The Sources of Capital Misallocation: Online Appendix

Joel M. David* Venky Venkateswaran†

*joeldavi@usc.edu.
†vvenkate@stern.nyu.edu.

A Baseline Model

This appendix provides detailed derivations and proofs for our baseline model.

A.1 Solution

The first order condition and envelope conditions associated with (3) are, respectively,

$$\begin{aligned} T^K_{it+1}\left(1-\beta\left(1-\delta\right)\right)+\Phi_1\left(K_{it+1},K_{it}\right) &= \beta\mathbb{E}_{it}\left[\mathcal{V}_1\left(K_{it+1},\mathcal{I}_{it+1}\right)\right] \\ \mathcal{V}_1\left(K_{it},\mathcal{I}_{it}\right) &= \Pi_1\left(K_{it},A_{it}\right)-\Phi_2\left(K_{it+1},K_{it}\right) \end{aligned}$$

and combining yields the Euler equation

$$\mathbb{E}_{it}\left[\beta\Pi_1\left(K_{it+1},A_{it+1}\right)-\beta\Phi_2\left(K_{it+2},K_{it+1}\right)-T^K_{it+1}\left(1-\beta\left(1-\delta\right)\right)-\Phi_1\left(K_{it+1},K_{it}\right)\right]=0$$

where

$$\begin{aligned} \Pi_1\left(K_{it+1},A_{it+1}\right) &= \alpha G A_{it+1}K_{it+1}^{\alpha-1} \\ \Phi_1\left(K_{it+1},K_{it}\right) &= \hat{\xi}\left(\frac{K_{it+1}}{K_{it}}-\left(1-\delta\right)\right) \\ \Phi_2\left(K_{it+1},K_{it}\right) &= -\hat{\xi}\left(\frac{K_{it+1}}{K_{it}}-\left(1-\delta\right)\right)\frac{K_{it+1}}{K_{it}}+\frac{\hat{\xi}}{2}\left(\frac{K_{it+1}}{K_{it}}-\left(1-\delta\right)\right)^2 \\ &= \frac{\hat{\xi}}{2}\left(1-\delta\right)^2-\frac{\hat{\xi}}{2}\left(\frac{K_{it+1}}{K_{it}}\right)^2 \end{aligned}$$

In the undistorted $\left(\bar{T}^K=1\right)$ non-stochastic steady state, these are equal to

$$\begin{aligned} \bar{\Phi}_1 &= \hat{\xi}\delta \\ \bar{\Phi}_2 &= \frac{\hat{\xi}}{2}\left(1-\delta\right)^2-\frac{\hat{\xi}}{2} \\ \bar{\Pi}_1 &= \alpha\bar{G}\bar{A}\bar{K}^{\alpha-1} \end{aligned}$$

Log-linearizing the Euler equation around this point yields

$$\mathbb{E}_{it}\left[\beta\bar{\Pi}_1\pi_{1,it+1}-\beta\bar{\Phi}_2\phi_{2,it+1}-\tau^K_{it+1}\left(1-\beta\left(1-\delta\right)\right)-\bar{\Phi}_1\phi_{1,it}\right]=0$$

where $\tau^K_{it+1} = \log T^K_{it+1}$ and

$$\begin{aligned}
\bar{\Pi}_1 \pi_{1,it+1} &\approx \alpha \bar{G} \bar{A} \bar{K}^{\alpha-1} \left(a_{it+1} + (\alpha - 1) k_{it+1} \right) \\
\bar{\Phi}_1 \phi_{1,it} &\approx \hat{\xi} \left(k_{it+1} - k_{it} \right) \\
\bar{\Phi}_2 \phi_{2,it+1} &\approx -\hat{\xi} \left(k_{it+2} - k_{it+1} \right)
\end{aligned}$$

Rearranging gives

$$k_{it+1} \left((1+\beta)\xi + 1 - \alpha \right) = \mathbb{E}_{it} \left[a_{it+1} + \tau_{it+1} \right] + \beta \xi \mathbb{E}_{it} \left[k_{it+2} \right] + \xi k_{it}$$

where

$$\xi = \frac{\hat{\xi}}{\beta \bar{\Pi}_1}, \qquad \tau_{it+1} = -\frac{1 - \beta (1 - \delta)}{\beta \bar{\Pi}_1} \tau^K_{it+1}$$

which is expression (4) in the text. Using the steady state Euler equation,

$$\beta(\bar{\Pi}_1 + 1 - \delta) - \beta \bar{\Phi}_2 = 1 + \bar{\Phi}_1 \quad \Rightarrow \quad \alpha \beta \bar{G} \bar{A} \bar{K}^{\alpha - 1} = 1 - \beta (1 - \delta) + \hat{\xi} \delta \left(1 - \beta \left(1 - \frac{\delta}{2} \right) \right)$$

we have

$$\begin{aligned}
\xi &= \frac{\hat{\xi}}{1 - \beta (1 - \delta) + \hat{\xi} \delta \left(1 - \beta \left(1 - \frac{\delta}{2} \right) \right)} \\
\tau_{it+1} &= -\frac{1 - \beta (1 - \delta)}{1 - \beta (1 - \delta) + \hat{\xi} \delta \left(1 - \beta \left(1 - \frac{\delta}{2} \right) \right)} \tau^K_{it+1}
\end{aligned} \tag{A.1}$$

To derive the investment policy function, we conjecture that it takes the form in (7). Then,

$$\begin{aligned}
k_{it+2} &= \psi_1 k_{it+1} + \psi_2 (1 + \gamma) \mathbb{E}_{it+1} a_{it+2} + \psi_3 \varepsilon_{it+2} + \psi_4 \chi_i \\
\mathbb{E}_{it} \left[k_{it+2} \right] &= \psi_1 k_{it+1} + \psi_2 (1 + \gamma) \rho \mathbb{E}_{it} \left[a_{it+1} \right] + \psi_4 \chi_i \\
&= \psi_1 \left(\psi_1 k_{it} + \psi_2 (1 + \gamma) \mathbb{E}_{it} \left[a_{it+1} \right] + \psi_3 \varepsilon_{it+1} + \psi_4 \chi_i \right) + \psi_2 (1 + \gamma) \rho \mathbb{E}_{it} \left[a_{it+1} \right] + \psi_4 \chi_i \\
&= \psi_1^2 k_{it} + (\psi_1 + \rho) \psi_2 (1 + \gamma) \mathbb{E}_{it} \left[a_{it+1} \right] + \psi_1 \psi_3 \varepsilon_{it+1} + \psi_4 (1 + \psi_1) \chi_i
\end{aligned}$$

where we have used $\mathbb{E}_{it} \left[\varepsilon_{it+2} \right] = 0$ and $\mathbb{E}_{it} \left[\mathbb{E}_{it+1} \left[a_{it+2} \right] \right] = \rho \mathbb{E}_{it} \left[a_{it+1} \right]$. Substituting and rearranging,

$$\begin{aligned}
& \left(1 + \beta \xi \psi_4 (1 + \psi_1) \right) \chi_i + \left(1 + \beta \xi \psi_1 \psi_3 \right) \varepsilon_{it+1} \\
+ \; & \left(1 + \beta \xi (\psi_1 + \rho) \psi_2 \right) (1 + \gamma) \mathbb{E}_{it} \left[a_{it+1} \right] + \xi \left(1 + \beta \psi_1^2 \right) k_{it} \\
= \; & \left((1 + \beta) \xi + 1 - \alpha \right) \left(\psi_1 k_{it} + \psi_2 (1 + \gamma) \mathbb{E}_{it} \left[a_{it+1} \right] + \psi_3 \varepsilon_{it+1} + \psi_4 \chi_i \right)
\end{aligned}$$

Finally, matching coefficients gives

$$
\begin{aligned}
\xi\left(\beta\psi_1^2+1\right) &= \psi_1\left((1+\beta)\xi+1-\alpha\right) \\
1+\beta\xi\left(\psi_1+\rho\right)\psi_2 &= \psi_2\left((1+\beta)\xi+1-\alpha\right) \Rightarrow \psi_2 = \frac{1}{1-\alpha+\beta\xi\left(1-\psi_1-\rho\right)+\xi} \\
1+\beta\xi\psi_1\psi_3 &= \psi_3\left((1+\beta)\xi+1-\alpha\right) \Rightarrow \psi_3 = \frac{1}{1-\alpha+\left(1-\psi_1\right)\beta\xi+\xi} \\
1+\beta\xi\psi_4\left(1+\psi_1\right) &= \psi_4\left((1+\beta)\xi+1-\alpha\right) \Rightarrow \psi_4 = \frac{1}{1-\alpha+\xi\left(1-\beta\psi_1\right)}
\end{aligned}
$$

A few lines of algebra yields the expressions in (8).

A.2 Aggregation

To derive aggregate TFP and output, substitute the firm's optimality condition for labor

$$N_{it} = \left(\frac{\alpha_2 Y^{\frac{1}{\theta}}}{W}\hat{A}_{it}K_{it}^{\alpha_1}\right)^{\frac{1}{1-\alpha_2}}$$

into the production function (1) to get

$$Y_{it} = \left(\frac{\alpha_2 Y^{\frac{1}{\theta}}}{W}\right)^{\frac{\hat{\alpha}_2}{1-\alpha_2}} \hat{A}_{it}^{\frac{\hat{\alpha}_2}{1-\alpha_2}} K_{it}^{\frac{\hat{\alpha}_1}{1-\alpha_2}}$$

and using the demand function, revenues are

$$P_{it}Y_{it} = Y^{\frac{1}{\theta}\frac{1}{1-\alpha_2}}\left(\frac{\alpha_2}{W}\right)^{\frac{\alpha_2}{1-\alpha_2}} A_{it}K_{it}^{\alpha}$$

Labor market clearing implies

$$\int N_{it}di = \int\left(\frac{\alpha_2 Y^{\frac{1}{\theta}}}{W}\right)^{\frac{1}{1-\alpha_2}} A_{it}K_{it}^{\alpha}di = N$$

so that

$$\left(\frac{\alpha_2}{W}\right)^{\frac{\alpha_2}{1-\alpha_2}} = \left(\frac{N}{\int A_{it}K_{it}^{\alpha}di}\frac{1}{Y^{\frac{1}{\theta}\frac{1}{1-\alpha_2}}}\right)^{\alpha_2} \Rightarrow P_{it}Y_{it} = Y^{\frac{1}{\theta}}\frac{A_{it}K_{it}^{\alpha}}{\left(\int A_{it}K_{it}^{\alpha}di\right)^{\alpha_2}}N^{\alpha_2}$$

By definition,

$$ARPK_{it} = \frac{A_{it}K_{it}^{\alpha-1}}{\left(\int A_{it}K_{it}^{\alpha}di\right)^{\alpha_2}}Y^{\frac{1}{\theta}}N^{\alpha_2}$$

so that

$$K_{it} = \left(\frac{Y^{\frac{1}{\theta}} A_{it}}{ARPK_{it}}\right)^{\frac{1}{1-\alpha}} \left(\frac{N}{\int A_{it} K_{it}^{\alpha} di}\right)^{\frac{\alpha_2}{1-\alpha}}$$

and capital market clearing implies

$$K = \int K_{it} di = \left(Y^{\frac{1}{\theta}}\right)^{\frac{1}{1-\alpha}} \left(\frac{N}{\int A_{it} K_{it}^{\alpha} di}\right)^{\frac{\alpha_2}{1-\alpha}} \int A_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{1}{1-\alpha}} di$$

The latter two equations give

$$K_{it}^{\alpha} = \left(\frac{A_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{1}{1-\alpha}}}{\int A_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{1}{1-\alpha}} di} K\right)^{\alpha}$$

Substituting into the expression for $P_{it} Y_{it}$ and rearranging, we can derive

$$P_{it} Y_{it} = \frac{\frac{A_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{\alpha}{1-\alpha}}}{\left(\int A_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{1}{1-\alpha}} di\right)^{\alpha}}}{\left(\frac{\int A_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{\alpha}{1-\alpha}} di}{\left(\int A_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{1}{1-\alpha}} di\right)^{\alpha}}\right)^{\alpha_2}} Y^{\frac{1}{\theta}} K^{\alpha_1} N^{\alpha_2}$$

Using the fact that $Y = \int P_{it} Y_{it} di$, we can derive

$$Y = \int P_{it} Y_{it} di = Y^{\frac{1}{\theta}} A K^{\alpha_1} N^{\alpha_2}$$

where

$$A = \left(\frac{\int A_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{\alpha}{1-\alpha}} di}{\left(\int A_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{1}{1-\alpha}} di\right)^{\alpha}}\right)^{1-\alpha_2}$$

or in logs,

$$a = (1-\alpha_2)\left[\log\left(\int A_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{\alpha}{1-\alpha}}\right) - \alpha \log\left(\int A_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{1}{1-\alpha}}\right)\right]$$

The first term inside brackets is equal to

$$\frac{1}{1-\alpha}\bar{a} - \frac{\alpha}{1-\alpha}\overline{arpk} + \frac{1}{2}\left(\frac{1}{1-\alpha}\right)^2 \sigma_a^2 + \frac{1}{2}\left(\frac{\alpha}{1-\alpha}\right)^2 \sigma_{arpk}^2 - \frac{\alpha}{(1-\alpha)^2}\sigma_{arpk,a}$$

and the second,

$$\frac{\alpha}{1-\alpha}\overline{a}-\frac{\alpha}{1-\alpha}\overline{arpk}+\frac{1}{2}\alpha\left(\frac{1}{1-\alpha}\right)^2\sigma_a^2+\frac{1}{2}\alpha\left(\frac{1}{1-\alpha}\right)^2\sigma_{arpk}^2-\frac{\alpha}{(1-\alpha)^2}\sigma_{arpk,a}$$

Combining,

$$a=(1-\alpha_2)\left[\overline{a}+\frac{1}{2}\frac{1}{1-\alpha}\sigma_a^2-\frac{1}{2}\frac{\alpha}{1-\alpha}\sigma_{arpk}^2\right]$$

and

$$\begin{aligned}
y &= \frac{1}{\theta}y+(1-\alpha_2)\,\bar{a}+\frac{1}{2}\frac{1-\alpha_2}{1-\alpha}\sigma_a^2-\frac{1}{2}\alpha\frac{1-\alpha_2}{1-\alpha}\sigma_{arpk}^2+\alpha_1 k+\alpha_2 n \\
&= \frac{\theta}{\theta-1}(1-\alpha_2)\,\bar{a}+\frac{\theta}{\theta-1}\frac{1}{2}\frac{1-\alpha_2}{1-\alpha}\sigma_a^2-\frac{\theta}{\theta-1}\frac{1}{2}\alpha\frac{1-\alpha_2}{1-\alpha}\sigma_{arpk}^2+\hat{\alpha}_1 k+\hat{\alpha}_2 n \\
&= a+\hat{\alpha}_1 k+\hat{\alpha}_2 n
\end{aligned}$$

where, using $a_{it}=\frac{1}{1-\alpha_2}\hat{a}_{it}$, $\sigma_a^2=\left(\frac{1}{1-\alpha_2}\right)^2\sigma_{\hat{a}}^2$ and $\alpha=\frac{\alpha_1}{1-\alpha_2}$,

$$\begin{aligned}
a &= \frac{\theta}{\theta-1}\bar{\hat{a}}+\frac{1}{2}\frac{\theta}{\theta-1}\frac{1}{1-\alpha_1-\alpha_2}\sigma_{\hat{a}}^2-\frac{1}{2}\left(\theta\hat{\alpha}_1+\hat{\alpha}_2\right)\hat{\alpha}_1\sigma_{arpk}^2 \\
&= a^*-\frac{1}{2}\left(\theta\hat{\alpha}_1+\hat{\alpha}_2\right)\hat{\alpha}_1\sigma_{arpk}^2
\end{aligned}$$

which is equation (9) in the text.

