

NBER WORKING PAPER SERIES

PRODUCTIVITY VOLATILITY AND THE MISALLOCATION OF RESOURCES
IN DEVELOPING ECONOMIES

Allan Collard-Wexler
John Asker
Jan De Loecker

Working Paper 17175
<http://www.nber.org/papers/w17175>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2011

We would like to thank Dave Backus, for introducing us to the World Bank Data, and Gian Luca Clementi, Alessandro Gavazza, Pete Klenow, Richard Rogerson, Daniel Xu, and participants in the NYU Stern Micro Lunch and the MIT IO seminar for their comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2011 by Allan Collard-Wexler, John Asker, and Jan De Loecker. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Productivity Volatility and the Misallocation of Resources in Developing Economies
Allan Collard-Wexler, John Asker, and Jan De Loecker
NBER Working Paper No. 17175
June 2011
JEL No. D24,L2,O47

ABSTRACT

We investigate the role of dynamic production inputs and their associated adjustment costs in shaping the dispersion of total factor productivity (TFP) and static measures of capital misallocation within a country. Using data on 5,010 establishments in 33 developing countries from the World Bank's Enterprise Research Data, we find that countries exhibiting greater time-series volatility of productivity are also characterized by greater cross-sectional dispersion in productivity. Volatility in TFP explains one quarter to one third of cross-country productivity dispersion. We document a similar relationship between productivity volatility and the dispersion of the marginal revenue product of capital (static capital misallocation). We then use a standard model of investment with adjustment costs, parameterized using numbers calibrated to U.S. data, to show that increasing the volatility of productivity to the level observed in these developing economies can quantitatively replicate the observed relationship between static misallocation and volatility observed in the data. We find that sixty-one percent of the static capital misallocation in the data is captured by the model's prediction. Our findings suggest that the dynamic process governing productivity shocks is a first-order determinant of differences in misallocation and, hence, income across countries.

Allan Collard-Wexler
Stern School of Business
New York University
44 West Fourth Street
KMC-7-80
New York, NY 10012
and NBER
wexler@nyu.edu

Jan De Loecker
Department of Economics
307 Fisher Hall
Princeton University
Princeton, NJ 08544-1021
and NBER
jdeloeck@princeton.edu

John Asker
Stern School of Business
Department of Economics, Suite 7-79
New York University
44 West 4th Street
New York, NY 10012
and NBER
jasker@stern.nyu.edu

1 Introduction

It is well-documented that there exist large differences in performance, typically measured by total factor productivity (TFP), across producers within narrowly defined industries at any given point in time.¹ In addition, the extent of this dispersion varies considerably across countries, particularly when comparing countries at different stages of economic development. A growing literature considers the welfare effect of this TFP dispersion. Most prominently, Hsieh and Klenow (2009) have put forward a framework that identifies the degree of misallocation of resources from the variation in marginal products of inputs across producers. They find that if producers in the manufacturing sectors of India and China had the same degree of misallocation as the manufacturing sector in the United States, output would increase by thirty and sixty percent, respectively.² Their conclusions are derived by viewing the micro data on production and input choices through the lens of a static model of the economy, coupled with parameters of production and demand for the U.S. economy. Under this setting, optimal input choices are directly informative about the extent to which factors of production are misallocated. Much of the work following this paper has tried to find mechanisms to explain why TFP differences do not get eliminated by reallocation of resources.³

In this paper, we start our analysis by establishing an empirical relationship between the dynamic process governing changes in productivity over time and cross-sectional measures of productivity dispersion and (static) capital misallocation. Data from the World Bank's Enterprise Data on 5,010 establishments in 33 developing countries shows the cross-sectional dispersion of TFP is correlated with the time-series volatility of TFP. We find that countries exhibiting greater time-series volatility of productivity are also characterized by greater cross-sectional dispersion in productivity.

After establishing this result in the data, we focus on a model that provides an economic basis for this observed link. In particular, this model examines the role of dynamic production inputs, capital in our case, and the costs associated with adjusting those inputs in shaping the dispersion of TFP in a given country. We construct a variant of a standard dynamic investment model with adjustment costs where producers face a stochastic process for their underlying (heterogeneous) productivity. We parametrize this investment model using estimates obtained by Bloom (2009) on U.S. data. This, together with the estimated evolution process of establishment-level productivity (obtained using our rich micro data), allows us to compute implications for the dispersion and volatility of the marginal revenue product of capital, the volatility of capital, and the dispersion of productivity.

We show that a parsimonious model of the country-specific productivity process has the ability to explain a large amount of the variation in productivity dispersion, as well as dispersion in the static marginal revenue product of capital across countries. We provide

¹See Bartelsman and Doms (2000), Bartelsman, Haltiwanger, and Scarpetta (2009) and references therein.

²At the industry level, undoing misallocation has been shown empirically to have first-order welfare effects. A well known example is Olley and Pakes (1996)'s study of productivity growth in the telecommunications equipment industry. They find that a large fraction of aggregate productivity growth in the industry is accounted for by the reallocation of output to more productive firms. Many papers also find the Olley and Pakes (1996) covariance term between output and productivity plays a key role in accounting for aggregate productivity growth. See Bartelsman, Haltiwanger, and Scarpetta (2009).

³See for instance Restuccia and Rogerson (2008), Collard-Wexler (2009), Midrigan and Xu (2009), and Moll (2010) for some recent work.

evidence supporting the claim that the dynamic process of productivity is important, both empirically and theoretically, in determining the patterns observed in the cross-section. In particular, we find that sixty-one percent of the observed capital misallocation in the data is captured by the model's prediction.

More importantly, we make the point that the welfare implication of productivity dispersion depends crucially on whether the productivity process is a primitive. If the process is exogenous, then our model provides an optimal capital allocation and, therefore, a socially optimal dispersion in the marginal revenue product of capital. To the extent that producers in developing countries seem to be doing at least as well as this benchmark productivity dispersion, differences across countries are welfare irrelevant (conditional on adjustment costs also being exogenous).

Hence, by missing the dynamics in capital adjustment, large measures of productivity and marginal product dispersion may appear inefficient from a static perspective but are not necessarily inefficient from a dynamic perspective. Clearly, we are not the first to make this point, but our paper goes beyond this by empirically quantifying the extent to which differences in dispersion can be generated from a dynamic model of investment. Emphasizing the potential importance of productivity dynamics is the fact that, in our empirical model, the only difference between countries is the stochastic process of productivity shocks.

The remainder of the paper is organized as follows. In section 2, we briefly describe the data and present some preliminary findings on the empirical relationship between dispersion and volatility. Section 3 introduces the investment model and discusses its empirical implications. In the final section, we discuss conclusions and a few outstanding issues.

2 Data and preliminary analysis

In this section, we establish the statistical link between productivity dispersion and productivity volatility. First, we introduce the rich micro panel dataset, the World Bank's Enterprise Research Data, which we use to investigate misallocation. Our introduction is brief and we refer to Appendix A for more details. We then establish the correlation between productivity dispersion and volatility.

2.1 Production data

These data were collected by the World Bank across 41 countries and many different industries between 2002 and 2006. The main advantage of this dataset is the uniform data collection protocol across countries. Indeed, without this uniformity, it would be more difficult to argue that differences in measurement error are not responsible for the differences we observe across countries. Standard output and input measures are reported in a harmonized fashion. In particular, we observe sales, intermediate inputs, various measures of capital, and employment during (and up to) a three-year period, which allows us to compute changes in productivity and capital.

While there are 71,789 establishments surveyed, only 5,010 can be used, since we need two years of information on sales, assets, and employment to compute productivity and the change in productivity. Out of the 41 countries in the data, only 33 have usable establishment-level observations.⁴ To make magnitudes comparable across countries, the

⁴The limited number of establishments with productivity data is primarily an issue of survey design: For many

data are converted from local currency units into U.S. dollars, using the World Bank’s measure of purchasing power parity (PA.NUS.PPP).⁵

2.2 Measuring productivity and summary statistics

Before we report a few important summary statistics of the data, we discuss the measurement of productivity (TFP), productivity dispersion, and misallocation.

2.2.1 Modeling preliminaries

To guide measurement we begin by providing an explicit model of productivity, in the context of a profit maximizing establishment (we assume that establishments operate as autonomous units so, for our purposes, firms and establishment are synonymous). An establishment i , in country c , in time t , produces output Q_{ict} using the following technology:⁶

$$Q_{ict} = A_{ict} K_{ict}^{\alpha_K} L_{ict}^{\alpha_L} \quad (1)$$

where K_{ict} is the capital input and L_{ict} is the labor input; and the demand curve for the establishment’s product is given by a constant elasticity of demand curve:

$$Q_{ict} = B_{ict} P_{ict}^{-\epsilon} \quad (2)$$

Combining these two equations, we obtain an expression for the revenue generating production function:

$$S_{ict} = Y_{ict} K_{ict}^{\beta_K} L_{ict}^{\beta_L} \quad (3)$$

where $Y_{ict} = A_{ict}^{1-\frac{1}{\epsilon}} B_{ict}^{\frac{1}{\epsilon}}$, and $\beta_X = \alpha_X (1 - \frac{1}{\epsilon})$ such that $X \in \{K, L\}$.

We now need to impute the value of β_L . Cost minimization implies that the labor cost share (i.e., the share of the wage bill in costs) can be expressed as

$$\beta_L = \frac{\epsilon}{\epsilon - 1} \frac{w_{ct} L_{ict}}{S_{ict}} \quad (4)$$

We allow β_L to vary at the level of a sector within a country to allow for flexibility in the production function. Thus, our approach to measuring productivity is to compute:

$$y_{ict} = s_{ict} - \beta_K k_{ict} - \beta_L l_{ict} \quad (5)$$

We define TFP as $y_{ict} = \ln(Y_{ict})$.⁷ We measure the value of β_L using the cost share (as derived above) and simply recover the capital share from the constant returns to scale

years and countries, the World Bank does not collect historical data on capital stock.

