 39 industries in total, representative of China's overall industrial structure.

Four key variables are collected from both dataset: investment expenditure $I_{i,t}$, capital stock $K_{i,t}$, sales revenue $Y_{i,t}$ and gross profit $\pi_{i,t}$. The Appendix explains how these variables are collected, deflated and cleaned from the data.

One potential concern in comparing the UK and Chinese results is that firms in our UK sample are on average larger than those in our Chinese sample. The mean and median number of employees are 4856 and 1102 in UK, and are 2011 and 1055 in China. One possibility is that the firm-level data from UK might be consolidated across several plants within the firm. To make sure the comparability across these two samples, we assume that the UK data are aggregated over two plants and the Chinese data are from a single plant so that plants in UK and in China have similar size in terms of average number of employees.

Assumption 6 ***Aggregation:*** *Each UK firm is aggregated over* 2 *plants. For each plant j of firm i in period t, the law of motion for $Z_{j,i,t}$ is given by*

$$\begin{aligned}\log Z_{j,i,t} &= \mu t + z_{j,i,t} \\ z_{j,i,t} &= \rho z_{j,i,t-1} + \frac{1}{\sqrt{2}} e_{i,t} + \frac{1}{\sqrt{2}} e_{j,i,t}\end{aligned}$$

where $e_{i,t} \overset{i.i.d.}{\sim} N(0,\sigma^2)$, $e_{j,i,t} \overset{i.i.d.}{\sim} N(0,\sigma^2)$, $z_{j,i,0}=0$, $e_{i,t} \perp e_{j,i,t}$, *and* $e_{1,i,t} \perp e_{2,i,t}$.

That is plants $(j=1,2)$ in the same firm i share the common type $(\tau_i, \beta_i, \eta_i)$ and a common firm-level TFPR shock $e_{i,t}$, but also have a plant-specific TFPR shock $e_{j,i,t}$. The $\frac{1}{\sqrt{2}}$ equal weight of firm-level shock and plant-level shock implies that the overall level of uncertainty is still σ for a firm with two plants, which is comparable to the level of uncertainty from a single-plant firm in China.

4 Structural Estimation

This section estimates the structural parameters in the model and in the empirical specification using the MSM. Readers who are focused on the simulated effects of distortions may move on directly to Section 4.7.

4.1 Method of Simulated Moments

The MSM 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. To be specific, following Gouriéroux and Monfort (1996), the **MSM estimator** Θ^* solves the minimal quadratic distance problem,

$$\widehat{\Theta}^* = \arg\min_{\Theta} \left(\widehat{\Phi}^D - \frac{1}{S}\sum_{s=1}^{S} \widehat{\Phi}_s^M(\Theta) \right)' \Omega \left(\widehat{\Phi}^D - \frac{1}{S}\sum_{s=1}^{S} \widehat{\Phi}_s^M(\Theta) \right) \tag{15}$$

where Θ is the vector of parameters of interest; $\widehat{\Phi}^D$is a set of empirical moments estimated from an empirical dataset; $\widehat{\Phi}^M(\Theta)$ 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 MSM procedure provides a global specification test of the overidentifying restrictions of the model:

$$\begin{aligned} OI &= \frac{NS}{1+S} \left(\widehat{\Phi}^D - \frac{1}{S}\sum_{s=1}^{S} \widehat{\Phi}_s^M(\Theta) \right)' \Omega^* \left(\widehat{\Phi}^D - \frac{1}{S}\sum_{s=1}^{S} \widehat{\Phi}_s^M(\Theta) \right) \\ &\sim \chi^2 \left[\dim\left(\widehat{\Phi}\right) - \dim(\Theta) \right]. \end{aligned} \tag{16}$$

4.2 Parameters

4.2.1 Parameters to Estimate

Table 1 lists the set of parameters to estimate (Θ). It includes the key parameter characterizing the magnitude of distortions σ_τ; mean and standard deviation of the log capital share $\mu_{\log\beta}$ and $\sigma_{\log\beta}$; mean and standard deviation of the log inverse of demand elasticity $\mu_{\log\eta}$ and $\sigma_{\log\eta}$; three parameters measuring the magnitude of different capital adjustment costs $\left(b^q, b^i, b^f\right)$; growth rate of the TFPR μ; standard deviation of the TFPR shocks or the level of uncertainty σ; and the standard deviations of the measurement errors in capital stock, sales revenue and gross profit σ_{meK}, σ_{meY}, and $\sigma_{me\pi}$.

4.2.2 Predetermined Parameters

In addition to these 13 structural parameters, the depreciation rate δ and the discount rate r also affect the investment decision through the Jorgensonian user cost of capital

J. We pin down the depreciation rate for UK and China by the differences between the average sales growth rate $\Delta \log Y_{i,t}$ and log investment rate $\log(1 + I_{i,t}/K_{i,t})$ from the data,[3] which are 0.08 and 0.04, respectively. As for the discount rate, Bai, Hsieh and Qian (2005) have found that in China the aggregate rate of return to capital is around 14% taking into account market power. We therefore set $r = 0.14$ for China. To maintain comparability, we impose $r = 0.10$ for UK so that the overall Jorgensonian user cost of capital J is about the same across these two samples.

To determine the serial correlation in the TFPR, we follow Cooper and Haltiwanger (2006) by estimating the following dynamic panel data model,[4]

$$\log Y_{i,t} = \alpha + \rho \log Y_{i,t-1} + (1 - \gamma_i) \log \widehat{K}_{i,t} - \rho (1 - \gamma_i) \log \widehat{K}_{i,t-1} + u_i + v_{i,t} \qquad (17)$$

We estimate this equation using system GMM (Blundell and Bond, 1998) and allow for a complete set of year dummies to capture the aggregate shocks. ρ is estimated at 0.878 for UK and 0.558 for China. The estimate of ρ for UK is very close to the value 0.885 found in Cooper and Haltiwanger (2006), while a substantially lower estimate for China may reflect the attenuation bias due to measurement errors in the sales data. Therefore we impose $\rho = 0.878$ for both UK and China in estimating the investment model. A later section considers the sensitivity of the estimates to imposing different values for δ, r and ρ.

4.3 Moments

4.3.1 Choice of Moments

Table 2 lists the set of moments to match $\left(\widehat{\Phi}^D\right)$. This includes means (*mean*), between-group standard deviations (*bsd*), within-group standard deviations (*wsd*), coefficients of skewness (*skew*), and serial correlations (*scorr*) for profit-to-sales ratio $\pi_{i,t}/Y_{i,t}$, log sales-to-capital ratio $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$, investment rate $I_{i,t}/K_{i,t}$ and sales growth rate $\Delta \log Y_{i,t}$; the cross correlation between profit-to-sales ratio and log sales-to-capital ratio, and two cross correlations capturing how investment rate and sales growth rate response to a proxy for $Z_{i,t}/K_{i,t}$.

The choice of the moments is guided by two principles. First, this is a relatively comprehensive set of moments which characterize the distribution and dynamics of

---

[3] This is because by capital accumulation fomular Equation (5), $\log\left(1 + \frac{I_{i,t}}{K_{i,t}}\right) = \log\left(\frac{K_{i,t}+I_{i,t}}{K_{i,t}}\right) = \log\left(\frac{\widehat{K}_{i,t}}{\widehat{K}_{i,t-1}(1-\delta)}\right) = \triangle \log \widehat{K}_{i,t} - \log(1-\delta) \simeq \triangle \log \widehat{K}_{i,t} + \delta$. Bloom (2000) shows that despite the presence of adjustment costs, when a firm is on its balanced growth path, the gap between friction and frictionless capital stock is bounded. Therefore in the long run both capital stock and sales will grow at the same rate as TFPR, $\Delta \log Y_{i,T} \equiv \lim_{T\to\infty} \frac{1}{T} \ln(Y_{i,T+t}/Y_{i,t}) = \lim_{T\to\infty} \frac{1}{T} \ln\left(\widehat{K}_{i,T+t}/\widehat{K}_{i,t}\right) = \triangle \log \widehat{K}_{i,T} = \mu$. This implies that the difference between $\log\left(1 + \frac{I_{i,t}}{K_{i,t}}\right)$ and $\Delta \log Y_{i,t}$ is governed by depreciation rate δ.

[4] This equation is derived by taking logs on both sides of the sales equation $Y_{i,t} = \frac{\gamma_i}{\eta_i} Z_{i,t}^{\gamma_i} \widehat{K}_{i,t}^{1-\gamma_i}$, quasi differencing and replacing $\log Z_{i,t}$ and $\rho \log Z_{i,t-1}$ using the AR(1) structure specified in Equation (3).

key model 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.

4.3.2 Illustration for Identification

We discuss the second point by illustrating how these moments vary with underlying parameters in different panels of Table 3.[5] We start with a model in which there is no capital adjustment costs, unobserved heterogeneities, and measurement errors and label it as Model A in Table 3.1. In this baseline model there is virtually no variation in $\pi_{i,t}/Y_{i,t}$ and $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$. Furthermore, $I_{i,t}/K_{i,t}$ and $\Delta\log Y_{i,t}$ are very volatile, negatively serially correlated and fully response to proxy for $Z_{i,t}/K_{i,t}$.

