COLUMBIA | SIPA
Center on Global Economic Governance

CGEG-CDEP WORKING PAPER SERIES

CGEG-CDEP WP No. 51

Resource Misallocation in European Firms: The Role of Constraints, Firm Characteristics and Managerial Decisions

Yuriy Gorodnichenko, Debora Revoltella, Jan Svejnar, and Christoph T. Weiss

June 2019

COLUMBIA | SIPA
Center for Development Economics and Policy

Resource Misallocation in European Firms: The Role of Constraints, Firm Characteristics and Managerial Decisions

Yuriy Gorodnichenko
University of California, Berkeley and NBER

Debora Revoltella
European Investment Bank

Jan Svejnar
Columbia University, CERGE-EI, CEPR and IZA

Christoph T. Weiss
European Investment Bank

This draft: June 6, 2019

Abstract: Using a new survey, we document high dispersion of marginal revenue products across firms in the European Union (EU). To interpret this dispersion, we develop a portable framework to quantify gains from better allocation of resources. We demonstrate that apart from direct measures of distortions, firm characteristics, such as demographics, quality of inputs, utilization of resources, and dynamic adjustment of inputs, are predictors of the marginal revenue products of capital and labor. We emphasize that some firm characteristics may reflect compensating differentials rather than constraints and the effect of constraints on the dispersion of marginal products may hence be smaller than has been assumed in the literature. We show that cross-country differences in the dispersion of marginal products in the EU are largely due to differences in how the business, institutional and policy environment translates firm characteristics into outcomes rather than to the differences in firm characteristics *per se*. Removing distortions could raise EU aggregate productivity by 40 percent or more.

JEL codes: O12, O47, O52, D22, D24

Keywords: Marginal products, resource allocation, firm-specific factors, economic growth.

Acknowledgement: We thank Albert Park and seminar participants in Bristol and at the 2018 ASSA meeting (Philadelphia), the CompNet 7th Annual Conference (Halle) and the conference "Investment in the new monetary and financial environment" (Paris) for helpful comments. In preparing this paper, Yuriy Gorodnichenko and Jan Svejnar worked under a grant from the European Investment Bank. The views expressed in this paper are those of the authors and do not necessarily reflect those of the European Investment Bank.

I. Introduction

Although the large cross-country differentials in income per capita have been the subject of much research, accounting for sources of this dispersion has proven to be difficult. The most important factor appears to be differences in "productivity", which Moses Abramovitz called a measure of our ignorance. In an attempt to explain productivity differences within and across countries, recent research pioneered by Hsieh and Klenow (2009) emphasizes the importance of firm-level misallocation of resources for aggregate economic outcomes. It is based on the insight that if there is a dispersion of marginal revenue products of inputs across firms, the economy may achieve considerable productivity – and hence output – gains by reallocating capital from firms with low marginal revenue product of capital (MRPK) to firms with high MRPK and, similarly, from firms with low marginal revenue product of labor (MRPL) to firms with high MRPL. This concept is reflected in the textbook outcome when cost-minimizing firms face identical input prices in a perfectly functioning spot market economy and MRPK and MRPL are equalized across firms.

The recent slowdown in productivity growth in the United States (US), European Union (EU) and other developed economies has generated a sense of urgency among policymakers and academics to identify impediments to productivity increases and to find ways to spur economic growth. Although a number of explanations has been put forth, rising misallocation of resources in European countries could be one of the culprits (see, e.g., Gopinath et al., 2017).[1]

While existing research has been successful in measuring the dispersion of marginal products and assessing potential gains from better allocation of resources, little is known about *why* firms have different marginal products. To a large extent, the lack of research on this question has been brought about by data limitations. In particular, research in this area typically uses census-type data to calculate MRPK or MRPL for firms in *one* economy. But census-type data usually contain only income statements, balance sheet information about capital, and basic data on employment. As a result, researchers usually do not have complete firm-level information as well as rich, consistently measured cross-country variation to tell why a given dispersion of MRPK or MRPL exists. Furthermore, the lack of exogenous variation in potential explanatory variables limits the scope of possible inferences or requires strong identification assumptions.

[1] Consistent with this view, the dispersion of MRPK and MRPL within individual European countries and within individual economic sectors has been trending up since the mid-1990s according to data in Orbis, a popular source of firm-level data.

In this paper we aim to make progress along several dimensions. First, we use a new, large cross-country survey of firms: the Investment Survey carried out by the European Investment Bank (EIBIS). This survey has been administered annually since 2016 to a stratified random sample of firms in each of the 28 EU countries and it is designed to be representative of different sectors and firm sizes in each EU country. EIBIS contains information about the behavior and constraints of firms – e.g., how firms obtain capital and whether the quantity is sufficient, whether their capital stock is state-of-the-art, and information about capacity utilization, rates of innovation, access to infrastructure, and foreign presence in management. Survey responses are also matched to administrative data of the firm (e.g., balance sheet information). Importantly, the design and implementation of the survey is consistent across countries and sectors, which is critical for understanding cross-country and cross-industry variation in the dispersion of marginal revenue products.

Second, informed by theory, we develop an empirical framework to quantify the contribution of various forces to the dispersion of marginal revenue products across firms and map the contribution to potential productivity gains. In particular, we show that under general conditions one can use marginal R^2 to obtain an *upper bound* for possible gains from removing a friction by estimating equilibrium relationships (optimality conditions) in a regression framework, thus linking the misallocation literature to Mincer (1958) and subsequent work studying earnings inequality. This framework does not rely on exogenous variation in frictions or other predictors of MRPK and MRPL which makes our approach highly portable.

Third, we examine the extent to which the dispersion of marginal products is related to firm-level (as opposed to country- or sector-level) characteristics and we compute associated productivity gains. We note that while the existing literature treats the dispersion of marginal products as reflecting barriers and distortions, this may not always be the case. Some dispersion may reflects optimizing behavior of firms (e.g., compensating differentials in the labor market), in which case it is economically rational from the standpoint of the firms and thus should be interpreted as a "cost" to aggregate productivity. While we cannot always establish which of these phenomena is consistent with the data, we present a range of estimates consistent with various interpretations. Relatedly, we perform the Machado-Mata decomposition to construct counterfactual distributions of MRPK and MRPL for each country on the assumption that it has estimated coefficients or values of explanatory variables from another country (e.g., Greece and Germany). This decomposition exercise allows us to understand better whether observed dispersion in MRPK and MRPL is brought about by cross-country differences in firm characteristics or cross-country differences in how the business, institutional and policy environment guides the allocation of resources across heterogeneous firms.

We document that there is a sizable dispersion of marginal products measured across all the firms in our sample. Our estimates indicate that in terms of labor allocation firms are more segmented across countries than industries, as seen in the fact that differences in the levels of MRPL are higher across countries than across industries. The opposite is true for capital. This suggests that national regulations and language barriers could play an important part in the efficiency of resource allocation within the EU, particularly when labor is concerned. When we exploit detailed firm-level information in EIBIS, we find that the significant association between marginal products and firm characteristics is predominantly driven by variables measuring firm demographics, quality of inputs, utilization of resources, and dynamic adjustment of inputs. In contrast, the contribution of direct measures of "barriers and constraints" to cross-sectional variation in MRPK and MRPL seems to be modest. Using the Machado-Mata decomposition we document that cross-country variation in the within-country dispersion of marginal revenue products is largely brought about by differences in the regression coefficients—reflecting how a country's business, institutional and policy environment "prices" firm characteristics—rather than by differences in the ("endowments" of) firm characteristics. This result is important because it provides large-scale *microeconomic* evidence that institutions matter. In short, if one took the 28 EU countries as a single market where marginal products ought to be equalized, then the current state of Europe is very far from that. We estimate that removing distortions to allocation of resources across EU firms would raise productivity by 40 percent or more.

Our work is related to several strands of research. First, we contribute to the rapidly growing literature measuring misallocation of resources (see e.g., Restuccia and Rogerson, 2008, Hsieh and Klenow, 2009, Bartelsman et al., 2013); and also Restuccia and Rogerson, 2013, 2017, for surveys). In particular, we document new facts about the allocation of capital and labor in the 28 EU countries. Since EIBIS data are consistent across countries, our analysis is particularly well suited for cross-country comparisons.

Second, we provide new insights into the nascent literature on *sources* of observed dispersion in marginal products. For example, consistent with Asker et al. (2014), we show that dynamic adjustment of inputs is an important factor in accounting for cross-sectional variation in marginal products. However, we also document that other firm characteristics and various measures of distortions have predictive power for marginal revenue products. In contrast to previous work using *country-level* measures of distortions (e.g., Gamberoni et al. 2016, Kalemli-Ozcan and Sorensen 2012) or tight theoretical restrictions (e.g., Joel and Venkateswaran 2017), we use the richness of our survey to utilize *firm-level* information about various constraints and characteristics to account for cross-sectional variation in marginal revenue products with minimal restrictions. As a result, we can move beyond comparison of raw dispersion across countries and estimate the contribution of specific factors to misallocation, as well as compute associated productivity losses.

Third, by comparing administrative data to survey data, we contribute to recent efforts to assess the importance of measurement errors in observed marginal products (e.g., Bils, Klenow, and Ruane, 2017). In particular, we document high consistency of responses in survey data of EIBIS and Orbis (census-like) administrative data and hence show that surveys can be a useful source of information so that applied work does not necessarily have to use only data with census-like coverage.

Finally, we relate a large literature on dispersion of earnings across workers (see, e.g., Heckman et al., 2006) to the studies of dispersion of marginal products across firms. We show that many of the tools developed to understand dispersion of earnings can be employed to understand the dispersion of marginal products across firms.

At the policy level, our estimates provide important information for the ongoing debate about the need to remove distortions and obstacles to the EU single market (see European Parliament, 2019). Launched in 1993, the EU single market initiative allows free mobility of people, capital, goods and services within the EU. Yet, the perception is that persistency of frictions and distortions has prevented full exploitation of potential benefits, with major costs of mis-allocation.

The remainder of the paper is structured as follows. In Section II we present a dynamic model of a profit maximizing firm that yields steady state conditions for MRPK and MRPL. We use these conditions in Section III to formulate our estimating equations. In Section IV we describe EIBIS and Orbis data sets and we present our hypotheses related to the explanatory variables from EIBIS. We present our empirical estimates in Section V and draw conclusions in Section VI.

II. Theoretical Framework

To motivate our empirical analysis, consider a Cobb-Douglass production function, isoelastic demand function, and additively separable quadratic adjustment costs. Firm i's profit at time t is given by

$$\pi_{it} = G_{it}\left[(U_{it}K_{it})^{\alpha}(E_{it}L_{it})^{\beta}(\Delta_{it}X_{it})^{\omega}\right]^{1-\frac{1}{\sigma}} - R_{it}(U_{it})K_{it} - W_{it}(E_{it})L_{it} - P_{it}^{X}(\Delta_{it})X_{it}$$
$$-\frac{\phi_K}{2}\times\left(\frac{K_{it}}{K_{i,t-1}}-1\right)^2 R_{it}(U_{it})K_{i,t-1} - \frac{\phi_L}{2}\times\left(\frac{L_{it}}{L_{i,t-1}}-1\right)^2 W_{it}(E_{it})L_{i,t-1}$$

where $\gamma = \alpha + \beta + \omega$ reflects returns to scale in production, K_{it} is capital, L_{it} is labor, X_{it} is an intermediate input, U_{it} is a measure of capital utilization (or quality), E_{it} is a measure of labor effort (this can also capture efficiency wages or labor quality), Δ_{it} is a measure of intermediate input quality (or utilization), $R_{it}(U_{it})$ is the price schedule for the price of capital as a function of capital utilization (or quality), $W_{it}(E_{it})$ is the price schedule for the price of labor as a function labor effort (or quality), $P_{it}^{X}(\Delta_{it})$ is the price schedule for the price of intermediate input as a function of its quality (or utilization), ϕ_K and ϕ_L capture the size of adjustment costs (these could be stochastic and firm specific), G_{it} is a combination of productivity and demand shifters, and σ is the elasticity of demand. The price schedules could be modelled as $R_{it}(U_{it}) = R_t^{base} \times U_{it}^{\psi_K}/\psi_K \times \xi_{it}^{R}$, $W_{it}(E_{it}) = W_t^{base} \times E_{it}^{\psi_L}/\psi_L \times \xi_{it}^{W}$, and $P_{it}^{X}(\Delta_{it}) = P_t^{X,base} \times \Delta_{it}^{\psi_X}/\psi_X \times \xi_{it}^{X}$, where ψ_K, ψ_L, and ψ_X are slopes of the respective supply schedules, $R_t^{base}, W_t^{base}, P_t^{X,base}$ are market prices for the base quality/utilization of capital, labor and intermediate input, and $\xi_{it}^{R}, \xi_{it}^{W}, \xi_{it}^{X}$ are random shocks (structural distortions) to the schedule. We assume that firms rent capital, but similar expressions can be derived for the case when firms own capital.

Firms are assumed to maximize the present value of their profits

$$\Pi_{it} = \sum_{\tau=t}^{\infty}\left(\prod_{s=t}^{\tau}(1+r_s)\right)^{-1}\pi_{i\tau}$$

where r is the market interest rate which we assume to be constant across firms (e.g., the marginal or representative investor is the same across firms).