To compute the effect on output, notice that the aggregate production function is

$$y=\hat{\alpha}_1 k+\hat{\alpha}_2 n+a$$

so that

$$\begin{aligned}
\frac{dy}{d\sigma_{arpk}^2} &= \hat{\alpha}_1\frac{dk}{da}\frac{da}{d\sigma_{arpk}^2}+\frac{da}{d\sigma_{arpk}^2} \\
&= \frac{da}{d\sigma_{arpk}^2}\left(1+\hat{\alpha}_1\frac{dk}{da}\right)
\end{aligned}$$

In the stationary equilibrium, the aggregate marginal product of capital must be a constant, denote it by $\bar{R}$, i.e., $\log\hat{\alpha}_1+y-k=\bar{r}$ so that

$$k=\frac{1}{1-\hat{\alpha}_1}\left(\log\hat{\alpha}_1+\hat{\alpha}_2 n+a-\bar{r}\right)$$

and

$$\frac{dk}{da} = \frac{1}{1-\hat{\alpha}_1}$$

Combining,

$$\frac{dy}{d\sigma^2_{arpk}} = \frac{da}{d\sigma^2_{arpk}}\left(1+\frac{\hat{\alpha}_1}{1-\hat{\alpha}_1}\right) = \frac{da}{d\sigma^2_{arpk}}\frac{1}{1-\hat{\alpha}_1}$$

A.3 Identification

In this appendix, we derive analytical expressions for the four moments in the random walk case, i.e., when $\rho = 1$, and prove Proposition 1.

Moments. From expression (7), we have the firm's investment policy function

$$k_{it+1} = \psi_1 k_{it} + \psi_2 (1+\gamma)\,\mathbb{E}_{it}[a_{it+1}] + \psi_3 \varepsilon_{it+1} + \psi_4 \chi_i$$

and substituting for the expectation,

$$k_{it+1} = \psi_1 k_{it} + \psi_2 (1+\gamma)(a_{it} + \phi(\mu_{it+1} + e_{it+1})) + \psi_3 \varepsilon_{it+1} + \psi_4 \chi_i$$

where $\phi = \frac{\mathbb{V}}{\sigma_e^2}$ so that $1-\phi = \frac{\mathbb{V}}{\sigma_\mu^2}$. Then,

$$\Delta k_{it+1} = \psi_1 \Delta k_{it} + \psi_2 (1+\gamma)((1-\phi)\mu_{it} + \phi\mu_{it+1} + \phi(e_{it+1} - e_{it})) + \psi_3 (\varepsilon_{it+1} - \varepsilon_{it})$$

We will use the fact that

$$\begin{aligned}
\text{cov}(\Delta k_{it+1}, \mu_{it+1}) &= \psi_2 (1+\gamma)\,\phi\sigma_\mu^2 \\
\text{cov}(\Delta k_{it+1}, e_{it+1}) &= \psi_2 (1+\gamma)\,\phi\sigma_e^2 \\
\text{cov}(\Delta k_{it+1}, \varepsilon_{it+1}) &= \psi_3 \sigma_\varepsilon^2
\end{aligned}$$

Now,

$$\begin{aligned}
\text{var}(\Delta k_{it+1}) &= \psi_1^2 \text{var}(\Delta k_{it}) + \psi_2^2 (1+\gamma)^2 (1-\phi)^2 \sigma_\mu^2 \\
&+ \psi_2^2 (1+\gamma)^2 \phi^2 \sigma_\mu^2 + 2\psi_2^2 (1+\gamma)^2 \phi^2 \sigma_e^2 + 2\psi_3^2 \sigma_\varepsilon^2 \\
&+ 2\psi_1\psi_2 (1+\gamma)(1-\phi)\,\text{cov}(\Delta k_{it}, \mu_{it}) - 2\psi_1\psi_2 (1+\gamma)\,\phi\text{cov}(\Delta k_{it}, e_{it}) \\
&- 2\psi_1\psi_3 \text{cov}(\Delta k_{it}, \varepsilon_{it})
\end{aligned}$$

where substituting, rearranging and using the fact that the moments are stationary gives

$$\sigma_k^2 \equiv \text{var}\left(\Delta k_{it}\right) = \frac{\left(1+\gamma\right)^2 \psi_2^2 \sigma_\mu^2 + 2\left(1-\psi_1\right)\psi_3^2 \sigma_\varepsilon^2}{1-\psi_1^2}$$

which can be rearranged to yield expression (10). Next,

$$\begin{aligned}
\text{cov}\left(\Delta k_{it+1}, \Delta k_{it}\right) &= \psi_1 \text{var}\left(\Delta k_{it}\right) + \psi_2\left(1+\gamma\right)\left(1-\phi\right)\text{cov}\left(\Delta k_{it}, \mu_{it}\right) \\
&- \psi_2\left(1+\gamma\right)\phi \text{cov}\left(\Delta k_{it}, e_{it}\right) - \psi_3 \text{cov}\left(\Delta k_{it}, \varepsilon_{it}\right) \\
&= \psi_1 \text{var}\left(\Delta k_{it}\right) - \psi_3 \text{cov}\left(\Delta k_{it}, \varepsilon_{it}\right) \\
&= \psi_1 \sigma_k^2 - \psi_3^2 \sigma_\varepsilon^2
\end{aligned}$$

so that

$$\rho_{k,k_{-1}} \equiv \text{corr}\left(\Delta k_{it}, \Delta k_{it-1}\right) = \psi_1 - \psi_3^2 \frac{\sigma_\varepsilon^2}{\sigma_k^2}$$

which is expression (11). Similarly,

$$\begin{aligned}
\text{cov}\left(\Delta k_{it+1}, \Delta a_{it}\right) &= \text{cov}\left(\Delta k_{it+1}, \mu_{it}\right) \\
&= \psi_1 \text{cov}\left(\Delta k_{it}, \mu_{it}\right) + \psi_2\left(1+\gamma\right)\left(1-\phi\right)\sigma_\mu^2 \\
&= \psi_1 \psi_2\left(1+\gamma\right)\phi\sigma_\mu^2 + \psi_2\left(1+\gamma\right)\left(1-\phi\right)\sigma_\mu^2 \\
&= \left(1-\phi\left(1-\psi_1\right)\right)\psi_2\left(1+\gamma\right)\sigma_\mu^2
\end{aligned}$$

and from here it is straightforward to derive

$$\rho_{k,a_{-1}} \equiv \text{corr}\left(\Delta k_{it}, \Delta a_{it-1}\right) = \left[\frac{\mathbb{V}}{\sigma_\mu^2}\left(1-\psi_1\right) + \psi_1\right] \frac{\sigma_\mu \psi_2\left(1+\gamma\right)}{\sigma_k}$$

as in expression (12).

Finally,

$$arpk_{it} = p_{it} + y_{it} - k_{it} = \text{Const} + a_{it} + \alpha k_{it} - k_{it} = \text{Const} + a_{it} - \left(1-\alpha\right)k_{it}$$

so that

$$\Delta arpk_{it} = \Delta a_{it} - \left(1-\alpha\right)\Delta k_{it} = \mu_{it} - \left(1-\alpha\right)\Delta k_{it}$$

which implies

$$\text{cov}\left(\Delta arpk_{it}, \mu_{it}\right) = \left(1 - \left(1-\alpha\right)\left(1+\gamma\right)\psi_2\phi\right)\sigma_\mu^2$$

and

$$\lambda_{arpk,a} \equiv \frac{\text{cov}\left(\Delta arpk_{it}, \mu_{it}\right)}{\sigma_\mu^2} = 1 - (1-\alpha)(1+\gamma)\psi_2\phi$$
$$= 1 - (1-\alpha)(1+\gamma)\psi_2\left(1 - \frac{\mathbb{V}}{\sigma_\mu^2}\right)$$

which is expression (13).

To see that the correlation $\rho_{arpk,a}$ is decreasing in σ_ε^2, we derive

$$\begin{aligned}
\text{var}\left(\Delta arpk_{it}\right) &= \sigma_\mu^2 + (1-\alpha)^2\sigma_k^2 - 2(1-\alpha)\,\text{cov}\left(\Delta k_{it}, \mu_{it}\right) \\
&= \sigma_\mu^2 + (1-\alpha)^2\left(\frac{\psi_2^2(1+\gamma)^2\sigma_\mu^2 + 2(1-\psi_1)\psi_3^2\sigma_\varepsilon^2}{1-\psi_1^2}\right) - 2(1-\alpha)\psi_2(1+\gamma)\phi\sigma_\mu^2 \\
&= \frac{1}{1-\psi_1^2}\left(\left(\left(1-\psi_1^2\right)(1-2(1-\alpha)(1+\gamma)\psi_2\phi) + (1-\alpha)^2(1+\gamma)^2\psi_2^2\right)\sigma_\mu^2\right) \\
&+ \frac{1}{1-\psi_1^2}\left(2(1-\alpha)^2(1-\psi_1)\psi_3^2\sigma_\varepsilon^2\right)
\end{aligned}$$

so

$$\rho_{arpk,a} = \frac{(1-(1-\alpha)(1+\gamma)\psi_2\phi)\,\sigma_\mu\sqrt{1-\psi_1^2}}{\sqrt{\left(\left(1-\psi_1^2\right)(1-2(1-\alpha)(1+\gamma)\psi_2\phi) + (1-\alpha)^2(1+\gamma)^2\psi_2^2\right)\sigma_\mu^2 + 2(1-\alpha)^2(1-\psi_1)\psi_3^2\sigma_\varepsilon^2}}$$

Proof of Proposition 1. Write the variance of investment as

$$\sigma_k^2 = \psi_1^2\sigma_k^2 + (1+\gamma)^2\psi_2^2\sigma_\mu^2 + 2(1-\psi_1)\psi_3^2\sigma_\varepsilon^2$$

We can rewrite the last term as a function of an observable moment, the autocovariance of investment, which is given by

$$\sigma_{k,k_{-1}} = \psi_1\sigma_k^2 - \psi_3^2\sigma_\varepsilon^2\ . \tag{A.2}$$

Substituting,

$$\sigma_k^2 = \psi_1^2\sigma_k^2 + (1+\gamma)^2\psi_2^2\sigma_\mu^2 + 2(1-\psi_1)\left(\psi_1\sigma_k^2 - \sigma_{k,k_{-1}}\right) \tag{A.3}$$

To eliminate the second term, use the equation for $\lambda_{arpk,a}$ to solve for

$$(1+\gamma)\psi_2\phi = \frac{1-\lambda_{arpk,a}}{1-\alpha} = \tilde{\lambda} \tag{A.4}$$

where $\tilde{\lambda}$ is a decreasing function of $\lambda_{arpk,a}$ that depends only on the known parameter α. Substituting into the expression for the covariance of investment with the lagged shock, $\sigma_{k,a_{-1}}$, and rearranging yields

$$(1+\gamma)\,\psi_2 = \frac{\sigma_{k,a_{-1}}}{\sigma_\mu^2} + \tilde{\lambda}\,(1-\psi_1) \tag{A.5}$$

which is an equation in ψ_1 and observable moments. Substituting into (A.3) gives

$$\sigma_k^2 = \psi_1^2\sigma_k^2 + \left(\frac{\sigma_{k,a_{-1}}}{\sigma_\mu^2} + \tilde{\lambda}\,(1-\psi_1)\right)^2 \sigma_\mu^2 + 2\,(1-\psi_1)\left(\psi_1\sigma_k^2 - \sigma_{k,k_{-1}}\right)$$

and rearranging, we can derive

$$0 = \left(\hat{\lambda}^2 - 1\right)(1-\psi_1)^2 + 2\left(\hat{\lambda}\rho_{k,a_{-1}} - \rho_{k,k_{-1}}\right)(1-\psi_1) + \rho_{k,a_{-1}}^2 \tag{A.6}$$

where

$$\hat{\lambda} = \frac{\sigma_\mu}{\sigma_k}\tilde{\lambda} = \frac{\sigma_\mu}{\sigma_k}\left(\frac{1-\lambda_{arpk,a}}{1-\alpha}\right)$$

Equation (A.6) represents a quadratic equation in a single unknown, $1-\psi_1$, or equivalently, in ψ_1. The solution features one positive root and one negative. The positive root corresponds to the true ψ_1 that represents the solution to the firm's investment policy. The value of ψ_1 pins down the adjustment cost parameter ξ as well as ψ_2 and ψ_3. We can then back out γ from (A.5), ϕ (and so $\mathbb{V}$) from (A.4) and finally, σ_ε^2 from (A.2).

□

B Data

Our Chinese data are from the Annual Surveys of Industrial Production conducted by the National Bureau of Statistics. The data span the period 1998-2009 and are built into a panel following quite closely the method outlined in Brandt et al. (2014). We measure the capital stock as the value of fixed assets and calculate investment as the change in the capital stock relative to the preceding period. We construct firm productivity, a_{it}, as the log of value-added less α multiplied by the log of the capital stock and (the log of) the average product of capital, $arpk_{it}$ as the log of value-added less the log of the capital stock. We compute value-added from revenues using a share of intermediates of 0.5 (our data does not include a direct measure of value-added in all years). Investment growth and changes in productivity are the first differences of the investment and productivity series (in logs) respectively.

To extract the firm-specific variation in our variables, we regress each on a year by time fixed-effect and work with the residual. Industries are defined at the 4-digit level. This eliminates the industry-wide component of each series common to all firms in an industry and time period (as well the aggregate component common across all firms) and leaves only the idiosyncratic variation. To estimate the parameters governing firm productivity, i.e., the persistence ρ and variance of the innovations σ_μ^2, we perform the autoregression implied by (5), again including industry by year controls. We eliminate duplicate observations (firms with multiple observations within a single year) and trim the 3% tails of each series. We additionally exclude observations with excessively high variability in investment (investment rates over 100%). Our final sample in China consists of 797,047 firm-year observations.

Our US data are from Compustat North America and also spans the period 1998-2009. We measure the capital stock using gross property, plant and equipment. We treat the data in exactly the same manner as just described for the set of Chinese firms. We additionally eliminate firms that are not incorporated in the US and/or do not report in US dollars. Our final sample in the US consists of 34,260 firm-year observations.

Table B.1 reports a number of summary statistics from one year of our data, 2009: the number of firms (with available data on sales), the share of GDP they account for, and average sales and capital.

Table B.1: Sample Statistics 2009

	No. of Firms	Share of GDP	Avg. Sales ($M)	Avg. Capital ($M)
China	303623	0.65	21.51	8.08
US	6177	0.45	2099.33	1811.35

Materials and labor expenses. For the analyses in Section IV.A, labor input is measured as the wage bill. The wage bill is directly reported in the Chinese data. For the US, we follow, e.g., Keller and Yeaple (2009) and impute a measure of the wage bill as the number of employees multiplied by the average industry wage, calculated using data from the NBER-CES Manufacturing Industry Database (available at `http://www.nber.org/nberces/`; the average industry wage is calculated as total industry-wide payroll divided by total employees). Expenditures on intermediate inputs are reported in the Chinese data. In the US, we construct a measure of intermediates following the method in, e.g., De Loecker and Eeckhout (2017) and İmrohoroğlu and Tüzel (2014). Specifically, intermediate expenditures are calculated as total expenses less labor expenses, where total expenses are calculated as sales less operating income (before depreciation and amortization, Compustat series OIBDP) and labor expenses are measured as described earlier. We can then calculate all the series used in Section IV.A, i.e., the raw and 'markup-adjusted' average revenue products of capital, labor and materials (the inverse of materials' share of revenues). We isolate the firm-specific variation in these series following a similar procedure as described above, i.e., by extracting a full set of industry by time fixed-effects and working with the residual. We trim the 1% tails of each series.