Keeping other parameters constant, Table 3.1 illustrates the moments when $b^q = 0.2$, $b^i = 0.2$ and $b^f = 0.02$, respectively. Comparison with Model A shows which moments are informative about capital adjustment costs. Overall the presence of these adjustment costs mainly affects $I_{i,t}/K_{i,t}$ and $\Delta\log Y_{i,t}$, that is to make them less dispersed, more right skewed, more persistent, and less responsive to the proxy for $Z_{i,t}/K_{i,t}$. However, these effects are much stronger on $I_{i,t}/K_{i,t}$ than on $\Delta\log Y_{i,t}$. The rational comes from the fact that $\Delta\log Y_{i,t}$ is a linear combination of $\Delta\log Z_{i,t}$ and $\Delta\log\widehat{K}_{i,t}$. The growth rate $\Delta\log Z_{i,t}$ depends on the serial correlation ρ, the growth rate μ and the standard deviation of idiosyncratic shocks σ. Instead, the growth rate $\Delta\log\widehat{K}_{i,t}$, which is the investment rate net of depreciation rate, depends on both the stochastic process (ρ,μ,σ) and the adjustment costs $\left(b^q, b^i, b^f\right)$. Therefore using investment rate $I_{i,t}/K_{i,t}$ and sales growth rate $\Delta\log Y_{i,t}$ jointly could distinguish the adjustment costs from the stochastic process. Among the three forms of adjustment costs themselves, b^q and b^i will both make $I_{i,t}/K_{i,t}$ more persistent, while b^i and b^f will both make $I_{i,t}/K_{i,t}$ more right skewed. This distinguishes different forms of adjustment costs from each other. Finally, the presence of these adjustment costs also makes $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$ more dispersed. However, the magnitude is relatively small and the effect is mainly on within-group standard deviation. We then simulate a model when all three forms of adjustment costs are in present and call it Model B.

Table 3.2 illustrates the moments when $\sigma_\tau = 0.5$, $\sigma_{\log\beta} = 0.5$ and $\sigma_{\log\eta} = 0.5$, respectively. Comparison with Model B shows which moments are informative about unobserved heterogeneities. Overall the presence of these unobserved heterogeneities mainly affects $\pi_{i,t}/Y_{i,t}$ and $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$. Across column (1) to (3), we find that all three forms of unobserved heterogeneities will generate large between-group standard deviation, high serial correlation in $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$ and low cross correlation $corr\left(I_{i,t}/K_{i,t}, \log\left(Y_{i,t}/K_{i,t}\right)\right)$ and $corr\left(\Delta\log Y_{i,t}, \log\left(Y_{i,t}/K_{i,t}\right)\right)$. However, only $\sigma_{\log\beta}$ and $\sigma_{\log\eta}$ will cause between-group standard deviation and serial correlation in $\pi_{i,t}/Y_{i,t}$. This distinguish our fundamental parameter of interest σ_τ from other forms of unob-

[5] In all the simulations reported in Table 3, we impose $r = 0.12$, $\delta = 0.06$, $\mu_{\log\beta} = \mu_{\log\eta} = -2.30$, $\rho = 0.88$, $\mu = 0.05$ and $\sigma = 0.30$, simulate $S = 10$ pathes for a panel of 10000 firms and 26 years, and calculate moments using data in the last 6 years.

served heterogeneities. Among $\sigma_{\log\beta}$ and $\sigma_{\log\eta}$, as indicated by Equations (4) and (14),

$$corr\left(\pi_{i,t}/Y_{i,t}, \log\left(Y_{i,t}/\widehat{K}_{i,t}\right)\right)\begin{cases} >0, \text{ if } \sigma_{\log\eta}>0 \text{ and } \sigma_{\log\beta}=0 \\ <0, \text{ if } \sigma_{\log\eta}=0 \text{ and } \sigma_{\log\beta}<0 \end{cases}$$

This implies that the correlation between $\pi_{i,t}/Y_{i,t}$ and $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$ provides additional information to distinguish $\sigma_{\log\eta}$ from $\sigma_{\log\beta}$. In the last column of Table 3.2, we then simulate a model with both adjustment costs and unobserved heterogeneities and label it as Model C.

Using Model C as benchmark, Table 3.3 illustrates which moments are informative about measurement errors by simulating $\sigma_{meK}=0.2$, $\sigma_{meY}=0.2$ and $\sigma_{me\pi}=0.2$, respectively. The common finding is that whenever a variable is contaminated with measurement errors, it will become more dispersed. However, the effects are mainly on the within-group standard deviation instead of the between-group standard deviation. Furthermore, measurement errors also substantially reduce the serial correlation of the variables. These two facts distinguish the transitory measurement errors we have modeled from the capital adjustment costs and unobserved heterogeneities. Among the three measurement errors, σ_{meK} will only affect moments on $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$ and $I_{i,t}/K_{i,t}$; σ_{meY} will only affect moments on $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$, $\pi_{i,t}/Y_{i,t}$ and $\Delta\log Y_{i,t}$; while $\sigma_{me\pi}$ 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.

4.3.3 Features of Empirical Moments

With these parameters-moments properties in mind, we now take a closer look at the features of empirical moments. The left panel of Table 2 reports the value of these empirical moments and their standard errors estimated from the UK data, while the right panel presents the corresponding values for the Chinese sample.

On average firms in China have similar $\pi_{i,t}/Y_{i,t}$ and $I_{i,t}/K_{i,t}$ as in UK, but a lower $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$ and a higher $\Delta\log Y_{i,t}$, both of which are consistent with a lower depreciation rate δ imposed in China than in UK. In both countries, the between-group standard deviation for $\pi_{i,t}/Y_{i,t}$ and $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$ are much larger than the within-group counterparts, which highlights the importance of unobserved heterogeneities. Although $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$ is indeed more dispersed in China than in UK, one cannot infer more severe distortions in China simply from this difference, since $\pi_{i,t}/Y_{i,t}$ is also more dispersed in China than in UK. Both $I_{i,t}/K_{i,t}$ and $\Delta\log Y_{i,t}$ are more volatile in China than in UK, which may imply a higher level of uncertainty or more measurement errors in China. In both samples, $\pi_{i,t}/Y_{i,t}$ is positively skewed and $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$ is much more symmetric, which are consistent with the log-normality assumption for $(1+\tau_i)$, β_i and η_i. $I_{i,t}/K_{i,t}$ is positively skewed and $\Delta\log Y_{i,t}$ is much more symmetric. This highlights the importance of b^i or b^f in both samples. The high serial correlation for both $\pi_{i,t}/Y_{i,t}$ and $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$ is another indicator for the importance of unobserved heterogeneities. $I_{i,t}/K_{i,t}$ is positively correlated in both samples, which is consistent with the presence of b^q. $\Delta \log Y_{i,t}$ is positively correlated in UK but negatively correlated in China, which implies that σ_{meY} might be greater in the Chinese sample. The low correlation between $I_{i,t}/K_{i,t}$, $\Delta \log Y_{i,t}$ and the proxy of $Z_{i,t}/K_{i,t}$ is consistent with the importance of capital adjustment costs, but is also consistent with large heterogeneities in the data.

4.4 Empirical Results

Table 4 presents our estimation results.[6] For each country, the first column reports the optimal estimates of the structural parameters and the second column lists the corresponding numerical standard errors. Simulated moments at these optimal estimates are listed in the lower panel of Table 4 to compare with their empirical counterparts. Overall the model has provided a close fit to the large set of the moments, with an OI value equal to 826 for UK and 465 for China.

In both UK and China, σ_τ is estimated to be significantly different from zero, and is significantly larger in China than in UK. This implies that heterogeneity in capital goods prices does exist in both countries, but is indeed greater in China than in UK.

The estimated mean and standard deviation for $\log \beta_i$ implies that capital share β_i varies from 0.054 to 0.145 with a median at 0.088 in UK, and varies from 0.049 to 0.211 with a median at 0.102 in China. Both the dispersion and median value of our estimates are very close to those found in the empirical literature estimating capital share in a three factor model based on sales revenue, for example, Jorgenson, Gollop and Fraumeni (1987).[7] The estimated mean and standard deviation for $\log \eta_i$ implies that the demand elasticity $1/\eta_i$ varies from 5 to 20 with a median around 10 in UK, and varies from 4.5 to 30 with a median around 11 in China. The significant and substantial estimates for $\sigma_{\log \beta}$ and $\sigma_{\log \eta}$ highlight the fact that unobserved heterogeneities in production technology and market power are the salient features of firm-level data.

Two out of three forms of capital adjustment costs are found to be quantitatively important. In particular, $\widehat{b_q} = 0.098$ in UK and $\widehat{b_q} = 0.280$ in China. These quadratic adjustment costs imply an investment friction, which increases the user cost of capital by 1.23% for UK and 3.44% for China. Substantial level of irreversibility is estimated for both countries, which implies that the resale price of capital is about 30~40% lower than the purchase price of capital.