Let $S_{it} \equiv G_{it}\left[(U_{it}K_{it})^{\alpha}(E_{it}L_{it})^{\beta}(\Delta_{it}X_{it})^{\omega}\right]^{1-\frac{1}{\sigma}}$ be the firm revenue (sales). The optimality conditions for capital and labor are, respectively,

$$MRPK_{it} \equiv (1-\sigma^{-1})\alpha\frac{S_{it}}{K_{it}} = R_{it}(U_{it})\left\{1+\phi_K\times\left(\frac{K_{it}}{K_{i,t-1}}-1\right)-\frac{\phi_K}{1+r_{t+1}}\times\left(\frac{K_{i,t+1}}{K_{it}}-1\right)\right\} \quad (1')$$

$$MRPL_{it} \equiv (1-\sigma^{-1})\beta \frac{S_{it}}{L_{it}} = W_{it}(E_{it}) \left\{ 1 + \phi_L \times \left(\frac{L_{it}}{L_{i,t-1}} - 1 \right) - \frac{\phi_L}{1 + r_{t+1}} \times \left(\frac{L_{i,t+1}}{L_{it}} - 1 \right) \right\} \quad (1'')$$

$$MRPX_{it} \equiv (1-\sigma^{-1})\omega \frac{S_{it}}{X_{it}} = P_{it}(\Delta_{it}) \quad (1''')$$

Note that in a steady state when adjustment costs are zero, the costs of capital and labor are given by

$$R_i(U_i)K_i = \left((1-\sigma^{-1})\alpha \frac{S_i}{K_i} \right) K_i = (1-\sigma^{-1})\alpha S_i$$

$$W_i(E_i)L_i = \left((1-\sigma^{-1})\beta \frac{S_i}{L_i} \right) L_i = (1-\sigma^{-1})\beta S_i$$

where we drop the time index to underscore that this is a steady state. Hence, the steady-state cost shares for capital and labor are

$$s_i^K = \frac{R_i(U_i)K_i}{R_i(U_i)K_i + W_i(E_i)L_i + P_i^X X_i} = \frac{\alpha}{\alpha+\beta+\omega} = \frac{\alpha}{\gamma} \Leftrightarrow \alpha = \gamma s^K,$$

$$s_i^L = \frac{W_i(E_i)L_i}{R_i(U_i)K_i + W_i(E_i)L_i + P_i^X X_i} = \frac{\beta}{\alpha+\beta+\omega} = \frac{\beta}{\gamma} \Leftrightarrow \beta = \gamma s^L.$$

In the same spirit, $\omega = \gamma s^X$. We use these expressions to replace β, α, and ω in the expressions for marginal revenue products[2] to obtain

$$MRPK_{it} = (1-\sigma^{-1})\gamma s^K \frac{S_{it}}{K_{it}},$$

$$MRPL_{it} = (1-\sigma^{-1})\gamma s^L \frac{S_{it}}{L_{it}},$$

$$MRPX_{it} = (1-\sigma^{-1})\gamma s^X \frac{S_{it}}{X_{it}}.$$

Since markup $\mu = (\sigma-1)/\sigma$,

$$(1-\sigma^{-1})\gamma = \frac{1}{\mu}\gamma = (1 - s_\pi) \approx 1$$

[2] Consistent with Hsieh and Klenow (2009), our assumptions imply that we can measure marginal revenue products with average revenue products.

given that the share of pure economic profit in total revenue s_π is approximately zero in the data (e.g., Basu and Fernald 1997). Hence, we can further simplify the expressions for marginal revenue products to obtain

$$MRPK_{it} \approx s^K \frac{S_{it}}{K_{it}},$$

$$MRPL_{it} \approx s^L \frac{S_{it}}{L_{it}},$$

$$MRPX_{it} \approx s^X \frac{S_{it}}{X_{it}}.$$

Although it is conventional to define marginal products for physical units (e.g., number of employees and/or hours worked), capital is typically measured in dollars such as the replacement value of capital or the book value of fixed assets. In other words, we have

$$\widetilde{MRPK}_{it} \equiv (1-\sigma^{-1})\alpha \frac{S_{it}}{\tilde{R}_{it} K_{it}} \approx \frac{R_{it}(U_{it})}{\tilde{R}_{it}} \left\{1 + \phi_K \left(\frac{K_{it}}{K_{i,t-1}} - 1\right) - \frac{\phi_K}{1+r_{t+1}} \times \left(\frac{K_{i,t+1}}{K_{it}} - 1\right)\right\}$$

where $\tilde{R}_{it}$ is a measure of the capital price used in constructing the replacement value or the balance sheet value of fixed assets. In the case of replacement value of capital, we may have $R_{it}(U_{it}) \approx \tilde{R}_{it}$. With the balance sheet value of fixed asset, $\tilde{R}_{it}$ likely reflects the historical price rather than the current market price. Given technical change and inflation, the difference between the market and historical prices can be large, especially for assets bought a long time ago (e.g., buildings).[3] We are fortunate to have proxy information that enables us to try to correct for this effect. In particular, from EIBIS we know the share of capital (including machinery, equipment and ICT) that the management considers to be "state-of-the-art", which presumably means capital that has been obtained recently. Thus, for firms with a large share of state-of-the-art capital we can expect $R_{it}(U_{it}) \approx \tilde{R}_{it}$.

---

[3] For example, suppose that capital is bought at time t_0 and, for simplicity also, that capital does not depreciate, so that the balance sheet value is $p_\tau K_\tau$ at the time of purchase. $p_{t_0} K_{t_0}$ is also the balance-sheet value of fixed assets. The market price of capital at time t is given by $p_t = p_{t_0}\left(\frac{\Pi}{A}\right)^{t-t_0}$, where Π and A are the gross rates of inflation and technical change, respectively. Hence,

$$(1-\sigma)\alpha \frac{S_t}{p_{t_0} K_{t_0}} = (1-\sigma)\alpha \frac{S_t}{p_t K_{t_0}} \times \left(\frac{A}{\Pi}\right)^{t-t_0}.$$

If $\Pi > A$, a large share of state-of-the-art capital means a lower $\widetilde{MRPK}_{it}$ measured with the balance-sheet value of fixed assets. With depreciation, we obtain similar results but in this case the outcome also depends on whether the book value of capital depreciates faster on paper or *de facto*.

To make the connection to the misallocation literature, we consider the following canonical model where firm $i \in [0,1]$ maximizes profit

$$\max \tau_{it}^{Y} P_{it} Y_{it} - \tau_{it}^{K} R_t K_{it} - \tau_{it}^{L} W_t L_{it} - \tau_{it}^{X} P_t^X X_{it}$$

subject to the demand constraint $Y_{it} = Y_t \left(\frac{P_{it}}{P_t}\right)^{-\sigma}$ and production function: $Y_{it} = A_{it} K_{it}^{\alpha} L_{it}^{\beta} X_{it}^{\omega}$, where Y_{it} is output of firm i, Y_t is aggregate output, P_{it} is the price of firm i's output, P_t is the price index, K_{it} is capital, L_{it} is labor, X_{it} is materials (intermediate input), A_{it} is productivity, $\tau^Y, \tau^K, \tau^L, \tau^X$ are distortions in product and input market (no distortion corresponds to $\tau = 1$). Note that in this setting, firms face the same factor prices R_t, W_t, P_t^X. We show in Appendix B that optimality conditions for inputs are

$$MRPK_{it} \equiv (1-\sigma^{-1})\alpha \frac{P_{it} Y_{it}}{K_{it}} = \frac{\tau_{it}^{K}}{\tau_{it}^{Y}} R_t, \tag{2'}$$

$$MRPL_{it} \equiv (1-\sigma^{-1})\beta \frac{P_{it} Y_{it}}{L_{it}} = \frac{\tau_{it}^{L}}{\tau_{it}^{Y}} W_t, \tag{2''}$$

$$MRPX_{it} \equiv (1-\sigma^{-1})\omega \frac{P_{it} Y_{it}}{X_{it}} = \frac{\tau_{it}^{X}}{\tau_{it}^{Y}} P_t^X. \tag{2'''}$$

When we compare equations (1) with equations (2), we note that we can define "reduced-form" distortions τ as functions of structural distortions (e.g., ξ) and various compensating differentials for quality, utilization, and adjustment costs:

$$\frac{\tau_{it}^{K}}{\tau_{it}^{Y}} R_t = R_t^{base} \times \frac{U_{it}^{\psi_K}}{\psi_K} \times \xi_{it}^{R} \times \left\{1 + \phi_K \times \left(\frac{K_{it}}{K_{i,t-1}} - 1\right) - \frac{\phi_K}{1 + r_{t+1}} \times \left(\frac{K_{i,t+1}}{K_{it}} - 1\right)\right\},$$

$$\frac{\tau_{it}^{L}}{\tau_{it}^{Y}} W_t = W_t^{base} \times \frac{E_{it}^{\psi_L}}{\psi_L} \times \xi_{it}^{W} \times \left\{1 + \phi_L \times \left(\frac{L_{it}}{L_{i,t-1}} - 1\right) - \frac{\phi_L}{1 + r_{t+1}} \times \left(\frac{L_{i,t+1}}{L_{it}} - 1\right)\right\},$$

$$\frac{\tau_{it}^{X}}{\tau_{it}^{Y}} P_t^X = P_t^{X,base} \times \frac{\Delta_{it}^{\psi_X}}{\psi_X} \times \xi_{it}^{X}.$$

These expressions lead us to conclude that variation in marginal revenue products across firms that we attribute to distortions τ may reflect differences in adjustment costs, as well as input quality, utilization rates, and taxes or regulation. If one adjusted inputs for quality and/or accounted for adjustment costs and if the price schedules were the same across firms, then marginal revenue products for *effective* units of capital ($K_{it} U_{it}$), labor ($E_{it} L_{it}$), and intermediate inputs ($\Delta_{it} X_{it}$) should be equalized across firms and the cross-sectional dispersion ought to be zero. To the extent these adjustments and corrections are not possible, we may interpret variation in marginal revenue products as stemming from distortions. In short, marginal revenue products are functions of distortions and compensating differentials.

If dispersion in marginal revenue products is due to distortions, Hsieh and Klenow (2009) offer a simple approach to assess potential gains from a better allocation of resources. In particular, if distortions are log normally distributed $\log(\tau_{it}^{Y}) \sim N(0, V_{\tau Y})$, $\log(\tau_{it}^{K}) \sim N(0, V_{\tau K})$, $\log(\tau_{it}^{L}) \sim N(0, V_{\tau L})$, $\log(\tau_{it}^{X}) \sim N(0, V_{\tau X})$ and are uncorrelated, the loss in aggregate productivity from the distortions under constant returns to scale in production is given by (see Appendix B)

$$loss = -\left\{\frac{\alpha(1-\alpha)}{2} + \frac{\alpha^2\sigma}{2}\right\} V_{\tau K} - \left\{\frac{\beta(1-\beta)}{2} + \frac{\beta^2\sigma}{2}\right\} V_{\tau L} - \left\{\frac{\omega(1-\omega)}{2} + \frac{\omega^2\sigma}{2}\right\} V_{\tau X} - \frac{\sigma}{2} V_{\tau Y} + t.i.d. \quad (3)$$

where $t.i.d.$ captures terms independent of distortions.

III. The Econometric Framework and Identification

From our derivations it follows that the data analogue of the marginal revenue product of capital (the left-hand side of equation (1)) is $\log MRPK_{ijct} = \log\left(s_{jct}^{K} \frac{Y_{ijct}}{K_{ijct}}\right)$, where subscripts i, j, c, t index firms, sectors, countries, and time, respectively. Our discussion in Section II also makes it clear that $\log MRPK_{ijct}$ is a function of distortions, input quality, utilization, and other variables (the right-hand side of equation (1)), which after further linearization may be expressed as

$$\log MRPK_{ijct} = \psi_c + \kappa_j + \lambda_t + \boldsymbol{X}_{ijct}\boldsymbol{b} + \epsilon_{ijct} \quad (4)$$

where ψ_c is the set of country fixed effects, κ_j is the set of industry fixed effects, λ_t is the set of year fixed effects, $\boldsymbol{X}_{ijct}$ is the vector of explanatory variables (defined below), and ϵ_{ijct} is the error term that captures unexplained variation in $MRPK$. By combining equation (4) with the empirical measurement of $MRPK$, we obtain an empirical "Mincerian-type" specification.

An analogous specification and approach is used for other inputs. One can also estimate a more flexible specification with country × sector × year fixed effects η_{jct}:[4]

$$\log MRPK_{ijct} = \eta_{jct} + \boldsymbol{X}_{ijct}\boldsymbol{b} + \epsilon_{ijct}. \quad (4')$$

[4] If one is not willing to make an approximation with $s_{\pi} \approx 0$, then one may need to assume that elasticity of demand σ_{jc} and returns to scale in production γ_{jc} are constant across countries, industry or country/industry cells so that fixed effects absorb variation in σ_{jc} and γ_{jc}.

In this specification, one would expect that a significant part of the overall variation in firm-specific MRPK and MRPL will be absorbed by these country × sector × year fixed effects η_{jct} and that a smaller share of total variation will be explained by the vector $\boldsymbol{X}_{ijct}$.

In estimating equation (4) and similar specifications, we generate several important "outputs". First, we obtain estimates of $\boldsymbol{b}$ and hence can evaluate how the explanatory variables $\boldsymbol{X}_{isc}$ *predict* MRPK and MRPL. Second, we can use ϵ_{ijct} to compute a "residual" measure of dispersion in MRPL and MRPK across countries to assess whether some cross-country variation can be rationalized by differences in observable firm characteristics. Third, we can construct counterfactual distributions of MRPK and MRPL for a given country if it had coefficients $\boldsymbol{b}$ or endowments $\boldsymbol{X}$ from another country.