C Computation and Estimation of Baseline Model

In this appendix, we provide details of our numerical estimation procedure and results. We estimate the model via method of moments using the following procedure. For a given set of parameters, we compute the cross-sectional moments of interest using the steady state distribution. To do so, we cast the law of motion (7) in matrix form:

$$
\begin{aligned}
BX_{it} &= CX_{it-1} + DU_{it} \qquad \text{where} \\
X_{it} &= \begin{bmatrix} k_{it} \\ \iota_{it} \\ a_{it} \\ \mathbb{E}_{it-1}a_{it} \end{bmatrix} \qquad U_{it} = \begin{bmatrix} \mu_{it} \\ e_{it} \\ \varepsilon_{it} \\ \chi_i \end{bmatrix} \\
B &= \begin{bmatrix} 1 & 0 & 0 & -\psi_2(1+\gamma) \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad C = \begin{bmatrix} \psi_1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & \rho & 0 \\ 0 & 0 & \rho & 0 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 0 & \psi_3 & \psi_4 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1-\frac{\mathbb{V}}{\sigma_\mu^2} & 1-\frac{\mathbb{V}}{\sigma_\mu^2} & 0 & 0 \end{bmatrix}.
\end{aligned}
$$

Pre-multiplying by B^{-1} yields

$$X_{it} = B^{-1}CX_{it-1} + B^{-1}DU_{it} = \tilde{C}X_{it-1} + \tilde{D}U_{it} \ .$$

The steady state covariance matrix of X_{it}, denoted Σ_X, is then obtained by solving the Lyapunov equation:

$$\Sigma_X = \tilde{C}\Sigma_X\tilde{C}' + \tilde{D}\Sigma_U\tilde{D}' \ ,$$

where Σ_U denotes the covariance matrix of U_{it}. It is straightforward to compute other second moments. For example, to obtain the covariance matrix of $\Delta X_{it} = X_{it} - X_{it-1}$, note that

$$\Delta X_{it} = (\tilde{C}-I)X_{it-1} + \tilde{D}U_{it} \quad \Rightarrow \quad \Sigma_{\Delta X} = (\tilde{C}-I)\Sigma_X(\tilde{C}-I)' + \tilde{D}\Sigma_U\tilde{D}' \ .$$

We then use a non-linear solver to search over the parameter vector $(\xi, \mathbb{V}, \gamma, \sigma_\varepsilon^2, \sigma_\chi^2)$ to minimize the equally-weighted distance between the model and data values for the targeted moments.

Table C.1 displays the details of our baseline estimation. In the top panel, we report the target moments computed from the data, along with standard errors (in parentheses) and the simulated model counterparts. The estimated parameter vector is shown in the bottom panel along with standard errors and confidence intervals. The model is able to match the full set of target moments quite closely in both countries (in China, the fit is essentially exact) and the standard errors and confidence intervals indicate that the estimates are quite precise.

Standard errors and confidence intervals are calculated using the following bootstrap procedure: we draw 1,000 random samples (with replacement) from the data (these are block-boostraps, i.e., we resample entire histories of firms). For each re-sampled dataset, we re-calculate the target moments and re-estimate the model parameters. Standard errors are computed as the standard deviations of the resulting distribution of estimated moments and parameters. 95% confidence intervals for the parameters are computed as the 2.5th and 97.5th percentiles of the distributions of the parameter estimates.

D Interactions Between Factors

In the main text (specifically, Table 3), we measured the contribution of each factor in isolation, i.e., setting all other forces to zero. The top panel of Table D.1 reproduces those estimates (labeled 'In isolation') and compares them to the case where all the other factors are held fixed at their estimated levels (labeled 'Joint'). The table shows some evidence of interactions, but since adjustment and informational frictions are modest, the numbers are quite similar under both approaches.

E Labor Market Distortions

This appendix presents two tractable versions of our model with labor market distortions. In the first, these are modeled as firm-specific 'taxes' with an arbitrary correlation structure. The second describes the environment from Section V.B in the main text, where all the factors acting on investment – adjustment, informational and other – are assumed to apply to the labor decision as well. Under both specifications, the profit function takes the same form as in the baseline analysis with suitably re-defined productivity and curvature. This implies that our identification arguments and empirical strategy go through exactly. More importantly, quantifying the sources of *arpk* dispersion still requires only data on value-added and capital as before.

E.1 Firm-Specific Labor Taxes

Here, we introduce labor market distortions in the form of firm-specific taxes on the cost of labor. These distortions change the interpretation of the profitability shifter, A_{it}, which has implications for the correct measurement strategy of this term. But, apart from this re-interpretation, they do not change our estimates/conclusions about the drivers of *arpk* dispersion.

Table C.1: Baseline Esimation - Model Fit

Moments	ρ	σ^2_μ	$\rho_{\iota,a_{-1}}$	$\rho_{\iota,\iota_{-1}}$	$\rho_{arpk,a}$	σ^2_ι	σ^2_{arpk}
China							
Data	0.914	0.146	0.287	−0.359	0.757	0.143	0.922
	(0.001)	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)	(0.002)
Model-Simulated			0.287	−0.375	0.737	0.124	0.914
US							
Data	0.933	0.078	0.126	−0.297	0.547	0.056	0.450
	(0.002)	(0.001)	(0.007)	(0.006)	(0.011)	(0.001)	(0.008)
Model-Simulated			0.126	−0.297	0.547	0.056	0.450
Parameter Estimates			ξ	$\mathbb{V}$	γ	σ^2_ε	σ^2_χ
China			0.132	0.095	−0.704	0.000	0.410
			(0.003)	(0.000)	(0.003)	(0.000)	(0.002)
			[0.127,0.137]	[0.095,0.096]	[−0.709,−0.698]	[0.000,0.000]	[0.406,0.413]
US			1.382	0.033	−0.328	0.029	0.292
			(0.163)	(0.001)	(0.024)	(0.009)	(0.006)
			[1.099,1.724]	[0.031,0.034]	[−0.372,−0.279]	[0.015,0.049]	[0.279,0.304]

Notes: Top panel: rows labeled 'Data' display the empirical values of the moments. Values in parentheses are bootstrapped standard errors. Rows labeled 'Model-Simulated' display the value of the moments from the model simulation at the estimated parameter values. Bottom panel displays the parameter estimates. Values in parentheses are bootstrapped standard errors. Values in brackets are 95% confidence intervals.

Table D.1: Interactions Between Factors - US

	Adj Costs	Uncertainty	Other Factors: Correlated	Transitory	Permanent
In isolation					
$\Delta\sigma^2_{arpk}$	0.05	0.03	0.06	0.03	0.29
$\frac{\Delta\sigma^2_{arpk}}{\sigma^2_{arpk}}$	10.8%	7.3%	14.4%	6.3%	64.7%
Joint					
$\Delta\sigma^2_{arpk}$	0.04	0.03	0.08	0.00	0.29
$\frac{\Delta\sigma^2_{arpk}}{\sigma^2_{arpk}}$	8.0%	5.7%	17.4%	0.3%	64.7%

With firm-specific labor taxes, denoted T^N_{it}, the firm's problem becomes

$$\begin{aligned}\mathcal{V}\left(K_{it},\mathcal{I}_{it}\right) = \max_{N_{it},K_{it+1}} \quad & \mathbb{E}_{it}\left[Y^{\frac{1}{\theta}}\hat{A}_{it}K^{\alpha_1}_{it}N^{\alpha_2}_{it} - WT^N_{it}N_{it} - T^K_{it+1}K_{it+1}\left(1-\beta\left(1-\delta\right)\right) - \Phi\left(K_{it+1},K_{it}\right)\right] \\ + \quad & \beta\mathbb{E}_{it}\left[\mathcal{V}\left(K_{it+1},\mathcal{I}_{it+1}\right)\right] \ .\end{aligned}$$

The labor choice satisfies the first order condition:

$$N_{it} = \left(\alpha_2\frac{Y^{\frac{1}{\theta}}\hat{A}_{it}K^{\alpha_1}_{it}}{WT^N_{it}}\right)^{\frac{1}{1-\alpha_2}} \ .$$

Substituting, we can derive operating profits (value-added net of total wages) as

$$\begin{aligned}P_{it}Y_{it} - WT^N_{it}N_{it} &= Y^{\frac{1}{\theta}}_t\hat{A}_{it}K^{\alpha_1}_{it}\left(\alpha_2 Y^{\frac{1}{\theta}}_t\frac{\hat{A}_{it}K^{\alpha_1}_{it}}{WT^N_{it}}\right)^{\frac{\alpha_2}{1-\alpha_2}} - WT^N_{it}\left(\alpha_2 Y^{\frac{1}{\theta}}_t\frac{\hat{A}_{it}K^{\alpha_1}_{it}}{WT^N_{it}}\right)^{\frac{1}{1-\alpha_2}} \\ &= \left(1-\alpha_2\right)\left(\frac{\alpha_2}{W}\right)^{\frac{\alpha_2}{1-\alpha_2}}Y^{\frac{1}{\theta}\frac{1}{1-\alpha_2}}\frac{\hat{A}^{\frac{1}{1-\alpha_2}}_{it}}{\left(T^N_{it}\right)^{\frac{\alpha_2}{1-\alpha_2}}}K^{\frac{\alpha_1}{1-\alpha_2}}_{it} \\ &= GA_{it}K^{\alpha}_{it} \ ,\end{aligned}$$

where

$$A_{it} \equiv \left(\frac{\hat{A}_{it}}{\left(T^N_{it}\right)^{\alpha_2}}\right)^{\frac{1}{1-\alpha_2}} \ . \tag{E.1}$$

Thus, the profit function (and therefore, the firm's investment problem) takes the same form as in the baseline version, except that A_{it} now incorporates the effect of the labor distortion as well. With this re-interpretation, our identification strategy remains valid, so long as A_{it} is correctly measured.

Measuring profitability. Recall that our empirical strategy in the baseline analysis measured profitability shocks using $a_{it} = va_{it} - \alpha k_{it}$. Expression (E.1) shows that this is the correct measure of profitability even with distortions to labor. This result implies that, apart from issues of interpretation, our quantitative analysis is not affected at all. In other words, our empirical strategy requires neither data on labor inputs nor taking a stand on the extent of labor distortions.

This is not the case for an alternative strategy that directly estimates the true productivity, $\hat{a}_{it} \equiv \log \hat{A}_{it} = va_{it} - \alpha_1 k_{it} - \alpha_2 n_{it}$, and uses it to construct the implied profitability term as $a_{it} = \frac{1}{1-\alpha_2}\hat{a}_{it}$, i.e., without controlling for labor distortions. It is easy to see that when there are firm-specific labor distortions, this approach leads to an incorrect measure of a_{it}, and therefore, to biased estimates of (ρ, σ_μ^2) and the other parameters. In particular, consider the empirically relevant case where the labor distortion is positively correlated with productivity: using a measure of A_{it} inferred from the estimated $\hat{A}_{it}$ without adjusting for T_{it}^N will overstate the variability in profitability. Quantitatively, this bias can be very large: in our data, this strategy produces estimates of (ρ, σ_μ) of $(0.90, 0.28)$ and $(0.88, 0.35)$ for the US and China, respectively, compared to our baseline estimates of $(0.93, 0.08)$ and $(0.91, 0.15)$. In other words, it overstates the volatility of shocks by a factor of almost 3.

This is essentially the strategy followed by Asker et al. (2014) and contributes to the difference between our estimates and theirs. To get a sense of the magnitude, a model with only convex adjustment costs estimated to match the variability of investment growth would yield an adjustment cost parameter, ξ, that is about 3 times higher in both countries under this more volatile shock process than under the baseline process (i.e., using a profitability measure that does account for labor distortions). This, in turn, would imply *arpk* dispersion from adjustment costs alone that exceeds the total observed dispersion in the US (and is about 60% of the total in China).

E.2 Frictional Labor

Here, we provide detailed derivations for the case of frictional labor in Section V.B.

E.2.1 Model Solution

When labor is chosen under the same frictions as capital, the value function takes the form

$$\begin{aligned}\mathcal{V}\left(K_{it},N_{it},\mathcal{I}_{it}\right)=\max_{K_{it+1},N_{it+1}} \quad & \mathbb{E}_{it}\left[Y^{\frac{1}{\theta}}\hat{A}_{it}K_{it}^{\alpha_1}N_{it}^{\alpha_2}\right] \\ - \quad & \mathbb{E}_{it}\left[T_{it+1}K_{it+1}\left(1-\beta\left(1-\delta\right)\right)+\Phi\left(K_{it+1},K_{it}\right)\right] \\ - \quad & \mathbb{E}_{it}\left[T_{it+1}WN_{it+1}\left(1-\beta\left(1-\delta\right)\right)+W\Phi\left(N_{it+1},N_{it}\right)\right] \\ + \quad & \mathbb{E}_{it}\left[\beta\mathcal{V}\left(K_{it+1},N_{it+1},\mathcal{I}_{it+1}\right)\right]\end{aligned} \tag{E.2}$$

where the adjustment cost function $\Phi(\cdot)$ is as defined in (2). Because the firm makes a one-time payment to hire incremental labor, the cost of labor W is now to be interpreted as the present discounted value of wages. Capital and labor are both subject to the same adjustment frictions, the same distortions, denoted T_{it+1}, and are chosen under the same information set, though the cost of labor adjustment is denominated in labor units.

The first order and envelope conditions yield two Euler equations:

$$\begin{aligned}\mathbb{E}_{it}\left[T_{it+1}\left(1-\beta\left(1-\delta\right)\right)+\Phi_1\left(K_{it+1},K_{it}\right)\right] &= \mathbb{E}_{it}\left[\beta\alpha_1 Y^{\frac{1}{\theta}}\hat{A}_{it+1}K_{it+1}^{\alpha_1-1}N_{it+1}^{\alpha_2}-\beta\Phi_2\left(K_{it+2},K_{it+1}\right)\right] \\ W\mathbb{E}_{it}\left[T_{it+1}\left(1-\beta\left(1-\delta\right)\right)+\Phi_1\left(N_{it+1},N_{it}\right)\right] &= \mathbb{E}_{it}\left[\beta\alpha_2 Y^{\frac{1}{\theta}}\hat{A}_{it+1}K_{it+1}^{\alpha_1}N_{it+1}^{\alpha_2-1}-\beta W\Phi_2\left(N_{it+2},N_{it+1}\right)\right]\end{aligned}$$

To show that this setup reduces to a Bellman equation of the same form as (3), we guess – and verify – that there exists a constant η such that the firm's labor policy takes the form $N_{it+1}=\eta K_{it+1}$.