⁵Note that we account for differences in the rate of inflation across countries by using a year-specific measure of PPP. Since productivity is a ratio, these PPP conversions get netted out in many specifications, but they are useful when, for instance, we use controls for establishment size.

⁶If materials, M , (or other inputs) are a concern, we can expand the production function to be $Q_{ict} = \min(A_{ict} K_{ict}^{\alpha_K} L_{ict}^{\alpha_L}, f(M))$, such that $f(M)$ is monotonic and increasing. Note that this Leontief structure means that we can ignore materials. Tables 8 and 9 in Appendix A.3 run the analysis in this paper using value-added and gross output measures of productivity and find the same qualitative results, although changes in the data required changes the composition of the sample we can use.

⁷We use lower cases to denote variables in logs throughout the paper. In much of the literature, such as Foster, Haltiwanger, and Syverson (2008), this measure of TFP is referred to as revenue productivity.

assumption, or $\beta_K = 1 - \beta_L$. In order to compute the labor cost share, we need to assign a value to the elasticity parameter, ϵ . We follow Bloom (2009) and set it equal to four.⁸

2.2.2 Summary statistics

The 33 countries that have usable data are shown in Table 1, together with the dispersion of TFP, computed as the standard deviation in y_{ict} . That is, taking Indonesia as an example, 438 establishments have sufficient data for our purposes, and the standard deviation in y_{it} across these Indonesian establishments is 1.23. The interpretation of this number is as follows: The mean productivity level in the data (y_{ict}) is 0.7. If all countries were to have the same average level of productivity, an Indonesian establishment with productivity one standard deviation higher than the average would be generating 3.4 times more output than average, despite using the same inputs. The comparable multiples for Poland and Tanzania are 1.7 and 4, respectively. These large differences between countries in productivity dispersion echo the findings that Hsieh and Klenow (2009) report in their comparison of India, China, and the United States.

Table 2 presents summary statistics at the establishment, country, and country-industry level in panels 1, 2, and 3, respectively. The first panel shows considerable variation in establishment size. Average (logged) sales are 7.1, which translates to 1,211 USD with a standard deviation of 3.1. This sets the point located one standard deviation higher than the mean, at 26,903 USD. The largest establishment has total sales approaching 5 billion USD. Fifty two percent of establishments report no investment (measured by Zero Investment), which explains why the number of observations for Log Investment is 2,389, rather than 5,010.⁹

On average sales do not change substantially from year to year; the mean change in log sales is 0.1. However, the standard deviation of this change is larger, and economically significant, at 0.6. Likewise, productivity changes for some establishments are substantial from year to year, as is reflected in a standard deviation of productivity change of 0.6, which is striking given that the standard deviation of productivity is 1.1.

The bottom panels present aggregate statistics at the country and country-industry levels. The standard deviation of productivity within a country ranges from 0.42 to 1.49, which are economically significant differences. Moreover, the standard deviation in the *change* in productivity within a country varies between 0.2 and 1.1. Similar statistics are shown at the country-industry level of aggregation, where we find a standard deviation of productivity between 0 and 3.8, and a standard deviation of the change in productivity between 0.4 and 2. As one would expect, cross-country-industry differences in productivity dispersion and volatility are somewhat larger than cross-country differences.

2.3 Dispersion and volatility

After measuring TFP for each establishment using data on sales and input usage, we construct the standard deviation of y_{ict} as a measure of TFP dispersion in each country. We rely

⁸Alternatively, we could estimate the output elasticity directly from production data. We follow the standard in this literature and rely on cost shares to compute TFP and thereby avoid the issues surrounding identification of output elasticities (in our case, across many industries and countries).

⁹The large share of zero investment data is a well known feature of production data in developing countries. See, for instance, Levinsohn and Petrin (2003).

on the standard deviation of $(y_{ict} - y_{ict-1})$ as a measure of TFP volatility for each country. We then examine the correlation between these measures of a country’s TFP dispersion and TFP volatility. The result of this process is shown in Figure 1 (depicting specification I in Table 3).

Figure 1 illustrates the positive correlation between TFP dispersion and TFP volatility. Indeed, cross-country variation in TFP volatility explains 26 percent of the cross-country variation in TFP dispersion in an OLS regression with a constant as the only other regressor.

Table 3 presents regressions of productivity dispersion on productivity volatility, measured as described in the previous two paragraphs. Specification I (depicted in Figure 1) shows the OLS regression, using observations at the country level, weighted by the number of productivity observations per country. This weighting is used to give more importance to countries whose measurements of productivity dispersion and productivity volatility are relatively precise. In other words, this weighting by observations per country is a simple proxy for the optimal weighting matrix, the inverse of the covariance matrix of the estimated productivity dispersion and productivity volatility. In this specification, 26 percent of the (appropriately weighted) variation in within-country productivity dispersion is accounted for by productivity volatility. Specification II shows the results from an unweighted regression. Across both specifications, we find a coefficient of around 0.57, with standard errors of less than 0.25. Thus, there is strong statistical support for the hypothesis that dispersion and volatility are related.

In specifications III and IV the unit of observation changes from the country to the establishment. The standard deviation of y_{ict} is common for all establishments in a country, but we now control for establishment size using total assets and the industrial activity of the establishment. The coefficients are equivalent to those found without these controls. The standard errors are also comparable. The results from these regressions eliminate the concern that dispersion and volatility are co-generated by a third variable, such as a country’s industrial composition or the size of plants within a country. We ran additional robustness checks using different measures of productivity and dispersion, and these are presented in the appendix. These rely on specifications that identify effects using country-industry level variation and, hence, have country and industry fixed effects. We find that the relationship between dispersion and volatility presented here is robust.

2.4 Capital misallocation and volatility

Productivity dispersion is economically relevant, to the extent that it reflects movements away from an optimal feasible resource allocation. This is most often examined in the context of productive inefficiency within an economy by inspecting differences in the static marginal revenue product of capital (MRPK) across establishments. The MRPK should, in the absence of adjustment costs (or other frictions), be equal across establishments. Building on the model in section 2.2.1 the static marginal revenue product of capital is given by:

$$MRPK = \frac{\partial S_{ict}}{\partial K_{ict}} = \beta_K \frac{S_{ict}}{K_{ict}} \quad (6)$$

Thus, the dispersion (measured in standard deviations) of $\log(MRPK)$ is:

$$\text{Std.}(\log(MRPK)) = \text{Std.}(\log(S_{ict}) - \log(K_{ict})) = \text{Std.}(s_{ict} - k_{ict}) \quad (7)$$

We use this as our measure of dispersion of the marginal revenue product of capital.

Table 4 presents regressions of static misallocation, $\text{Std.}(s_{ict} - k_{ict})$, on productivity volatility, $\text{Std.}(y_{ict} - y_{ict-1})$. We use the same controls and estimation procedures as before and, as such, the only difference between tables 3 and 4 is the dependent variable. The coefficients in each specification are 0.65, 0.74, 0.63, and 0.66, respectively. All coefficients are statistically significant. Moreover, the R^2 is 0.23 in specification I, where no other controls are included. This increases to 0.3 when industry fixed-effects and log assets are included. Thus, a substantial fraction of cross-country differences in misallocation can be attributed to differences in country-specific productivity volatility. This suggests the existence of an important link between the volatility of productivity in a country and the extent of (static) capital misallocation in that economy.

As with productivity dispersion, we report additional robustness checks in the appendix, using different measures of productivity. Again, we find that these results are robust.¹⁰

3 Structural analysis: Model and empirics

The previous section set the stage by establishing a central fact in the data. In this section, we investigate the quantitative and qualitative implications of a link between dispersion and volatility in a more structural setting, employing a calibrated model.

3.1 A dynamic investment model

We begin by articulating a dynamic investment model that allows us to examine the link between productivity volatility and dispersion in both the static marginal revenue product of capital and productivity. Then we take the model to the data and assess the extent to which the model has explanatory power.

We take the framework developed in section 2.2.1 and embed it in a dynamic framework. Our model follows, and builds on, a standard model of investment used in the work of Bloom (2009), Cooper and Haltiwanger (2006), Dixit and Pindyck (1994), and Caballero and Pindyck (1996). Taking the structure in section 2.2.1 as given, we begin by noting that establishments can hire labor in each period for a wage w . Thus, we can optimize out labor, conditional on Y_{ict} and K_{ict} . This leads to a ‘period-profit’ (ignoring capital costs for the moment) of

$$\pi_{ict} = Y_{ict}^{\frac{1}{1-\beta_L}} K_{ict}^{\frac{\beta_K}{1-\beta_L}} \left(\frac{\beta_L}{w}\right)^{\frac{1}{1-\beta_L}} \left[\left(\frac{\beta_L}{w}\right)^{\beta_L} - w \right] \quad (8)$$

Capital depreciates at rate δ so $K_{ict+1} = (1 - \delta)K_{ict} + I_{ict}$ where I_{ict} denotes investment. These investment decisions are affected by a one-period time to build and a cost of investment $C(I_{ict}, K_{ict}, Y_{ict})$.¹¹ We employ the cost function used in Bloom (2009). It is composed of: 1) a fixed disruption cost of investing; 2) capital irreversibilities, which are captured by a wedge between the purchase and sale price for capital; and 3) a convex adjustment cost expressed as a function of the percent investment rate and, therefore,

¹⁰We have run, but do not report, the full set of robustness checks reported for productivity dispersion.

¹¹This time to build assumption is, in itself, a friction that we can easily shut down by allowing investment to become productive within a period. As an indication of the economic effect of adjustment costs, if we set these to zero, then dispersion in the MRPK is reduced by half.