Note that estimates of $\boldsymbol{b}$ are not causal and hence interpretation of $\boldsymbol{b}$ is not straightforward. Finding convincing instrumental variables for many regressors in $\boldsymbol{X}$ is an unsurmountable challenge. As a result, our empirical strategy is based on two different insights that do not rely on exogeneity of $\boldsymbol{X}$. First, we know that under the null hypothesis of no misallocation, none of the regressors should have *predictive* power for marginal revenue products. This simple test does not rely on the exogeneity of regressors.

Second and more importantly, we are interested in how much variation in marginal revenue products is explained by a given regressor or a set of regressors. Basic econometrics implies that estimating equation (4) and similar specifications with OLS is likely to overstate the contribution of a given regressor.[5] Intuitively, some of the attributed variation in OLS estimates will be due to potentially omitted variables that may confound the OLS estimated relationship between marginal revenue products and regressors.[6] Thus, (marginal) R^2 in an OLS estimate of specification (4) provides an *upper bound* on how much variation in marginal revenue products can be due to a given friction or a given compensating differential measured in $\boldsymbol{X}$. As we argue below, productivity gains from removing a friction are based on (marginal) R^2 associated with the friction. Because larger R^2s are *ceteris paribus* associated with greater productivity gains and OLS yields an upper bound for R^2, we likely provide an *upper bound* for productivity gains from a better allocation of resources across firms. While having an upper bound may be only partially informative, our analysis does not rely on a structural interpretation of $\boldsymbol{b}$ and thus opens a number of opportunities. For example, one does not

[5] For example, instrumental variable estimators have (weakly) lower R^2 than OLS.
[6] To address this issue of potentially confounding factors, we include many control variables in specification (4).

have to restrict the analysis only to variables with well-identified, exogenous variation. Likewise, one does not have to impose tight theoretical restrictions to achieve identification.

Equation (2) makes it clear that we have fewer observables (marginal revenue products) than distortions ($\tau^Y, \tau^K, \tau^L, \tau^X$). Hence, to identify distortions from the observables, we need to impose a restriction. We follow Hsieh and Klenow (2009) and impose $\tau_{it}^L = 1$ for all i and t.[7] Under this assumption, one can show (Appendix B) that[8]

$$\log(\tau_{it}^Y) = constant - \log(MRPL_{it}),$$

$$\log(\tau_{it}^K) = constant + \log(MRPK_{it}) - \log(MRPL_{it}).$$

Hence, $V_{\tau Y}$ can be estimated with $var\,(\log MRPL_i)$ and $V_{\tau K}$ can be estimated with $var(\log MRPK_i - \log MRPL_i)$.

Because the variance of distortions is directly mapped to dispersion of marginal revenue products, there is a simple way to quantify the productivity gain from "removing" a friction. Consider specification (4) with marginal revenue product of capital as the dependent variable. We can quantify the contribution of a given friction to the variation in $MRPK$ across firms by using the marginal R^2 associated with the friction – i.e., the increase in R^2 when a regressor measuring the friction is added to some baseline regression. Because in our model $V_{\tau Y} = var\,(\log MRPL_i)$, it follows that the change in $V_{\tau Y}$ brought about by removing the friction is $var(\log MRPL_i) \times (marginal\ \ R^2)$. Likewise, we can compute the change in $V_{\tau K}$ as $var(\log MRPK_i - \log MRPL_i) \times (marginal\ \ R^2)$, where $(\log MRPK_i - \log MRPL_i)$ is the dependent variable in a regression.[9] Thus, we measure the productivity gains from removing a friction with

$$loss = -\left\{\frac{\alpha(1-\alpha)}{2} + \frac{\alpha^2\sigma}{2}\right\} \times var(\log MRPK_i - \log MRPL_i) \times (marginal\ \ R^2)$$
$$-\frac{\sigma}{2} \times var(\log MRPL_i) \times (marginal\ \ R^2). \qquad (3')$$

[7] The results are qualitatively similar when we use an alternative assumption that $\tau_{it}^K = 1$ for i and t.

[8] Because we do not have a measure of material cost in EIBIS, we cannot recover a distortion in inputs. However, we know that this distortion has a non-negative variance and hence this distortion will lower aggregate productivity and output. Hence, by ignoring this distortion, we likely understate gains from improving the allocation of resources across firms.

[9] Classical measurement error may affect the level of dispersion in marginal revenue products. However, because classical measurement error is additive in terms of variances, it does not influence the variance contribution attributed to a friction.

IV. Data

The main data source for our analysis is the EIB Investment Survey (EIBIS). We next provide information on its design and implementation. We also compare EIBIS responses to the administrative data of the surveyed firms, as collected in the Orbis database. Once we establish consistency across the survey and administrative data, we describe survey questions that we use in the empirical analysis to account for the variation in MRPK and MRPL across firms.

A. The EIB Investment Survey (EIBIS)

EIBIS is an annual firm-level survey conducted by the market research company Ipsos MORI on behalf of the European Investment Bank (see Ipsos (2017) for a detailed review of the survey). The first wave of EIBIS was administrated in 2016, targeting firms in the 28 EU member states with the objective of being representative in each country of different size classes and sectors. The sampling targeted head offices.[10] Eligible respondents were senior persons with responsibility for investment decisions and their financing. The respondent could be the owner, finance manager, finance director or head of accounts, Chief Financial Officer (CFO) or Chief Executive Officer (CEO).

The sample was stratified disproportionally by country, industry group (sector) and size-class, and stratified proportionally by region within the country. The minimum number of employees of all enterprises is five, with full-time and part-time employees being counted as one employee and employees working less than twelve hours per week being excluded. The Orbis dataset of Bureau van Dijk was used as the sampling frame in all countries. Brutscher and Coali (2019) provide evidence on the representativeness of the data for the business population of interest (namely enterprises with five or more employees).

The fieldwork for the first wave started in July 2016 and continued until November 2016. The vast majority of interviews were conducted in August and September 2016. The interview was administrated by telephone using computer-assisting telephone interviewing (CATI). Responses refer to fiscal year 2015 and the response rate was approximately thirteen percent. The resulting sample consists of 12,483 non-financial enterprises in the 28 EU member states in NACE categories C to J (industrial firms). The sample size varies with the size of the population and ranges from 150 enterprises in Cyprus and Luxembourg to 600 in France, Germany, Italy, and the UK. Because the

[10] An enterprise is defined as a company trading as its own legal entity. As such, branches were excluded from the target population. However, the definition is broader than a typical enterprise survey given that some company subsidiaries are their own legal entities.

sampling frame as well as the resulting samples may be not representative, Ipsos MORI constructed weights to correct for possible imbalances. Specifically, firms are weighted to make them representative of the EU economy based on country, sector and firm size (employment) where the population distribution is reported by the Structural Business Statistics (SBS) in Eurostat.

The second (2017) and third (2018) waves have similar properties and were conducted between April and September of 2017 and 2018, respectively. At the end of each wave, firms are invited to participate in the next wave of the survey so that EIBIS has a panel component. Approximately 2,000 firms participated in all three waves and approximately 4,500 firms participated in two waves.

EIBIS is a rich source of information with a number of unique characteristics. First, EIBIS collects basic information on firms (e.g., number of employees, value of fixed assets and sales) that is matched to administrative data.[11] This feature allows us to cross-check survey responses against data from administrative sources and hence assess the quality of the survey data. Second, EIBIS gathers data on expectations and perceptions of firms' management (e.g., perceived barriers for operations and plans for future investment), as well as statistics that are often not available in standard official sources (e.g., the quality of capital stock, utilization rate, sources of financing). These variables can inform us directly about the sources of variation in marginal revenue products across firms and thus advance the analysis beyond studies that are based only on income statements and balance sheets. Third, EIBIS data are collected for a large number of firms in a *consistent* manner across *many* countries and industries, thus permitting us to carry out a comparative analysis of resource allocation in various institutional settings. Using these unique matched data, we explore the relationship between MRPK, as well as MRPL, and a large number of explanatory variables at the *firm* level. To this end, we use questions on the firms' demographics, capacity utilization, quality of the capital stock, obstacles to long-term investment, investment plans, investment rate, employment growth, and sources of finance.

[11] The data on each firm from EIBIS was merged with the corresponding data inOrbis. The matching was done by Ipsos-Mori, which in turn provided the anonymized matched data to the EIB. This means that EIBIS does not have the name, the address, the contact details or any additional individual information that could identify the firms in the final sample. Not every firm in EIBIS has complete information in Orbis (e.g., Orbis may have missing information on employment while EIBIS does not).

B. Comparison with Administrative Data

The Orbis database is a popular source of administrate data for cross-country analyses at the firm level.[12] As mentioned above, we use the Orbis data to check the validity of EIBIS responses. In particular, we match firms responses with the administrative data of the interviewed firms and compare cross-firm dispersion of the logarithm of sales, fixed assets and employment in EIBIS and in the administrative data as collected in Orbis for the same firms by country (Table 1) and industry (Table 2). In columns (4), (7) and (10) of the two tables we also report correlations between the responses in EIBIS and the administrative data in Orbis.

As may be seen from the tables, we observe a high degree of consistency in the two sources of data. For example, the correlation between log employment in EIBIS and Orbis is 0.91. The dispersion of the survey responses across firms is on average slightly larger than the dispersion in the administrative data, which is consistent with small noise (measurement error) in the survey responses.[13] We conclude that EIBIS provides satisfactory quality of firm-level data and therefore that the survey responses are suitable for our analysis.

C. Dispersion of marginal revenue products

We report descriptive statistics for the full EIBIS sample in Table 3. The key statistic for our analysis is the dispersion of marginal revenue products. We observe a sizable dispersion across firms in the EU: the standard deviation is 1.43 for $\log MRPK$, 1.19 for $\log MRPL$, and 1.63 for $(\log MRPL - \log MRPK)$.[14] For comparison, the dispersion of marginal value-added product across establishments ("plants") for the U.S. is 0.98 for capital (Table 2 in Asker et al., 2014) and 0.58 for labor (Table 1 in Bartelsman et al., 2013).

Note that there are three potentially confounding sources of differences between our statistics and statistics reported for the U.S. First, EIBIS does not collect information on the cost of intermediate inputs. We therefore use sales to compute marginal revenue products for capital and labor, while previous studies used value added. Using EIBIS firms matched to the Orbis database (which has information on sales and value added), we find that the standard deviation of $\log MRPK$ based on sales

---

[12] See Kalemli-Ozcan et al. (2015) for a detailed analysis of the advantages and disadvantages of using this dataset.
[13] Note that compared to the data on fixed assets, data on employment are available for fewer firms in the Orbis database.
[14] We find similar magnitudes when we use robust methods to estimate standard deviation. For example, when we employ median absolute deviation (MAD) to estimate $st.dev. = 1.48 \times MAD$ where $MAD = median|x_i - \tilde{x}|$, $\tilde{x} = median(x_i)$ and x_i is a random variable, we find that the standard deviation is1.49 for $\log MRPK$, 1.14 for $\log MRPL$ and 1.68 for $(\log MRPL - \log MRPK)$.

is approximately 0.16 log points higher than the standard deviation of $\log MRPK$ based on value added. On the other hand, the standard deviation of $\log MRPL$ based on sales is approximately 0.21 log points lower than the standard deviation of $\log MRPL$ based on value added. Thus, using revenue rather than value added does not appear to explain the difference between the EU and US.

Second, our analysis in based on survey responses while US studies rely on administrative data. One might hence hypothesize that there is a US-EU difference in dispersion because survey data are more likely to have measurement error. As we discussed above, however, the survey responses in EIBIS are broadly consistent with the administrative data in Orbis.[15] To further explore the quantitative significance of measurement error, we exploit the panel component of EIBIS. In particular, we compute the average $\log MRPK$ and $\log MRPL$ across years for a given firm and then compute cross-sectional dispersion for these averages. Of course, taking averages attenuates not only measurement errors (as we use repeated measurements) but also transitory factors (e.g., adjustment costs, high-frequency variation in demand) and the reduction in dispersion hence likely overstates the role of measurement errors. With this caveat in mind, we find that the standard deviation of average $\log MRPK$ is 8 percent lower for firms participating in two waves of the survey and 12 percent lower for firms participating in three waves. The corresponding figures for $\log MRPL$ are 4 percent for two-wave firms and 1 percent for three-wave firms. These results suggest that measurement error can rationalize only for a portion of the EU-US difference in dispersion in marginal revenue products.

Third, the unit of analysis in EIBIS is either a firm or a subsidiary which is a (weakly) larger unit than an establishment. Kehrig and Vincent (2017) document that approximately two-thirds of the variance in marginal value-added product of capital across establishments happens across establishments within a firm; that is, the variance across firms is approximately one-third of the variance across establishments. Hence, the figures for the US likely *understate* the true difference between the US and EU.