Under this conjecture, we can rewrite the firm's problem in (E.2) as

$$\begin{aligned}\tilde{\mathcal{V}}\left(K_{it},\mathcal{I}_{it}\right)=\max_{K_{it+1}} \quad & \mathbb{E}_{it}\left[\frac{\eta^{\alpha_2}}{1+W\eta}Y^{\frac{1}{\theta}}\hat{A}_{it}K_{it}^{\alpha_1+\alpha_2}-T_{it+1}K_{it+1}\left(1-\beta\left(1-\delta\right)\right)\right] \\ + \quad & \mathbb{E}_{it}\left[-\Phi\left(K_{it+1},K_{it}\right)+\beta\tilde{\mathcal{V}}\left(K_{it+1},\mathcal{I}_{it+1}\right)\right]\end{aligned}$$

Let $\{K_{it}^*\}$ be the solution to this problem. By definition, it must satisfy:

$$\begin{aligned}\mathbb{E}_{it}\left[T_{it+1}\left(1-\beta\left(1-\delta\right)\right)+\Phi_1\left(K_{it+1}^*,K_{it}^*\right)\right] &= \mathbb{E}_{it}\left[\beta\frac{\left(\alpha_1+\alpha_2\right)Y^{\frac{1}{\theta}}\hat{A}_{it+1}K_{it+1}^{*\alpha_1+\alpha_2-1}\eta^{\alpha_2}}{1+W\eta}\right] \\ &- \mathbb{E}_{it}\left[\beta\Phi_2\left(K_{it+2}^*,K_{it+1}^*\right)\right]\end{aligned} \tag{E.3}$$

Now substitute the conjecture $N_{it}^*=\eta K_{it}^*$ into the optimality condition for labor from the original problem and rearrange to get:

$$\begin{aligned}\mathbb{E}_{it}\left[T_{it+1}\left(1-\beta\left(1-\delta\right)\right)+\Phi_1\left(K^*_{it+1},K^*_{it}\right)\right] &= \mathbb{E}_{it}\left[\beta\frac{\alpha_2 Y^{\frac{1}{\theta}}\hat{A}_{it+1}K^{*\alpha_1+\alpha_2-1}_{it+1}\eta^{\alpha_2}}{W\eta}\right] \\ &- \mathbb{E}_{it}\left[\beta\Phi_2\left(K^*_{it+2},K^*_{it+1}\right)\right]\end{aligned} \tag{E.4}$$

If η satisfies

$$\frac{\alpha_1+\alpha_2}{1+W\eta}=\frac{\alpha_2}{W\eta} \qquad \Rightarrow \qquad W\eta=\frac{\alpha_2}{\alpha_1} \tag{E.5}$$

then $(E.4)$ is identical to $(E.3)$. In other words, under $(E.5)$, the sequence $\{K^*_{it},N^*_{it}\}$ satisfies the optimality condition for labor from the original problem. It is straightforward to verify that this also implies that $\{K^*_{it},N^*_{it}\}$ satisfy the optimality condition for capital from the original problem:

$$\begin{aligned}\mathbb{E}_{it}\left[T_{it+1}\left(1-\beta\left(1-\delta\right)\right)+\Phi_1\left(K^*_{it+1},K^*_{it}\right)\right] &= \mathbb{E}_{it}\left[\beta\alpha_1 Y^{\frac{1}{\theta}}\hat{A}_{it+1}K^{*\alpha_1+\alpha_2-1}_{it+1}\eta^{\alpha_2}-\beta\Phi_2\left(K^*_{it+2},K^*_{it+1}\right)\right] \\ &= \mathbb{E}_{it}\left[\beta\frac{\alpha_2 Y^{\frac{1}{\theta}}\hat{A}_{it+1}K^{*\alpha_1+\alpha_2-1}_{it+1}\eta^{\alpha_2}}{W\eta}-\beta\Phi_2\left(K^*_{it+2},K^*_{it+1}\right)\right]\end{aligned}$$

Thus, this version can be solved following the same steps as the baseline setup. The firm's problem takes the same form as expression (3), with $\alpha=\alpha_1+\alpha_2$, $G=\frac{\eta^{\alpha_2}Y^{\frac{1}{\theta}}}{1+W\eta}$ and $A_{it}=\hat{A}_{it}$.

E.2.2 Aggregation

To derive aggregate output and TFP for this case, we use $N_{it}=\eta K_{it}$ where $\eta=\frac{\alpha_2}{\alpha_1 W}$. Substituting into the revenue function gives

$$P_{it}Y_{it}=Y^{\frac{1}{\theta}}\hat{A}_{it}\eta^{\alpha_2}K^{\alpha_1+\alpha_2}_{it}=Y^{\frac{1}{\theta}}\hat{A}_{it}\eta^{\alpha_2}K^{\alpha}_{it}$$

By definition,

$$ARPK_{it}=Y^{\frac{1}{\theta}}\hat{A}_{it}\eta^{\alpha_2}K^{\alpha-1}_{it}$$

so that

$$K_{it}=\left(\frac{Y^{\frac{1}{\theta}}\hat{A}_{it}\eta^{\alpha_2}}{ARPK_{it}}\right)^{\frac{1}{1-\alpha}}$$

so that

$$
\begin{aligned}
P_{it} Y_{it} &= Y^{\frac{1}{\theta}} \eta^{\alpha_2} \hat{A}_{it} \left(\frac{Y^{\frac{1}{\theta}} \eta^{\alpha_2} \hat{A}_{it}}{ARPK_{it}} \right)^{\frac{\alpha}{1-\alpha}} \\
&= Y^{\frac{1}{\theta}\frac{1}{1-\alpha}} \eta^{\frac{\alpha_2}{1-\alpha}} \hat{A}_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{\alpha}{1-\alpha}}
\end{aligned}
$$

and

$$
Y = \int P_{it} Y_{it} di = Y^{\frac{1}{\theta}\frac{1}{1-\alpha}} \eta^{\frac{\alpha_2}{1-\alpha}} \int \hat{A}_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{\alpha}{1-\alpha}} di
$$

or, rearranging,

$$
Y = Y^{\frac{1}{\theta}\frac{\hat{\alpha}_1+\hat{\alpha}_2}{1-\alpha}} \eta^{\frac{\hat{\alpha}_2}{1-\alpha}} \left(\int \hat{A}_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{\alpha}{1-\alpha}} di \right)^{\frac{\theta}{\theta-1}}
$$

Capital market clearing implies

$$
K = \int K_{it} di = Y_t^{\frac{1}{\theta}\frac{1}{1-\alpha}} \eta^{\frac{\alpha_2}{1-\alpha}} \int \hat{A}_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{1}{1-\alpha}} di
$$

so that

$$
K^{\hat{\alpha}_1} N^{\hat{\alpha}_2} = Y_t^{\frac{1}{\theta}\frac{\hat{\alpha}_1+\hat{\alpha}_2}{1-\alpha}} \eta^{\hat{\alpha}_2+\frac{\alpha_2}{1-\alpha}(\hat{\alpha}_1+\hat{\alpha}_2)} \left(\int \hat{A}_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{1}{1-\alpha}} di \right)^{\hat{\alpha}_1+\hat{\alpha}_2}
$$

Aggregate TFP is

$$
A = \frac{Y}{K^{\hat{\alpha}_1} N^{\hat{\alpha}_2}} = \frac{\left(\int \hat{A}_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{\alpha}{1-\alpha}} di \right)^{\frac{\theta}{\theta-1}}}{\left(\int \hat{A}_{it}^{\frac{1}{1-\alpha}} ARPK_{it}^{-\frac{1}{1-\alpha}} di \right)^{\hat{\alpha}_1+\hat{\alpha}_2}}
$$

Following similar steps as in the baseline case, we can derive

$$
a = a^* - \frac{1}{2} \frac{\theta}{\theta - 1} \frac{\alpha}{1-\alpha} \sigma^2_{arpk}
$$

Under constant returns to scale in production, this simplifies to

$$
a = a^* - \frac{1}{2} \theta \sigma^2_{arpk}
$$

The output effects of σ^2_{arpk} are the same as in the baseline case.

F Heterogeneity in Markups/Technologies

Baseline approach. The firm's cost minimization problem is

$$\min_{K_{it},N_{it},M_{it}} R_t T_{it}^K K_{it} + W_t T_{it}^N N_{it} + P_{it}^M M_{it} \quad s.t. \quad Y_{it} \leq K_{it}^{\hat{\alpha}_{it}} N_{it}^{\hat{\zeta}-\hat{\alpha}_{it}} M_{it}^{1-\hat{\zeta}}$$

The first order condition on M_{it} gives

$$P_{it}^M = \left(1-\hat{\zeta}\right) \frac{Y_{it}}{M_{it}} MC_{it} \quad \Rightarrow \quad \frac{P_{it}^M M_{it}}{P_{it} Y_{it}} = \left(1-\hat{\zeta}\right) \frac{MC_{it}}{P_{it}}$$

where MC_{it} is the Lagrange multiplier on the constraint (i.e., the marginal cost). Rearranging gives expression (14). In logs,

$$\log \frac{P_{it}}{MC_{it}} = \log\left(1-\hat{\zeta}\right) + \log \frac{P_{it} Y_{it}}{P_{it}^M M_{it}} \quad \Rightarrow \quad \sigma^2 \left(\log \frac{P_{it}}{MC_{it}}\right) = \sigma^2 \left(\log \frac{P_{it} Y_{it}}{P_{it}^M M_{it}}\right)$$

Similarly, the optimality conditions for K_{it} and N_{it} yield:

$$\begin{aligned} \log \frac{P_{it} Y_{it}}{K_{it}} &= \log \frac{P_{it}}{MC_{it}} - \log \hat{\alpha}_{it} + \tau_{it}^K + \text{Constant} \\ \log \frac{P_{it} Y_{it}}{N_{it}} &= \log \frac{P_{it}}{MC_{it}} - \log\left(\hat{\zeta} - \hat{\alpha}_{it}\right) + \tau_{it}^N + \text{Constant} \end{aligned}$$

Log-linearizing around the average $\hat{\alpha}_{it}$, denote it $\bar{\alpha}$, and ignoring constants yields $\log\left(\hat{\zeta} - \hat{\alpha}_{it}\right) \approx -\frac{\bar{\alpha}}{\hat{\zeta}-\bar{\alpha}} \log \hat{\alpha}_{it}$. Substituting gives expression (17).

Proof of Proposition 2. Assuming $\log \hat{\alpha}_{it}$ is uncorrelated with τ_{it}^K and τ_{it}^N,

$$cov\left(\widetilde{arpk}_{it}, \widetilde{arpn}_{it}\right) = -\frac{\bar{\alpha}}{\hat{\zeta}-\bar{\alpha}} \sigma_{\log\hat{\alpha}}^2 + \text{cov}\left(\tau_{it}^k, \tau_{it}^n\right) \tag{F.1}$$

$$\sigma_{\widetilde{arpk}}^2 = \sigma_{\log\hat{\alpha}}^2 + \sigma_{\tau^k}^2 \tag{F.2}$$

$$\sigma_{\widetilde{arpn}}^2 = \left(\frac{\bar{\alpha}}{\hat{\zeta}-\bar{\alpha}}\right)^2 \sigma_{\log\hat{\alpha}}^2 + \sigma_{\tau^n}^2 \tag{F.3}$$

From here, we can solve for the correlation of the distortions:

$$\rho\left(\tau_{it}^K, \tau_{it}^N\right) = \frac{\text{cov}\left(\widetilde{arpk}_{it}, \widetilde{arpn}_{it}\right) + \frac{\bar{\alpha}}{\hat{\zeta}-\bar{\alpha}} \sigma_{\log\hat{\alpha}}^2}{\sqrt{\sigma_{\widetilde{arpk}}^2 - \sigma_{\log\hat{\alpha}}^2} \sqrt{\sigma_{\widetilde{arpn}}^2 - \left(\frac{\bar{\alpha}}{\hat{\zeta}-\bar{\alpha}}\right)^2 \sigma_{\log\hat{\alpha}}^2}}$$

which is increasing in $\sigma_{\log\hat{\alpha}}^2$. An upper bound for $\sigma_{\log\hat{\alpha}}^2$, denoted $\bar{\sigma}_{\log\hat{\alpha}}^2$, is where $\rho\left(\tau_{it}^K, \tau_{it}^N\right) = 1$, and substituting and rearranging gives

$$\bar{\sigma}_{\hat{\alpha}}^2 = \frac{\sigma^2_{\widetilde{arpk}}\sigma^2_{\widetilde{arpn}} - \text{cov}\left(\widetilde{arpk}_{it}, \widetilde{arpn}_{it}\right)^2}{2\frac{\bar{\alpha}}{\hat{\zeta}-\bar{\alpha}}\text{cov}\left(\widetilde{arpk}_{it}, \widetilde{arpn}_{it}\right) + \left(\frac{\bar{\alpha}}{\hat{\zeta}-\bar{\alpha}}\right)^2 \sigma^2_{\widetilde{arpk}} + \sigma^2_{\widetilde{arpn}}}$$

□

Heterogeneous materials elasticities. We now allow for heterogeneity in $\hat{\zeta}_{it}$, so the cost minimization problem becomes

$$\min_{K_{it},N_{it},M_{it}} \quad R_t T^K_{it} K_{it} + W_t T^N_{it} N_{it} + P^M_{it} M_{it} \quad s.t. \quad Y_{it} \leq K^{\hat{\alpha}_{it}}_{it} N^{\hat{\zeta}_{it}-\hat{\alpha}_{it}}_{it} M^{1-\hat{\zeta}_{it}}_{it}$$

The first order conditions give the optimal average products of inputs (after some rearranging):

$$\begin{aligned}
ARPK_{it} &\equiv \frac{P_{it}Y_{it}}{K_{it}} = \frac{P_{it}}{MC_{it}}\frac{1}{\hat{\alpha}_{it}}T^K_{it}R_t \\
ARPN_{it} &\equiv \frac{P_{it}Y_{it}}{N_{it}} = \frac{P_{it}}{MC_{it}}\frac{1}{\hat{\zeta}_{it} - \hat{\alpha}_{it}}T^N_{it}W_t \\
ARPM_{it} &\equiv \frac{P_{it}Y_{it}}{P^M_{it}M_{it}} = \frac{P_{it}}{MC_{it}}\frac{1}{1 - \hat{\zeta}_{it}}
\end{aligned}$$

or in logs:

$$\begin{aligned}
arpk_{it} &= \varphi_{it} - \log\hat{\alpha}_{it} + \tau^K_{it} + \text{Constant} \\
arpn_{it} &= \varphi_{it} - \log\left(\hat{\zeta}_{it} - \hat{\alpha}_{it}\right) + \tau^N_{it} + \text{Constant} \\
&\approx \varphi_{it} - \frac{\bar{\zeta}}{\bar{\zeta} - \bar{\alpha}}\log\hat{\zeta}_{it} + \frac{\bar{\alpha}}{\bar{\zeta} - \bar{\alpha}}\log\hat{\alpha}_{it} + \tau^N_{it} + \text{Constant} \\
arpm_{it} &= \varphi_{it} - \log\left(1 - \hat{\zeta}_{it}\right) \\
&\approx \varphi_{it} + \frac{\bar{\zeta}}{1 - \bar{\zeta}}\log\hat{\zeta}_{it} + \text{Constant}
\end{aligned}$$

where, to ease notation, we define $\varphi_{it} \equiv \log\frac{P_{it}}{MC_{it}}$ as the log markup and the approximations reflect log-linearizations around the average elasticities, denoted $\bar{\alpha}$ and $\bar{\zeta}$.