For the stochastic process, the model finds $\widehat{\mu} = 0.029$ in UK and 0.065 in China, while $\widehat{\sigma} = 0.251$ in UK and 0.338 in China. These estimates indicate that on average

[6] Wu (2009) discusses the details on how to draw the optimal weighting matrix and calculate numerical standard errors.

[7] Among the 28 U.S. manufacturing industries they estimated by production function regression over intermediate input, labor input and capital input, the capital share estimate varies from 0.0486 (apparel and other fabricated textile products) to 0.333 (tobacco) with median at 0.098 (electric machinery and equipment supplies). Such estimate for β should be distinguished from the one in an aggregate production function for value added with capital and labor inputs only, where they find a capital share of 0.385 for the U.S. in such an aggregate model.

the TFPR grows faster in China than in UK, however, firms in China also face a higher level of uncertainty. Such estimates are consistent with our prior based on the economic growth experience for both countries over the corresponding sample periods.

The model estimates significant measurement errors in capital stock and gross profit for both UK and China, and in sales revenue for China. As a double check for the magnitude of measurement errors, we estimate Equation (17) based on $\log \pi_{i,t}$ instead of $\log Y_{i,t}$. The corresponding GMM estimator for the AR(1) coefficient ρ is 0.557 for UK and 0.506 for China, in contrast to 0.878 for UK and 0.558 for China as we reported earlier. This is consistent with our structural estimates that $\sigma_{me\pi}$ is substantial and much larger than σ_{meY} in UK, while both $\sigma_{me\pi}$ and σ_{meY} are substantial and $\sigma_{me\pi}$ is slightly larger than σ_{meY} in China.

To summarize, these empirical results are consistent with the existence of significant capital market distortions. The investment model has been applied to both UK and China to control for potential model misspecification. Using UK as benchmark, we detect much more severe capital market distortions in China, even after controlling for more heterogenous production technology and market competition, more investment frictions and more volatile stochastic process, and more measurement errors in the data in China.

4.5 Specification Tests

There are three new features in this paper compared with the existing distortion literature: unobserved heterogeneities in capital share and inverse of demand elasticity, capital adjustment costs and measurement errors. This subsection reports specification tests for three restricted models, in order to understand the effect of missing each feature on the estimates of distortions. We illustrate the results for China in Table 5, where the preferred full model is listed as benchmark. Similar patterns are found for UK.

Column (1) shows the results of imposing no unobserved heterogeneities in capital share and inverse of demand elasticity, that is $\sigma_{\log\beta} = \sigma_{\log\eta} = 0$. As a sharp contrast to the benchmark, the estimated σ_τ increases significantly from 0.706 to 1.000 in this restricted model. The model also severely overestimates capital adjustment costs, level of uncertainty and measurement errors. Although the model can still fit the general features of $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$, it fails to match the pattern in the dispersion and persistence of $\pi_{i,t}/Y_{i,t}$. As a result, the overall fit of this restricted model degenerates enormously.

Column (2) 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. However, the estimate for σ_τ is still within the 95% confidence interval of the benchmark result. This is because we have used information on variables both in levels, namely $\pi_{i,t}/Y_{i,t}$ and $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$ and in growth rates, namely $I_{i,t}/K_{i,t}$ and $\Delta \log Y_{i,t}$. 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 production, demand and user cost of capital.

Column (3) 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 all the variables. Although it tends to overestimate capital adjustment costs and the level of uncertainty, the estimate for σ_τ is also within the 95% confidence interval of 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.

To summarize, 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. Consequently, a model without the unobserved heterogeneities in capital share 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.

4.6 Robustness Tests

Table 6 presents results for four robustness checks. Recall the benchmark model has imposed $\delta = 0.04$, $r = 0.14$, $\rho = 0.878$ and no aggregation for China. Columns (1) and (2) show the results for the same model but imposing $\delta = 0.03$ and 0.05, respectively. Columns (3) and (4) test the sensitivity of imposing $r = 0.12$ and $r = 0.16$, respectively. These alternative values of depreciation rate and discount rate allow us to investigate the importance of Jorgensonian user cost of capital. Column (5) imposes a lower value of serial correlation $\rho = 0.80$. This aims to test whether the empirical results in China is affected by imposing a value for ρ as high as that in UK. Column (6) assumes aggregation over two-plants for China so as to study whether the different estimates on distortions between UK and China is driven by the aggregation for UK.

Consistent with the relatively large standard errors for capital adjustment costs in the benchmark model, we find the estimates for b^q and b^i vary across different models, but they are all still significantly different from zero. The estimates for the growth rate μ is relatively sensitive to the choice of depreciation rate δ, but are in the right directions as one may expect. Nevertheless, across columns (1) to (6), the estimates for other parameters, in particular, for σ_τ, are not significantly different from those in our benchmark model. Therefore the key finding of our empirical exercises does not depend on our choice on Jorgensonian user cost of capital, serial correlation and aggregation within the reasonable ranges that we have considered.

4.7 Counterfactual Simulations

The estimated structural model provides a useful framework to quantify the effects of distortions on aggregate TFPR. Table 7.1 and 7.2 simulate such effects for UK and China respectively according to Equation (9). Since there are heterogeneities in both capital share and inverse of demand elasticity, these effects are simulated for different type of firms and the average effects are reported in the last rows.

In UK, on average the actual aggregate TFPR is 9.8% lower than the first-best benchmark, among which 0.5% is due to the investment frictions while 9.3% is caused by capital market distortions. All else being equal, the losses in aggregate TFPR increase monotonically with $1-\gamma$, the capital elasticity in the sales revenue or gross profit function. Intuitively, firms with larger capital share in production function and less market power in product market demand more capital stock hence suffer most from capital market distortions. Table 7.2 shows that similar qualitative patterns are found in China. However, quantitatively, the aggregate TFPR loss is as large as 40.9%, while the vast majority 37.5% is due to the capital market distortions, only 2.6% is due to the investment frictions.

Our finding that investment frictions 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 distortions on aggregate TFPR loss is much smaller than what have been estimated in Hshieh and Klenow (2009). Nevertheless, a 9.3% aggregate TFPR loss in UK and a 37.5% aggregate TFPR loss in China are still substantial. Had we not controlled for potential unobserved heterogeneities in capital share and inverse of demand elasticity, the estimated σ_τ would have implied a doubled magnitude.

We can also apply the estimated structural model as a laboratory, where controlled experiments can be conducted to investigate the hypothetical questions, such as, how aggregate TFPR losses would differ, if firms in China face distortions and frictions to the extent as in UK? Table 7.3 thus simulates the effects of distortions and frictions for China by imposing its σ_τ, b^q, and b^i to be the corresponding values estimated from UK. Table 7.4 reports gains from the improvement by comparing the quantitative effects in Table 7.2 with those in Table 7.3. We find that averaging across different type of firms, the aggregate TFPR in China would enhance by 23.4%, if these Chinese firms had been operating in an environment such as UK. And 91.9% of such improvement is from the reduction of capital market distortions.

5 A Generalized MRPK Approach

The counterfactual simulations highlight the substantial aggregate TFPR loss due to idiosyncratic distortions in the capital market. This naturally leads to the following question: what have caused such distortions? Since the distortions we model are unobservable, it motives us to link such distortions with some observable firm characteristics and check whether our model predictions are consistent with common institutional knowledge.

Our starting point is that when there are idiosyncratic distortions in capital market, all else being equal, $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$ is higher for a firm facing unfavorable distortions and lower for a firm benefiting from favorable distortions. Of course, all else is not equal. This is because, unobserved heterogeneities in capital share and inverse of demand elasticity, the presence of capital adjustment costs and measurement errors will all drive the dispersion in $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$. Our empirical exercises indeed find significant evidence for all these features. However, both capital adjustment costs and measurement errors mainly cause within-group dispersion in $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$, while distortions and heterogeneities in capital share and inverse of demand elasticity can both cause between-group dispersion in $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$. The good news is that heterogeneities in capital share and inverse of demand elasticity will also cause between-group dispersion in $\pi_{i,t}/Y_{i,t}$.

This implies that we can focus on the average log sales-to-capital ratio over time for each firm $\frac{1}{T}\sum_{t=1}^{T}\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$ to filter the effects of capital adjustment costs and measurement errors, and use the average log profit-to-sales ratio over time for each firm $\frac{1}{T}\sum_{t=1}^{T}\log\left(\pi_{i,t}/Y_{i,t}\right)$ to account for heterogeneities in capital share and inverse of demand elasticity. We then test whether the differences in the residual across firms are statistically related with some firm characteristics. In other words, we do not have to discard the MRPK approach completely, if panel data on capital stock, sales revenue and gross profit are available. To formalize this "generalized MRPK approach", consider the following reduced-form specification

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\log\left(Y_{i,t}/\widehat{K}_{i,t}\right) &= \log J-\log\beta_i\left(1-\eta_i\right)+\log\left(1+\tau_i\right)\\ &= \alpha_0+\alpha_1\frac{1}{T}\sum_{t=1}^{T}\log\left(\pi_{i,t}/Y_{i,t}\right)+\zeta_i \qquad (18)\\ &= \alpha_0+\alpha_1\frac{1}{T}\sum_{t=1}^{T}\log\left(\pi_{i,t}/Y_{i,t}\right)+\theta_1X_{1i}+\cdots+\theta_pX_{pi}+\epsilon_i\end{aligned}$$

where $(X_{1i},\cdots,X_{pi})$ contains a set of p observable firm characteristics. A positive estimator for θ_j implies a higher capital goods price or unfavorable distortions for firms with characteristics j.