While contrasting the EU and US figures highlights challenges of cross-country comparisons, it is clear that qualitatively the greater dispersion of marginal revenue products (misallocation of resources) in the EU relative to the US is consistent with lower aggregate

[15] In agreement with high consistency in measures of employment, capital, and sales across data sources, we observe that measured dispersion of marginal revenue products is similar in EIBIS and Orbis. For example, for the sample of EIBIS firms that are matched to Orbis firms with non-missing data (6,432 firms in 2015), the standard deviation of $\log MRPK$ is 1.44 in Orbis and 1.37 in EIBIS. The corresponding figures for $\log MRPL$ are 1.07 in Orbis and 1.30 in EIBIS. Consistent with some measurement error in survey responses, the dispersion of MRPL is somewhat larger in the survey than in the administrative data.

productivity in the EU relative to the US (e.g., van Ark et al. 2008). Fortunately, EIBIS covers many EU member states and we can exploit the consistency of measurement across countries within the survey to examine whether the greater dispersion of marginal revenue products is associated with lower aggregate productivity. Figure 1 demonstrates that as predicted by theory, there is a robust negative correlation between dispersion and productivity. This fact not only provides additional credibility to survey-based measures of marginal revenue products but also gives further motivation for studying dispersion as a potential source of cross-country differences in productivity.

D. EIBIS Variables Used in the Regression Analysis

We consider several blocks of variables available in the survey to construct vector $\boldsymbol{X}$ in equation (4). We next discuss the possible relationships between these variables and the marginal revenue products. The choice of variables is motivated by previous work and is constrained by data availability. Descriptive statistics related to these variables are presented in Table 3.[16]

Demographics of the firm

Employment size – Garicano et al. (2016), Bento and Restuccia (2017) and other studies document that various size-based policies can introduce distortions to the scale of operation and, hence, allocation of resources (e.g., firms in low-productivity countries are systematically smaller than firms in high-productivity countries). These distortions may arise for a variety of reasons. For example, firms with a large labor force may have local monopsony power in the labor market and hence a higher MRPL than they would in a competitive setting. The effect of this on MRPK depends on whether labor and capital are substitutes or complements. Average log employment in our sample is 4.84 (approximately 125 employees).

Firm age – One may expect that on average firms that have been longer in existence have older and probably lower quality capital. They may have also amortized their old capital stock and deploy it even if its marginal product is low. The hypothesis is hence that older firms will have lower MRPK. The effect of firm age on MRPL may be positive if older firms have workers with more firm-specific human capital but the effect depends on whether labor and capital are substitutes or complements in production. At the same time, since age may be also a proxy for productivity (and hence various survival or selection effects), both MRPK and MRPL may be higher for older firms.

[16] Descriptive statistics by survey wave are in Appendix Table 4.

Subsidiary status – Subsidiaries may have access to cheap intra-group capital, resulting in a lower optimal MRPK, or they may be rationed and monitored for efficient use of capital by the parent company, resulting in a higher MRPK. Subsidiaries may also have a higher quality capital, resulting in higher MRPK. As regards labor, subsidiaries of foreign firms tend to pay higher wages than local firms (see e.g., Lipsey, 2003, and Malchow-Møller, 2013). One may hence expect that their MRPL will be higher than that of other firms. Approximately a third of firms in our sample are subsidiaries.

Exporter status – Being more exposed to competition, exporters are relatively more likely to employ high-quality, and hence more expensive inputs (see e.g., Verhoogen 2008). About one-half of firms in our sample are exporters.

<u>Utilization and quality of inputs:</u>

Quality of capital – A higher quality of capital, measured by a greater share of "machinery and equipment (including ICT) that are state-of-the-art", and a higher proportion of "commercial buildings that satisfy high energy efficiency standards" are expected to have a positive effect on MRPK if they represent an upward shift in the MRPK curve or a negative effect if they constitute a movement along the MRPK curve. The average shares of "machinery and equipment that are state-of-the-art" and "energy efficient buildings" are 41 and 36 percent, respectively.

Capacity utilization – Firms operating at (or even above) maximum capacity may be expected to have high MRPK and MRPL as all machinery, equipment and labor are used to the fullest extent and there is demand for more. By the same token, firms with low capacity utilization are expected to have low marginal product of (idle) machinery, equipment and labor. In the survey, 44 percent of the firms report operating at maximum capacity and 6 percent above maximum capacity.

<u>Obstacles to investment</u>

The variables included in this cluster are answers of firms' top management to questions about constraints on investment. When asked about a specific potential constraint, a respondent could choose a "major obstacle", "minor obstacle", or "not an obstacle at all". The list of constraints includes:

Demand for products or services – Deficient demand as an obstacle to investment may be expected to result in lower MRPK and MRPL as existing capital and labor are adequate or more than adequate to satisfy product demand. On the other hand, this effect may already have been taken into account given that capacity utilization is controlled for.

Availability of staff with the right skills – To the extent that the firm cannot obtain a sufficient number of appropriately qualified employees to expand production, the marginal product of labor may be expected to be high.

Energy costs – The effects of high energy costs on MRPK and MRPL depends on whether energy is a complement or substitute to capital and labor.

Access to digital infrastructure – Similarly, the effects of a limited access to digital infrastructure on MRPK and MRPL depends on whether it is a complement or substitute to capital and labor.

Availability of adequate transport infrastructure – The effects of a limited availability of adequate transport infrastructure on MRPK and MRPL also depends on whether it is a complement or substitute to capital and labor.

Labor market regulation – The effects on MRPL will be positive if the firm uses less labor, and negative if the regulation results in excess employment in the firm. The effect on MRPK depends on whether the two inputs are substitutes or complements.

Business regulations and taxation – The effects of business regulations as an obstacle to investment is a priori unclear, depending on what form the regulations take.

Availability of finance – If access to finance is an obstacle to investment, one may expect MRPK to be higher than if the availability of finance is not a constraint.

Uncertainty about future – If uncertainty about future is an obstacle to investment, MRPK may be expected to be higher than in the absence of uncertainty.

For each obstacle, approximately 20-40 percent of firms report it to be a major obstacle and another 30 percent regard it as a minor obstacle.

Dynamic adjustment

Firms are exposed to a variety of shocks and with adjustment costs it may take time and resources for firms to reoptimize factor allocation. Although EIBIS data does not have a large panel component yet, the survey has questions about firms' current and previous choices – an aspect that enables us to examine the dynamics of inflows and outflows of capital and labor.[17] The variables included in this cluster are:

Investment – Investment increases the amount of capital used and should result in a lower MRPK as the firm experiences diminishing returns to capital (movement down along the MRPK curve). While it is common to use investment rate (that is, investment normalized by capital stock or by sales), we use $\log(1 + investment)$. Our choice is motivated by the possible presence of measurement error in reported fixed assets and/or sales. Since these two variables appear on the left-hand side of equation (4), the conventional scaling of investment may introduce spurious correlations due to measurement errors. We use the log transformation to take care of the thick right tail in the volume of investment. We add one to the transformation to keep in the sample firms with zero investment.[18]

Employment growth over the past three years – This explanatory variable should have a negative effect on MRPL as the firm experiences diminishing returns to labor. The average employment growth for firms in our sample is 13 percent.

Investment over the past three years – This variable comes in the form of management's information about whether this investment was "too much", "too little" or "about the right amount." One would expect that "too much" results in a low MRPK as the firm experiences diminishing returns to capital, while "too little" goes the other way. Most firms (79 percent) report that the amount of investment was about right.

Investment plans for the next three years – Our derivations indicate that MRPK should be a function of not only current and past investment rates but also expected future investment. Thus, having information about firms' investment plans may be useful in explaining contemporaneous dispersion of MRPK across firms. A unique feature of the EIBIS data is that the survey asks firms to report their expected investment for the next three years. In particular, firms can report whether investment will be for "replacing capacity (existing buildings, machine, equipment and IT)", "expanding capacity for existing products and services", "developing or introducing new products, processes or services", or "do not have

[17] Since EIBIS does not have information about material costs, we assume implicitly that materials may be adjusted quickly.
[18] The results are similar when we also include an indicator variable equal to one if a firm reports positive investment, and zero otherwise (the baseline specification uses log(1+investment)). See Appendix Table 3 (analogue of Table 4).

investment planned." There is no *a priori* expectation as to which types of investment (replacing buildings, machinery, equipment, and IT versus expanding capacity for existing products and services) would enhance or diminish the effect of the investment rate variable. However, the response "developing or introducing new products, processes or services" may be expected to have a positive effect on MRPK as the firm expands into these new areas and needs time to accumulate the optimal capital stock. The most popular response is "replacing capacity" (35 percent).

Source of funding

Share of investment funded by internal and external finance – The standard model of a profit maximizing or cost minimizing firm yields the prediction that MRPK should be equal across firms if all firms face the same price of capital. In practice, firms may have different cost of capital depending on how old they are, how connected to capital markets they are, etc. In particular, a number of studies (e.g., Desai et al., 2004, Fama and French, 2002) document that the cost of external funds is higher than the cost of internal funds (or funds obtained within a business group). EIBIS asks firms with positive investment to report the source of their funds to pay for their investment (internal, external, intra-group).

Credit constraint – this indicator variable is equal to one if a firm was rejected in its loan application, was discouraged from applying for a loan, or received a loan that was too small or too expensive. Holding everything else constant, a credit constrained firm should have a shortage of capital and likely substitute capital with other inputs thus making MRPK high.

Data filters and additional data

To minimize potentially adverse effects of extreme observations, we winsorize continuous variables at the top and bottom one percent. For firms with missing information for a given variable, we impute the average value of that variable in the industry-country cell. For each variable, we create a corresponding indicator variable taking value one if the values were imputed. We include these indicator variables as additional regressors but do not report their estimated coefficients in the regression tables. We estimate the cost shares for labor and materials using data from the Industrial Analysis section of the OECD's Structural Analysis Database (STAN) or from Eurostat national accounts that are available at the level of the country, year and industry (two-digit NACE classification).

V. Empirical Analysis

In this section, we present four sets of results. First, we explore the extent to which firm characteristics predict log(MRPK) and log(MRPL). Second, we use our estimates to quantify productivity gains from better allocation of resources. Third, we consider how adjustment for observed firm characteristics can influence measures of cross-sectional dispersion in MRPK and MRPL and hence potentially reduce inefficiencies in resource allocation. Fourth, we assess whether observed cross-country dispersion in MRPK and MRPL is due to differences in firm characteristics ("endowments" as reflected in the values of the explanatory variables) or due to differences in how these characteristics are "priced" (i.e., in how regression coefficients – reflecting business, institutional and policy environment – affect MRPK and MRPL).

A. Regression results

Our preferred specification for the regression analysis is equation (4') in which we enter as regressors variables $\boldsymbol{X}$ together with country × industry × year fixed effects. The estimated coefficients for this specification are reported in Table 4. We re-iterate that we do not interpret the estimated relationships as causal. We estimate equilibrium relationships and estimated coefficients may therefore have signs and magnitudes potentially inconsistent with priors built on causal relationships between the variables. For example, we may observe a positive association between a marginal product and a constraint because the constraint is only binding for the more advanced firms. While this is a limitation, our analysis has important benefits. Recall that if $\boldsymbol{X}$ does not predict the variation in marginal revenue products across firms, under certain conditions one can use "raw" marginal revenue products to compute productivity losses from the dispersion of marginal revenue products across firms. On the other hand, if $\boldsymbol{X}$ predicts a sizable fraction of the variation in marginal revenue products, then the dispersion of "raw" marginal revenue products is potentially not the appropriate indicator for productivity calculations. Moreover, in our explanatory analysis we assess the potential of $\boldsymbol{X}$ to predict the variation of marginal revenue products in the data which likely provides an upper bound on the magnitude of causal effects and thus (marginal) R^2 is an informative statistic.

Whether the variables in vector $\boldsymbol{X}$ reflect genuine distortions (e.g., undesirable regulations) or compensating differentials (e.g., quality of inputs or intensity of effort) influences how one should interpret the relatively high R^2s. If the variables measure distortions, then our estimates suggest that by removing distortions one can achieve considerable productivity gains. On the other hand, if variables in $\boldsymbol{X}$ measure compensating differentials, then R^2s indicate adjustments one should make *before* calculating productivity losses. In other words, the observed dispersion may overstate inefficiency and hence productivity losses. To illustrate this point, we later classify $\boldsymbol{X}$ into "distortions" and "compensating differentials", although as we emphasized above, the interpretation of estimated coefficients is tentative and the issue ought to be tackled systematically in future research.

Turning to the estimated coefficients of equation (4) in Table 4, we find that for the "demographics" block of variables MRPK is positively related to the age of firm early on, rising 22 percent (log points) in firms that are 10-19 years old and 34 percent (log points) in firms that are 20+ years old. The effect of firm age on MRPL is estimated to increase by 4 percent after the first nine years of existence and slightly to increase over time thereafter, reaching 10 percent for firms that are 20+ years old. The higher MRPL in older firms could be related to a greater accumulation of firm-specific human capital over time. Perhaps unsurprisingly, employment size is positively associated with MRPK and negatively associated with MRPL. Subsidiary status is positively related to MRPK and MRPL; the estimated coefficients are 35 percent (log points) for MRPK and 14 percent for MRPL, respectively. Export status is also associated with higher MRPK (coefficient is 14 log points) and MRPL (coefficient is 24 log points). This finding is consistent with these firms being more exposed to competition and hence relatively more careful in avoiding excessive amounts of inputs and employing more high-quality inputs.[19]

For the "quality of inputs" block of regressors, we observe that firms reporting to have a greater share of "state-of-the-art machinery and equipment" and a higher proportion of "energy efficient buildings" are estimated to have a significantly lower MRPK, suggesting that this indicator captures a movement along the MRPK curve rather than an upward shift in the MRPK curve. The effect on MRPL is positive, both for state-of-the-art machinery and equipment as well as energy efficiency.