$C(I_{ict}, K_{ict}, Y_{ict})$:

$$C_K^F 1(I_{ict} \neq 0) \pi(Y_{ict}, K_{ict}) + I_{ict}^+ - (1 - C_K^P) I_{ict}^- + C_K^Q K_{ict} \left(\frac{I_{ict}}{K_{ict}} \right)^2 \quad (9)$$

Next, let Y_{ict} follow an AR(1) process given by

$$\ln(Y_{ict}) = \mu + \rho_c \ln(Y_{ict-1}) + \sigma_c \nu_{ict} \quad (10)$$

where $\nu_{ict} \sim \mathcal{N}(0, 1)$ is an i.i.d. standard normal random variable. Note that we allow the volatility of productivity, as measured by σ_c , and the persistence coefficient, ρ_c to vary from one country to another. Indeed, when we present results from computing our model, the volatility and persistence parameters (σ_c, ρ_c) are the only difference between countries.¹²

An establishment's value function V is given by the Bellman equation:

$$V(Y_{ict}, K_{ict}) = \max_{I_{ict}} \pi(Y_{ict}, K_{ict}) - C(I_{ict}, K_{ict}, Y_{ict}) + \beta \int_{Y_{ict+1}} V(Y_{ict+1}, \delta K_{ict} + I_{ict}) \phi(Y_{ict+1} | Y_{ict}, \rho_c, \mu, \sigma_c) dY_{ict+1} \quad (11)$$

and, thus, an establishment's policy function $I^*(Y_{ict}, K_{ict})$ is just the investment level that maximizes the establishment's continuation value.

Note that since there is neither entry nor exit in this model, there is no truncation of the productivity distribution.¹³ Thus, given the AR(1) structure above, the cross-sectional standard deviation of productivity is mechanically given by the ergodic distribution of Y_{it} . Hence,

$$\text{Std.}(y_{ict}) = \frac{\sigma_c}{\sqrt{1 - \rho_c^2}} \quad (12)$$

where, as earlier, $y_{ict} = \ln(Y_{ict})$.

We analyze the model using computation. The parameters we use are found in table 5. Parameters for the capital share, elasticity of demand, depreciation rate, and discount rate are taken from Bloom (2009). We also follow Bloom in adopting his estimated parameters for adjustments costs. These adjustment cost parameters are estimated from US Compustat data, focusing on large U.S. firms. Bloom's estimation procedure uses a model in which investment decisions are made each month, with the model's predictions aggregated to the year-level to form moments in which to fit the data. Modeling decisions on a monthly level is an attractive approach, as the model incorporates the likely time aggregation embedded in annual data. We follow this approach in computing the model and interpret a period in the model as equivalent to a month in data.¹⁴

¹²Note that the specification in equation (10) rules out aggregate-level shocks to productivity growth. However, a regression of changes in productivity on country-year dummies yields an R-squared of only 4 percent. Thus, there is only a small aggregate component to productivity change.

¹³The absence of entry and exit is a consequence of the decreasing returns to scale in the revenue equation (yielded by constant returns to scale in the production function and an elastic demand curve) and the absence of fixed costs, which make it profitable for any establishment to operate at a small enough scale.

¹⁴This interpretation requires transforming the AR(1) process into its monthly equivalent. After noting that the sum of normal random variables with the same mean is distributed normally, this reduces to a straightforward algebraic exercise.

We also need to set a value for the wage term, w , which takes the role of a scaling parameter in the per-period profit function. We set it to 0.3, simply to make sure per-period profits are greater than zero.¹⁵

The last set of parameters we need to fix are the σ_c and ρ_c terms in the AR(1) process, which governs the evolution of productivity over time. In the subsequent section 3.3.1, we estimate this process using the establishment data from the World Bank Enterprise Survey used earlier. For the moment, however, we merely note that the range of σ_c observed in data lies in the interval $[0.22, 1.12]$. As a result, we compute the model for values of σ_c between 0.1 and 1.4. For ρ_c we pick three values found in the data, the first is 0.865, which is the median value across countries (corresponding to Tanzania), and 0.754 and 0.968, the are the values closest to the 10th and 90th percentile values (i.e. a low value, corresponding to Moldova, and a high value, corresponding to Nicaragua). For more details, see the subsequent discussion in section 3.3.1 and Tables 6 and 7.

We compute the optimal investment policies for the value function in equation (11). We solve this model using a discretized version of the state space (Y_{it}, K_{it}) . Specifically, we use a grid of capital states ranging from log capital 3 to log capital equal to 20, in increments of 0.1. Moreover, we use a grid of productivity $y_{it} = \log(Y_{it})$ with 50 grid points, whose transition matrix and grid points are computed using Tauchen (1986)'s method. The model is solved in MATLAB using policy iteration with a sparse transition matrix (since there are 8,500 states). Using the computed optimal policies, we simulate the evolution of a country, or industry, for 10,000 plants over 1,000 periods. We use the output from the 1,000th and 988th periods to compute the reported results (corresponding to years t and $t - 1$; recall that we interpret a period as a month).

3.2 Computational results

Figure 2 shows the output of the model. Panel 1 puts values of σ_c on the horizontal axis, and computed values of $\text{Std.}[s_{ict} - k_{ict}]$ are on the vertical axis. That is, it examines the way dispersion in the static marginal revenue product of capital changes as σ_c , the volatility of productivity, changes. Panels 2 through 5 have the same format, showing the computed dispersion in productivity ($\text{Std.}[y_{ict}]$), the computed volatility in the marginal revenue product of capital over time ($\text{Std.}[(s_{ict} - k_{ict}) - (s_{ict-1} - k_{ict-1})]$), the volatility in establishments' capital over time ($\text{Std.}[k_{ict} - k_{ict-1}]$), and the proportion of establishments with zero investment over a year, respectively.

Panel 2 is the most mechanical of the relationships reported in Figure 2. As noted above, the dispersion in productivity is given by $\frac{\sigma_c}{\sqrt{1-\rho_c^2}}$. Hence, any deviations away from a straight line in the simulation merely reflect variance introduced by simulation error. A further implication of this is that, if ρ_c and σ_c were constant over countries, there would be no cross-country differences in productivity dispersion.

Panel 1 contains the dispersion of the static marginal revenue product of capital ($\text{Std.}[s_{ict} - k_{ict}]$). This dispersion reflects the optimal investment choices of establishments facing different productivity shocks over time and, hence, different state variables. To make the effect of this clear, note that if all plants had the same capital stock, this graph would replicate the relationship found in Panel 2. Yet the relationship between $\text{Std.}[s_{ict} - k_{ict}]$ and σ_c is not

¹⁵It can be shown that different values of w that yield positive profits give the same dispersion and misallocation statistics. Indeed, the adjustment cost model of Bloom (2009) is purposefully homogenous of degree one.

linear and has a slope change at $\sigma_c = 0.4$ for $\rho_c = 0.968$, at $\sigma_c = 0.8$ for $\rho_c = 0.865$, and at $\sigma_c = 1.1$ for $\rho_c = 0.754$. At this point, the relationship between $\text{Std.}[s_{ict} - k_{ict}]$ and σ_c becomes steeper.

To see why this is happening, examine Panel 4. Panel 4 shows the relationship between $\text{Std.}[k_{ict} - k_{ict-1}]$ and σ_c . As volatility increases, plants will engage in more investment and disinvestment. Since greater volatility leads to larger changes in productivity, it is natural that plants respond by altering their capital stock more frequently. However, for at least some values of the state space, plants begin to reduce their response to productivity shocks after σ_c reaches 0.4 for $\rho_c = 0.968$, 0.8 for $\rho_c = 0.865$ and 1.1 for $\rho_c = 0.754$. At these high levels of volatility, current productivity is a weaker signal of the future marginal revenue product of capital. In the limit, where the productivity process is an i.i.d. draw, current productivity provides no information about future profitability. Firms would choose an optimal level of capital and stick to it forever, resulting in no variance in investment across establishments. Thus, the “hump-shaped” response of capital adjustments to volatility is due to the changing trade-off in determining the value of investment today, between the size of shocks experienced today and the likelihood that they will be swamped by future shocks.

Panel 5 shows the fraction of plants with zero investment in the current year as a function of volatility, σ_c . This panel mirrors the effect of volatility on dispersion in capital adjustment found in Panel 4: As volatility increases, establishments at first respond by investing and disinvesting more often, but above $\sigma_c = 0.4, 0.8$ or 1.1 , the frequency of adjustment falls (for at least some states).

The results in panels 4 and 5 help explain the relationship between misallocation and volatility in Panel 1. As volatility increases above 0.4, 0.8 or 1.1, the capital adjustment mechanism starts to shut down, and this speeds up the dispersion of the static marginal revenue product of capital.

Finally, Panel 3 shows the relationship between the standard deviation of the *change* in $[s_{ict} - k_{ict}]$ and σ_c . This relationship is essentially linear and is driven by year-to-year changes in productivity, rather than large year-to-year changes in capital stock.

3.3 Empirical implications

We now confront the model with the data. Doing this requires that the σ_c and ρ_c parameters in the AR(1) process, which governs the evolution of productivity over time, be estimated. We estimate this using the World Bank Enterprise Survey data, resulting in a σ_c and ρ_c parameter for each country. We then run the model for each country, using the country’s estimated σ_c and ρ_c in the AR(1) productivity process, as well as the other model parameters described above and in Table 5. This generates the prediction for each country that we take to the data.

It is worth emphasizing from the outset that the only part of the model estimated using country-specific data are the σ_c and ρ_c parameters in the AR(1) process. All other parameters are taken from either standard practice in handling U.S. establishment data, or (in the case of the adjustment cost parameters) from models estimated using U.S. data. Thus, the only source of differences between countries in the model’s predictions is the different parameterizations on the productivity process. Thus, to the extent that the model fits the data, this reflects the impact of cross-country differences in the productivity process in accounting for differences in countries’ cross-sectional dispersion in the static marginal revenue

product of capital (static capital misallocation) and the other variables of interest.