Table 8 presents the empirical results for Equation (18) for our Chinese sample. Variables in upper case are dummy variables which indicate firm size, ownership, location and whether the chief manager of the firm is a member of the communist party. The baseline group is a large, collective-owned firm in Jilin province whose chief manager is a member of the communist party. Variables in lower case are continues variables which include the age of the firm, the age and years of education of the chief manager to control potential ability bias.

According to Table 8, all else being equal, a small firm faces an actual capital goods price that is 22.3% higher than a large firm while the price premium is 6.1% for a medium size firm. This is consistent with the typical findings in the large literature on capital market imperfections, for example, Fazzari, Hubbard and Peterson (1988), that small firms tend to face higher user cost of capital than large firms. Similar to

Dollar and Wei (2007), we find the private owned firms face the highest capital goods price across all types of ownership, which is 67.1% higher than those of their collective counterpart and 31.4% higher than those of state-owned firms. Compared with firms in Jilin province, on average firms in Jiangsu province have a capital goods price that is 34.4% higher. Such significant regional disparity is consistent with the finding in Brandt, Tombe and Zhu (2010). There is also significant evidence on the return to education, in the sense that on average an additional year of education for the chief manager will reduce the capital goods price for the firm by 4.9%. Finally, a firm with a chief manager who is a member of the communist party pays a capital goods price that is 12.0% lower than otherwise. This implies that there is indeed significant economic benefit on the communist party membership in China, as emphasized in Li, Meng, Wang and Zhou (2008).

If one takes the residual in the log sales-to-capital ratio as an indicator of idiosyncratic policy and institution distortions, the generalized MRPK approach indicates that small, private-owned firms without political connection in east China face unfavorable distortions compared with large, non-private-owned firms with political connection in other areas in China, and the magnitude of such distortions is substantial. Therefore economic reforms that channel capital towards small, private-owned firms in a more market oriented environment will stimulate economic growth even without further capital accumulation.

6 Alternative Explanations

6.1 Potential Heterogeneities in r and δ

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 and assume a common Jorgensonian user cost of capital J. Since $J = \frac{r+\delta}{1+r}$, at the other extreme, one could attribute all the unobserved heterogeneity in the user cost of capital to firm-specific discount rate r_i and/or depreciation rate δ_i and assume a common capital goods price across firms.

In theory, a firm-specific discount rate may arise due to either a firm-specific market beta if there is systematic risk, or a firm-specific consumption beta if investors cannot fully diversify firm-specific risk. However, distinguishing firm-specific market beta and consumption beta from firm-specific capital goods prices requires additional information on firm value, which is either unavailable or poorly measured. Nevertheless, if one believes the capital market should be relatively undistorted for UK firms, one may interpret the estimated heterogeneity in the user cost of capital in UK as the result of heterogeneity in the discount rate. Taking UK as a benchmark, the significantly larger heterogeneity in the user cost of capital in China still imply the importance of capital market distortions in China.

To check to what extent the heterogeneity in the user cost of capital may be explained by firm-specific depreciation rate δ_i, we assume δ_i follows a uniform distribution with mean $\overline{\delta}$ and standard deviation σ_δ, that is $\delta_i \sim U\left(\overline{\delta}, \sigma_\delta\right)$. Different from the firm-specific discount rate, it is possible to identify σ_δ using our existing set of moments. This is because the differences between the between-group standard deviations of investment rate $\left(\frac{I_{i,t}}{K_{i,t}} \simeq \mu + \delta_i\right)$ and sales growth rate $(\Delta \log Y_{i,t} \simeq \mu)$, can be attributed to potential heterogeneity in δ_i.

Table 9 reports the empirical results when we assume $\overline{\delta} = 0.08$ for UK and 0.04 for China and estimate σ_δ together with other structural parameters. We find by allowing for potential heterogeneity in δ_i does reduce our point estimates for σ_τ both in UK and in China; however, such estimates are still within the 95% confidence interval of our benchmark estimates for σ_τ. Furthermore, although introducing this additional dimension of heterogeneity improves the overall fit of the model, the estimates for σ_δ are not locally significant. This implies that heterogeneity in depreciation rate can only account for a small and insignificant part in the large unobserved heterogeneity in the user cost of capital.

6.2 Potential Distortions in Product and Labor Markets

Since the focus point of this paper is capital market distortions, we have assumed no distortion in the product and labor markets. We now consider how our estimated capital market distortions may be affected if there are potential distortions in other markets.

Suppose there are product market distortions in the form of firm-specific tax rate $\left(\tau_i^Y < 0\right)$ or subsidy rate $\left(\tau_i^Y > 0\right)$ on sales, which implies a new short-run gross profit maximization problem

$$\pi_{i,t} = \max_{M_{i,t},\, L_{i,t}} \left\{\left(1+\tau_i^Y\right) P_{i,t} Q_{i,t} - m_t M_{i,t} - w_t L_{i,t}\right\}$$

Then the intra-temporal and inter-temporal optimalities become

$$\frac{\pi_{i,t}}{Y_{i,t}} = \left(1+\tau_i^Y\right)\left[\beta_i(1-\eta_i)+\eta_i\right]$$

and

$$\frac{Y_{i,t}}{\widehat{K}_{i,t}} \simeq \frac{J_t\left(1+\tau_i^K\right)\left(1+b^q \frac{I_{i,t}}{K_{i,t}}\right)}{\left(1+\tau_i^Y\right)\beta_i\left(1-\eta_i\right)}$$

From the view of econometricians, one cannot separate the heterogeneity in τ_i^Y from those in β_i and η_i. One intuitive example is that, when firm A has higher profit-to-sales ratio than firm B, one cannot tell whether such difference is because firm A faces favorable product market distortion or because firm A has more market power by nature. At one extreme, by assuming no heterogeneities in production technology and market power, Hsieh and Klenow (2009) attributes such difference completely to product market distortions; at the other extreme, by assuming no product market distortions, we attributes such difference completely to unobserved heterogeneities in production technology and market power. Without additional information, one cannot tell which assumption is a better description of reality. Therefore we take a more conservative strategy by allowing for heterogeneities in β_i and η_i, which implicitly absorb potential product market distortions in our framework. Although we cannot tell how severe the product market distortions are, we are able to at least get an uncontaminated estimate on capital market distortions.

Suppose there are labor market distortions in the form of firm-specific wage rate

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

where w_t is the average wage rate in the economy and τ_i^L is a firm-specific component due to various labor market distortions. The short-run gross profit maximization problem now becomes

$$\pi_{i,t} = \max_{M_{i,t},\, L_{i,t}} \left\{P_{i,t}Q_{i,t} - m_t M_{i,t} - w_t \left(1 + \tau_i^L\right) L_{i,t}\right\}$$

with the first-order-conditions for optimal material input and labor input as following

$$\begin{aligned} \frac{m_t M_{i,t}}{Y_{i,t}} &= \alpha_i(1 - \eta_i) \\ \frac{w_t L_{i,t}}{Y_{i,t}} &= \left(1 + \tau_i^L\right)(1 - \alpha_i - \beta_i)(1 - \eta_i) \end{aligned}$$

If we assume $\log \alpha_i \overset{i.i.d}{\sim} N\left(\mu_{\log \alpha}, \sigma^2_{\log \alpha}\right)$ and $\log\left(1 + \tau_i^L\right) \overset{i.i.d}{\sim} N\left(0, \sigma^2_{\tau^L}\right)$, and take β_i and η_i as given from capital market distortions estimation, we are able to estimate $\mu_{\log \alpha}$ and $\sigma_{\log \alpha}$ from the mean and between-group standard deviation of material cost share $\frac{m_t M_{i,t}}{Y_{i,t}}$ and back out σ_{τ^L} from the between-group standard deviation of labor cost share $\frac{w_t L_{i,t}}{Y_{i,t}}$.

One potential concern is that firms pay different wage rates because the quality of labor is different across firms. To control for such possibility, we define effective number of employees of firm i in year t as

$$L^e_{i,t} = e^{(b_t E_{i,t})} L_{i,t}$$

where b_t is the time-varying Mincerian rate of return to education and $E_{i,t}$ is educational attainment measured as years of schooling. We adopt estimates for b_t from Zhang, Zhao, Park and Song (2005), which vary from 6.7% to 9.9% over our sample period. We assume 6 years of education for employees whose educational attainment is below junior high school; 12 years of education for those whose have finished senior high school or professional senior high school; and 16 years of education for those who have achieved a college degree or above.

Table 10 reports the empirical results. The estimate $\widehat{\mu}_{\log \alpha} = -0.607$ implies that the median material share in the production function equals to 0.545, while there is also substantial heterogeneity as reflected by $\widehat{\sigma}_{\log \alpha} = 0.456$. Even after taking into account potential differences in the quality of labor, $\widehat{\sigma}_{\tau^L}$ is estimated to be 0.229 and is significantly different from zero. Such evidence is consistent with the hypothesis of labor market distortions in China, although the magnitude of such distortions is much smaller than that of capital market distortions.