For the "utilization" block, we find that capacity utilization has a strong positive effect on both MRPK and MRPL, consistent with the expectation that marginal products of inputs are high when high product demand requires machinery, equipment and labor to be used to the fullest possible extent and "beyond". Correspondingly, the estimated coefficients also suggest that firms with low capacity utilization have low marginal product of (idle) machinery, equipment and labor.

---

[19] The results are qualitatively similar when we restrict the sample to firms that participated in all three waves of the survey (Appendix Table 2).

In the “obstacles for investment” block, management’s responses to questions about obstacles to investment paint a nuanced picture: perceived obstacles appear to have differential relationships with marginal revenue products. For example, facing demand for products or services as an obstacle is positively associated with MRPK, but not associated with MRPL. With capacity utilization already being controlled for, deficient demand may signal that firms are cautious in augmenting input use and prefer to keep MRPK high. Having availability of staff with the right skills as an obstacle to investment has a positive association with MRPK and a negative association with MRPL. The effect of energy costs as an obstacle is negative on both MRPK and MRPL.

Facing access to transport or digital infrastructure as an obstacle is generally associated with a positive coefficient on MRPL, suggesting that these types of infrastructure are complements to labor. The availability of adequate infrastructure as a constraint does not appear to have an association with MRPK. Labor market regulation as an obstacle to investment has a negative association with MRPL and no material correlation with MRPK. Business regulations and taxation has a positive association with MRPL, while it has some negative correlation with MRPK. Unavailability of finance is negatively related to both MRPK and MRPL. Uncertainty about future has a positive correlation with both MRPK and MRPL.

In the “adjustment” block, investment has a strong negative association with MRPK and a positive association with MRPL. These associations are consistent with movements along the MRPK curve and a shift in the MRPL curve: as investment increases the amount of capital used, it should result in a lower MRPK as the firm experiences diminishing returns to capital (movement down along the MRPK curve) and a higher MRPL as labor becomes relatively scarcer. Symmetrically, we find that employment growth in the last three years is associated with a higher MRPK and lower MRPL. Thus, a change in employment appears to be consistent with moving along the MRPL curve and a shift up in the MRPK curve (labor and capital being substitutes). Too little or too much investment in the past is associated with lower MRPK and MRPL. Consistent with standard adjustment costs, future investment into capital reduces MRPK currently and has no material association with MRPL (except for the negative association with investment to replace capacity).

In the “source of funds” block, the “credit constrained” status is negatively correlated with MRPL and MRPK. Although one could have expected that being credit constrained would lead to a higher MRPK, one should note that firms may be denied credit because of their poor fundamentals. If this latter effect dominates, we should observe the negative correlation between the "credit constrained" indicator and MRPK. Using internal funding rather than external funding to pay for capital is associated with a high MRPK. This finding is consistent with the view that firms using internal funds are more likely to be capital constrained. At the same time, a high share of internal funds is positively associated with MRPL suggesting that there could be selection effects similar to a "credit constraint." Using intra-group funding is negatively correlated with MRPK and positively correlated with MRPL, indicating that these funds may indeed reflect a lower cost of capital and result in firms substituting labor with capital.

Our analysis of partial correlations suggests that the significant cross-sectional association between marginal products and firm characteristics varies across blocks of variables. For example, variables measuring firm demographics, dynamic adjustment of inputs, and source of funds appear to have robust predictive power. On the other hand, the contribution of "constraint" variables to the variation in MRPK and MRPL is modest, with some coefficients not being statistically significant. To quantify this observation, in Table 5 (panels A and B) we present marginal R^2s for the blocks of variables, that is, by how much R^2 increases after a given block of variables is added to various fixed effects. In line with the results in Table 4, we observe for MRPK and MRPL that marginal R^2s are the largest for variables in the "adjustment" and "demographics" blocks and relatively low for variables in the "obstacles for investment" block.

For illustration purposes, we next lump these blocks of variables into two groups. In the first group we include "quality of capital," "capacity utilization," and "adjustment." We interpret this group as compensating differentials because they could be argued to reflect firm policies. The second group includes "firm demographics," "obstacles to investment," and "source of funds," which we interpret as *constraints and distortions* because they reflect predetermined factors and business environment. We see in Table 5 that in terms of marginal R^2 the predictive power is similar for the two groups of variables. Conditional on accepting this classification of variables, one can reach two important conclusions. First, the "raw" dispersion in marginal products is likely to overstate the extent of misallocation since some variation is likely to be brought about by heterogeneity in the "quality" of inputs. Second, "distortions" are likely to be substantial and removing them may lead to significant gains in productivity.[20]

[20] We find similar results (Appendix Table 1) when we estimate equation (4') using a "between" regression (that is, regression is estimated on average (across years) values of regressors and regressands). This specification likely reduces the importance of transitory factors such as measurement errors and adjustment costs.

B. Productivity Gains

Equation (3') provides a straightforward approach to measure potential gains from improved allocation of resources. We use $\bar{s}^K$ (equal to 0.19 in the data) to parametrize α. We follow Hsieh and Klenow (2009) and calibrate $\sigma = 3$ which likely yields a conservative estimate of productivity losses due to misallocation. With these parameter values, the weight on $V_{\tau Y}$ is 1.50 and the weight on $V_{\tau K}$ is 0.13.

We carry out calculations for several policy scenarios. First, assume that the policymakers would eliminate the dispersion in marginal revenue products brought about by the "distortions" group. To have a level of variation in marginal revenue products, we use the dispersion from Table 3: $var(\log MRPL) = 1.19^2 \approx 1.42$ and $var(\log MRPL - \log MRPK) = 1.63^2 \approx 2.66$. In Table 5 we report marginal R^2 for various specifications without and with fixed effects. Clearly, the marginal R^2 is decreasing with the richness of fixed effects. To have an upper bound on productivity gains, we consider marginal R^2 without fixed effects (column (1)). The marginal R^2s of the "distortions group" are 0.186 for $\log MRPL$ (row 17, Table 5) and 0.134 for $(\log MRPL - \log MRPK)$ (row 26, Table 5). It follows that the gain in productivity is $1.5 \times 1.42 \times 0.186 + 0.13 \times 2.66 \times 0.134 = 0.442$, which is reported in Panel D (row 35, column 1) of Table 5. In other words, removing distortions can raise aggregate productivity by more than 40 percent.

Second, consider the possibility that *all* variables in $\boldsymbol{X}$ capture distortions. In this case, the marginal R^2s are 0.289 for $\log MRPL$ (row 18, Table 5) and 0.289 for $(\log MRPL - \log MRPK)$ (row 27, Table 5). With these marginal R^2s, the gain is $1.5 \times 1.42 \times 0.289 + 0.13 \times 2.66 \times 0.289 = 0.715$, which is reported in Panel D (row 36, column 1) of Table 5. Interestingly, variables in the "adjustment" block contribute to productivity gains (row 32) as much as variables in the "demographics" block (row 28). In short, treating all variables in $\boldsymbol{X}$ as distortions increases the magnitude of potential gains by a half relative to the first scenario.

These results suggest that the EU has potential to considerably increase its productivity by improving allocation of resources. Obviously, the magnitude of the gains depends on the interpretation of variables collected in $\boldsymbol{X}$ but our approach is highly portable and can provide an upper-bound estimate for any variable of interest. Indeed, comparing dispersion of marginal revenue products across countries may be the first step in identifying the problem, but our approach permits one to identify *which* factors are likely most limiting. Given that frictions are often best measured with surveys (e.g., firms may report the importance of various barriers, regulation, etc. in a survey), EIBIS and similar efforts can provide a key input for policymakers.

C. CROSS-COUNTRY AND CROSS-INDUSTRY DIFFERENCES

In our data, there is also considerable cross-country variation in the *average* marginal revenue products—0.33 for log (MRPK) and 0.70 for log (MRPL) —but this variation is small relative to the within-country variation in MRPK and MRPL. In Figure 2, we show the estimated dispersion in MRPK (Panel A) and MRPL (Panel B) within countries, measured as the within-country standard deviation in the logarithm of MRPK and MRPL, respectively. We present the dispersion in "raw" marginal revenue products and in marginal revenue products adjusted for various groups of observed characteristics (just variables $\boldsymbol{X}$, variables $\boldsymbol{X}$ plus country, industry and year fixed effects, and variables $\boldsymbol{X}$ plus country × industry × year fixed effects) in a cross-country regression given by equation (4). As may be seen in Figure 3, the dispersion of MRPK and MRPL is highly correlated at the country level.

There is considerable dispersion in the raw MRPK and MRPL in both the more and less advanced economies. Note that in Figure 3 the dispersion of raw marginal products is particularly high in smaller countries such as Malta (MT), Luxembourg (LU) and Cyprus (CY). Among the larger countries, Germany (DE) is the country with the lowest raw dispersion of marginal revenue products.

If one takes the view that some of the dispersion is due to compensating differentials rather than distortions, then one may for instance start cross-country comparisons by using the red bars in Figure 2 (MRPK and MRPL adjusted for observed firm characteristics $\boldsymbol{X}$, with no fixed effects included). Although using $\boldsymbol{X}$ reduces the cross-sectional dispersion, it generally preserves the ranking of the countries. Adding country, industry and year fixed effects further reduces the levels of dispersion and the ranking of countries is generally preserved, although the ranking for some countries shifts (e.g., Romania (RO) is similar to Slovenia (SI) in terms of "raw" MRPK dispersion, but after this adjustment Romania becomes more similar to the Netherlands (NL)). Introducing country × industry × year fixed effects not only reduces the level of dispersion, it also attenuates differences across countries. For example, France and Italy have rather different dispersion of "raw" MRPK but they have similar dispersion of MRPK after adjustment for the controls and country × industry fixed effects. Depending on the interpretation, these results suggest either that removing distortion can reduce cross-county differences in the allocation of capital and labor, and thus bring about improvements in productivity, or that the observed cross-country differences in raw dispersions are misleading and after adjusting for compensating differentials these differences become smaller.

The quantitative importance of country, industry and year fixed effects or the interaction terms country × industry × year raises an important identification challenge. In particular, fixed effects can absorb not only cross-country/industry compensating differentials for quality of inputs but also barriers for capital and/or labor flows across countries and industries. While it is beyond the scope of this paper to resolve this identification problem, we can provide some leads for discussion and future research.

Table 6 reports R^2 for regression (4) with various sets of fixed effects and no controls $\boldsymbol{X}$. Country fixed effects alone account for $R^2 = 0.170$ for MRPK and $R^2 = 0.445$ for MRPL. Industry (2-digit level) fixed effects alone account for $R^2 = 0.239$ for MRPK and $R^2 = 0.268$ for MRPL. Time fixed effects have little explanatory power. The combined contribution of country, industry and year fixed effects (column 5) is $R^2 = 0.275$ for MRPK and $R^2 = 0.611$ for MRPL. To the extent that fixed effects embody distortions or compensating differentials common to countries or industries, these patterns suggest (for MRPL) either that moving a worker from one country to another is "costlier" than moving the worker from one industry to another – that is, countries are more segmented than industries and therefore differences in levels of MRPL are higher across countries than across industries and these differences are reflected in fixed effects – or that quality differences across workers are larger between countries than between industries. Indeed, the R^2 in the regression with country, industry and year fixed effects is similar to the R^2 with country fixed effects only, which suggests that industry is not likely to be the main driver of MRPL dispersion across countries. This is also consistent with empirical evidence that labor supply to an industry is more elastic than to a country. On the other hand, for capital the increment in the R^2 with country, industry and year fixed effects relative to the regression with no fixed effect is approximately equal to the sum of R^2 increments in the regression with country fixed effects and in regression with industry fixed effects relative to the regression with no fixed effects. Since the increment is somewhat larger for the regression with the industry fixed effect than country fixed effect, the interpretation is that moving a unit of capital from one country to another is "cheaper/easier" than moving it from one industry to another, or that quality differences in capital are smaller between countries than between industries.

Finally, there is a large increase in the R^2 when we introduce country × industry × year fixed effects: R^2 is 0.492 for MRPK and 0.736 for MRPL (column 6 of Table 6). Again, these results are consistent with two explanations. First, there is an additional barrier to move a worker or a unit of capital across countries *and* industries relative to moving a worker or a unit of capital across countries but within an industry or across industries within a country. Second, there is an additional quality difference when workers or capital are compared across industries *and* countries.

Irrespective of which view is taken, it is clear that there are quantitatively important complementarities in industry and country attributes.

If one interprets country and/or industry fixed effects as capturing barriers and distortions, then the EU is rather fragmented economically. This interpretation suggests that the EU can achieve considerable gains in productivity. For example, removing inequality in average marginal revenue products across countries (i.e., making the country fixed effects be all identical) could raise productivity by 102 percentage log points (row 4, column 1 of Table 6) according to the Hsieh-Klenow framework (equation (3')). Removing barriers between industries *and* countries (i.e., making the country × industry fixed effects all identical) could raise productivity by at least 176 percentage log points (row 4, column 5 of Table 6).

D. Machado-Mata Decomposition

While our analysis so far is helpful for understanding what factors can predict MRPK and MRPL, it is also useful to understand whether the cross-country differences in dispersion are brought about by differences in firm characteristics or by the way how these characteristics are translated into differences in marginal revenue products. To address this question, we carry out a Machado and Mata (2005) decomposition of the variance in MRPK and MRPL.[21] We start by using Germany and Greece as two polar cases: $\sigma(\log(MRPK))$ is 0.92 for Germany and 1.64 for Greece, while $\sigma(\log(MRPL))$ is 0.61 for Germany and 0.91for Greece. We decompose the distributions of MRPK and MRPL, respectively, into effects that are due to the values of the explanatory variables $\boldsymbol{X}$ ("endowments") and effects that are due to the coefficients $\boldsymbol{b}$ ("prices") on these variables. This decomposition permits us to assess whether the cross-country differences in the dispersion of marginal revenue products are due to differences in endowments of observed firm characteristics $\boldsymbol{X}$ or to how the business environment, institutions and policies translate ("price") these characteristics via $\boldsymbol{b}$ into outcomes.