There are three categories of firm-specific variation in the average product of inputs: markups, φ_{it}, input elasticities in production, $\hat{\alpha}_{it}$ and $\hat{\zeta}_{it}$, and distortions, τ^K_{it} and τ^N_{it}. We make the following key assumption: this variation is independent across categories. Within categories, however, we allow for arbitrary correlations. In other words, the covariances $\sigma_{\log\hat{\alpha},\log\hat{\zeta}}$ and σ_{τ^K,τ^N} are unrestricted but the other covariances are set to zero.[1] We can derive the following expressions for the second moments of the average products of the three inputs:

$$\sigma^2_{arpk} = \sigma^2_{\varphi} + \sigma^2_{\log\hat{\alpha}} + \sigma^2_{\tau^K} \tag{F.4}$$

$$\sigma^2_{arpn} = \sigma^2_{\varphi} + \left(\frac{\bar{\zeta}}{\bar{\zeta}-\bar{\alpha}}\right)^2 \sigma^2_{\log\hat{\zeta}} + \left(\frac{\bar{\alpha}}{\bar{\zeta}-\bar{\alpha}}\right)^2 \sigma^2_{\log\hat{\alpha}} - \frac{2\bar{\alpha}\bar{\zeta}}{\left(\bar{\zeta}-\bar{\alpha}\right)^2}\sigma_{\log\hat{\alpha},\log\hat{\zeta}} + \sigma^2_{\tau^N} \tag{F.5}$$

$$\sigma_{arpk,arpn} = \sigma^2_{\varphi} + \frac{\bar{\zeta}}{\bar{\zeta}-\bar{\alpha}}\sigma_{\log\hat{\alpha},\log\hat{\zeta}} - \frac{\bar{\alpha}}{\bar{\zeta}-\bar{\alpha}}\sigma^2_{\log\hat{\alpha}} + \sigma_{\tau^K,\tau^N} \tag{F.6}$$

$$\sigma^2_{arpm} = \sigma^2_{\varphi} + \left(\frac{\bar{\zeta}}{1-\bar{\zeta}}\right)^2 \sigma^2_{\log\hat{\zeta}} \tag{F.7}$$

$$\sigma_{arpk,arpm} = \sigma^2_{\varphi} - \frac{\bar{\zeta}}{1-\bar{\zeta}}\sigma_{\log\hat{\alpha},\log\hat{\zeta}} \tag{F.8}$$

$$\sigma_{arpn,arpm} = \sigma^2_{\varphi} - \frac{\bar{\zeta}^2}{\left(\bar{\zeta}-\bar{\alpha}\right)\left(1-\bar{\zeta}\right)}\sigma^2_{\log\hat{\zeta}} + \frac{\bar{\alpha}\bar{\zeta}}{\left(\bar{\zeta}-\bar{\alpha}\right)\left(1-\bar{\zeta}\right)}\sigma_{\log\hat{\alpha},\log\hat{\zeta}} \tag{F.9}$$

The following result states that we can identify the dispersion in (log) markups and materials elasticities from the second moments of *arpm*.

Lemma 1. *The parameters $\sigma^2_{\varphi}, \sigma^2_{\log\hat{\zeta}}$ and $\sigma_{\log\hat{\alpha},\log\hat{\zeta}}$ are uniquely identified by $\sigma^2_{arpm}, \sigma_{arpn,arpm}$ and $\sigma_{arpk,arpm}$.*

Proof. Rearrange (F.9) to derive

$$\frac{\bar{\zeta}}{1-\bar{\zeta}}\sigma_{\log\hat{\alpha},\log\hat{\zeta}} = \frac{\bar{\zeta}-\bar{\alpha}}{\bar{\alpha}}\left(\sigma_{arpn,arpm} - \sigma^2_{\varphi}\right) + \frac{1-\bar{\zeta}}{\bar{\alpha}}\left(\frac{\bar{\zeta}}{1-\bar{\zeta}}\right)^2 \sigma^2_{\log\hat{\zeta}}$$

Substituting into (F.8) gives:

$$\sigma^2_{\varphi} = \frac{\bar{\alpha}}{\bar{\zeta}}\sigma_{arpk,arpm} + \frac{\bar{\zeta}-\bar{\alpha}}{\bar{\zeta}}\sigma_{arpn,arpm} + \frac{1-\bar{\zeta}}{\bar{\zeta}}\left(\frac{\bar{\zeta}}{1-\bar{\zeta}}\right)^2 \sigma^2_{\log\hat{\zeta}}$$

and then into (F.7) yields:

$$\left(\frac{\bar{\zeta}}{1-\bar{\zeta}}\right)^2 \sigma^2_{\log\hat{\zeta}} = \bar{\zeta}\sigma^2_{arpm} - \bar{\alpha}\sigma_{arpk,arpm} - \left(\bar{\zeta}-\bar{\alpha}\right)\sigma_{arpn,arpm}$$

This equation pins down $\sigma^2_{\log\hat{\zeta}}$. Given this, the other two equations yield σ^2_{φ} and $\sigma_{\log\hat{\alpha},\log\hat{\zeta}}$. □

Intuitively, the greater the positive covariation in the average revenue products of the three inputs, the lower (higher) the variation in the output elasticity of materials (markups). Using

[1] The fact that materials shares were relatively uncorrelated with size provides some justification for this assumption.

these estimates, we can appropriately adjust the second moments of $arpk$ and $arpn$ and apply the logic of Proposition 2 to derive an upper bound for the variation in $\hat{\alpha}_{it}$. We state this result formally in the following proposition:

Proposition 1. *The dispersion in* $\log \hat{\alpha}_{it}$ *satisfies*

$$\sigma^2(\log \hat{\alpha}_{it}) \leq \frac{\tilde{\sigma}^2_{arpk}\tilde{\sigma}^2_{arpn} - \tilde{cov}\left(arpk, arpn\right)^2}{2\frac{\bar{\alpha}}{\hat{\zeta}-\bar{\alpha}}\tilde{cov}\left(arpk, arpn\right) + \left(\frac{\bar{\alpha}}{\hat{\zeta}-\bar{\alpha}}\right)^2 \tilde{\sigma}^2_{arpk} + \tilde{\sigma}^2_{arpn}} . \tag{F.10}$$

where

$$\begin{aligned}
\tilde{\sigma}^2_{arpk} &\equiv \sigma^2_{arpk} - \sigma^2_{\varphi} \\
\tilde{\sigma}^2_{arpn} &\equiv \sigma^2_{arpn} - \sigma^2_{\varphi} - \left(\frac{\bar{\zeta}}{\bar{\zeta}-\bar{\alpha}}\right)^2 \sigma^2_{\log\hat{\zeta}} + \frac{2\bar{\alpha}\bar{\zeta}}{\left(\bar{\zeta}-\bar{\alpha}\right)^2}\sigma_{\log\hat{\alpha},\log\hat{\zeta}} \\
\tilde{cov}\left(arpk, arpn\right) &\equiv \sigma_{arpk,arpn} - \sigma^2_{\varphi} - \frac{\bar{\zeta}}{\bar{\zeta}-\bar{\alpha}}\sigma_{\log\hat{\alpha},\log\hat{\zeta}} .
\end{aligned}$$

Proof. Note that

$$\begin{aligned}
\tilde{\sigma}^2_{arpk} &= \sigma^2_{\log\hat{\alpha}} + \sigma^2_{\tau^K} \\
\tilde{\sigma}^2_{arpn} &= \left(\frac{\bar{\alpha}}{\bar{\zeta}-\bar{\alpha}}\right)^2 \sigma^2_{\log\hat{\alpha}} + \sigma^2_{\tau^N} \\
\tilde{cov}\left(arpk, arpn\right) &= -\frac{\bar{\alpha}}{\bar{\zeta}-\bar{\alpha}}\sigma^2_{\log\hat{\alpha}} + \mathrm{cov}\left(\tau^K_{it}, \tau^N_{it}\right) .
\end{aligned}$$

which are identical to expressions (F.1), (F.2) and (F.3). The proof is then the same as for Proposition 2. □

In Table F.1, we report the results from applying this methodology to our datasets. Comparing them to the results in Table 4 reveals that allowing for unobserved variation in the output elasticity of materials slightly attenuates the contribution of markup dispersion and raises the upper bound for the effects of technology heterogeneity, but the values are extremely close. One reason why the two approaches are so similar is that the variation in the materials elasticity (across firms within the same industry), $\sigma^2_{\log\hat{\zeta}}$, is estimated to be very small in both countries: 0.007 in the US and 0.018 in China.

Table F.1: Heterogeneous Markups and Technologies with Firm-Specific Materials Elasticities

	China			US		
Covariance matrix	$arpk_{it}$	$arpn_{it}$	$arpm_{it}$	$arpk_{it}$	$arpn_{it}$	$arpm_{it}$
$arpk_{it}$	1.30			0.41		
$arpn_{it}$	0.33	0.69		0.10	0.20	
$arpm_{it}$	0.01	0.01	0.05	0.03	0.05	0.06
Estimated $\Delta\sigma^2_{arpk}$	Level	Share		Level	Share	
Dispersion in Markups	0.03	(2.4%)		0.05	(11.8%)	
Dispersion in $\log \hat{\alpha}_{it}$	0.31	(24.1%)		0.21	(51.6%)	
Total	0.34	(26.5%)		0.26	(63.4%)	

G Size-Dependent Policies

In this appendix, we explore the relationship between size- and productivity-dependent factors. First, note that our empirical strategy can be thought of as essentially recovering the law of motion for k_{it} – in particular, the coefficients ψ_1, $\psi_2(1+\gamma)$, ψ_3 and ψ_4. Importantly, these estimates are invariant to assumptions about γ_k, which only affects the mapping from these coefficients to the underlying structural parameters. For example, suppose we assume $\gamma_k = 0$. Then, given our values for (α, β, δ), the estimated ψ_1 identifies the adjustment cost parameter ξ. Next, the value of ξ can be used to pin down ψ_2, allowing us to recover γ from the estimated $\psi_2(1+\gamma)$. This procedure can be applied for any given γ_k as well. Since the estimated ψ_1 and $\psi_2(1+\gamma)$ do not change, for any γ_k, the adjustment cost parameter becomes, from (8),

$$\xi = \psi_1 \frac{1-\alpha-\gamma_k}{\beta\psi_1^2 + 1 - \psi_1(1+\beta)} .$$

The next step is the same as before: the estimated ξ implies a value for ψ_2, which then allows us to back out γ from the estimated $\psi_2(1+\gamma)$. Table 5 applies this procedure for various values of γ_k to trace out a set of parameters that are observationally equivalent, i.e., that cannot be distinguished using only data on capital and value-added.

H Financial Frictions

Including the liquidity cost, the firm's problem can be written as

$$\begin{aligned} \mathcal{V}(K_{it}, B_{it}, \mathcal{I}_{it}) = \max_{B_{it+1}, K_{it+1}} \quad & \mathbb{E}_{it}\left[\Pi(K_{it}, A_{it}) + RB_{it} - B_{it+1} - T^K_{it+1}K_{it+1}(1-\beta(1-\delta))\right] \\ - \quad & \Phi(K_{it+1}, K_{it}) - \Upsilon(K_{it+1}, B_{it+1}) + \beta\mathbb{E}_{it}\left[\mathcal{V}(K_{it+1}, B_{it+1}, \mathcal{I}_{it+1})\right] \end{aligned}$$

The first order conditions are given by

$$
\begin{aligned}
\mathbb{E}_{it}\left[\beta\Pi_1\left(K_{it+1},A_{it+1}\right)-\beta\Phi_2\left(K_{it+2},K_{it+1}\right)\right] &= T^K_{it+1}\left(1-\beta\left(1-\delta\right)\right)+\Phi_1\left(K_{it+1},K_{it}\right)+\Upsilon_1\left(K_{it+1},B_{it+1}\right) \\
-\Upsilon_2\left(K_{it+1},B_{it+1}\right)+\beta R &= 1
\end{aligned}
$$

Note that

$$
\Upsilon_2\left(K_{it+1},B_{it+1}\right)=-\hat{\nu}\omega_2\frac{K_{it+1}^{\omega_1}}{B_{it+1}^{\omega_2+1}},\qquad \Upsilon_1\left(K_{it+1},B_{it+1}\right)=\hat{\nu}\omega_1\frac{K_{it+1}^{\omega_1-1}}{B_{it+1}^{\omega_2}}
$$

Using the FOC for B_{it+1}

$$
\begin{aligned}
1 &= \hat{\nu}\omega_2\frac{K_{it+1}^{\omega_1}}{B_{it+1}^{\omega_2+1}}+\beta R \quad\Rightarrow\quad B_{it+1}=\left(\frac{\hat{\nu}\omega_2}{1-\beta R}\right)^{\frac{1}{\omega_2+1}}K_{it+1}^{\frac{\omega_1}{\omega_2+1}} \\
\Upsilon_1\left(K_{it+1},B_{it+1}\right) &= \hat{\nu}\omega_1\frac{K_{it+1}^{\omega_1-1}}{B_{it+1}^{\omega_2}}=\hat{\nu}\omega_1\frac{K_{it+1}^{\omega_1-1}}{\left(\frac{\hat{\nu}\omega_2}{1-\beta R}\right)^{\frac{\omega_2}{\omega_2+1}}K_{it+1}^{\frac{\omega_2\omega_1}{\omega_2+1}}}=\left(\frac{\hat{\nu}}{\omega_2^{\omega_2}}\right)^{\frac{1}{\omega_2+1}}\omega_1\left(1-\beta R\right)^{\frac{\omega_2}{\omega_2+1}}K_{it+1}^{\frac{\omega_1-(\omega_2+1)}{\omega_2+1}} \\
&= \nu\left(1-\beta R\right)^{\frac{\omega_2}{\omega_2+1}}K_{it+1}^{\omega}\ ,
\end{aligned}
$$

where

$$
\begin{aligned}
\nu &\equiv \left(\frac{\hat{\nu}}{\omega_2^{\omega_2}}\right)^{\frac{1}{\omega_2+1}}\omega_1 \\
\omega &\equiv \frac{\omega_1-(\omega_2+1)}{\omega_2+1}.
\end{aligned}
$$

Log-linearizing,

$$
\begin{aligned}
\bar{\Upsilon}_1+\bar{\Upsilon}_1\upsilon_{1t+1} &\approx \nu\left(1-\beta R\right)^{\frac{\omega_2}{\omega_2+1}}\bar{K}^{\omega}+\nu\left(1-\beta R\right)^{\frac{\omega_2}{\omega_2+1}}\bar{K}^{\omega}\omega k_{it+1} \\
\bar{\Upsilon}_1\upsilon_{1t+1} &\approx \nu\left(1-\beta R\right)^{\frac{\omega_2}{\omega_2+1}}\bar{K}^{\omega}\omega k_{it+1}\ .
\end{aligned}
$$

Substituting into the FOC,

$$
\begin{aligned}
&\mathbb{E}_{it}\left[\alpha\beta\bar{G}\bar{A}\bar{K}^{\alpha-1}\left(a_{it+1}+\left(\alpha-1\right)k_{it+1}\right)+\beta\hat{\xi}\left(k_{it+2}-k_{it+1}\right)-\tau^K_{it+1}\left(1-\beta\left(1-\delta\right)\right)\right] \\
=\ &\hat{\xi}\left(k_{it+1}-k_{it}\right)+\nu\left(1-\beta R\right)^{\frac{\omega_2}{\omega_2+1}}\bar{K}^{\omega}\omega k_{it+1}\ ,
\end{aligned}
$$

or

$$
k_{it+1}\left(\left(1+\beta\right)\xi+1-\alpha-\gamma_k\right)=\mathbb{E}_{it}\left[a_{it+1}+\tau_{it+1}\right]+\beta\xi\mathbb{E}_{it}\left[k_{it+2}\right]+\xi k_{it}\ ,
$$

where

$$\gamma_k = -\frac{\nu(1-\beta R)^{\frac{\omega_2}{\omega_2+1}}\omega\bar{K}^{\omega}}{\alpha\beta\bar{G}\bar{A}\bar{K}^{\alpha-1}} = -\frac{\nu(1-\beta R)^{\frac{\omega_2}{\omega_2+1}}\omega\bar{K}^{\omega}}{\nu(1-\beta R)^{\frac{\omega_2}{\omega_2+1}}\bar{K}^{\omega}+1-\beta(1-\delta)+\hat{\xi}\delta\left(1-\beta\left(1-\frac{\delta}{2}\right)\right)}$$
$$= -\frac{\omega\bar{\Upsilon}_1}{\bar{\Upsilon}_1+\kappa}$$

where we have substituted in from the steady state Euler equations and $\kappa \equiv 1-\beta(1-\delta)+\hat{\xi}\delta\left(1-\beta\left(1-\frac{\delta}{2}\right)\right)$.