7 Conclusion

This paper investigates three important questions on capital market distortions. First, we find there is significant evidence of capital market distortions in both UK and China, with the magnitude in China much greater than that in UK. Second, such distortions imply substantial aggregate TFPR losses, which are around 9% in UK and 38% in China. Finally, small, private-owned firms without political connection in east China faces unfavorable distortions.

Although this is not the first paper in inferring capital market distortions from dispersion in log sales-to-capital ratio, we contribute to this MRPK approach by an identification strategy, which accounts for unobserved heterogeneities in production technology and market power, capital adjustment costs and potential measurement errors in the data. We employ a structural econometric approach to highlight the importance of missing potential unobserved heterogeneities. This allows us to develop a generalized MRPK approach in linking the "residual" in log sales-to-capital ratio with observable firm characteristics.

Like most research in this line, we have modelled capital market distortions through unobserved heterogeneity in capital goods prices. By nature such unobserved heterogeneity is assumed to be time-invariant. When longer panels are available, it would be both important and interesting to study how capital market distortions have been evolving, especially for a country like China where new economic reforms constantly take place. Furthermore, the investment model estimated in this paper is for existing firms only hence neglects entry and exit. Intuitively, taking into account this extensive margin may even enlarge the magnitude estimated in this paper, if one assumes deferred entry and exit due to idiosyncratic distortions. Such extensions are beyond the scope of this paper and it is important to develop future works along these two lines.

Data Appendix

In the Chinese sample, we use annual gross investment (B108) as a measure of $I_{i,t}$. Using annual average net book value of tangible fixed capital assets (B11) and investment expenditure of each year (B108), $K_{i,t}$ is constructed according to Equation (5) with a depreciation rate of 0.04. We use price indices of investment in fixed assets by province from the *China Statistic Yearbook* to deflate the capital stock series. $Y_{i,t}$ is defined as sales revenue of products (B31) plus changes in the inventory of finished products (B4). Since China has experienced a period from high inflation to deflation during 1994 to 1999, the survey explicitly asked the annual percentage change in the price of its main product for each firm and in each year (B140). We use this information to deflate the sales series. Gross profit $\pi_{i,t}$ is constructed by subtracting the total costs of products sold (B32) from the sales revenue (B31). A measure for net profit is constructed using the bottom-up method. That is we start with the net profit after depreciation, interest and tax (B43), to recover the net profit before interest, tax and depreciation by topping up depreciation (B10), interest (B41) and tax (B38).

In the UK sample, $I_{i,t}$ is defined as total new fixed assets (DS435) less sales of fixed assets (DS423). $K_{i,t}$ is constructed by applying a perpetual inventory procedure with a depreciation rate of 0.08. The starting value was based on the net book value of the tangible fixed capital assets (DS339) in the first observation within our sample period, adjusted for previous inflation. Subsequent values were obtained using accounts data on investment and asset sales and an aggregate series for investment goods prices. $Y_{i,t}$ is the value of total sales (DS104) deflated by the aggregate GDP deflator. There is no information on gross profit in the UK dataset. Instead we construct a net profit before interest, tax and depreciation by adding depreciation (DS136) back to the operating profit (DS137), which is the net profit before interest, tax and after depreciation. We then match the gross profit based on net profit for UK, using the relationship between gross profit and net profit from China.

In both samples, to exclude outliers, we trim the top and bottom 1% of observations for $\pi_{i,t}/Y_{i,t}$ and $\log\left(Y_{i,t}/\widehat{K}_{i,t}\right)$; and restrict $I_{i,t}/K_{i,t}$ and $\Delta \log Y_{i,t}$ within the range of [-1, 1].

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

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

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

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

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

[14] 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.

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

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

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

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

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

[20] 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.

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

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

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

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

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

[26] 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.

Figure 1. Investment Policies

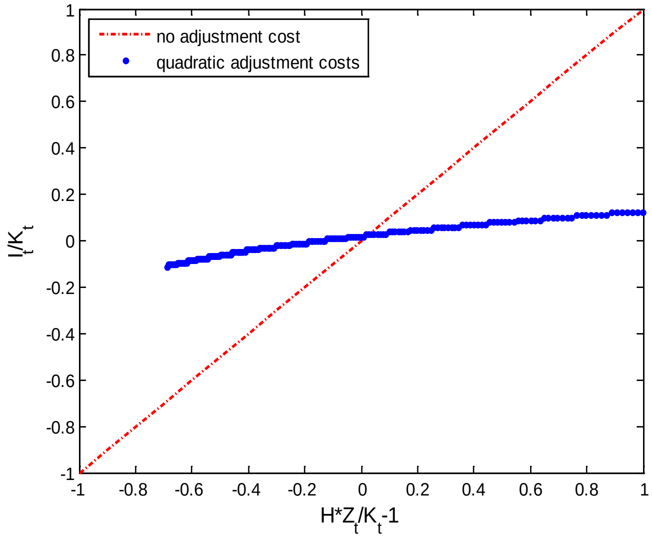

Figure 1.1. Quadratic Adjustment Costs

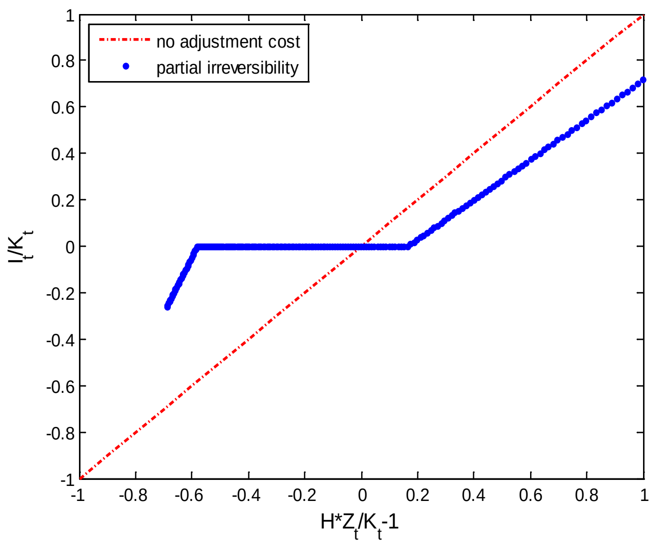

Figure 1.2. Irreversibility

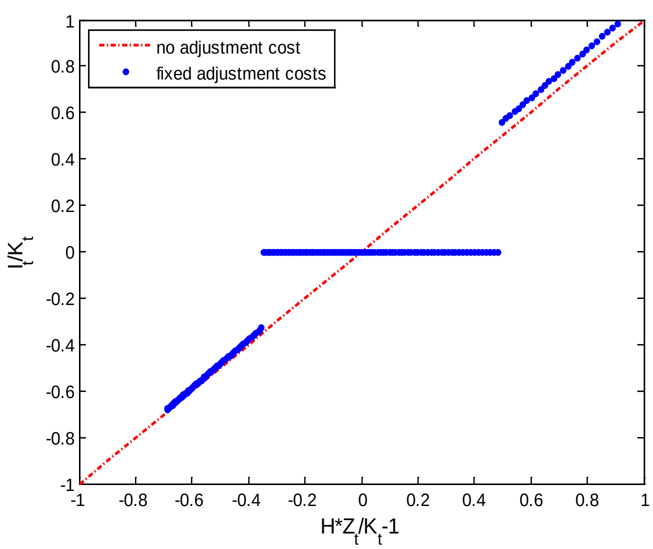

Figure 1.3. Fixed Adjustment Costs

Table 1. Parameters to Estimate

Parameters	Definition
σ_{τ}	standard deviation of heterogeneities in distortions
$\mu_{log\beta}$	mean of log capital share in production function
$\sigma_{log\beta}$	standard deviation of heterogeneities in log capital share
$\mu_{log\eta}$	mean of log 1/demand elasticity
$\sigma_{log\eta}$	standard deviation of heterogeneities in log 1/demand elasticity
b^q	quadratic adjustment costs
b^i	irreversibility
b^f	fixed adjustment costs
μ	mean of TFPR growth rate
σ	standard deviation of TFPR shocks
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 gross profit