[21] This decomposition is implemented as in Gorodnichenko and Sabirianova Peter (2007). For country c we make B = 10,000 independent random draws (with replacement) from the distribution of firm characteristics X so that we generate samples $\{X_{bc}\}_{b=1}^{B}$. We also make B independent random draws (with replacement) from the distribution of quantile regressions $Q_{c\theta}(\log MPRK_{ic}|X_{ic}) = X_{ic}\gamma_{c\theta}$ estimated for each country c and quantile θ separately. Thus, we obtain $\{\gamma_{cb}\}_{b=1}^{B}$. Coefficients $\gamma_{c\theta}$ can be interpreted as prices for observable characteristics of firms. Machado and Mata (2005) show that the generated sequence $\{X_{bc}\gamma_{cb}\}_{b=1}^{B}$ reproduces the distribution of the original series of $\log MPRK_{ic}$. We can also combine $\{X_{bc}\}_{b=1}^{B}$ for country c with $\{\gamma_{db}\}_{b=1}^{B}$ for country d to construct a counterfactual distribution of $\log MPRK_{ic}$ if observables from country c were priced as in country d. Since the number of firms per industry is relatively small for any given country, we use 1-digit industry fixed effects rather than 2-digit industry fixed effects as in Table 3.

In Figure 4, we depict the distribution of Greek MRPK in Panel A and Greek MRPL in Panel B. In each panel, we show the actual distribution using Greek ***X*** and ***b*** (solid black line), as well as a counterfactual distribution using Greek ***X*** and German ***b*** (long-dash, blue line) and a counterfactual distribution using German ***X*** and Greek ***b*** (short-dash, red line). Using Greek ***X*** and German ***b*** results in a less dispersed distribution of both MRPK and MRPL, suggesting that German business, institutional and policy environment would increase the efficiency of Greek firms by reducing the dispersion of marginal products of capital and labor across firms. In other words, German "prices" help increase the equalization of returns across firms. Indeed, the standard deviation of this counterfactual distribution is much closer to the actual distribution of marginal revenue products in Germany (e.g., for MRPK the counterfactual standard deviation for Greece is 0.94 rather than 1.66).

When we use German ***X*** and Greek ***b***, the distribution of MRPK is more dispersed and shifts to the right. The latter is consistent with German firms having characteristics associated with high levels of productivity. The former suggests that the dispersion of firm characteristics in Germany is greater than the corresponding dispersion in Greece which, when combined with the Greek business, institutional and policy environment ("prices"), results in a wider dispersion of marginal products than is actually observed in Greece. Interestingly, using German ***X*** and Greek ***b*** does not generate large differences in the mean or dispersion of MRPL. This pattern suggests that differences in firm characteristics are not likely to be a key determinant of German vs. Greek differences in the dispersion of MRPL. In contrast, using German ***b*** with Greek ***X*** not only reduces dispersion of MRPL but also increases the mean value of MRPL.

Our decomposition exercise suggests that German business, institutional and policy environment is the main reason for the smaller dispersion of marginal revenue products in Germany relative to Greece. We generalize this result by showing in Table 7 for each EU country the standard deviation of MRPK and MRPL when we use the country's own ***X*** and ***b*** (column 1) as compared to using (a) German ***X*** or ***b*** (columns 2 and 3 for MRPK and columns 7 and 8 for MRPL) and (b) Greek ***X*** or ***b*** (columns 4 and 5 for MRPK and columns 9 and 10 for MRPL). We find that using German ***b*** tends to reduce the dispersion of MRPK for most countries, while using German ***X*** tends to increase it. This suggests that relative to other countries Germany has more diverse firm characteristics but the business, institutional and policy environment is relatively effective in ensuring that marginal returns are not very different across firms. In contrast, other countries have relatively more homogenous firm characteristics or, at least, have more homogeneity for characteristics with large variation in "prices" (that is, steep slopes in ***X***). Core EU countries, such as France and Denmark, exhibit relatively little sensitivity to using German ***X*** or ***b***, while countries of the EU periphery, such as Portugal and Ireland, show relatively large movements in the counterfactual dispersions of marginal revenue products.

As may also be seen in Table 7, when we combine Greek ***b*** with ***X*** for a given country, the counterfactual distributions tend to increase considerably, as they did in the Germany and Greece comparison. Similarly, using Greek ***X*** with ***b*** for a given country tends to increase (but to a smaller extent) the dispersion of marginal revenue products across firms. These results suggest that the Greek business, policy and institutional environment would be relatively ineffective in reducing the dispersion of marginal returns across firms.

In sum, while there is heterogeneity in the allocation of firm characteristics across countries, the primary source of cross-country differences in the dispersion of marginal revenue products is how these characteristics are converted into outcomes ("priced") via the business, institutional and policy environment. In particular, we observe that Germany and similar countries are more effective in equalizing returns even across heterogeneous firms than Greece and similar countries.

VI. Concluding remarks

Misallocation of resources is often seen as an important reason for the slowdown in productivity growth in Europe, the Unites States and other advanced economies. Using data from the unique EIB Investment Survey (EIBIS) of firms in the 28 EU countries, we go beyond existing studies by using firm-level data to explain *why* there is variation in marginal revenue products. Apart from presenting new cross-country evidence on allocation of resources, we propose a novel approach to quantify potential productivity gains from better allocations. This approach does not rely on exogenous variation in measured frictions (or compensating differentials) so that researchers can apply it in a wide range of settings.

Using a simple dynamic theoretical framework as a guide, we find that there is a sizable dispersion of marginal products across the firms in our sample. If one took the 28 EU countries as a single market, where marginal products ought to be equalized, then the current state of Europe is very far from that. Our calculations suggest that by removing frictions in the EU could increase EU productivity by more than 40 percent. Thus, we find large "costs of non-Europe" induced by frictions and distortions related to incomplete integration of the EU single market, which still persist 26 years after its inception.

Much of the overall dispersion in marginal products could be attributed to fixed differences among countries or sectors/industries. For example, if one removed the dispersion in marginal products across countries (i.e., make the country fixed effects in the regression analysis all identical), EU productivity could rise by 102 percentage log points. We also find that the significant association between marginal products and firm characteristics is predominantly driven by variables measuring firm demographics, quality of inputs, utilization of resources, and dynamic adjustment of inputs. In contrast, the contribution of direct measures of "barriers and constraints" to cross-sectional variation in MRPK and MRPL seems to be modest. Finally, we show that cross-country variation in the within-country dispersion of marginal revenue products is largely rationalized by differences in how a country's business, institutional and policy environment translates firm characteristics into outcomes than by differences in firm characteristics *per se*.

Our work contributes to the growing literature measuring misallocation of resources, provides new insights into the nascent literature on *sources* of observed dispersion in marginal products, documents that various firm characteristics and measures of distortions have predictive power for marginal revenue products, contributes to recent efforts to assess the importance of measurement errors in observed marginal products, and relates a large literature on the dispersion of earnings across workers to the studies of dispersion of marginal products across firms. Future research should make progress by further combining administrative and survey data to reduce measurement errors, generate direct measures of distortions and compensating differentials, and improve identification of causal effects.

References

Asker, John, Allan Collard-Wexler, and Jan De Loecker (2014). "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy* 122(5): 1013-1063.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta (2013). "Cross-country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review* 103(1): 305–334.

Basu, Susanto, and John G. Fernald (1997). "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy* 105(2): 249-283.

Bento, Pedro, and Diego Restuccia (2017). "Misallocation, Establishment Size, and Productivity." *American Economic Journal: Macroeconomics* 9 (3): 267-303.

Bils, Mark, Peter J. Klenow, and Cian Ruane (2017). "Misallocation or Mismeasurement?" mimeo.

Brutscher, Philipp-Bastian, and Andrea Coali (2019). "EIB Group Survey on Investment and Investment Finance – A Data Quality Assessment," EIB Working Paper 2019/06.

David, Joel M., and Venky Venkateswaran (2017). "The Sources of Capital Misallocation," NBER Working Paper 23129.

Desai, Mihir A., C. Fritz Foley, and James R. Hines (2004). "A Multinational Perspective on Capital Structure Choice and Internal Capital Markets." *Journal of Finance* 59(6): 2451-2487.European Parliament, "Europe's Two Trillion Euro Dividend: Mapping the Cost of Non-Europe, 2019-24", European Added Value Unit, Dir C., DG European Parliamentary Research Service, April 2019, 248 pp.

Fama, Eugene, and Kenneth R. French (2002). "Testing Trade-off and Pecking Order Predictions about Dividends and Debt." *Review of Financial Studies* 15(1): 1-33.

Gamberoni, Elisa, Claire Giordano, and Paloma Lopez-Garcia (2016). "Capital and Labour (Mis)allocation in the Euro Area: Some Stylized Facts and Determinants." ECB Working Paper No. 1981.

Garicano, Luis, Claire Lelarge, and John Van Reenen (2016). "Firm Size Distortions and the Productivity Distribution: Evidence from France." *American Economic Review* 106(11): 3439-3479.

Gopinath, Gita, Sebnem Kalemli-Ozcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez (2017). "Capital Allocation and Productivity in South Europe," *Quarterly Journal of Economics*, 132(4): 1915-1967.

Gorodnichenko, Yuriy, and Klara Sabirianova Peter (2007). "Public Sector Pay and Corruption: Measuring Bribery from Micro Data," *Journal of Public Economics* 91(5-6): 963-991.

Heckman James J., Lance J. Lochner, and Petra E. Todd (2006). "Fifty Years of Mincer Earnings Regressions," in Hanushek, E. and F. Welch (eds.), *Handbook of the Economics of Education*, Vol. 1, Elsevier, The Netherlands.

Hsieh, Chang-Tai, and Peter J. Klenow (2009). "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics* 124(4): 1403-1448.

Ipsos (2017). EIB Group Survey of Investment and Investment Finance. Technical Report.

Kalemli-Ozcan, Sebnem, and Bent E. Sørensen (2012). “Misallocation, Property Rights, and Access to Finance: Evidence from within and across Africa,” in S. Edwards, S. Johnson and D.N. Weil (eds.), *African Successes, Volume III: Modernization and Development*, NBER Conference Report, University of Chicago Press, Chicago, IL.

Kalemli-Ozcan, Sebnem, Bent Sorensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yesiltas (2015). “How to Construct Nationally Representative Firm Level data from the ORBIS Global Database,” NBER Working Paper 21558.

Kehrig, Matthias, and Nicolas Vincent (2017) “Do Firms Mitigate or Magnify Capital Misallocation? Evidence from Plant-Level Data.” US Census Bureau Center for Economic Studies Paper No. CES-WP-17-14.

Lipsey, R. E. (2003) “Foreign Direct Investment and the Operations of Multinational Firms: Concepts, History, and Data,” ch. 10 in E. K. Choi and J. Harrigan (eds.), *Handbook of International Trade*, Blackwell Publishing Ltd, Oxford, UK.

Machado, José A. F., and José Mata (2005). “Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression,” *Journal of Applied Econometrics* 20(4): 445-465.

Malchow-Møller, N., Markusen, J. R. and Schjerning, B. (2013), “Foreign Firms, Domestic Wages,” *The Scandinavian Journal of Economics*, 115: 292–325.

Mincer, Jacob (1958). “Investment in Human Capital and Personal Income Distribution,” *Journal of Political Economy* 66 (4): 281-302.

Restuccia, Diego, and Richard Rogerson (2013). “Misallocation and Productivity,” *Review of Economic Dynamics* 16(1): 1-10.

Restuccia, Diego, and Richard Rogerson (2017). “The Causes and Costs of Misallocation,” *Journal of Economic Perspectives* 31(3): 151-74.

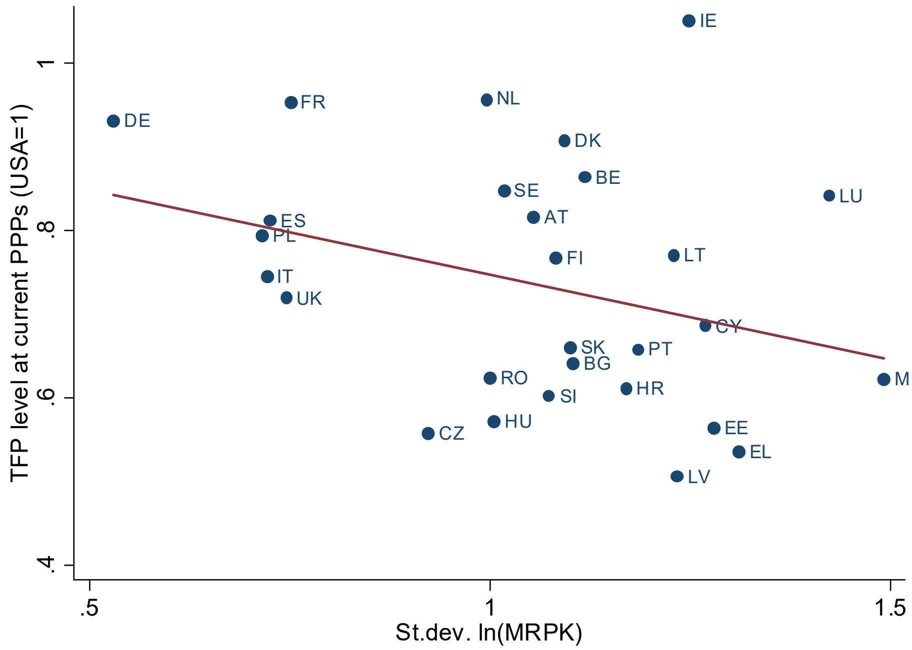
Restuccia, Diego, and Richard. Rogerson (2008). “Policy Distortions and Aggregate Productivity with Heterogenous Establishments,” *Review of Economic Dynamics* 11(4): 707-720.

van Ark, Bart, Mary O'Mahoney, and Marcel P. Timmer (2008). “The Productivity Gap between Europe and the United States: Trends and Causes.” *Journal of Economic Perspectives* 22(1): 25-44.