3.3.1 Estimating the productivity AR(1) process

The first step is to estimate the AR(1) process used to model the stochastic productivity process over time. Tables 6 and 7, summarize the results of this exercise. Specification I, in Table 6, shows a simple specification in which the process is modeled as

$$y_{ict} = \mu + \rho y_{ict-1} + \sigma v_{ict} \quad (13)$$

such that ρ_c and σ do not vary across countries. Both ρ and σ are significant, with $\rho = 0.83$ and $\sigma = 0.6$. The unit of observation is the establishment. The data is a short panel, as we only have two years of data per establishment. Identification of the AR(1) relies on the assumption that ρ and σ are constant over time and across establishments.¹⁶ This allows us to identify the model using cross-sectional variation in establishment-specific productivity pairs, $\langle y_{ict}, y_{ict-1} \rangle$.

Specifications II through IV estimate the model with a variety of robustness checks on the base specification, examining the sensitivity of the parameters with respect to establishment size and to non-linear specification of the law of motion of productivity. The specification we use to generate the model's predictions is specification V, which provides country specific estimates of both ρ_c and σ_c . These estimates are presented in Table 9, together with the country-specific estimates of σ_c from specifications II and IV. As can be seen, the parameters do not vary in any meaningful way across specifications. Moreover, we find statistically significant differences in the process for productivity in different countries.

3.3.2 The fit of the model to the data

To assess the fit of the model, we compute the sum of squared errors, scaled by the sum of the squared 'dependent' variable (data). That is, if the data is a vector \mathbf{x} that is predicted by a variable $\hat{\mathbf{x}}$, then we compute

$$S^2 = 1 - \frac{(\mathbf{x} - \hat{\mathbf{x}})' (\mathbf{x} - \hat{\mathbf{x}})}{\mathbf{x}' \mathbf{x}} \quad (14)$$

as our measure of fit. This measure of fit is closely related to the uncentered R^2 measure of fit familiar from regression analysis. However, because our model's prediction does not come from a regression, but rather from a parameterized model, nothing in the structure restricts S^2 to lie between 0 and 1, though, by definition, it must be less than or equal to one. That said, to map our measure of fit into a context equivalent to the R^2 , it is correct to interpret S^2 as the proportion of the observed data captured by the model's prediction, with the caveat that it is possible for this number to be negative.¹⁷

¹⁶This restriction is only driven by the data, and our framework could handle various forms of time-specific persistence and volatility if the data had a longer time dimension. We have estimated this model on a longer panel (of about 7-12 years) for two countries, India and Slovenia, and find that the AR(1) coefficient is stable over time. We refer to De Loecker (2007) and Goldberg, Khandelwal, Pavcnik, and Topalova (2009) for more details on the underlying data sources.

¹⁷We use an uncentered measure of fit, as our model does not incorporate anything analogous to the estimated constant commonly found in a regression specification.

Figure 3 shows the model’s prediction for each country, plotted against the measure of the variable of interest found in the data.¹⁸ Circles indicate countries, and circle size is proportional to the number of establishments per country. Each country is plotted using a (x, y) coordinate, where the x-coordinate indicates the model’s prediction and the y-coordinate indicates the value in the data. The closer the country lies to the 45° line, the more accurate the model’s prediction.

Panel 1 shows the dispersion in the static marginal revenue product of capital ($\text{Std.}[s_{ict} - k_{ict}]$) for each country. The S^2 of the model is 0.61. This indicates a good fit, with 61 percent of the observed static capital misallocation captured by the model’s prediction. Of some interest is the fact that the model over-predicts noticeably for a series of countries with model predictions greater than 2. These are (from right to left) Moldova, Guyana, Costa Rica, Madagascar and Chile.

Panel 2 shows the dispersion in productivity ($\text{Std.}[y_{ict}]$) for each country. The S^2 of the model is 0.69. Again, this indicates a good fit, with 69 percent of the observed productivity dispersion captured by the model’s prediction. As in Panel 1, the model over predicts noticeably for some countries. The three countries with model predictions greater than 2 are (from right to left) Guyana, Moldova and Costa Rica. Note that the model’s misprediction in Panel 2 is a test of the AR(1) specification of the productivity process, rather than a failure of the economic model, per se. In particular, from equation (12), the model prediction is invariant to the level of adjustment cost.¹⁹

Panel 3 shows the volatility in the marginal revenue product of capital over time for each country ($\text{Std.}[(s_{ict} - k_{ict}) - (s_{ict-1} - k_{ict-1})]$). The S^2 of the model is 0.28. The data lie on a line below the 45° line. As discussed below (and reported in table 8), if the adjustment costs are set to zero this line rotates up to close to the 45° line.

Panels 4 and 5 show the the volatility in establishments’ capital over time ($\text{Std.}[k_{ict} - k_{ict-1}]$) and the proportion of establishments with zero investment over a single year, respectively. The S^2 ’s are 0.74 and 0.82, respectively.

In table 8 we summarize the fit of the full model (illustrated in figure 3). We compare it to two alternative models: one in which the adjustment costs are set to zero; and another in which the ρ_c and σ_c terms are set to 0.83 and 0.6 (resp.) for all countries (i.e. we use specification I in table 6 for the AR(1) process).

Table 8 indicates that a zero adjustment cost model does a poor job of capturing the variation in the change in capital, but improves on the fit of variation in the change in static MRPK. It does comparably to the full model in fitting the dispersion moments (panels 1 and 2), which reflects under-prediction that matches the over-prediction of the full model. Taken together, these results suggest improving model fit by calibrating the adjustment cost using some combination of the moments examined in panels 1 through 5. Table 8 suggests that the outcome of such an exercise would be adjustment costs that decrease relative to those in Bloom 2009, and (by construction) somewhat improved model fit.²⁰

Table 8 also shows that eliminating country heterogeneity in ρ_c and σ_c improves fit on

¹⁸El Salvador is excluded since its estimated ρ_c exceeds 1. This means that producing a stationary distribution in the simulation is not feasible.

¹⁹Because the variables of interest here are standard deviations of logs, it is useful to direct the reader to the range of the underlying economic activity covered by the data, as noted in our discussion of Table 1 in section 2

²⁰The likely literal implication, that developing countries have lower adjustment costs than the US is almost surely misleading. Rather, such a finding would probably reflect heterogeneity in the capital stock used in different countries.

the dispersion moments (panels 1 and 2) but decreases fit on the change moments (panels 3 and 4). Whether the specification with, or without, country heterogeneity in ρ_c and σ_c is preferred, the conclusion is the same: taking a dynamic investment model calibrated to US data, and replacing the productivity process parameters with estimates from developing country data, generates model predictions that capture an economically significant amount of the dispersion in the static MRPK and productivity, in addition to several other moments, in these developing countries. Thus, the economic structure, suggested by the reduced form data work in section 2 and developed in the model in this section, seems to have value in organizing the data.

3.4 Discussion

The model does a surprisingly good job of fitting the data, despite the fact that all parameters other than the ρ_c and σ_c terms in the AR(1) productivity process are taken from studies fitting U.S. data. This has a series of implications for our thinking about resource allocation, productivity differences across countries, and the welfare and policy implications therein.

First, the dynamic process underlying productivity is important in explaining cross-sectional patterns of productivity and capital allocation. The importance of this is found in several places in the existing literature. Our contribution is to demonstrate the empirical importance of this mechanism. Our findings suggest that the dynamic process governing productivity shocks is a first-order determinant of differences in productivity and, hence, income across countries.

Second, the conclusions one draws regarding welfare and policy depend on the model one has in mind regarding this dynamic process. If one has the view that the productivity process is an exogenous, or primitive, feature of the model, then our findings suggest that, in an aggregate sense, the establishments in the countries we studied are acting much as the social planner in our model would have them act (assuming that the social planner takes the capital adjustment costs as given). This suggests that there are few welfare implications for differences in productivity dispersion and static misallocation across countries. On the other hand, if the productivity process can be affected by government policy, then there may be welfare dividends to policy interventions aimed at moving toward some socially-optimal productivity process. However, characterization of what this optimal process is likely requires a more subtle modeling approach than that offered here. At any rate, this paper sharpens the debate on the role of policy interventions that are geared at eliminating resource misallocation in developing countries by offering an alternative view that is strongly supported by the data.

4 Conclusions and future work

We have focused on the adjustment costs in capital, coupled with productivity shock processes, in order to reinterpret the large dispersion in marginal revenue product of capital. In doing so, we shut down many other economically relevant features of firms' environment, including, for instance, the heterogeneity in market power. This keeps our model parsimonious and makes the approach in this paper directly comparable with the approach taken in the existant literature on cross-country productivity differences. A natural, alternative,

starting point would be to include additional heterogeneity in market power and interpret the differences in marginal revenue product differently, i.e. as a reflection of differences in market power which vary over time. See De Loecker and Warzynski (2009) for a detailed illustration of this approach.²¹ We note this to underscore the fact that observed productivity differences can have many underlying drivers. We focus on just one.

The primary contribution of this paper is to establish the link between the dynamic process governing productivity changes over time and cross-sectional measures of productivity dispersion and (static) capital misallocation. In particular, we show that a parsimonious model of the country-specific productivity process explains much of the variation in the dispersion of productivity and the dispersion of the static marginal revenue product of capital across countries. We provide evidence to support the claim that the dynamic process of productivity is important, both empirically and theoretically, in determining the patterns observed in the cross-section.