I Robustness

I.1 Computation and Estimation of Non-Convex Model

This appendix provides details of our analysis of non-convex adjustment costs from Section V.A. Since we can no longer rely on the perturbation approach, we solve the model non-linearly using value function iteration and estimate the parameters via simulated method of moments.

Estimation details. Our estimation uses the following procedure. For a given parameter vector $\left(\hat{\xi},\hat{\xi}_f,\mathbb{V},\gamma.\sigma_{\varepsilon}^2,\sigma_{\chi}^2\right)$, we solve for the value and policy functions using a standard iterative procedure and discretized grids for the state variables.[2] We then use these solutions to simulate time paths (10,000 periods) for firm-level capital and productivity. We discard the first 5,000 periods and compute the moments of interest using the remaining observations. We then search over the parameter vector to minimize the equally-weighted sum of squared deviations between the simulated values of the six target moments and their empirical counterparts.

Model fit. We report the fit of the estimated model in row 'All factors baseline' of Table I.1. The left panel displays the parameter estimates and the right panel the simulated moments (the top row labeled 'Data' reports the empirical values of the moments). The set of targeted moments is marked in bold italics. The table shows that the model matches the targeted moments quite well (the fit is almost exact in the US; the model slightly undershoots the variance of investment growth and the correlation of *arpk* with a in China, but is still fairly close on those dimensions).

The last four columns of the table contain four additional moments not explicitly targeted in the estimation – the autocorrelation of investment (in levels), denoted $\rho_{k,k_{-1}}$, the correlation

[2] We use a relatively fine grid for capital (at intervals of 0.025 log points). For the productivity process, we use 21 grid points, spanning 6 standard deviations (±3 standard deviations on either side of the mean). We have verified that our results are not particularly sensitive to these choices.

of investment with productivity, $\rho_{k,a}$, and investment 'spikes,' defined as the fraction of observations with (gross) investment rates above 20%, $spike^{+}$, or less than -20%, $spike^{-}$. These are the moments targeted in Cooper and Haltiwanger (2006). The set of moments examined in the table were broadly chosen to encompass those from our estimation and additional moments considered in previous influential studies of adjustment costs, namely, Cooper and Haltiwanger (2006) and Asker et al. (2014) (the latter paper targets the variability of investment, inaction and spikes, where the latter two moments are defined in the same way as here).

Turning to the non-targeted moments, the model somewhat over-predicts the serial correlation of investment as well as its correlation with productivity. The model also overshoots a bit on the fraction of positive investment spikes. To explore the extent to which these deviations matter for our main conclusions, we estimated two alternative versions of the model. The first replaces inaction as a target with $spike^{+}$ and $spike^{-}$.[3] The results, reported in row 'All factors spikes', yield estimates of the adjustment costs that are only slightly higher than the baseline values in both countries. In the second exercise, we targeted the serial correlation of investment in levels (rather than growth rates). As with the first exercise, the parameter estimates (reported in row 'All factors $\rho_{k,k_{-1}}$') change only slightly. Importantly, across both exercises, the contribution of the various factors to *arpk* dispersion (not reported in the table) are almost unchanged. In sum, these exercises reveal that (i) while our relatively simple specification of adjustment costs and other distortionary factors can reconcile an extremely broad set of investment moments, it struggles to exactly match all moments simultaneously, but (ii) despite this, our conclusions about the sources of dispersion in *arpk* are quite robust to the precise choice of moments.[4]

The role of other factors. Explicitly allowing for other distortionary factors plays a key role in our analysis. It significantly contributes to our ability to simultaneously match various data moments and is the primary reason for the difference between our estimates of adjustment costs and those in previous studies.

To show this more clearly, we estimated three alternative versions of our model with only adjustment costs.[5] The first is estimated by targeting the same moments as do Cooper and Haltiwanger (2006), namely, the serial correlation of investment, its correlation with productiv-

[3] This is in line with Cooper and Haltiwanger (2006), who argue that inaction may be poorly measured in the micro-data and use these moments instead.

[4] We also estimated a version where all 10 moments are targeted together. Unsurprisingly, the model cannot exactly match all of them, but the best-fit parameter estimates are quite similar to the baseline.

[5] Formally, all the parameters except the two governing adjustment costs ($\hat{\xi}$ and $\hat{\xi}_f$) are set to 0. We then search over $\hat{\xi}$ and $\hat{\xi}_f$ to minimize the equally-weighted distance between the model-implied and data values for the target moments.

Table I.1: Non-Convex Adjustment Costs - Alternative Approaches

	$\hat{\xi}$	$\hat{\xi}_f$	$\mathbb{V}$	γ	σ^2_ε	σ^2_χ	$\rho_{\iota,\iota_{-1}}$	σ^2_ι	$\rho_{arpk,a}$	$\rho_{\iota,a_{-1}}$	σ^2_{arpk}	$inact$	$\rho_{k,k_{-1}}$	$\rho_{k,a}$	$spike^+$	$spike^-$
US																
Data							−0.30	0.06	0.55	0.13	0.45	0.18	0.25	0.06	0.26	0.10
All factors baseline	0.25	0.002	0.03	−0.30	0.02	0.29	***-0.30***	***0.05***	***0.55***	***0.13***	***0.45***	***0.18***	0.46	0.17	0.31	0.08
All factors spikes	0.31	0.007	0.03	−0.27	0.02	0.29	***-0.29***	***0.04***	***0.54***	***0.13***	***0.45***	0.30	0.50	0.20	***0.30***	***0.07***
All factors $\rho_{k,k_{-1}}$	0.34	0.003	0.04	−0.32	0.16	0.26	−0.42	***0.09***	***0.59***	***0.13***	***0.45***	***0.18***	***0.26***	0.17	0.33	0.10
AC only I	0.11	0.130	-	-	-	-	−0.42	0.26	0.57	−0.26	0.06	0.70	***0.24***	***0.03***	***0.23***	***0.07***
AC only II	0.10	0.120	-	-	-	-	−0.40	0.24	0.57	−0.27	0.05	0.34	***0.27***	***0.03***	***0.21***	***0.06***
AC only III	0.65	0.003	-	-	-	-	−0.18	***0.03***	0.75	−0.22	0.11	***0.18***	0.66	0.20	***0.30***	***0.07***
China																
Data							−0.36	0.14	0.76	0.29	0.92	0.20	0.04	−0.06	0.27	0.15
All factors baseline	0.07	0.002	0.09	−0.64	0.00	0.44	***-0.38***	***0.11***	***0.71***	***0.28***	***0.92***	***0.21***	0.24	0.17	0.35	0.13
All factors spikes	0.09	0.001	0.09	−0.66	0.00	0.41	***-0.38***	***0.09***	***0.73***	***0.28***	***0.92***	0.18	0.25	0.17	***0.33***	***0.11***
All factors $\rho_{k,k_{-1}}$	0.05	0.002	0.10	−0.70	0.01	0.38	−0.47	***0.16***	***0.76***	***0.26***	***0.92***	***0.23***	***0.05***	0.14	0.36	0.15
AC only I	0.01	0.200	-	-	-	-	−0.47	2.23	0.51	−0.39	0.05	0.60	***0.05***	***-0.09***	***0.28***	***0.12***
AC only II	0.01	0.190	-	-	-	-	−0.47	2.28	0.51	−0.40	0.04	0.24	***0.05***	***-0.10***	***0.27***	***0.11***
AC only III	0.80	0.012	-	-	-	-	−0.17	***0.05***	0.88	−0.20	0.28	***0.20***	0.67	0.26	***0.34***	***0.13***

Notes: Rows labeled 'All factors' contain estimations of the full model. Rows labeled 'AC only' contain estimations of adjustment cost-only models. Targeted moments are in bold italics. 'All factors baseline' replicates our baseline analysis from Section V.A. 'All factors spikes' targets investment spikes. 'All factors $\rho_{k,k_{-1}}$' targets the serial correlation of investment in levels. 'AC only I' targets the moments from Cooper and Haltiwanger (2006). 'AC only II' targets the same moments, but only subjects investment rates greater than than 5% to the fixed adjustment cost. 'AC only III' targets moments along the lines of Asker et al. (2014).

ity and the fractions of positive and negative spikes.[6] The results, presented in row 'AC Only I' in Table I.1, show much larger fixed adjustment costs (and lower convex costs) compared to all the variants of our estimation in the first three rows. Intuitively, as the only offsetting force, large non-convex costs are necessary to match the relatively modest serial correlation of investment observed in the data. However, this comes at the expense of counterfactually high values for inaction and investment variability. For example, in the US, this version of the model predicts an inaction rate of 70% and a variance of investment growth of 0.26, compared to their empirical values of 18% and 0.06, respectively. In other words, an adjustment cost-only model estimated to match only the serial correlation of investment and investment spikes struggles to match the modest degrees of both inaction and investment variability observed in the data.

It is possible to partly fix some of these counterfactual implications using a more complicated specification of the adjustment cost function. For example, assuming that only large investments are subject to the fixed cost helps reduce the degree of inaction.[7] The row labeled 'AC Only II' shows results from such a modification, where the fixed cost is incurred only for investment rates greater than 5% in absolute value. As expected, this brings the predicted value for inaction much closer to the data, particularly in China (predicted inaction remains excessively high in the US, 34% compared to 18% in the data), but has little effect on the variability of investment growth, which remains counterfactually high. More importantly for our purposes, neither version of the adjustment cost-only model generates significant dispersion in *arpk* (indeed the second specification that better fits the data actually *reduces* the implied dispersion from adjustment costs) – in the US, the predicted σ^2_{arpk} is only about 13% of the observed level in the data. This fraction is even lower in China.

The final exercise, displayed in row 'AC Only III', targets the variability of investment growth along with inaction and spikes together. This is similar to the strategy in Asker et al. (2014) (with the caveat that they target the variance of investment in levels). This produces substantially higher estimates for the convex cost in both countries:[8] Intuitively, a large convex component is necessary to match the extremely low variability of investment. However, these estimates imply a counterfactually high serial correlation. In China, for example, the predicted $\rho_{k,k_{-1}}$ is 0.67, compared to the empirical value of only 0.04 (the corresponding values in the US are 0.66 compared to 0.25). These findings are precisely in line with the logic presented in

[6]See section 4.1.2 of that paper. The moments and resulting parameter estimates are not directly comparable since the set of firms is quite different – Cooper and Haltiwanger (2006) work with data on US manufacturing firms from the Longitudinal Research Database.

[7]This is along the lines of the specification in Khan and Thomas (2008), who also point out the inability of standard adjustment cost models to simultaneously match both inaction and spikes in firm/establishment-level data.

[8]The results for the fixed component are more mixed – the estimate is very close to our baseline value in the US and is significantly higher in China, though well below the previous two adjustment cost-only estimations.

Section II and are analogous to our discussion of the adjustment cost results in Section III.D – the fact that the data show only modest serial correlations of investment/investment growth rates limits the potential for convex adjustment frictions; an estimation strategy ignoring this moment (and targeting the volatility of investment) can lead to substantial upward bias in the estimates of convex costs.

In contrast to these adjustment cost-only specifications, our baseline model is able to capture a broader set of data patterns precisely because of the inclusion of other factors influencing investment. To gain some intuition for how they help, we turn to the formulae for the variance and serial correlation of investment derived in Section 3 for the random walk case:

$$\begin{aligned} \sigma_k^2 &= \left(\frac{\psi_2^2}{1-\psi_1^2}\right)(1+\gamma)^2\sigma_\mu^2 + \frac{2}{1+\psi_1}\psi_3^2\sigma_\varepsilon^2 \\ \rho_{k,k-1} &= \psi_1 - \frac{\psi_3^2\sigma_\varepsilon^2}{\sigma_k^2} , \end{aligned}$$

where ψ_1, ψ_2, ψ_3 are composite parameters independent of distortions. These expressions show that, *ceteris paribus*, more severe correlated distortions (i.e., more negative γ) reduce both the volatility and serial correlation of investment (the latter through the effects on σ_k^2). Intuitively, correlated distortions lessen the influence of the persistent productivity process on investment, reducing the serial correlation. Uncorrelated factors (higher σ_ε^2) also make investment less serially correlated, but more volatile. Quantitatively, the first effect is much larger.[9] As a result, the model can match both of these moments without resorting to large non-convex costs (and the associated counterfactual implications).

In sum, the exercises in this appendix emphasize one of the main messages of our analysis: examining a broad set of investment moments imposes additional discipline on the magnitude of the various forces (including adjustment costs). In both countries, the data show that investment/investment growth is (i) neither particularly volatile (ii) nor highly autocorrelated, but (iii) there are large and extremely extremely persistent deviations of firm-level capital from its 'efficient' level. These patterns seem hard to rationalize with standard specifications of adjustment costs alone and lead us to find a significant role for other factors, particularly when it comes to explaining the dispersion in σ_{arpk}^2.

I.2 Alternative Stochastic Processes

In this section, we analyze the implications of alternative, richer stochastic processes for firm-level productivity and distortions.

[9] For example, at our baseline estimates in the US, the coefficient on $\psi_3^2\sigma_\varepsilon^2$ in the expression for σ_k^2 is $\frac{2}{1+\psi_1} \approx 1.2$ while the coefficient in the expression for $\rho_{k,k-1}$ is $\frac{1}{\sigma_k^2} \approx 25$.

Fixed-effects in productivity. First, we generalize the process on productivity in equation (5) to include firm-level fixed-effects. Specifically, we assume:

$$\begin{aligned} a_{it} &= \bar{a}_i + \hat{a}_{it}, \qquad \bar{a}_{it} \sim \mathcal{N}\left(0, \sigma^2_{\bar{a}}\right) \\ \hat{a}_{it} &= \rho \hat{a}_{it-1} + \mu_{it}, \quad \mu_{it} \sim \mathcal{N}\left(0, \sigma^2_{\mu}\right) \end{aligned} \tag{I.1}$$

Now, productivity is composed of both an AR(1) component, $\hat{a}_{it}$, as in equation (5) and a firm fixed-effect, $\bar{a}_i$, with cross-sectional variance $\sigma^2_{\bar{a}}$.

We can show that the three parameters $(\rho, \sigma^2_\mu, \sigma^2_{\bar{a}})$ are uniquely identified by (i) the serial correlation of productivity growth along with (ii) the coefficient from a simple autoregression of a_{it} on a_{it-1} and (iii) the residuals from that regression, which we denote $\sigma^2_{\hat{\mu}}$ to distinguish it from σ^2_μ, which is the true, unobserved variance of the innovations in the process (the latter two moments are the same that were used to identify the baseline process). Specifically, we derive:

$$\begin{aligned} \rho_{\Delta a, \Delta a_{-1}} &= \frac{\rho - 1}{2} \\ \rho_{a, a_{-1}} &= \frac{\sigma^2_{\bar{a}} + \rho \sigma^2_{\hat{a}}}{\sigma^2_{\bar{a}} + \sigma^2_{\hat{a}}} \\ \sigma^2_{\hat{\mu}} &= \left(1 - \rho_{a,a_{-1}}\right)^2 \sigma^2_{\bar{a}} + \left(\rho - \rho_{a,a_{-1}}\right)^2 \sigma^2_{\hat{a}} + \sigma^2_\mu \end{aligned}$$

where $\sigma^2_{\hat{a}} = \frac{\sigma^2_\mu}{1-\rho^2}$. The first equation identifies ρ directly. The second two represent two equations in two unknowns, which can be solved for $\sigma^2_{\bar{a}}$ and σ^2_μ. We can then re-estimate the other parameters of the model using this richer process for a_{it} (the remaining moments are unchanged).