Table 2. Empirical Moments for UK and China

	UK		China	
Number of Firms	629		701	
Number of Years	9		6	
Median No. of Employees	1102		1055	
Mean No. of Employees	4856		2011	
Set of Moments	**moments**	**s.e.**	**moments**	**s.e.**
mean(π/Y)	0.210	0.002	0.231	0.004
mean(log(Y/Khat))	0.748	0.022	0.588	0.031
mean(I/K)	0.125	0.003	0.123	0.004
mean(ΔlogY)	0.030	0.003	0.076	0.004
bsd(π/Y)	0.070	0.002	0.104	0.002
wsd(π/Y)	0.031	0.001	0.054	0.001
bsd(log(Y/Khat))	0.539	0.017	0.828	0.020
wsd(log(Y/Khat))	0.216	0.005	0.362	0.010
bsd(I/K)	0.095	0.004	0.129	0.006
wsd(I/K)	0.103	0.003	0.108	0.004
bsd(ΔlogY)	0.082	0.004	0.104	0.004
wsd(ΔlogY)	0.139	0.003	0.277	0.005
skew(π/Y)	0.522	0.074	0.706	0.057
skew(log(Y/Khat))	0.009	0.137	-0.117	0.105
skew(I/K)	2.723	0.096	2.575	0.099
skew(dlogY)	0.585	0.109	-0.105	0.045
scorr(π/Y)	0.880	0.008	0.799	0.012
scorr(log(Y/Khat))	0.939	0.005	0.904	0.008
scorr(I/K)	0.392	0.023	0.491	0.032
scorr(ΔlogY)	0.216	0.018	-0.310	0.018
corr(π/Y, log(Y/Khat))	-0.245	0.029	-0.223	0.028
corr(I/K, log(Y/K))	0.139	0.022	0.246	0.024
corr(ΔlogY, log(Y/K))	0.150	0.016	0.191	0.015

Table 3.1. Illustration for Identification of Capital Adjustment Costs

Parameters	**Model A**	**col (1)**	**col (2)**	**col (3)**	**Model B**
	$\boldsymbol{b^q = 0.0}$	$b^q = 0.2$	$b^q = 0.0$	$b^q = 0.0$	$b^q = 0.2$
	$\boldsymbol{b^i = 0.0}$	$b^i = 0.0$	$b^i = 0.2$	$b^i = 0.0$	$b^i = 0.2$
	$\boldsymbol{b^f = 0.0}$	$b^f = 0.0$	$b^f = 0.0$	$b^f = 0.02$	$b^f = 0.02$
Set of Moments					
mean(π/Y)	0.190	0.190	0.190	0.190	0.190
mean(log(Y/Khat))	0.576	0.575	0.541	0.581	0.589
mean(I/K)	0.171	**0.128**	**0.136**	0.171	0.129
mean(ΔlogY)	0.050	0.050	0.050	0.050	0.050
bsd(π/Y)	0.000	0.000	0.000	0.000	0.000
wsd(π/Y)	0.000	0.000	0.000	0.000	0.000
bsd(log(Y/Khat))	0.002	**0.067**	**0.089**	**0.038**	0.094
wsd(log(Y/Khat))	0.002	**0.086**	**0.095**	**0.070**	0.100
bsd(I/K)	0.127	**0.084**	**0.100**	0.135	0.083
wsd(I/K)	0.342	**0.124**	**0.200**	0.367	0.140
bsd(ΔlogY)	0.119	**0.099**	**0.099**	0.115	0.095
wsd(ΔlogY)	0.280	**0.191**	**0.202**	0.258	0.189
skew(π/Y)	0.000	0.000	0.000	0.000	0.000
skew(log(Y/Khat))	0.000	0.023	-1.526	-0.049	-1.112
skew(I/K)	0.981	**0.357**	**2.311**	**1.743**	0.898
skew(dlogY)	0.000	-0.010	**0.634**	**0.353**	0.290
scorr(π/Y)	N.A.	N.A.	N.A.	N.A.	N.A.
scorr(log(Y/Khat))	N.A.	0.501	0.619	0.193	0.618
scorr(I/K)	-0.064	**0.448**	**0.135**	-0.079	0.289
scorr(ΔlogY)	-0.066	**0.064**	0.019	-0.026	0.036
corr(π/Y, log(Y/Khat))	N.A.	N.A.	N.A.	N.A.	N.A.
corr(I/K, log(Y/K))	0.977	0.986	**0.874**	**0.933**	0.903
corr(ΔlogY, log(Y/K))	1.000	**0.889**	**0.864**	**0.966**	0.834

Table 3.2. Illustration for Identification of Unobserved Heterogeneities

Parameters	**Model B**	col (1)	col (2)	col (3)	Model C
	$\boldsymbol{\sigma_{\tau} = 0.0}$	$\sigma_{\tau} = 0.5$	$\sigma_{\tau} = 0.0$	$\sigma_{\tau} = 0.0$	$\sigma_{\tau} = 0.5$
	$\boldsymbol{\sigma_{\log\beta} = 0.0}$	$\sigma_{\log\beta} = 0.0$	$\sigma_{\log\beta} = 0.5$	$\sigma_{\log\beta} = 0.0$	$\sigma_{\log\beta} = 0.5$
	$\boldsymbol{\sigma_{\log\eta} = 0.0}$	$\sigma_{\log\eta} = 0.0$	$\sigma_{\log\eta} = 0.0$	$\sigma_{\log\eta} = 0.5$	$\sigma_{\log\eta} = 0.5$
Set of Moments					
mean(π/Y)	0.190	0.190	0.202	0.202	0.213
mean(log(Y/Khat))	0.589	0.591	0.589	0.606	0.608
mean(I/K)	0.129	0.129	0.129	0.129	0.130
mean(ΔlogY)	0.050	0.050	0.050	0.050	0.050
bsd(π/Y)	0.000	0.000	**0.049**	**0.049**	0.068
wsd(π/Y)	0.000	0.000	0.000	0.000	0.000
bsd(log(Y/Khat))	0.094	**0.496**	**0.502**	**0.112**	0.703
wsd(log(Y/Khat))	0.100	0.101	0.102	0.103	0.105
bsd(I/K)	0.083	0.083	0.083	0.083	0.083
wsd(I/K)	0.140	0.142	0.141	0.142	0.146
bsd(ΔlogY)	0.095	0.094	0.095	0.095	0.094
wsd(ΔlogY)	0.189	0.189	0.190	0.191	0.192
skew(π/Y)	0.000	0.000	0.344	0.344	0.117
skew(log(Y/Khat))	-1.112	**-0.029**	**0.065**	**-0.573**	0.022
skew(I/K)	0.898	1.037	0.935	0.951	1.142
skew(dlogY)	0.290	0.301	0.270	0.263	0.260
scorr(π/Y)	N.A.	N.A.	1.000	1.000	1.000
scorr(log(Y/Khat))	0.618	**0.972**	**0.972**	**0.669**	0.984
scorr(I/K)	0.289	0.269	0.279	0.274	0.243
scorr(ΔlogY)	0.036	0.034	0.033	0.031	0.025
corr(π/Y, log(Y/Khat))	N.A.	N.A.	**-0.948**	**0.365**	-0.428
corr(I/K, log(Y/K))	0.903	**0.423**	0.413	**0.873**	0.318
corr(ΔlogY, log(Y/K))	0.834	**0.378**	0.375	**0.806**	0.280

Table 3.3. Illustration for Identification of Measurement Errors

Parameters	**Model C**	col (1)	col (2)	col (3)	Model D
	σ_{meK} = 0.0	*σ_{meK} = 0.2*	*σ_{meK} = 0.0*	*σ_{meK} = 0.0*	*σ_{meK} = 0.2*
	σ_{meY} = 0.0	*σ_{meY} = 0.0*	*σ_{meY} = 0.2*	*σ_{meY} = 0.0*	*σ_{meY} = 0.2*
	$\sigma_{me\,\pi}$ = 0.0	*$\sigma_{me\,\pi}$ = 0.0*	*$\sigma_{me\,\pi}$ = 0.0*	*$\sigma_{me\,\pi}$ = 0.2*	*$\sigma_{me\,\pi}$ = 0.2*
Set of Moments					
mean(π/Y)	0.213	0.213	0.218	0.218	0.222
mean(log(Y/Khat))	0.608	0.607	0.608	0.608	0.606
mean(I/K)	0.130	0.132	0.130	0.130	0.132
mean(ΔlogY)	0.050	0.050	0.050	0.050	0.050
bsd(π/Y)	0.068	0.068	0.072	0.072	0.076
wsd(π/Y)	0.000	0.000	**0.042**	**0.042**	0.061
bsd(log(Y/Khat))	0.703	0.707	0.708	0.703	0.712
wsd(log(Y/Khat))	0.105	**0.196**	**0.211**	0.105	0.268
bsd(I/K)	0.083	**0.087**	0.083	0.083	0.087
wsd(I/K)	0.146	**0.154**	0.146	0.146	0.154
bsd(ΔlogY)	0.094	0.094	**0.110**	0.094	0.110
wsd(ΔlogY)	0.192	0.192	**0.337**	0.192	0.337
skew(π/Y)	0.117	0.117	**0.636**	**0.639**	0.982
skew(log(Y/Khat))	0.022	0.017	0.019	0.022	0.015
skew(I/K)	1.142	**1.354**	1.142	1.142	1.354
skew(dlogY)	0.260	0.260	**0.055**	0.260	0.055
scorr(π/Y)	1.000	1.000	**0.691**	**0.692**	0.523
scorr(log(Y/Khat))	0.984	**0.924**	**0.912**	0.984	0.860
scorr(I/K)	0.243	**0.228**	0.243	0.243	0.228
scorr(ΔlogY)	0.025	0.025	**-0.310**	0.025	-0.310
corr(π/Y, log(Y/Khat))	-0.428	**-0.414**	**-0.484**	**-0.355**	-0.408
corr(I/K, log(Y/K))	0.318	**0.336**	**0.307**	0.318	0.325
corr(ΔlogY, log(Y/K))	0.280	**0.270**	**0.310**	0.280	0.300