Our findings reinforce the point that the country-specific stochastic process of productivity is sufficient to explain a significant proportion of cross-country variation in productivity and static capital misallocation. It suggests that producers in countries that experience larger uncertainty in the operating environment (i.e., higher volatility in productivity) make different investment decisions than those producers active in less volatile environments. This leads to different levels of capital and output and, moreover, means that the welfare gains from policies inducing reallocation of factors of production are likely to be lower than otherwise implied by static models.

An alternative suite of policy options, aimed at making the productivity process more benign, may be attractive as a complement to the redistributive measures featured in the counterfactuals considered in other studies. It is likely that at least some component of the stochastic process of productivity is influenced by government policy. To the extent that this is true, our findings imply that, if government policies can provide a more predictable business environment, then this will benefit the economy and help producers allocate resources in more productive ways. This raises the issue of the sources of adjustment costs and productivity volatility, a topic on which we are silent in this paper. Our aim here is to merely cast light on the importance of dynamics in assessing the welfare relevance of productivity dispersion and in evaluating an appropriate policy response.

²¹De Loecker and Warzynski (2009) provide a way to obtain producer-level markups using standard production data, while allowing explicitly for dynamic inputs of production, such as capital. Pairing their approach with our framework could, in principal, allow for a decomposition of “static” and “dynamic” components of the perceived misallocation from the standard model. This we leave for future work.

A Appendix

In this Appendix we briefly discuss the data and elaborate on how we measure TFP in our data. It is important to stress that we measure TFP such that it is consistent with the theoretical model introduced in the paper. This implies that we need to explicitly allow for adjustment costs in capital and directly rely on monopolistic competition whereby markups are common across establishments.

We then examine the robustness of our central results to different productivity measure and different sources of identification. Finally, we discuss the measurement of the AR(1) process used in the dynamic investment model.

A.1 Data

The data is available from <http://www.enterprisesurveys.org>, accessed on December 15th, 2010. Extensive documentation is available from the same website.

The survey documentation describes the sampling universe as follows: “6. The population of industries to be included in the Enterprise Surveys and Indicator Surveys, the Universe of the study, includes the following list (according to ISIC, revision 3.1): all manufacturing sectors (group D), construction (group F), services (groups G and H), transport, storage, and communications (group I), and subsector 72 (from Group K). Also, to limit the surveys to the formal economy the sample frame for each country should include only establishments with five (5) or more employees. Fully government owned establishments are excluded as the Universe is defined as the non-agricultural private sector.” from page 3 in ‘Enterprise Survey and Indicator Surveys Sampling Methodology’ August 29th, 2009 at http://www.enterprisesurveys.org/Documents/Sampling_Note.pdf downloaded 23 April, 2011.

The survey used a stratified sampling procedure, in which establishments were sampled randomly within groups based on the establishment’s sector of activity, establishment size, and geographical location. The structure of the sampling leads to an oversampling of larger establishments (relative to random sampling of all establishments in the economy). The exact structure of the stratification varies by size of the economy in question. We have chosen to not do any sampling correction, preferring to maintain as much transparency as possible as to the mapping from data to findings, being mindful of the fact that we can use data from only 7% of the sampled establishments in any case (and there may also be non-random selection into this subset) and, most importantly, considering the absence of a well defined criterion that could be used to guide any such correction. It is an open question whether ideally we would weight by, say, contribution to GDP or weight each establishment equally. More likely, weighting by some measure of activity makes more sense for our purpose, but to the extent that, say, any Eritrean government statistics we would use to do this would have measurement error, this may merely contribute to attenuation bias in results. This is especially so, given that the sampling structure used in the surveys overweights large establishments and, hence, already moves in the direction of weighting by contribution to economic activity. In any case, the results in the paper are robust to controlling for differences in the size and industrial composition of establishments across countries.

The establishments in the data are drawn from the manufacturing, construction, services, and transport, storage, and communications sectors. As would be expected, the precise industry composition (defined at the 2 digit ISIC level) varies by country.

Establishments were surveyed between 2002 and 2006. The majority of establishments within a country were surveyed in the same year. The survey asked questions about activity in the current year and the previous 2 years. Thus, the panel data aspect of these data, relating to activity in year $t - 1$, come from the recollections and records of managers in year t .

A.2 Measurement

To measure TFP we require a measure of plant-level sales (Q_{ict}), employment (L_{ict}) and capital services (K_{ict}). We also rely on intermediate inputs (M_{ict}) in some of our TFP specifications, i.e. for TFP based on gross output and value added. We follow the standard practice and refer to Bartelsman, Haltiwanger, and Scarpetta (2009) for a very nice overview and discussion on the measurement of TFP and related variables using similar data sources.

A number of issues emerged in the calculation of labor and TFP, including that labor input is generally based on the number of employees or wage bill converted to the number of employees with no correction for hours worked. Moreover, sales and gross output data are not corrected for inventories, and the capital stock is based on book values. These are the standard data restrictions researchers face using this type of data.

We rely on the World Bank deflators to convert all monetary variables into USD. Sales is directly measured in the data, whereas labor is measured by the total number of workers active in a plant, or alternatively we convert the total wage bill of a plant into the number of workers using a plant-specific wage. The latter is corrected for aggregate wage trends using the median wage trend in a given industry-country pair. Finally, we rely on the book value of capital as measured by either total assets or net book value. We experimented with both measures and our results are invariant. When we consider a measure of value added we compute it by netting the sales variable from the use of intermediate inputs. The selection of these variables leaves us with 5,010 observations across countries and time.

Our base results rely on TFP measured as described in section 2.2.1. Relying on cost minimizing producers generates the well known FOC for labor, an input assumed to be freely chosen by the plant without adjustment cost. Thus,

$$\beta_L = \left(1 - \frac{1}{\epsilon}\right) \frac{wL}{S} \quad (15)$$

where we drop the relevant subscripts.²² Standard cost minimization dictates that the production function coefficients of variable inputs in production are equal to the expenditure share of the relevant input in total costs. We therefore directly measure an industry-country-time specific cost share for labor where we again rely on $\epsilon = 4$. Our assumption of constant returns to scale in production then gives us the capital coefficient by simply considering $1 - \beta_L$. Note that we hereby do not impose the static FOC on capital. We take the median labor share in a given industry-country at a given point in time, or

$$\beta_L^{sct} = \text{median}(\beta_L^{isct}) \quad (16)$$

²²Our framework can easily allow labor to be treated just like capital, and we would then rely on a gross output production function from the outset whereby we can solve for optimal intermediate input demand given input prices, productivity, labor and capital. However, given that we want to be able to compare our results directly to the existing literature we focus on a production function in labor and capital.

For many establishment-years in the data we can compute productivity straight from data. However, for some establishment-years we only observe the establishment's wage bill and not the number of workers. To address this issue we use the median country-industry wage, \tilde{w} , (imputed from observations with both the wage bill and the number of workers) as a deflator and apply it to the wage bill to give a measure of labor. That is, to compute L_{ict} we use $L_{ict} = \frac{wL_{ict}}{\tilde{w}}$. In what is presented in this paper we use this measure for all establishment-year observations.

Total factor productivity at the plant level is the log of deflated sales minus the weighted log of labor and capital, where the weights are the cost shares defined above.

$$y_{ict} = s_{ict} - \beta_L^{sct} l_{ict} - \beta_K^{sct} k_{ict} \quad (17)$$

Our alternative measures of TFP, using value added (y^{VA}) or gross output (y^{GO}), are obtained similarly:

$$y_{ict}^{VA} = va_{ict} - \beta_L^{sct} l_{ict} - \beta_K^{sct} k_{ict} \quad (18)$$

$$y_{ict}^{GO} = s_{ict} - \beta_L^{sct} l_{ict} - \beta_M^{sct} M_{ict} - \beta_K^{sct} k_{ict} \quad (19)$$

In the following subsection we show that our results are robust to use of these alternative measures of TFP. The trade-off in a cross-country analysis of plant-level productivity is between sample size and having a set of variables measured in a consistent and harmonized fashion across time and countries. Our three different measures of TFP are computed in order to verify the sensitivity of our main results to those underlying data issues.

A.3 Robustness of the productivity dispersion, static misallocation and volatility relationship

Tables 9 and 10 examine the robustness of specification I in tables 3 and 4 to alternate measures of TFP. The first column (TFP Leontief) reports the results from tables 3 and 4, while the second and third columns report results using gross output and value added measures of TFP respectively. As can be seen, little changes as the TFP measure changes: the productivity dispersion coefficient gets larger if gross output or value added is used; while the dispersion in the marginal revenue product of capital (misallocation) coefficient gets smaller if gross output or value added is used. In all cases the relationship remains statistically significant and the R-squared statistic is either comparable or larger as compared to using the TFP Leontief specification.

Table 11 examines the use of different measures of dispersion. In specification I we replicate specification I in table 3. In specification II we use the range between the 10th and 90th percentile as our measure of cross-sectional dispersion (in all specifications we use the standard error of the change in productivity as our measure of volatility: this will be changed in future version of the paper). In specification III we use the inter-quartile range. Specification III ceases to generate a significant relationship between dispersion and volatility.

Lastly, Table 12 runs regressions in which the unit of observation is a country-industry pair. In Columns III and IV, we include both industry and country fixed-effects (but not their interactions). This enables us to control for either industry or country specific factors that might artificially generate the volatility-dispersion relationship. For instance, in different countries it may be more or less straightforward to measure productivity. Including country

fixed-effects bars these country specific factors from explaining the dispersion-volatility relationship. We find lower coefficients in this regression, 0.41 without any controls and 0.34 when country and industry fixed-effects are added. These effects are precisely estimated, in part because there are far more country-industry pairs (236) than countries (33).