We report the results of this estimation in Table I.2. The first three columns display the parameters governing the process on productivity. The estimates for the fixed-effect, $\sigma^2_{\bar{a}}$, are significant in both countries – 0.29 and 0.33 in China and the US, respectively. Comparing the parameter estimates with those in Table 2 shows that (i) the persistence of the AR(1) component here is somewhat lower than under the baseline specification – 0.87 vs. 0.91 for China and 0.84 vs. 0.91 for the US and (ii) the volatility of the shocks, σ^2_μ, is almost unchanged in both countries.

We report the results for the other parameters in the remaining columns of Table I.2. The top panel shows the parameter estimates and the bottom panel the contribution of each factor to observed *arpk* dispersion. In both countries, the estimated adjustment costs, ξ, are slightly lower and the correlated distortion slightly higher (in absolute value) than under the baseline specification in Table 3. The values are almost unchanged for the other factors. The bottom panel of the table shows that our main conclusions regarding the sources of *arpk* dispersion

Table I.2: Estimates with Firm Fixed-Effects

Parameters	ρ	σ_μ^2	$\sigma_{\tilde{a}}^2$	ξ	$\mathbb{V}$	γ	σ_ε^2	σ_χ^2
China	0.87	0.14	0.29	0.12	0.10	-0.71	0.00	0.42
US	0.84	0.08	0.33	0.76	0.04	-0.38	0.00	0.30
$\frac{\Delta\sigma_{arpk}^2}{\sigma_{arpk}^2}$								
China				1.1%	10.3%	47.8%	0.0%	46.1%
US				6.7%	8.1%	18.9%	1.1%	65.9%

continue to hold.

Persistence in distortions. In our baseline setup, the transitory uncorrelated distortion was assumed to be an iid draw in each period. Here, we generalize that formulation and assume that it follows an AR(1) process:

$$\begin{aligned} \tau_{it} &= \gamma a_{it} + \hat{\tau}_{it} + \chi_i \\ \hat{\tau}_{it} &= \rho_\tau \hat{\tau}_{it-1} + \varepsilon_{it} \end{aligned} \tag{I.2}$$

Compared to the baseline case, there is now an additional parameter, ρ_τ, and therefore the estimation requires more moments. At the end of this appendix, we extend our analytical approach from Section II and prove identification of all the parameters under this more general process.[10] In particular, we show that adding the second-order serial correlation of *arpk* (in changes) is sufficient for identification of the new parameter, ρ_τ.[11] Guided by this result, we re-estimated the model adding both the first- and second-order serial correlations of *arpk* as target moments.

We report the results from this estimation in Table I.3. The estimates for ρ_τ are essentially zero in both countries, providing support for the baseline iid assumption. In other words, conditional on the fixed and correlated components, the remaining transitory piece of the distortion is extremely short-lived. The remaining parameters are quite close to their baseline values.[12]

The remainder of this appendix proves identification of the model parameters in the case

[10] As before, we cannot identify the fixed-effect, σ_χ^2 since second moments in levels are not well defined.

[11] The first-order serial correlation turns out to be a simple transformation of other target moments and thus does not contain new information.

[12] We also estimated another version of the model, where we assumed the uncorrelated component follows an AR(1) without fixed-effects, i.e., we set $\sigma_\chi^2 = 0$, and estimated ρ_τ along with the other parameters by targeting the same set of moments as in the baseline analysis. This yielded values of ρ_τ very close to one, again pointing to an extremely persistent component. Further, the estimated magnitude of the uncorrelated component in this case, i.e., $\frac{\sigma_\epsilon^2}{1-\rho_\tau^2}$, was quite close to the baseline estimate for σ_χ^2. The remaining parameters were almost identical in the two versions.

Table I.3: Persistence in Distortions

	New Moments		Parameter Estimates					
	$\rho_{arpk,arpk_{-1}}$	$\rho_{arpk,arpk_{-2}}$	ξ	$\mathbb{V}$	γ	σ^2_ε	ρ_τ	σ^2_χ
China	0.90	0.81	0.24	0.09	−0.69	0.00	0.00	0.40
US	0.91	0.83	1.00	0.03	−0.34	0.01	0.00	0.28

that productivity follows a random walk and distortions follow the process in (I.2).

Following the same steps as in Section I of the text, we can derive the firm's investment policy function under this more general structure:

$$k_{it+1} = \psi_1 k_{it} + \psi_2 (1+\gamma) \mathbb{E}_{it} [a_{it+1}] + \psi_3 \hat{\tau}_{it+1} + \psi_4 \chi_i$$

where ψ_1 solves the quadratic equation

$$\xi \left(\beta \psi_1^2 + 1\right) = \psi_1 \left((1+\beta)\xi + 1 - \alpha\right)$$

and

$$\begin{aligned} \psi_2 &= \frac{1}{1-\alpha-\beta\xi\psi_1+\xi} \\ \psi_3 &= \frac{1}{1-\alpha+\beta\xi(1-\psi_1-\rho_\tau)+\xi} \\ \psi_4 &= \frac{1}{1-\alpha+\xi(1-\beta\psi_1)} \end{aligned}$$

This law of motion is similar to (7) and (8) with a few modifications to the coefficients: ψ_1 and ψ_4 are the same as before, but ψ_2 here corresponds to the case where $\rho = 1$ and ψ_3 is generalized to allow for $\rho_\tau \neq 0$.

Investment is given by:

$$\begin{aligned} \Delta k_{it+1} &= \Delta k_{it} + \psi_2 (1+\gamma) \Delta \mathbb{E}_{it} [a_{it+1}] + \psi_3 \Delta \hat{\tau}_{it+1} \\ &= \psi_1 \Delta k_{it} + \psi_2 (1+\gamma) \left((1-\phi)\mu_{it} + \phi \mu_{it+1} + \phi (e_{it+1} - e_{it})\right) + \psi_3 \left((\rho_\tau - 1)\hat{\tau}_{it} + \varepsilon_{it+1}\right) \end{aligned}$$

where $1-\phi = \frac{\mathbb{V}}{\sigma^2_\mu}$. From here, we can derive the following four moments: the variance of investment, σ^2_k, the autocovariance of investment, $\sigma_{k,k_{-1}}$, the coefficient from a regression of $\Delta arpk_{it}$ on Δa_{it} ($\lambda_{arpk,a}$), and the covariance of investment with lagged innovations in productivity,

$\sigma_{k,a_{-1}}$:

$$\begin{aligned}
\sigma_k^2 &= \psi_1^2\sigma_k^2 + (1+\gamma)^2\psi_2^2\sigma_\mu^2 + \frac{2(1-\psi_1)\psi_3^2\sigma_\varepsilon^2}{(1+\rho_\tau)(1-\psi_1\rho_\tau)} && \text{(I.3)}\\
\sigma_{k,k_{-1}} &= \psi_1\sigma_k^2 - \frac{(1-\rho_\tau)\psi_3^2\sigma_\varepsilon^2}{(1+\rho_\tau)(1-\psi_1\rho_\tau)} && \text{(I.4)}\\
\lambda_{arpk,a} &= 1-(1-\alpha)(1+\gamma)\psi_2\phi && \text{(I.5)}\\
\sigma_{k,a_{-1}} &= (1-\phi(1-\psi_1))(1+\gamma)\psi_2\sigma_\mu^2 && \text{(I.6)}
\end{aligned}$$

Here, we have one new parameter, ρ_τ, and so will need an additional moment. It turns out that the first-order serial correlation of $arpk$ does not contain any additional information. To see this, we can derive

$$\begin{aligned}
\sigma_{arpk,arpk_{-1}} &\equiv \operatorname{cov}(\Delta arpk_{it}, \Delta arpk_{it-1})\\
&= \operatorname{cov}(\mu_{it}-(1-\alpha)\Delta k_{it}, \mu_{it-1}-(1-\alpha)\Delta k_{it-1})\\
&= (1-\alpha)^2\sigma_{k,k_{-1}} - (1-\alpha)\sigma_{k,a_{-1}}
\end{aligned}$$

which shows that the moment is a simple combination of two moments we have previously used.

Similarly, the variance of $\Delta arpk_{it}$ is

$$\begin{aligned}
\sigma_{arpk}^2 &\equiv \operatorname{var}(\Delta arpk_{it}) = \operatorname{var}(\mu_{it}-(1-\alpha)\Delta k_{it})\\
&= \sigma_\mu^2 + (1-\alpha)^2\sigma_k^2 - (1-\alpha)\lambda_{arpk,a}\sigma_\mu^2
\end{aligned}$$

which, again, is simply a combination of other moments we have already used.

However, the second-order serial correlation, $\sigma_{arpk,arpk_{-2}}$, does contain new information:

$$\begin{aligned}
\sigma_{arpk,arpk_{-2}} &\equiv \operatorname{cov}(\Delta arpk_{it}, \Delta arpk_{it-2})\\
&= \operatorname{cov}(\mu_{it}-(1-\alpha)\Delta k_{it}, \mu_{it-2}-(1-\alpha)\Delta k_{it-2})\\
&= (1-\alpha)^2\operatorname{cov}(\Delta k_{it}, \Delta k_{it-2})\\
&= (1-\alpha)^2\left(\psi_1\sigma_{k,k_{-1}} - \frac{\rho_\tau\psi_3^2\sigma_\varepsilon^2}{(1+\rho_\tau)(1-\psi_1\rho_\tau)}\right) && \text{(I.7)}
\end{aligned}$$

Substituting expressions (I.4)-(I.6) into (I.3), we obtain:

$$\sigma_k^2 = \psi_1^2\sigma_k^2 + \left(\frac{\sigma_{k,a_{-1}}}{\sigma_\mu^2} + \left(\frac{1-\lambda_{arpk,a}}{1-\alpha}\right)(1-\psi_1)\right)^2\sigma_\mu^2 + \frac{2\left(\psi_1\sigma_k^2-\sigma_{k,k_{-1}}\right)(1-\psi_1)}{1-\rho_\tau}$$

This is one equation in two unknowns, ψ_1 and ρ_τ. Next, substitute (I.4) into (I.7):

$$\sigma_{arpk,arpk_{-2}} = (1-\alpha)^2 \left(\psi_1 \sigma_{k,k_{-1}} - \frac{\rho_\tau}{1-\rho_\tau} \left(\psi_1 \sigma_k^2 - \sigma_{k,k_{-1}} \right) \right)$$

which gives a second equation in the two unknowns. The solution to the two equations yields ψ_1 and ρ_τ. From ψ_1, we can compute ξ and ψ_2. Along with ρ_τ, these give ψ_3. The remaining moment conditions yield the remaining parameters, γ, σ_ε^2 and $\mathbb{V}$.

I.3 Lower Elasticity of Substitution

There is no clear consensus on the appropriate value for the elasticity of substitution parameter, θ, which is set to 6 in our analysis. Estimates generally range between 3 and 10 (see, e.g., Broda and Weinstein (2006)). Studies on firm dynamics tend to use values closer to 6. For example, Cooper and Haltiwanger (2006) estimate a demand elasticity among US manufacturing firms of just about 6; the curvature parameter in Atkeson and Kehoe (2005) is 0.85, which corresponds to $\theta = 7$ in our setup. The literature on misallocation, following Hsieh and Klenow (2009), often uses a lower value, $\theta = 3$.

To investigate the robustness of our conclusions to this parameter, we re-did our analysis using $\theta = 3$. In conjunction with the production function elasticities, $\hat{\alpha}_1$ and $\hat{\alpha}_2$, reported in Table 1, this yields values of α of 0.4 in the US and 0.5 in China (compared to 0.62 and 0.71 in the baseline analysis). We have recomputed the target moments under these new values (recall that moments in productivity depend on the curvature parameter) – Table I.4 – and re-estimated the model targeting these moments – Table I.5.

The moments in Table I.4 show largely the same patterns as those in Table 2 and many of the point estimates change little – for example, investment growth in China is more correlated with lagged shocks, is more volatile and less serially correlated and shows a higher correlation between *arpk* and productivity (although this figure is somewhat lower in both countries than under the baseline α). The extent of the overall dispersion in the *arpk* is almost identical to the baseline, since this figure is independent of α (there is a negligible difference in this value for the US due to the trimming of outliers).

Table I.4: Moments – $\theta = 3$

	ρ	σ_μ^2	$\rho_{\iota,a_{-1}}$	$\rho_{\iota,\iota_{-1}}$	$\rho_{arpk,a}$	σ_ι^2	σ_{arpk}^2
China	0.92	0.13	0.21	−0.36	0.56	0.15	0.92
US	0.95	0.11	0.03	−0.30	0.28	0.06	0.46

The estimation results in Table I.5 also point to very similar patterns regarding the sources of *arpk* dispersion: adjustment/information frictions explain only a modest share, leaving a large role for other factors. The estimated adjustment costs are slightly higher in China and lower in the US. The opposite is true for the level of uncertainty. The estimates for correlated (permanent) factors are slightly smaller (larger) in both countries. Importantly, correlated factors are estimated to be much more severe in China than the US. Of course, θ also plays a significant role in determining the magnitude of aggregate productivity losses from a given amount of *arpk* dispersion, as expression (9) reveals. Thus, with $\theta = 3$, the implied TFP losses from all of the factors are smaller than the baseline. However, these losses remain substantial and differ across the two countries, totaling about 50% in China and 12% in the US.

Table I.5: Contributions to 'Misallocation' – $\theta = 3$

			Other Factors		
	Adjustment Costs	Uncertainty	Correlated	Transitory	Permanent
Parameters	ξ	$\mathbb{V}$	γ	σ^2_ε	σ^2_χ
China	0.21	0.08	−0.49	0.00	0.64
US	0.66	0.04	−0.10	0.00	0.38
$\Delta\sigma^2_{arpk}$					
China	0.01	0.08	0.24	0.00	0.64
US	0.02	0.04	0.01	0.00	0.38
$\frac{\Delta\sigma^2_{arpk}}{\sigma^2_{arpk}}$					
China	1.1%	8.3%	26.4%	0.0%	69.2%
US	4.3%	9.4%	1.4%	0.0%	82.5%
Δa					
China	0.01	0.04	0.12	0.00	0.32
US	0.01	0.01	0.00	0.00	0.10

Taken together, our findings in Table I.5 confirm that our main conclusions regarding the sources of *arpk* dispersion are not overly sensitive to the value of the elasticity of substitution. While the exact productivity costs of that dispersion does depend on this parameter (and more generally, on the extent of curvature), all the cases we have examined suggest they can be substantial.

I.4 Alternative Targets: Investment Moments

In this appendix, we re-estimate our model targeting the autocorrelation and variance of investment in levels, rather than growth rates. The values of these moments are 0.25 and 0.04, respectively, in the US and 0.04 and 0.08 in China. The other target moments are the same as in Table 2. Table I.6 reports the results. A comparison to Table 3 shows that the parameter estimates are quite close to the baseline, as are the contributions to *arpk* dispersion – adjustment costs and uncertainty account for between 15% and 20% of σ^2_{arpk} in the two countries, correlated factors play a large role in China and less so in the US, while fixed factors are quite significant in both countries.