Table 4. Empirical Results for UK and China

	UK		China	
Parameters	**estimate**	**s.e.**	**estimate**	**s.e.**
σ_{τ}	0.435	0.022	0.706	0.031
$\mu_{log\beta}$	-2.427	0.019	-2.281	0.032
$\sigma_{log\beta}$	0.406	0.025	0.595	0.033
$\mu_{log\eta}$	-2.255	0.017	-2.422	0.067
$\sigma_{log\eta}$	0.585	0.022	0.797	0.085
b^{q}	0.098	0.027	0.280	0.122
b^{i}	0.339	0.140	0.398	0.145
b^{f}	0.000	0.002	0.001	0.005
μ	0.029	0.002	0.065	0.002
σ	0.251	0.008	0.338	0.022
σ_{meK}	0.247	0.009	0.330	0.020
σ_{meY}	0.002	0.270	0.148	0.014
$\sigma_{me\,\pi}$	0.197	0.005	0.169	0.008
Set of Moments	**empirical**	**simulated**	**empirical**	**simulated**
mean(π/Y)	0.210	0.210	0.231	0.232
mean(log(Y/Khat))	0.748	0.743	0.588	0.553
mean(I/K)	0.125	0.125	0.123	0.124
mean(ΔlogY)	0.030	0.029	0.076	0.065
bsd(π/Y)	0.070	0.071	0.104	0.101
wsd(π/Y)	0.031	0.030	0.054	0.052
bsd(log(Y/Khat))	0.539	0.594	0.829	0.924
wsd(log(Y/Khat))	0.216	0.169	0.362	0.329
bsd(I/K)	0.095	0.033	0.129	0.078
wsd(I/K)	0.103	0.103	0.108	0.116
bsd(ΔlogY)	0.082	0.035	0.104	0.102
wsd(ΔlogY)	0.139	0.146	0.277	0.283
skew(π/Y)	0.522	0.468	0.706	0.679
skew(log(Y/Khat))	0.009	0.002	-0.117	0.012
skew(I/K)	2.723	1.239	2.575	2.053
skew(dlogY)	0.585	0.166	-0.105	0.046
scorr(π/Y)	0.880	0.838	0.799	0.749
scorr(log(Y/Khat))	0.939	0.926	0.904	0.874
scorr(I/K)	0.392	0.295	0.491	0.365
scorr(ΔlogY)	0.216	-0.011	-0.310	-0.242
corr(π/Y, log(Y/Khat))	-0.245	-0.184	-0.223	-0.303
corr(I/K, log(Y/K))	0.139	0.296	0.246	0.321
corr(ΔlogY, log(Y/K))	0.150	0.220	0.191	0.211
OI	826		465	

Table 5. Specification Tests

Parameters	benchmark	s.e.	col (1)	col (2)	col (3)
σ_{τ}	0.706	0.031	1.000	0.648	0.728
$\mu_{log\beta}$	-2.281	0.032	-2.004	-2.409	-2.464
$\sigma_{log\beta}$	0.595	0.033	**0.000**	0.615	0.465
$\mu_{log\eta}$	-2.422	0.067	-2.644	-2.314	-2.433
$\sigma_{log\eta}$	0.797	0.085	**0.000**	0.790	0.798
b^{q}	0.280	0.122	0.835	**0.000**	0.722
b^{i}	0.398	0.145	0.215	**0.000**	0.352
b^{f}	0.001	0.005	0.034	**0.000**	0.000
μ	0.065	0.002	0.061	0.080	0.069
σ	0.338	0.022	0.431	0.107	0.441
σ_{meK}	0.330	0.020	0.397	0.398	**0.000**
σ_{meY}	0.148	0.014	0.203	0.198	**0.000**
$\sigma_{me\pi}$	0.169	0.008	0.430	0.129	**0.000**
Set of Moments	**empirical**		**simulated**		
mean(π/Y)	0.231	0.232	0.220	0.232	0.201
mean(log(Y/Khat))	0.588	0.553	0.301	0.703	0.725
mean(I/K)	0.123	0.124	0.118	0.146	0.120
mean(ΔlogY)	0.076	0.065	0.059	0.080	0.068
bsd(π/Y)	0.104	0.101	0.045	0.105	0.085
wsd(π/Y)	0.054	0.052	0.101	0.055	0.000
bsd(log(Y/Khat))	0.829	0.924	0.951	0.914	0.843
wsd(log(Y/Khat))	0.362	0.329	0.397	0.372	0.192
bsd(I/K)	0.129	0.078	0.080	0.057	0.072
wsd(I/K)	0.108	0.116	0.115	0.147	0.086
bsd(ΔlogY)	0.104	0.102	0.107	0.069	0.116
wsd(ΔlogY)	0.277	0.283	0.332	0.293	0.250
skew(π/Y)	0.706	0.679	1.638	0.743	0.286
skew(log(Y/Khat))	-0.117	0.012	-0.004	-0.003	0.021
skew(I/K)	2.575	2.053	2.284	1.352	1.450
skew(dlogY)	-0.105	0.046	0.037	0.005	0.093
scorr(π/Y)	0.799	0.749	-0.001	0.742	1.000
scorr(log(Y/Khat))	0.904	0.874	0.830	0.830	0.967
scorr(I/K)	0.491	0.365	0.357	-0.050	0.545
scorr(ΔlogY)	-0.310	-0.242	-0.324	-0.442	-0.022
corr(π/Y, log(Y/Khat))	-0.223	-0.303	-0.079	-0.280	-0.163
corr(I/K, log(Y/K))	0.246	0.321	0.347	0.242	0.381
corr(ΔlogY, log(Y/K))	0.191	0.211	0.213	0.169	0.264
OI		465	5276	989	2883

Table 6. Robustness Tests

Parameters	benchmark	s.e.	r = 0.12	r = 0.16	δ = 0.03	δ = 0.05	ρ = 0.80	2 plants
σ_{τ}	0.706	0.031	0.682	0.707	0.735	0.713	0.649	0.662
$\mu_{log\beta}$	-2.281	0.032	-2.326	-2.230	-2.320	-2.284	-2.281	-2.319
$\sigma_{log\beta}$	0.595	0.033	0.581	0.600	0.596	0.600	0.615	0.622
$\mu_{log\eta}$	-2.422	0.067	-2.359	-2.435	-2.372	-2.417	-2.424	-2.415
$\sigma_{log\eta}$	0.797	0.085	0.782	0.800	0.800	0.796	0.800	0.793
b^q	0.280	0.122	0.177	0.279	0.235	0.190	0.173	0.150
b^i	0.398	0.145	0.367	0.252	0.307	0.296	0.333	0.305
b^f	0.001	0.005	0.000	0.000	0.000	0.000	0.000	0.000
μ	0.065	0.002	0.072	0.067	0.071	0.063	0.065	0.067
σ	0.338	0.022	0.308	0.335	0.334	0.312	0.333	0.292
σ_{meK}	0.330	0.020	0.306	0.331	0.333	0.312	0.318	0.291
σ_{meY}	0.148	0.014	0.153	0.157	0.147	0.160	0.148	0.150
$\sigma_{me\pi}$	0.169	0.008	0.172	0.168	0.184	0.172	0.176	0.171
Set of Moments	**empirical**				**simulated**			
mean(π/Y)	0.231	0.232	0.232	0.237	0.234	0.233	0.233	0.229
mean(log(Y/Khat))	0.588	0.553	0.492	0.592	0.534	0.600	0.546	0.505
mean(I/K)	0.123	0.124	0.132	0.126	0.119	0.134	0.123	0.125
mean(ΔlogY)	0.076	0.065	0.072	0.067	0.071	0.063	0.065	0.067
bsd(π/Y)	0.104	0.101	0.101	0.103	0.104	0.102	0.103	0.100
wsd(π/Y)	0.054	0.052	0.054	0.054	0.055	0.055	0.054	0.051
bsd(log(Y/Khat))	0.829	0.924	0.900	0.928	0.950	0.930	0.897	0.921
wsd(log(Y/Khat))	0.362	0.329	0.308	0.331	0.333	0.314	0.318	0.303
bsd(I/K)	0.129	0.078	0.078	0.079	0.078	0.080	0.070	0.073
wsd(I/K)	0.108	0.116	0.121	0.118	0.118	0.124	0.121	0.116
bsd(ΔlogY)	0.104	0.102	0.099	0.102	0.102	0.100	0.095	0.098
wsd(ΔlogY)	0.277	0.283	0.282	0.289	0.283	0.288	0.285	0.289
skew(π/Y)	0.706	0.679	0.715	0.690	0.728	0.716	0.691	0.666
skew(log(Y/Khat))	-0.117	0.012	0.008	0.010	0.011	0.008	0.011	0.009
skew(I/K)	2.575	2.053	1.928	2.056	2.181	1.991	2.047	2.068
skew(dlogY)	-0.105	0.046	0.045	0.043	0.050	0.046	0.052	0.048
scorr(π/Y)	0.799	0.749	0.738	0.738	0.736	0.732	0.744	0.752
scorr(log(Y/Khat))	0.904	0.874	0.881	0.872	0.877	0.883	0.873	0.891
scorr(I/K)	0.491	0.365	0.351	0.364	0.356	0.347	0.286	0.336
scorr(ΔlogY)	-0.310	-0.242	-0.258	-0.258	-0.239	-0.268	-0.264	-0.267
corr(π/Y, log(Y/Khat))	-0.223	-0.303	-0.276	-0.325	-0.266	-0.310	-0.329	-0.360
corr(I/K, log(Y/K))	0.246	0.321	0.311	0.319	0.313	0.304	0.311	0.297
corr(ΔlogY, log(Y/K))	0.191	0.211	0.215	0.212	0.208	0.211	0.213	0.211
OI		465	466	465	465	459	495	466