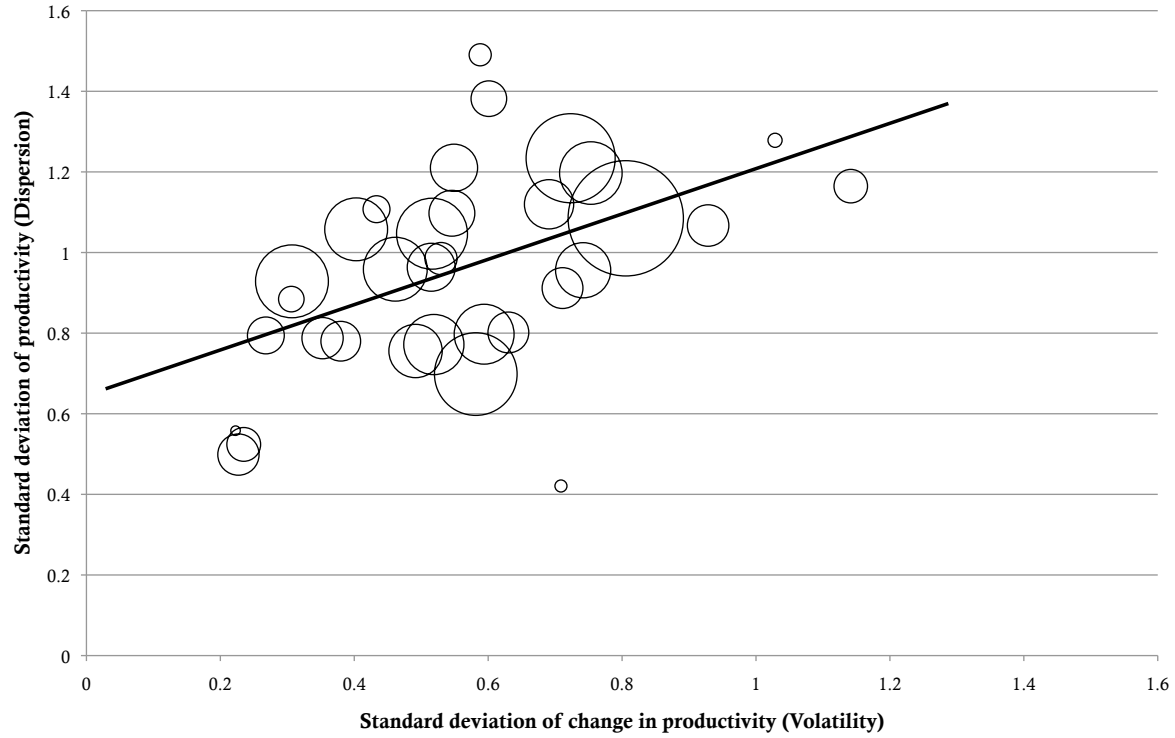
References

- BARTELSMAN, E., AND M. DOMS (2000): "Understanding productivity: Lessons from longitudinal microdata," *Journal of Economic Literature*, 38(3), 569–594.
- BARTELSMAN, E., J. HALTIWANGER, AND S. SCARPETTA (2009): "Cross-country differences in productivity: the role of allocation and selection," Discussion paper, NBER Working Paper 15490.
- BLOOM, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77(3), 623–685.
- CABALLERO, R., AND R. PINDYCK (1996): "Uncertainty, investment, and industry evolution," *International Economic Review*, 37(3), 641–662.
- COLLARD-WEXLER, A. (2009): "Productivity Dispersion and Plant Selection in the Ready-Mix Concrete Industry," *Working Paper, New York University*.
- COOPER, R., AND J. HALTIWANGER (2006): "On the nature of capital adjustment costs," *Review of Economic Studies*, 73(3), 611–633.
- DE LOECKER, J. (2007): "Do exports generate higher productivity? Evidence from Slovenia," *Journal of International Economics*, 73(1), 69–98.
- DE LOECKER, J., AND F. WARZYNSKI (2009): "Markups and firm-level export status," Discussion paper, NBER Working Paper 15198.
- DIXIT, A., AND R. PINDYCK (1994): *Investment Under Uncertainty*. Princeton University Press.
- FOSTER, L., J. HALTIWANGER, AND C. SYVERSON (2008): "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?," *American Economic Review*, 98(1), 394–425.
- GOLDBERG, P., A. KHANDELWAL, N. PAVCNIK, AND P. TOPALOVA (2009): "Trade liberalization and new imported inputs," *American Economic Review*, 99(2), 494–500.
- HSIEH, C.-T., AND P. J. KLENOW (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124(4), 1403–1448.
- LEVINSOHN, J., AND A. PETRIN (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70(2), 317–341.
- MIDRIGAN, V., AND D. Y. XU (2009): "Finance and Misallocation," Discussion paper, NYU.

- MOLL, B. (2010): “Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?,” Discussion paper, Princeton University.
- OLLEY, G. S., AND A. PAKES (1996): “The dynamics of productivity in the telecommunications equipment industry,” *Econometrica*, 64(6), 1263–1297.
- RESTUCCIA, D., AND R. ROGERSON (2008): “Policy distortions and aggregate productivity with heterogeneous establishments,” *Review of Economic Dynamics*, 11(4), 707 – 720.
- TAUCHEN, G. (1986): “Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions,” *Economics Letters*, 20(2), 177–81.

Tables and Figures

Figure 1: TFP Dispersion and TFP Volatility



Note: Circles indicate countries. Circle Size is proportional to the number of establishments per country. The straight line is the line-of-best-fit (computed using OLS with a constant term, as per specification I in Table 3). The horizontal axis indicates the value of the standard deviation of $[y_{ict} - y_{ict-1}]$ and the vertical axis indicates the standard deviation in y_{ict} , where $y_{ict} = \ln(Y_{ict})$, and Y_{ict} is defined as in equation (5).

Table 1: List of countries, with TFP dispersion and number of observed establishments

Region	Country	Standard Deviation of TFP	Establishments
North Africa			
	Morocco	0.70	377
Sub-Saharan Africa			
	Eritrea	0.42	8
	Zambia	0.76	157
	Madagascar	0.80	92
	South Africa	0.80	198
	Malawi	0.96	127
	Mauritius	1.16	62
	Ethiopia	1.20	216
	Tanzania	1.38	70
Central Asia			
	Kyrgyzstan	0.50	94
	Tajikistan	0.79	94
	Uzbekistan	0.91	92
Middle East			
	Oman	0.56	5
	Syria	0.99	57
South Asia			
	Bangladesh	1.12	134
	Sri Lanka	1.21	123
South East Asia			
	Philippines	0.93	293
	Indonesia	1.23	438
	Cambodia	1.28	11
Central America			
	El Salvador	0.77	200
	Guatemala	0.96	168
	Nicaragua	0.96	225
	Costa Rica	1.05	278
	Honduras	1.06	218
	Ecuador	1.10	116
South America			
	Brazil	0.78	87
	Peru	0.88	36
	Chile	1.08	734
	Guyana	1.49	27
Europe			
	Poland	0.52	63
	Moldova	0.79	75
	Lithuania	1.07	95
	Turkey	1.11	40

Table 2: Summary Statistics: World Bank Enterprise Research Data.

Establishment Level Data	Mean	Std. Dev.	Min.	Max.	N
Log Sales	7.1	3.1	-5.6	22.3	5010
Log Value Added	6.2	3.2	-8.3	21.7	4191
Log Materials	6.4	3.3	-7.2	21.5	4933
Log Capital	7	3.1	-6.5	22.5	5010
Log Wage Bill	5.3	3	-5.9	18.9	4096
Workers	261.3	840.1	1	23385	5010
Productivity	0.7	1.1	-5.8	5.9	5010
Log Investment	7.9	4	-5.1	25.4	2389
Zero Investment	0.5	0.5	0	1	5010
Change in Sales	0.1	0.6	-7.6	7.1	5010
Change in Value Added	0.1	0.7	-4.9	5.9	4021
Change in Capital	0.1	0.5	-4.4	8.7	5010
Change in Labor	0.2	0.7	-4.9	7.1	3991
Change in Productivity	0	0.6	-7.9	7.2	5010
Country Level Data					
Establishments per Country	151.8	147	5	734	33
Standard Deviation of Productivity (Dispersion)	0.9	0.3	0.4	1.5	33
Standard Deviation of Change in Productivity (Volatility)	0.6	0.2	0.2	1.1	33
Interquartile Range Deviation of Productivity	1	0.3	0	1.7	33
Interquartile Range of Change in Productivity	0.3	0.1	0.1	0.7	33
90-10 Range of Productivity	2.1	0.5	1.1	2.8	33
Country-Industry Level Data					
Establishments per Country-Industry	19	25.4	1	149	236
Standard Deviation of Productivity	0.9	0.5	0	3.8	236
Standard Deviation of Change in Productivity	0.5	0.4	0	2	236

Note: The data were drawn from the World Bank Enterprise Research Data. To be included in the final data set an establishment needed to have at least two years of information on sales, materials, assets, and salaries. We exclude establishments with productivity (y_{ict}) greater than 6 in absolute value to remove the effect of outliers. The results reported in the paper are qualitatively unchanged if the threshold on y_{ict} is set to be 2 or 9.

Table 3: Effect of time-series volatility on productivity dispersion

Specification	I	II (unweighted)	III	IV
Dependent Var:	Standard Deviation of y_{ict} , by country			
Std. $[y_{ict} - y_{ict-1}]$	0.57*** (0.20)	0.58*** (0.23)	0.56*** (0.19)	0.58*** (0.17)
Log Assets ($t - 1$)				-0.01 (0.01)
Constant	0.65*** (0.11)	0.63*** (0.13)	0.65*** (0.11)	0.67*** (0.10)
Industry FE			X	X
Establishments	5010	5010	5010	5010
Countries	33	33	33	33
R-Squared	.26	.24	.31	.32

Note: Column I and II run regressions on country level aggregates. Column I runs a weighted OLS with weights equal to the number of establishments per country, whereas Column II has equal weights for each country. Columns III and IV run regressions at the establishment level (where Std. y_{ict} and Std. $[y_{ict} - y_{ict-1}]$ only vary at the country level). Sampling error is accounted for using a bootstrap procedure where Std. y_{ict} and Std. $[y_{ict} - y_{ict-1}]$ are recomputed for each bootstrap replication (200 bootstrap replications are used). These standard errors are clustered by country by having the bootstrap resample countries rather than individual establishments.