Table I.6: Using Moments from Investment in Levels

			Other Factors		
	Adjustment Costs	Uncertainty	Correlated	Transitory	Permanent
Parameters	ξ	$\mathbb{V}$	γ	σ^2_{ε}	σ^2_{χ}
China	0.37	0.11	−0.72	0.02	0.38
US	1.77	0.04	−0.31	0.19	0.28
$\frac{\Delta\sigma^2_{arpk}}{\sigma^2_{arpk}}$					
China	4.3%	11.9%	48.9%	2.5%	40.8%
US	12.1%	8.1%	13.2%	42.4%	62.8%

I.5 Measurement of Capital

Our baseline analysis uses reported book values of firm-level capital stocks. Here, we use the perpetual inventory method to construct an alternative measure of capital for the US firms. To do this, we follow the approach in Eberly et al. (2012). Here, we briefly describe the procedure and refer the reader to that paper for more details. We use the book value of capital in the first year of our data as the starting value of the capital stock and use the recursion:

$$K_{it} = \left(K_{it-1}\frac{P_{Kt}}{P_{Kt-1}} + I_{it}\right)(1-\delta_j)$$

to estimate the capital stock in the following years, where I_t is measured as expenditures on property, plant and equipment, P_K is the implicit price deflator for nonresidential investment, obtained from the 2013 Economic Report of the President, Table 7, and δ_j is a four-digit industry-specific estimate of the depreciation rate. We calculate the useful life of capital goods in industry j as $L_j = \frac{1}{N_j}\sum_{N_j}\frac{PPENT_{it-1}+DEPR_{it-1}+I_{it}}{DEPR_{it}}$ where N_j is the number of firms in industry j, $PPENT$ is property, plant and equipment net of depreciation and $DEPR$ is depreciation and amortization. The implied depreciation rate for industry j is $\delta_j = \frac{2}{L_j}$. We use the average value for each industry over the sample period.

Table I.7 reports the estimation results. The parameters governing firm productivity, ρ and σ^2_{μ}, are quite close to the baseline values, as is the total amount of observed *arpk* dispersion, σ^2_{arpk}.[13] The autocorrelation of investment growth is somewhat higher and its volatility somewhat lower, which together lead to a higher estimate of the adjustment cost parameter, ξ. This is reflected in the higher contribution of these costs to *arpk* dispersion, which is about 27% of the total (compared to 11% in the baseline). The estimated degree of uncertainty is close to the baseline value. Together, these two forces account for about 33% of the observed *arpk* dispersion, compared to about 18% under our baseline calculations. Thus, our finding of a key role for other firm-specific factors continues to hold – these factors account for roughly two-thirds of σ^2_{arpk}. The largest component shows up as a permanent factor that is orthogonal to firm productivity. The time-varying correlated and uncorrelated components contribute only modestly.

Table I.7: Perpetual Inventory Method for Capital - US firms

Moments	ρ	σ^2_μ	$\rho_{\iota,a_{-1}}$	$\rho_{\iota,\iota_{-1}}$	$\rho_{arpk,a}$	σ^2_ι	σ^2_{arpk}
	0.94	0.07	0.15	-0.18	0.55	0.01	0.43
Parameters			ξ	$\mathbb{V}$	γ	σ^2_ε	σ^2_χ
			5.80	0.02	-0.17	0.05	0.26
Aggregate Effects							
$\Delta\sigma^2_{arpk}$			0.12	0.02	0.02	0.05	0.26
$\frac{\Delta\sigma^2_{arpk}}{\sigma^2_{arpk}}$			27.5%	5.7%	4.3%	12.8%	59.9%
Δa			0.05	0.01	0.01	0.02	0.11

Similar to the exercise in Appendix I.4, we have also re-estimated the model using this alternative measure of firm-level capital stocks and targeting the autocorrelation and variability of investment in levels, rather than growth rates. The results are reported in Table I.8. The estimates are broadly in line with those in Table I.7 and are extremely close to the baseline ones in Table 3. To see why, we have also computed the implied values of the autocorrelation and variance of investment using the parameter estimates from Table I.7. This gives values of 0.69 and 0.02, respectively, compared to the empirical values of 0.57 and 0.02. Because the estimation in Table I.7 already matches these (non-targeted) moments fairly closely, explicitly targeting them does not have a large effect.

I.6 Sectoral Analysis

In this appendix, we repeat our analysis for US firms at a disaggregated sectoral level, allowing for sector-specific structural parameters.

[13]Even in the last year of the sample, the correlation of the two capital stock measures exceeds 0.95.

Table I.8: Perpetual Inventory Capital and Investment in Levels - US

	ξ	$\mathbb{V}$	γ	σ^2_ε	σ^2_χ
Parameters	1.65	0.03	-0.32	0.00	0.28
$\frac{\Delta\sigma^2_{arpk}}{\sigma^2_{arpk}}$	12.0%	6.9%	14.0%	0.7%	64.3%

We begin by computing sector-specific α's (curvature in the profit function) using data on value-added and compensation of labor by sector from the Bureau of Economic Analysis, Annual Industry Accounts.[14] To match the SIC (or NAICS) classifications in Compustat, we compute labor's share of value-added for the 9 major sectors of the industrial classification – Agriculture, Forestry and Fishing; Mining; Construction; Manufacturing; Transportation, Communications and Utilities; Wholesale Trade; Retail Trade; Finance, Insurance and Real Estate; Services.[15] To translate these shares into a value of α, note that under our assumptions of monopolistic competition and constant returns to scale in production, labor's share of value-added is equal to $LS = \frac{\theta-1}{\theta}(1-\hat{\alpha}_1)$ where $1-\hat{\alpha}_1$ is the labor elasticity in the production function. Then, solving for $\hat{\alpha}_1$ and substituting into the definition of α, we have

$$\alpha = \frac{\alpha_1}{1-\alpha_2} = \frac{\frac{\theta-1}{\theta} - LS}{1-(1-\hat{\alpha}_1)\frac{\theta-1}{\theta}} = \frac{\frac{\theta-1}{\theta} - LS}{1-LS}$$

Implementing this procedure yields the values of α in the top panel of Table I.9.[16]

Next, we re-compute our cross-sectional moments for each sector, using the values of α to estimate firm-level productivities. We continue to control for time and industry fixed-effects to extract the firm-specific components of the series (there are multiple four-digit industries within each sector). We report the target moments in the first panel of Table I.9. We then estimate the model separately for each sector, allowing the structural parameters governing the various sources of *arpk* dispersion to vary across sectors. The resulting parameter estimates are presented in the second panel of the table and the implied contribution of each factor to *arpk* dispersion in the last two panels.

There is some heterogeneity across the sectors, both in the overall extent of *arpk* dispersion

[14] The data are available at `https://www.bea.gov/industry/iedguide.htm`.

[15] Most of these correspond one-for-one with sectors reported by the BEA data. There, Transportation and Utilities are reported separately, as are several subcategories of services, which we aggregate. The only sector we were unable to include from the BEA data was Information, as it does not line up one-for-one with an SIC or NAICS category. The shares are calculated as the average over the most recent period available, 1998-2011 (which roughly lines up with the period of the firm-level data, 1998-2009).

[16] We have also calculated this value for the entire US economy by summing across all the sectors reported by the BEA. This gives an aggregate labor share of 0.56 and an implied α of 0.62, exactly our baseline value.

as well as in the estimates for the underlying factors. For example, adjustment costs are largest in manufacturing, where they account for as much as 20% of the observed dispersion and are smallest in FIRE. But, overall, the main message from our baseline analysis continues to hold – adjustment and information frictions, although significant, do not create a lot of *arpk* dispersion, leaving a substantial role for other firm-specific factors. While the results point to some heterogeneity in the correlation structure of these factors, the permanent component seems to play a key role across all sectors.

J Estimates for Other Countries/Firms

In this appendix, we apply our empirical methodology to two additional countries for which we have firm-level data - Colombia and Mexico - as well as to publicly traded firms in China.

The Colombian data come from the Annual Manufacturers Survey (AMS) and span the years 1982-1998. The AMS contains plant-level data and covers plants with more than 10 employees, or sales above a certain threshold (around $35,000 in 1998, the last year of the data). We use data on output and capital, which includes buildings, structures, machinery and equipment. The construction of these variables is described in detail in Eslava et al. (2004). Plants are classified into industries defined at a 4-digit level. The Mexican data are from the Annual Industrial Survey over the years 1984-1990, which covers plants of the 3200 largest manufacturing firms. They are also at the plant-level. We use data on output and capital, which includes machinery and equipment, the value of current construction, land, transportation equipment and other fixed capital assets. A detailed description is in Tybout and Westbrook (1995). Plants are again classified into industries defined at a 4-digit level. Data on publicly traded Chinese firms are from Compustat Global. Due to a lack of a sufficient time-series for most firms, we focus on single cross-section for 2015 (the moments use data going back to 2012). Similarly, due to the sparse representation of many industries, we focus on those with at least 20 firms. For all the datasets, we compute the target moments following the same methodology as outlined in the main text of the paper. Our final samples consist of 44,909 and 3,208 plant-year observations for Colombia and Mexico, respectively, and 1,055 firms in China.

Table J.1 reports the moments and estimated parameter values for these sets of firms, as well as the share of *arpk* dispersion arising from each factor and the effects on aggregate productivity. The results are quite similar to those for Chinese manufacturing firms in Table 3 in the main text. The contribution of adjustment costs and uncertainty to observed *arpk* dispersion is rather limited, and that of uncorrelated transitory factors negligible - across these sets of firms, a large portion of the observed dispersion stems from correlated and permanent firm-specific factors.

References

Asker, J., A. Collard-Wexler, and J. De Loecker (2014): "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 122, 1013–1063.

Atkeson, A. and P. J. Kehoe (2005): "Modeling and Measuring Organization Capital," *Journal of Political Economy*, 113, 1026–1053.

Brandt, L., J. Van Biesebroeck, and Y. Zhang (2014): "Challenges of Working with the Chinese NBS Firm-Level Data," *China Economic Review*, 30, 339–352.

Broda, C. and D. E. Weinstein (2006): "Globalization and the Gains from Variety," *The Quarterly journal of economics*, 121, 541–585.

Cooper, R. W. and J. C. Haltiwanger (2006): "On the Nature of Capital Adjustment Costs," *The Review of Economic Studies*, 73, 611–633.

De Loecker, J. and J. Eeckhout (2017): "The Rise of Market Power and the Macroeconomic Implications," Tech. rep., National Bureau of Economic Research Working Paper 23687.

Eberly, J., S. Rebelo, and N. Vincent (2012): "What Explains the Lagged-Investment Effect?" *Journal of Monetary Economics*, 59, 370–380.

Eslava, M., J. Haltiwanger, A. Kugler, and M. Kugler (2004): "The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence from Colombia," *Journal of Development Economics*, 75, 333–371.

Hsieh, C. and P. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124, 1403–1448.

İmrohoroğlu, A. and Ş. Tüzel (2014): "Firm-Level Productivity, Risk, and Return," *Management Science*, 60, 2073–2090.

Keller, W. and S. R. Yeaple (2009): "Multinational Enterprises, International Trade, and Productivity Growth: Firm-Level Evidence from the United States," *The Review of Economics and Statistics*, 91, 821–831.

Khan, A. and J. K. Thomas (2008): "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica*, 76, 395–436.

Tybout, J. R. and M. D. Westbrook (1995): "Trade Liberalization and the Dimensions of Efficiency Change in Mexican Manufacturing Industries," *Journal of International Economics*, 39, 53–78.

Table I.9: Sector-Level Results

Moments	α	ρ	σ^2_μ	$\rho_{\iota,a_{-1}}$	$\rho_{\iota,\iota_{-1}}$	$\rho_{arpk,a}$	σ^2_ι	σ^2_{arpk}
Agr., Forestry and Fishing	0.77	0.92	0.11	0.13	−0.37	0.92	0.03	0.61
Mining	0.76	0.91	0.10	0.16	−0.29	0.74	0.07	0.35
Construction	0.49	0.93	0.15	0.17	−0.28	0.71	0.07	0.69
Manufacturing	0.59	0.94	0.08	0.10	−0.32	0.50	0.05	0.43
Trans., Comm. and Utilities	0.67	0.94	0.04	0.13	−0.32	0.58	0.03	0.38
Wholesale Trade	0.65	0.94	0.08	0.18	−0.31	0.67	0.05	0.57
Retail Trade	0.61	0.96	0.02	0.20	−0.30	0.25	0.02	0.20
Finance, Insurance and Real Estate	0.78	0.90	0.09	0.28	−0.32	0.77	0.07	0.61
Services	0.38	0.95	0.10	0.03	−0.28	0.31	0.08	0.53
Parameters				ξ	$\mathbb{V}$	γ	σ^2_ε	σ^2_χ
Agr., Forestry and Fishing				0.83	0.05	−0.78	0.01	0.09
Mining				0.49	0.04	−0.56	0.00	0.13
Construction				0.65	0.08	−0.50	0.00	0.32
Manufacturing				3.35	0.03	−0.17	0.18	0.28
Trans., Comm. and Utilities				0.55	0.02	−0.55	0.00	0.25
Wholesale Trade				0.55	0.04	−0.54	0.00	0.30
Retail Trade				1.97	0.01	−0.07	0.03	0.17
Finance, Insurance and Real Estate				0.18	0.06	−0.80	0.00	0.26
Services				0.81	0.04	−0.14	0.00	0.44
$\Delta\sigma^2_{arpk}$								
Agr., Forestry and Fishing				0.07	0.05	0.45	0.01	0.09
Mining				0.06	0.04	0.19	0.00	0.13
Construction				0.04	0.08	0.26	0.00	0.32
Manufacturing				0.09	0.03	0.02	0.18	0.28
Trans., Comm. and Utilities				0.01	0.02	0.10	0.00	0.25
Wholesale Trade				0.02	0.04	0.20	0.00	0.30
Retail Trade				0.02	0.01	0.00	0.03	0.17
Finance, Insurance and Real Estate				0.01	0.06	0.31	0.00	0.26
Services				0.02	0.04	0.02	0.00	0.44
$\frac{\Delta\sigma^2_{arpk}}{\sigma^2_{arpk}}$								
Agr., Forestry and Fishing				0.11	0.08	0.74	0.02	0.15
Mining				0.18	0.10	0.54	0.00	0.37
Construction				0.05	0.11	0.37	0.00	0.47
Manufacturing				0.21	0.07	0.05	0.41	0.63
Trans., Comm. and Utilities				0.03	0.05	0.26	0.01	0.65
Wholesale Trade				0.04	0.07	0.35	0.00	0.54
Retail Trade				0.08	0.05	0.01	0.14	0.85
Finance, Insurance and Real Estate				0.02	0.09	0.51	0.00	0.42
Services				0.05	0.07	0.04	0.00	0.83

Table J.1: Additional Countries/Firms

Moments	ρ	σ^2_μ	$\rho_{\iota,a_{-1}}$	$\rho_{\iota,\iota_{-1}}$	$\rho_{arpk,a}$	σ^2_ι	σ^2_{arpk}
Colombia	0.95	0.09	0.28	-0.35	0.61	0.07	0.98
Mexico	0.93	0.07	0.17	-0.39	0.69	0.02	0.79
China Compustat	0.96	0.04	0.30	-0.42	0.76	0.04	0.41
Parameters			ξ	$\mathbb{V}$	γ	σ^2_ε	σ^2_χ
Colombia			0.54	0.05	-0.55	0.01	0.60
Mexico			0.13	0.04	-0.82	0.00	0.42
China Compustat			0.15	0.03	-0.69	0.00	0.18
$\Delta\sigma^2_{arpk}$							
Colombia			0.02	0.05	0.30	0.01	0.60
Mexico			0.00	0.04	0.36	0.00	0.42
China Compustat			0.00	0.03	0.22	0.00	0.18
$\frac{\Delta\sigma^2_{arpk}}{\sigma^2_{arpk}}$							
Colombia			2.5%	5.6%	30.9%	0.7%	61.3%
Mexico			0.5%	4.9%	44.9%	0.0%	52.8%
China Compustat			0.8%	6.3%	54.0%	0.2%	43.7%
Δa							
Colombia			0.01	0.02	0.13	0.00	0.26
Mexico			0.00	0.02	0.16	0.00	0.18
China Compustat			0.00	0.02	0.19	0.00	0.16