7.1. Simulation for UK

type	β	η	1-γ	ΔlogTFPR overall	ΔlogTFPR frictions	ΔlogTFPR distortions
1	0.054	0.051	0.499	-0.099	-0.005	-0.094
2	0.054	0.105	0.315	-0.046	-0.002	-0.043
3	0.054	0.214	0.165	-0.020	-0.001	-0.019
4	0.088	0.051	0.620	-0.163	-0.009	-0.155
5	0.088	0.105	0.430	-0.075	-0.004	-0.071
6	0.088	0.214	0.245	-0.032	-0.002	-0.031
7	0.145	0.051	0.728	-0.268	-0.014	-0.254
8	0.145	0.105	0.553	-0.124	-0.007	-0.117
9	0.145	0.214	0.347	-0.053	-0.003	-0.050
average	0.096	0.124	0.433	-0.098	-0.005	-0.093

7.2. Simulation for China

type	β	η	1-γ	ΔlogTFPR overall	ΔlogTFPR frictions	ΔlogTFPR distortions
1	0.049	0.034	0.588	-0.382	-0.024	-0.359
2	0.049	0.089	0.337	-0.136	-0.009	-0.127
3	0.049	0.235	0.139	-0.043	-0.003	-0.040
4	0.102	0.034	0.747	-0.779	-0.045	-0.730
5	0.102	0.089	0.512	-0.281	-0.018	-0.264
6	0.102	0.235	0.250	-0.089	-0.006	-0.083
7	0.211	0.034	0.859	-1.209	-0.080	-1.051
8	0.211	0.089	0.685	-0.581	-0.035	-0.547
9	0.211	0.235	0.408	-0.185	-0.012	-0.173
average	0.121	0.119	0.502	-0.409	-0.026	-0.375

7.3. Simulation for China using UK as Counterfactual

type	β	η	1-γ	ΔlogTFPR overall	ΔlogTFPR frictions	ΔlogTFPR distortions
1	0.049	0.034	0.588	-0.149	-0.013	-0.136
2	0.049	0.089	0.337	-0.053	-0.005	-0.048
3	0.049	0.235	0.139	-0.017	-0.002	-0.015
4	0.102	0.034	0.747	-0.307	-0.026	-0.282
5	0.102	0.089	0.512	-0.110	-0.010	-0.100
6	0.102	0.235	0.250	-0.035	-0.003	-0.032
7	0.211	0.034	0.859	-0.611	-0.046	-0.556
8	0.211	0.089	0.685	-0.227	-0.020	-0.207
9	0.211	0.235	0.408	-0.072	-0.007	-0.065
average	0.121	0.119	0.502	-0.176	-0.015	-0.160

7.4. Gain from Improvement for China

type	β	η	1-γ	ΔΔlogTFPR overall	ΔΔlogTFPR frictions	ΔΔlogTFPR distortions
1	0.049	0.034	0.588	0.233	0.010	0.223
2	0.049	0.089	0.337	0.083	0.004	0.079
3	0.049	0.235	0.139	0.026	0.001	0.025
4	0.102	0.034	0.747	0.471	0.020	0.448
5	0.102	0.089	0.512	0.172	0.008	0.164
6	0.102	0.235	0.250	0.054	0.003	0.052
7	0.211	0.034	0.859	0.598	0.034	0.496
8	0.211	0.089	0.685	0.354	0.015	0.339
9	0.211	0.235	0.408	0.112	0.005	0.107
average	0.121	0.119	0.502	0.234	0.011	0.215

Table 8. Regression for $E_t[\log(Y_{i,t}/Khat_{i,t})]$ in China

	baseline	size	ownership	location	manager	full
$E_t[\log(\pi_{i,t}/Y_{i,t})]$	**-0.286**	**-0.296**	**-0.284**	**-0.248**	**-0.286**	**-0.256**
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
firm age	**0.008**	**0.009**	**0.003**	**0.006**	**0.008**	**0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
SMALL		**0.261**				**0.223**
		(0.031)				(0.031)
MEDIUM		**0.093**				**0.061**
		(0.023)				(0.023)
PRIVATE			**0.712**			**0.671**
			(0.117)			(0.125)
STATE			**0.306**			**0.357**
			(0.035)			(0.034)
SHAREHOLDING			**0.156**			**0.247**
			(0.038)			(0.042)
FOREIGN			**0.159**			**0.221**
			(0.060)			(0.060)
JIANGSU				**0.329**		**0.344**
				(0.027)		(0.026)
SHANXI				**0.119**		**0.134**
				(0.029)		(0.028)
SICHUAN				**0.080**		**0.120**
				(0.029)		(0.029)
manager age					**-0.006**	**-0.002**
					(0.002)	(0.002)
manager education					**-0.110**	**-0.049**
					(0.017)	(0.017)
PARTY MEMBER					**-0.079**	**-0.120**
					(0.057)	(0.058)

Note:

1. The baseline group is LARGE, COLLECTIVE, JILIN, and NOPARTYMEMBERSHIP.
2. Robust standard errors are reported in ( ).

Table 9. Potential Heterogeneity in Depreciation Rate

	UK		China	
Parameters	**estimate**	**s.e.**	**estimate**	**s.e.**
σ_{τ}	0.416	0.028	0.663	0.037
σ_{δ}	0.018	0.011	0.037	0.021
$\mu_{log\beta}$	-2.449	0.019	-2.324	0.031
$\sigma_{log\beta}$	0.418	0.029	0.609	0.037
$\mu_{log\eta}$	-2.214	0.015	-2.426	0.059
$\sigma_{log\eta}$	0.564	0.019	0.799	0.068
b^{q}	0.089	0.013	0.240	0.051
b^{i}	0.252	0.102	0.299	0.246
b^{f}	0.000	0.005	0.000	0.006
μ	0.032	0.002	0.068	0.002
σ	0.247	0.008	0.322	0.021
σ_{meK}	0.245	0.034	0.321	0.053
σ_{meY}	0.002	0.006	0.156	0.020
$\sigma_{me\,\pi}$	0.195	0.015	0.151	0.012
Set of Moments	**empirical**	**simulated**	**empirical**	**simulated**
mean(π/Y)	0.210	0.211	0.231	0.227
mean(log(Y/Khat))	0.748	0.762	0.588	0.573
mean(I/K)	0.125	0.129	0.123	0.128
mean(ΔlogY)	0.030	0.032	0.076	0.068
bsd(π/Y)	0.070	0.071	0.104	0.100
wsd(π/Y)	0.031	0.029	0.054	0.049
bsd(log(Y/Khat))	0.539	0.596	0.829	0.921
wsd(log(Y/Khat))	0.216	0.163	0.362	0.323
bsd(I/K)	0.095	0.046	0.129	0.086
wsd(I/K)	0.103	0.103	0.108	0.117
bsd(ΔlogY)	0.082	0.044	0.104	0.099
wsd(ΔlogY)	0.139	0.144	0.277	0.285
skew(π/Y)	0.522	0.458	0.706	0.663
skew(log(Y/Khat))	0.009	0.002	-0.117	0.007
skew(I/K)	2.723	1.248	2.575	2.002
skew(dlogY)	0.585	0.163	-0.105	0.031
scorr(π/Y)	0.880	0.839	0.799	0.766
scorr(log(Y/Khat))	0.939	0.928	0.904	0.876
scorr(I/K)	0.392	0.306	0.491	0.399
scorr(ΔlogY)	0.216	-0.018	-0.310	-0.265
corr(π/Y, log(Y/Khat))	-0.245	-0.194	-0.223	-0.315
corr(I/K, log(Y/K))	0.139	0.316	0.246	0.356
corr(ΔlogY, log(Y/K))	0.150	0.220	0.191	0.208
OI	813		450	

Table 10. Labor Market Distortions in China

Parameters	**estimate**	**s.e.**
σ_{τ}^{L}	0.229	0.045
$\mu_{log\ \alpha}$	-0.607	0.015
$\sigma_{log\alpha}$	0.456	0.016
Set of Moments	**empirical**	**simulated**
mean(mM/Y)	0.502	0.508
mean(wL^e/Y)	0.266	0.272
bsd(mM/Y)	0.204	0.203
bsd(wL^e/Y)	0.225	0.228
OI	1	