Table 4: Effect of time-series volatility on static misallocation

Specification	I	II (unweighted)	III	IV
Dependent Var:	Standard Deviation of $s_{ict} - k_{ict}$, by country			
Std. $[y_{ict} - y_{ict-1}]$	0.65*** (0.22)	0.74*** (0.22)	0.63*** (0.23)	0.66** (0.22)
Log Assets ($t - 1$)				-0.01 (0.01)
Constant	0.70*** (0.12)	0.65*** (0.13)	0.70*** (0.13)	0.74*** (0.13)
Industry FE			X	X
Establishments	5010	5010	5010	5010
Countries	33	33	33	33
R-Squared	.23	.33	.29	.30

Note: Column I and II run regressions on country level aggregates. Column I runs a weighted OLS with weights equal to the number of establishments per country, whereas Column II has equal weights for each country. Columns III and IV run regressions at the establishment level (where Std. $s_{ict} - k_{ict}$ and Std. $[y_{ict} - y_{ict-1}]$ only vary at the country level). Sampling error is accounted for using a bootstrap procedure where Std. y_{ict} and Std. $[y_{ict} - y_{ict-1}]$ are recomputed for each bootstrap replication (200 bootstrap replications are used). These standard errors are clustered by country by having the bootstrap resample countries rather than individual establishments.

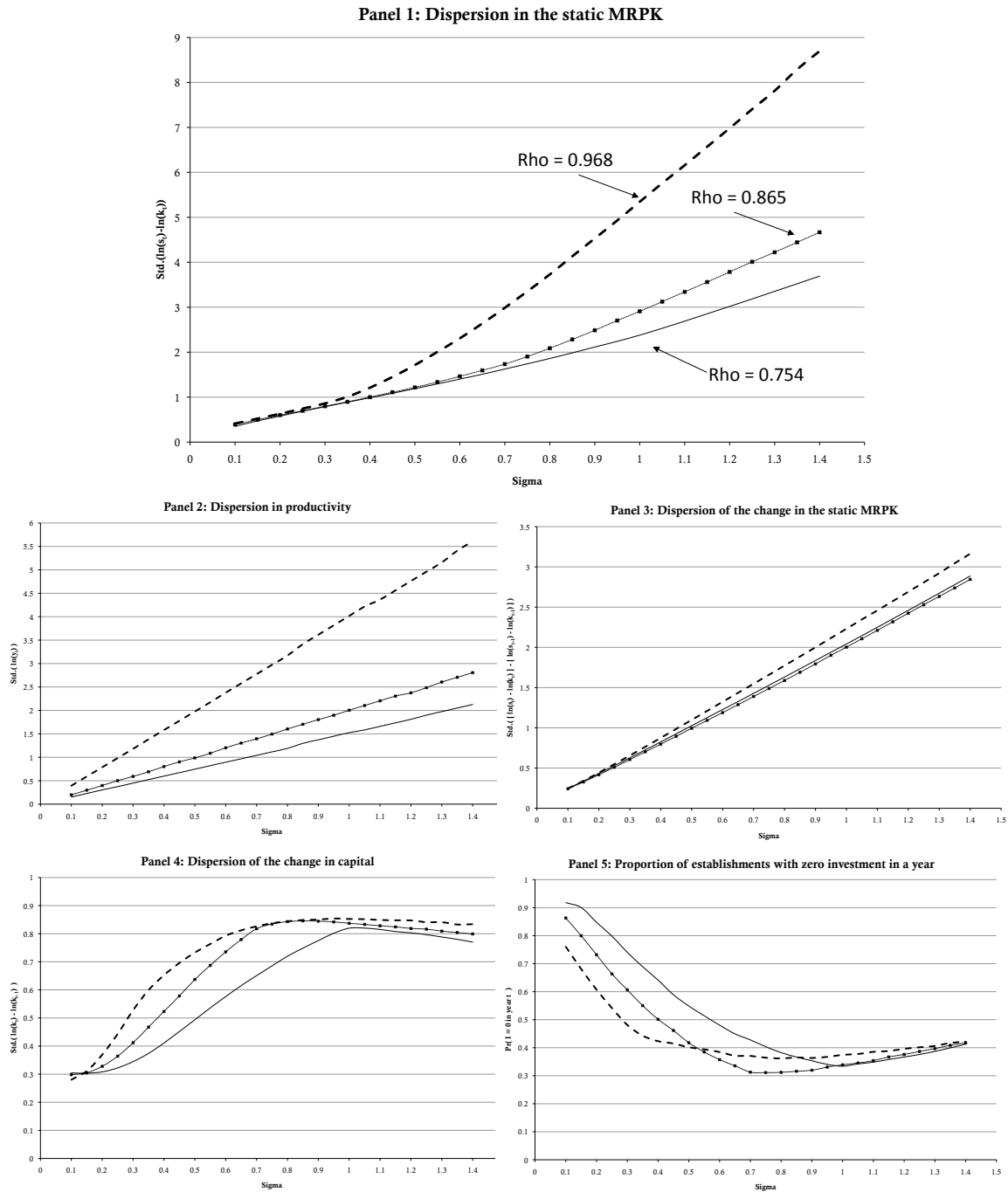
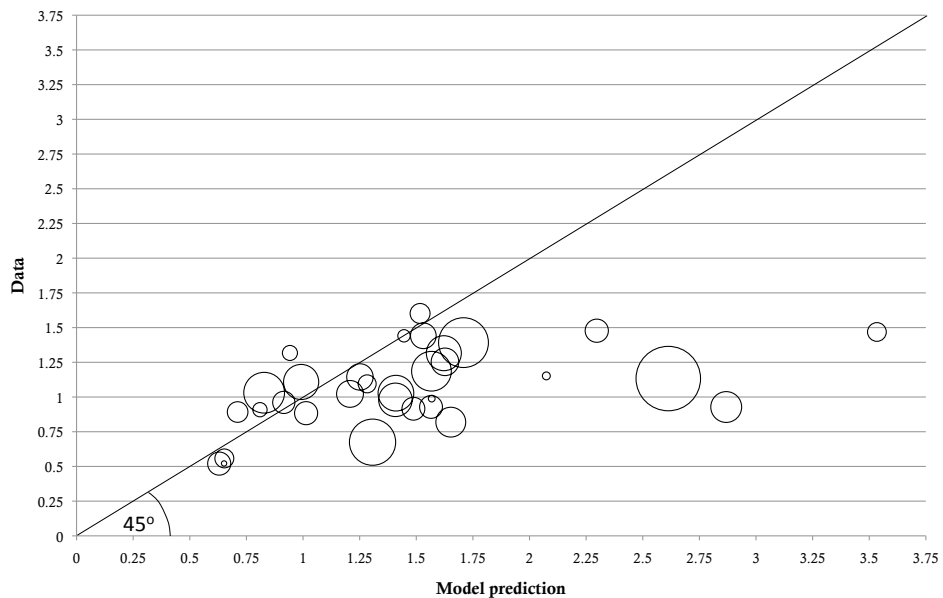


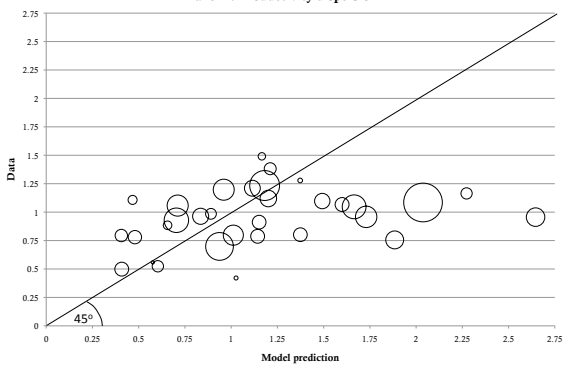
Figure 2: Model comparative statics, by σ_c [Sigma] (and ρ_c [Rho])

Note: This is output from the dynamic investment model, parametrized as per Table 5.

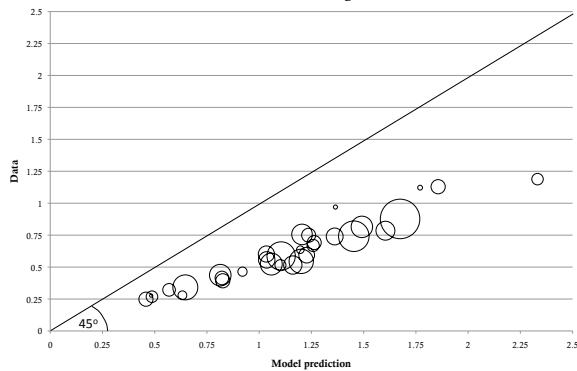
Panel 1: Dispersion in the static MRPK



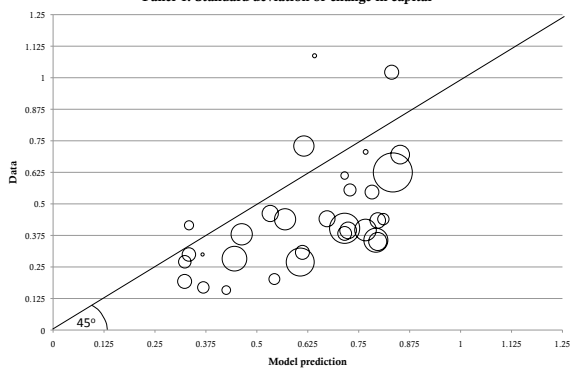
Panel 2: Productivity dispersion



Panel 3: Standard deviation of change in static MRPK



Panel 4: Standard deviation of change in capital



Panel 5: Proportion of establishments with zero investment

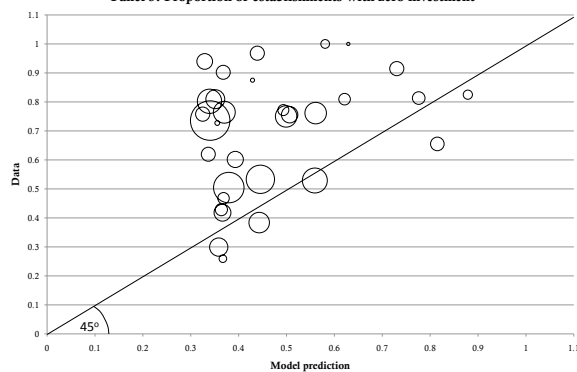


Figure 3: Correspondence between model prediction and data

Note: Circles indicate countries. Circle Size is proportional to the number of establishments per country. Each country is plotted using a (x, y) coordinate where the x-coordinate indicates the model prediction and the y-coordinate indicates the value in the data. The closer the country lies to the 45° line, the better the model prediction.