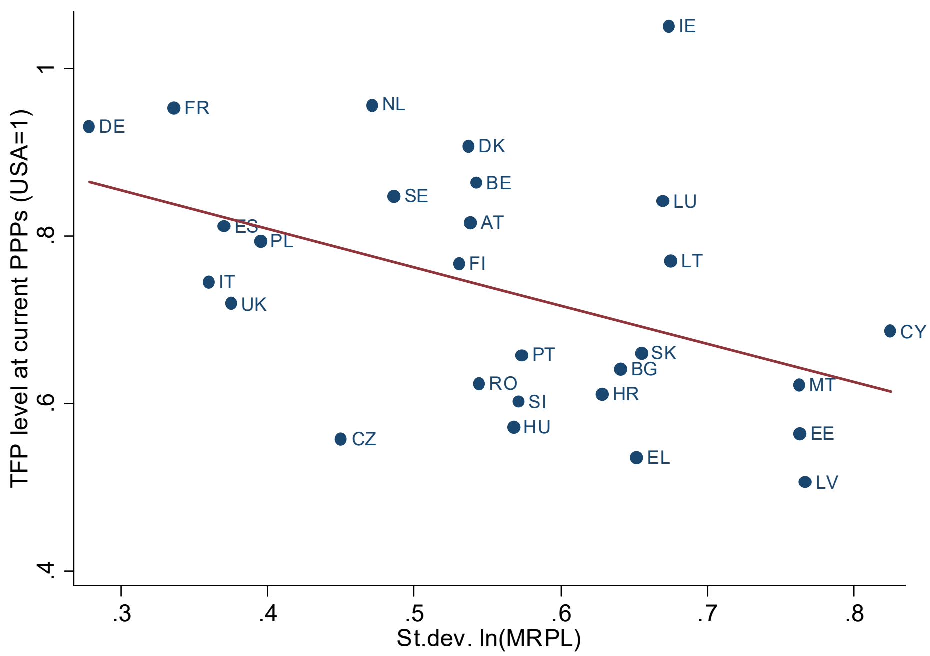
Verhoogen, Eric A. (2008). “Inequality in the Mexican Manufacturing Sector,” *Quarterly Journal of Economics* 123(2): 489-530.

Figure 1. Productivity and dispersion of marginal revenue products

Panel A. Marginal revenue product of capital (MRPK)



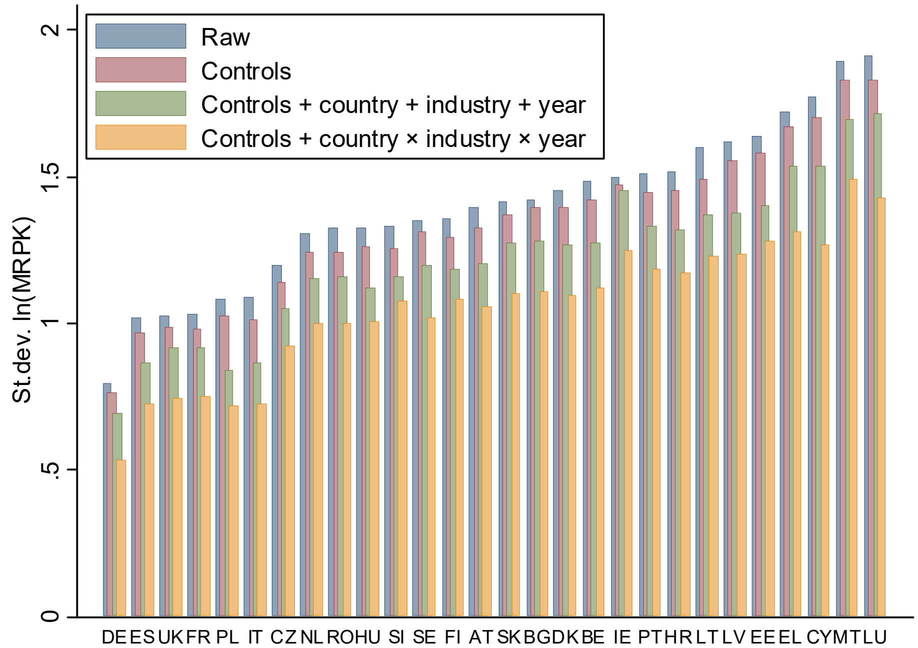
Panel B. Marginal revenue product of labor (MRPL)



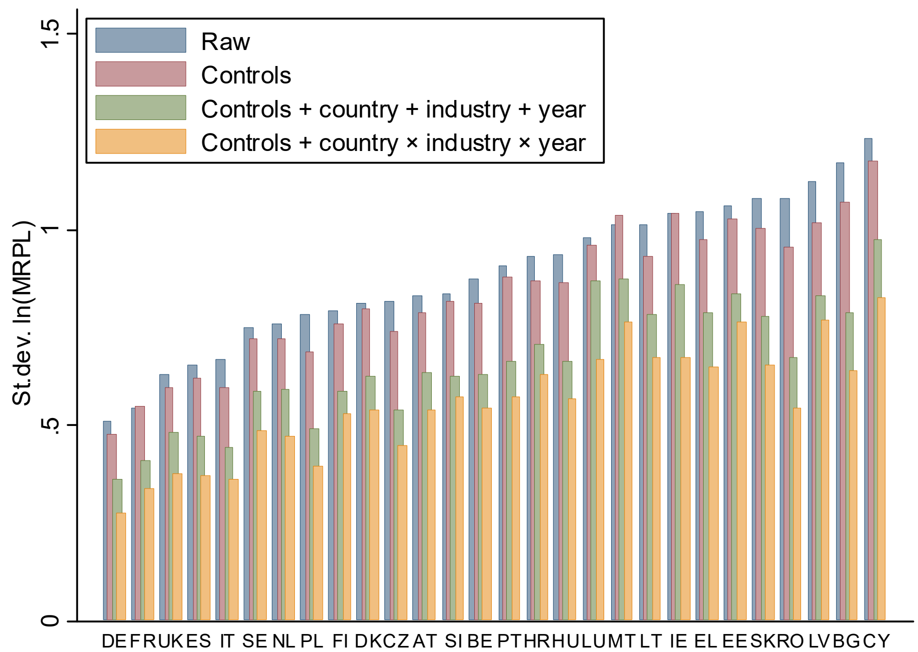
Note: TFP data is for year 2015 from Penn World Tables. Standard deviation of marginal revenue products is computed using EIBIS data. The red, solid line shows the fitted linear regression. The slope of the fitted relationship is -0.20 (s.e. 0.10) for top panel and -0.46 (s.e. 0.15) for top panel. Country codes: AT-Austria, BE-Belgium, BG-Bulgaria, CZ-Czech Republic, CY-Cyprus, DE-Germany, DK-Denmark, EE-Estonia, EL-Greece, ES-Spain, FI-Finland, FR-France, HR-Croatia, HU-Hungary, IE-Ireland, IT-Italy, LT-Lithuania, LU-Luxembourg, LV-Latvia, MT-Malta, NL-Netherlands, PL-Poland, PT-Portugal, RO-Romania, SE-Sweden, SI-Slovenia, SK-Slovakia, UK-United Kingdom.

Figure 2. Raw and residual dispersion of the marginal revenue products of capital and labor

Panel A. Marginal revenue product of capital (MRPK)

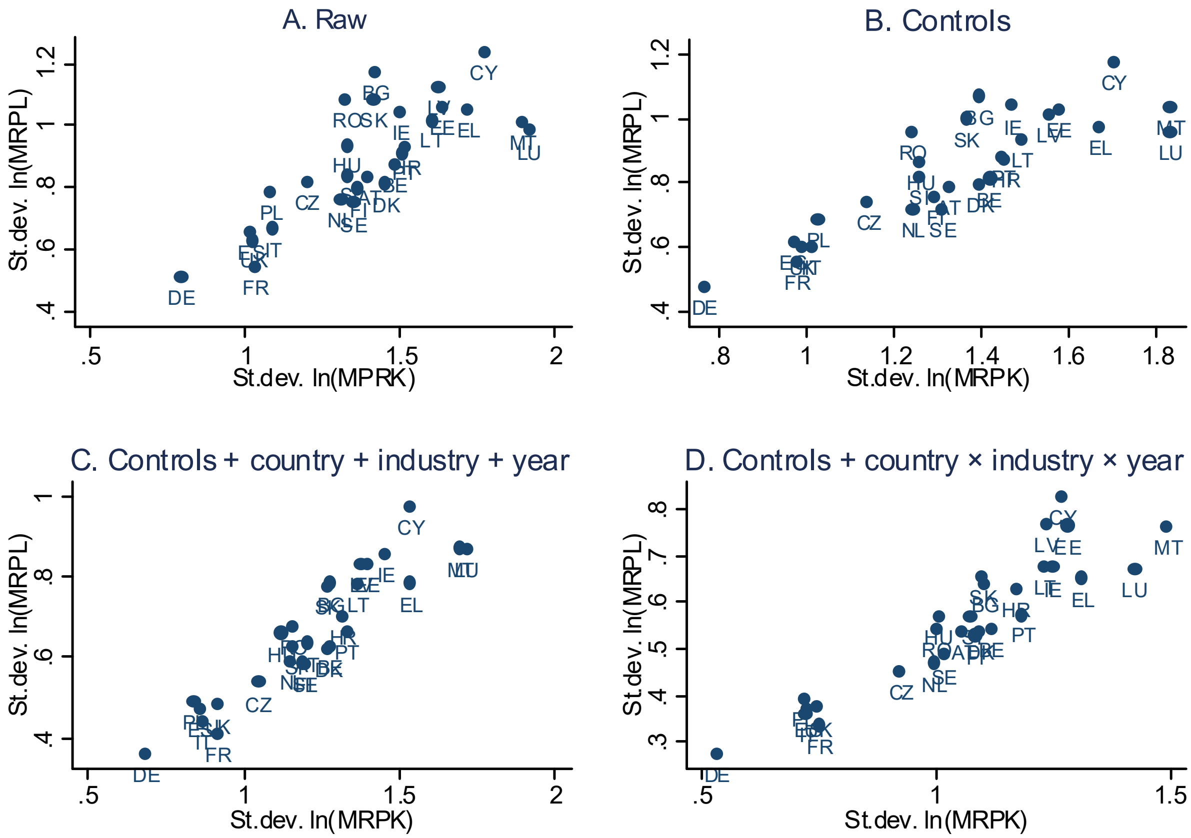


Panel B. Marginal revenue product of labor (MRPL)



Note: The figures show how adding different sets of controls accounts for the dispersion in MRPK and MRPL. "Raw" means no controls. "Controls" include the firm-level characteristics described in section IV. "Controls + country + industry + year" add fixed effects for countries, industries and years to firm-level characteristics (28 countries, industry at 2-digit NACE, 3 years). "Controls + country × industry × year" add the interactions country × industry × year to firm-level characteristics. Country codes: AT-Austria, BE-Belgium, BG-Bulgaria, CZ-Czech Republic, CY-Cyprus, DE-Germany, DK-Denmark, EE-Estonia, EL-Greece, ES-Spain, FI-Finland, FR-France, HR-Croatia, HU-Hungary, IE-Ireland, IT-Italy, LT-Lithuania, LU-Luxembourg, LV-Latvia, MT-Malta, NL-Netherlands, PL-Poland, PT-Portugal, RO-Romania, SE-Sweden, SI-Slovenia, SK-Slovakia, UK-United Kingdom.

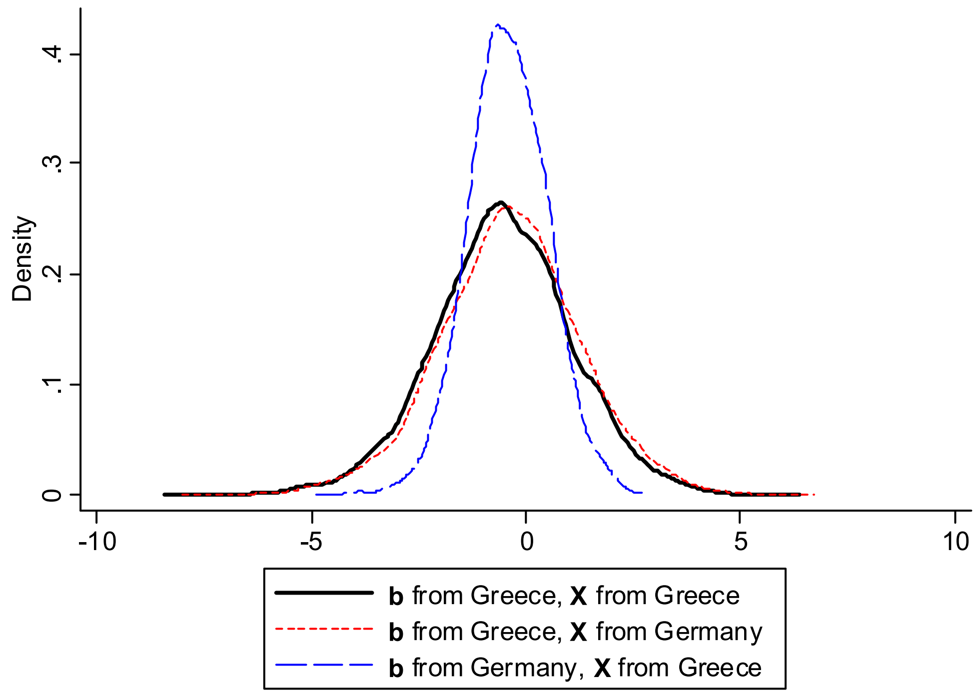
Figure 3. Association of the dispersion of the marginal revenue products of capital and labor.