Table 5: Baseline parameters

Parameter	Comments
$\alpha = \frac{1}{3}$ $\epsilon = -4$ $\delta = 10\%$ $\beta = \frac{1}{1+6.5\%}$ $C_K^P = 42.7\%$ $C_K^F = 1.1\%$ $C_K^Q = 0.996$	<p>} Values taken from Bloom (2009).</p>
$\rho_c \in \{0.754, 0.865, 0.968\}$ $\mu = 0.13$ $\sigma_c \in [0.1, 1.4]$	<p>} Values taken from specification V in Table 6. Values for ρ_c correspond to values for Moldova, Tanzania and Nicaragua respectively.</p>
$w = 0.3$	<p>Scaling parameter selected to ensure period profits are positive.</p>

Table 6: Time series process, AR(1), for productivity

Dependent Var: Productivity y_{it}	I	II	III	IV	V
y_{it-1}	0.83*** (0.03)	0.88*** (0.04)	0.87*** (0.02)	0.90*** (0.04)	0.84*** (0.01)
$(y_{it-1}) \bullet$ (Country Dummy) Var.					X 0.12
y_{it-1}^2		0.02 (0.02)		0.02 (0.02)	
y_{it-1}^3		-0.01 (0.00)		-0.01 (0.00)	
y_{it-1}^4		-0.00 (0.00)		-0.00 (0.00)	
Constant	0.15*** (0.04)	0.12*** (0.03)	0.13*** (0.03)	0.10*** (0.03)	0.13*** (0.03)
Variance σ					
Constant	0.60*** (0.04)	0.54*** (0.04)	0.66*** (0.00)	0.66*** (0.01)	0.66*** (0.00)
Country Specific Variance Var.			X 0.21	X 0.21	X 0.20
Log Assets		0.01 (0.01)			
Observations	5010	5010	5010	5010	5010
Countries	33	33	33	33	33
Log-Likelihood	-4515	-4468	-4055	-4026	-3985

Note: Standard Errors (in parentheses) clustered by country. ‘Var.’ indicates the variance of the set of parameters indicated in the row above.

Table 7: Country-specific σ_c 's and ρ_c 's, from Table 6

Country	Specification			
	III (σ_c)	IV (σ_c)	V (σ_c)	V (ρ_c)
Bangladesh	0.66	0.66	0.66	0.84
Brazil	0.40	0.40	0.38	0.63
Cambodia	0.91	0.91	0.85	0.79
Chile	0.76	0.73	0.75	0.94
CostaRica	0.54	0.53	0.53	0.95
Ecuador	0.58	0.62	0.56	0.92
El Salvador	0.51	0.50	0.51	1.23
Eritrea	0.67	0.67	0.66	0.78
Ethiopia	0.71	0.72	0.71	0.70
Guatemala	0.68	0.67	0.65	0.97
Guyana	0.60	0.61	0.58	0.87
Honduras	0.39	0.40	0.39	0.83
Indonesia	0.71	0.71	0.70	0.81
Kyrgyzstan	0.22	0.22	0.22	0.85
Lithuania	0.90	0.90	0.90	0.82
Madagascar	0.60	0.60	0.60	0.90
Malawi	0.50	0.50	0.50	0.80
Mauritius	1.11	1.12	1.10	0.87
Moldova	0.27	0.26	0.27	0.75
Morocco	0.55	0.55	0.54	0.83
Nicaragua	0.43	0.42	0.43	0.97
Oman	0.24	0.24	0.22	0.92
Peru	0.31	0.30	0.30	0.89
Philippines	0.30	0.31	0.30	0.90
Poland	0.23	0.22	0.22	0.93
South Africa	0.59	0.59	0.58	0.81
Sri Lanka	0.50	0.49	0.49	0.90
Syria	0.53	0.52	0.53	0.79
Tajikistan	0.36	0.35	0.35	0.96
Tanzania	0.62	0.62	0.61	0.86
Turkey	0.41	0.42	0.41	0.53
Uzbekistan	0.69	0.69	0.60	0.86
Zambia	0.47	0.47	0.47	0.97
Mean	0.57	0.57	0.56	0.85

Note: The AR(1) specification is:

$$y_{ict} = \mu + \rho_c y_{ict-1} + \sigma_c \eta_{ict}$$

Table 8: Model fit, reported as S^2 , of different specifications

	Full Model	Zero Adj. Costs	$(\rho_c, \sigma_c) = (\rho, \sigma)$
Panel 1: Dispersion in static MRPK	0.610	0.597	0.827
Panel 2: Productivity dispersion	0.686	0.682	0.920
Panel 3: Std. of change in static MRPK	0.284	0.938	0.107
Panel 4: Std. of change in capital	0.735	-10.715	0.597
Panel 5: Establishments with zero investment	0.821	0.304	0.746

Notes: “Full Model” has parameters as described in table 5 and (ρ_c, σ_c) as estimated in specification V of table 6. “Zero Adj. Costs” sets the adjustment cost parameters equal to zero. “ $(\rho_c, \sigma_c) = (\rho, \sigma)$ ” imposes the same productivity process on every country, drawn from specification I of Table 6.

Table 9: Appendix: Robustness checks: Productivity dispersion on volatility for different TFP measures

TFP measure: Dependent Var:	TFP Leontief Std. y_{ict}	TFP GO Std. y_{ict}	TFP VA Std. y_{ict}
Std. $[y_{ict} - y_{ict-1}]$	0.57*** (0.20)	0.94*** (0.09)	1.00*** (0.16)
Constant	0.65*** (0.13)	0.22*** (0.06)	0.39*** (0.11)
Observations	5010	4831	4571
Countries	33	33	33
R^2	0.29	0.78	0.57

Note: All columns show regressions on country level aggregates weighted by the number of establishments per country. Standard errors are clustered at the country level but do not account for sampling error in Std. y_{ict} and Std. $[y_{ict} - y_{ict-1}]$ for specifications other than TFP Leontief.

Table 10: Appendix: Robustness checks: Misallocation on volatility for different TFP measures

TFP measure:	TFP Leontief	TFP GO	TFP VA
Dependent Var:	Std. $[s_{ict} - k_{ict}]$	Std. $[s_{ict} - k_{ict}]$	Std. $[s_{ict} - k_{ict}]$
Std. $[y_{ict} - y_{ict-1}]$	0.65*** (0.22)	0.38*** (0.14)	0.59*** (0.18)
Constant	0.70*** (0.12)	0.85*** (0.09)	0.66*** (0.13)
Observations	5010	4831	4571
Countries	33	33	33
R^2	0.26	0.19	0.26

Note: All columns show regressions on country level aggregates weighted by the number of establishments per country. Standard errors are clustered at the country level but do not account for sampling error in Std. y_{ict} and Std. $[y_{ict} - y_{ict-1}]$ for specifications other than TFP Leontief.

Table 11: Appendix: Robustness checks: Different measures of productivity dispersion

Specification	I	II	III
Dependent Var:	Std. y_{ict}	90-10 percentile y_{ict}	IQR y_{ict}
Std. $[y_{ict} - y_{ict-1}]$	0.57*** (0.20)	0.69* (0.41)	0.22 (0.18)
Constant	0.65*** (0.11)	1.72*** (0.21)	0.93*** (0.11)
Observations	5010	5010	5010
Countries	33	33	33
R-Squared	.26	.08	.02

Note: All columns show regressions on country level aggregates weighted by the number of establishments per country. Sampling error is accounted for using a bootstrap procedure where the dependent variable and Std. $[y_{ict} - y_{ict-1}]$ are recomputed for each bootstrap replication (200 bootstrap replications are used). All TFP measures use TFP Leontief.

Table 12: Appendix: Industry-country productivity dispersion and productivity volatility

Specification	I	II	III	IV
Dependent Var:	Standard Deviation of y_{ict} , by industry-country			
Std. $[y_{ict} - y_{ict-1}]$	0.41*** (0.08)	0.41*** (0.07)	0.34*** (0.09)	0.34*** (0.09)
Log Capital		-0.00 (0.00)		0.00 (0.00)
Productivity		-0.03*** (0.01)		-0.01* (0.00)
Industry FE			X	X
Country FE			X	X
Constant	0.69*** (0.05)	0.71*** (0.05)	0.64*** (0.05)	0.65*** (0.05)
Observations	4983	4983	4983	4983
Country-Industries	236	236	236	236
R-Squared	.12	.13	.55	.55

Note: Standard errors are clustered by industry-country. Sampling error not accounted for. All TFP measures use TFP Leontief.