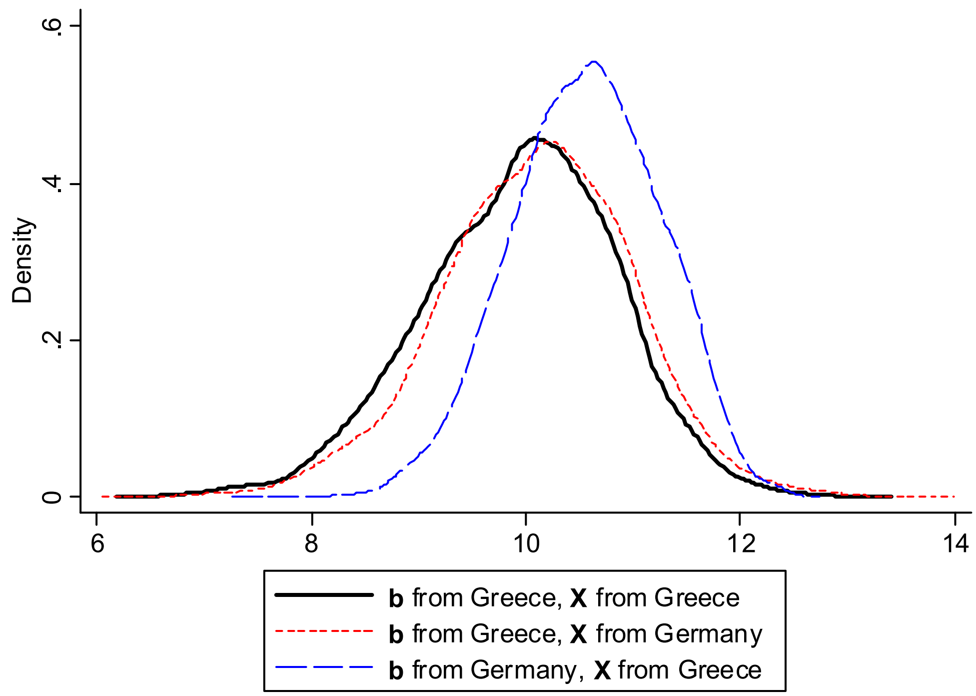
Note: The figures show the association between the dispersion of MRPK and MRPL across countries. "Raw" means no controls. "Controls" include the firm-level characteristics described in section IV. "Controls + country + industry" add fixed effects for countries, industries and years to firm-level characteristics (28 countries, industry at 2-digit NACE, 3 years). "Controls + country × industry × year" add the interactions country × industry × year to firm-level characteristics. Country codes: AT-Austria, BE-Belgium, BG-Bulgaria, CZ-Czech Republic, CY-Cyprus, DE-Germany, DK-Denmark, EE-Estonia, EL-Greece, ES-Spain, FI-Finland, FR-France, HR-Croatia, HU-Hungary, IE-Ireland, IT-Italy, LT-Lithuania, LU-Luxembourg, LV-Latvia, MT-Malta, NL-Netherlands, PL-Poland, PT-Portugal, RO-Romania, SE-Sweden, SI-Slovenia, SK-Slovakia, UK-United Kingdom.

Figure 4. Machado-Mata decomposition of the marginal revenue products of capital and labor for Greece

Panel A. Marginal revenue product of capital (MRPK)



Panel B. Marginal revenue product of labor (MRPL)



Note: The figures show actual and counterfactual distributions of the log marginal revenue product of capital (Panel A) and marginal revenue product of labor (Panel B).

Table 1. Dispersion of sales, fixed assets and employment in Orbis and EIB Investment Survey (EIBIS), by country.

		log(sales)			log(fixed assets)			log(employment)		
	Sample	St. dev.		Correl.	St. dev.		Correl.	St. dev.		Correl.
Country	size	Orbis	EIBIS	coeff.	Orbis	EIBIS	coeff.	Orbis	EIBIS	coeff.
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Austria	771	2.05	2.38	0.83	2.72	2.85	0.84	1.93	2.17	0.70
Belgium	1,111	1.83	2.21	0.86	2.90	2.74	0.87	1.95	2.18	0.85
Bulgaria	1,164	2.44	2.47	0.90	2.83	2.88	0.89	1.86	1.81	0.98
Cyprus	320	1.75	1.92	0.94	2.27	2.25	0.85	1.63	1.63	0.99
Czech Rep.	978	2.02	2.23	0.90	2.29	2.58	0.86	1.71	1.72	0.95
Germany	825	2.05	2.21	0.91	2.60	2.52	0.79	1.72	1.98	0.87
Denmark	1,105	2.68	2.52	0.80	3.27	3.01	0.91	2.16	2.36	0.88
Estonia	990	2.14	2.23	0.94	2.67	2.42	0.85	1.73	1.87	0.97
Greece	1,125	2.25	2.37	0.82	3.18	2.82	0.89	2.21	1.95	0.93
Spain	1,035	2.26	2.37	0.94	2.86	2.82	0.90	2.10	2.14	0.96
Finland	1,367	2.73	2.64	0.95	3.37	3.09	0.93	2.54	2.42	0.94
France	1,194	2.03	2.17	0.93	2.55	2.76	0.84	1.58	1.91	0.94
Croatia	1,218	2.13	2.18	0.82	2.87	2.74	0.86	1.88	1.93	0.96
Hungary	1,138	2.46	2.49	0.94	2.84	2.77	0.89	1.97	1.92	0.99
Ireland	845	1.76	2.10	0.97	2.75	2.47	0.68	1.71	1.99	0.69
Italy	1,361	2.44	2.54	0.96	3.03	2.96	0.90	2.12	2.18	0.96
Lithuania	977	1.98	2.05	0.94	2.40	2.68	0.86	1.65	1.50	0.97
Luxembourg	352	1.71	1.75	0.77	2.77	2.36	0.53	1.28	1.61	0.92
Latvia	995	2.31	2.37	0.80	2.96	3.13	0.87	1.71	1.82	0.86
Malta	415	2.00	1.81	0.17	2.47	2.24	0.66	0.64	1.50	0.95
Netherlands	1,128	1.63	2.26	0.97	3.03	2.52	0.71	1.99	2.05	0.88
Poland	1,086	1.84	2.01	0.90	2.38	2.38	0.85	2.11	1.77	0.97
Portugal	1,259	2.24	2.28	0.92	2.75	2.74	0.81	1.71	1.90	0.97
Romania	931	1.97	2.16	0.89	2.77	2.75	0.81	1.81	1.62	0.90
Sweden	1,143	2.29	2.31	0.92	3.06	2.67	0.85	2.04	2.07	0.93
Slovenia	1,104	2.18	2.05	0.94	2.65	2.46	0.91	1.95	1.80	0.95
Slovakia	832	2.03	2.34	0.95	2.49	2.62	0.82	1.74	1.83	0.95
UK	1,047	2.19	2.46	0.86	2.77	2.68	0.88	1.92	2.17	0.78
All countries	27,816	2.14	2.30	0.91	2.76	2.70	0.85	1.90	2.00	0.91

Note: Dispersion of the logarithm of sales, fixed assets and employment, by country and data source (ORBIS and EIB Investment Survey). Columns (4), (7) and (10) report correlation between the logarithm of sales, fixed assets and employment across the two data sources. All statistics are computed using sampling weights.

Table 2. Dispersion of sales, fixed assets and employment in Orbis and EIB Investment Survey (EIBIS), by industry.

NACE industry code	NACE industry name	Sample size	log(sales)			log(fixed assets)			log(employment)		
			St. dev.		Correl. coeff.	St. dev.		Correl. coeff.	St. dev.		Correl. coeff.
			Orbis	EIBIS		Orbis	EIBIS		Orbis	EIBIS	
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
10-12	food; beverages; tobacco	1,216	2.18	2.27	0.96	2.29	2.39	0.85	1.90	1.79	0.76
13-15	textiles; apparel; leather and related products	677	1.94	2.02	0.96	2.32	2.35	0.86	1.40	1.50	0.88
16-18	wood; paper; printing and recorded media	888	2.34	2.45	0.89	2.58	2.58	0.77	1.86	1.89	0.93
19-20	coke and refined petroleum; chemicals	315	1.95	1.86	0.92	3.03	2.11	0.78	1.60	1.54	0.94
21	pharmaceutical products	78	2.01	2.20	0.80	2.54	2.37	0.87	1.53	1.41	0.94
22-23	rubber and plastic products; mineral products	960	1.76	2.11	0.89	2.04	2.28	0.77	1.46	2.00	0.84
24-25	basic and fabricated metal products	1,535	2.06	1.94	0.95	2.32	2.17	0.88	1.58	1.53	0.95
26	computer, electronic and optical products	317	2.40	2.46	0.98	3.09	2.87	0.94	2.17	2.14	0.97
27	electrical equipment	375	1.91	2.02	0.96	2.19	2.13	0.87	1.81	2.21	0.96
28	machinery and equipment	931	1.97	2.08	0.94	2.37	2.23	0.90	1.75	1.87	0.93
29-30	motor vehicles; other transport equipment	335	2.06	1.91	0.88	1.94	2.11	0.87	1.54	1.56	0.88
31-33	furniture; other manuf.; repair and installation	763	2.13	2.23	0.91	2.46	2.39	0.82	1.86	1.90	0.97
35	electricity, gas, steam and air conditioning	565	2.70	2.78	0.88	2.99	3.16	0.91	2.16	2.20	0.92
36-39	water supply; sewerage and waste management	1,140	1.93	2.33	0.90	2.58	2.84	0.82	1.70	1.89	0.94
41	construction of buildings	2,040	2.51	2.59	0.93	2.77	2.63	0.78	1.89	1.85	0.91
42	civil engineering	1,026	2.41	2.41	0.93	2.49	2.59	0.79	1.96	2.00	0.93
43	specialised construction activities	3,210	2.00	1.92	0.94	2.29	2.23	0.73	1.65	1.58	0.95
45	wholesale and retail trade	755	2.15	2.47	0.83	2.22	2.61	0.81	1.61	1.99	0.92
46	wholesale trade, except of motor vehicles	2,962	2.13	2.12	0.94	2.56	2.41	0.82	1.79	1.91	0.86
47	retail trade, except of motor vehicles	1,804	2.53	2.53	0.93	2.82	2.83	0.82	2.30	2.28	0.97
49-53	transportation and storage	3,813	2.41	2.37	0.86	3.07	2.78	0.84	2.01	2.12	0.86
55-56	accommodation and food service activities	1,032	2.16	2.24	0.94	2.60	2.74	0.82	2.00	2.00	0.95
58-63	information and communication	1,021	2.41	2.43	0.96	3.38	2.90	0.84	1.91	1.88	0.96
64-99	other services	58	1.72	2.42	0.64	2.81	2.65	0.95	1.94	2.16	-0.25
10-99	all industries	27,816	2.20	2.25	0.92	2.58	2.53	0.83	1.84	1.92	0.90

Note: Dispersion of the logarithm of sales, fixed assets and employment, by country and data source (ORBIS and EIB Investment Survey). Columns (4), (7) and (10) report correlation between the logarithm of sales, fixed assets and employment across the two data sources. All statistics are computed using sampling weights.

Table 3. Descriptive statistics

Group of variables	Variable	Mean	St. dev.
Outcome variables	log(sales)	16.58	2.37
	log(fixed assets)	15.31	2.75
	log(employment)	4.84	2.01
	log(MRPK)	-0.42	1.43
	log(MRPL)	10.22	1.19
	log(MRPL) - log(MRPK)	10.65	1.63
Demographics	Firm age		
	less than 5 years	0.04	0.18
	5-9 years	0.08	0.26
	10-19 years	0.20	0.40
	20+ years	0.69	0.46
	Subsidiary	0.34	0.47
	Exporter	0.52	0.50
Quality of capital and other inputs	Share of state-of-the art machinery and equipment	0.41	0.32
	Share of high energy efficiency commercial building stock	0.36	0.34
Capacity utilization	above maximum capacity	0.06	0.24
	at maximum capacity	0.44	0.50
	somewhat below full capacity	0.40	0.49
	substantially below full capacity	0.08	0.27
Obstacles to investment	Demand for products or services		
	Major	0.24	0.43
	Minor	0.26	0.44
	Availability of staff with the right skills		
	Major	0.43	0.49
	Minor	0.31	0.46
	Energy costs		
	Major	0.24	0.42
	Minor	0.34	0.47
	Access to digital infrastructure		
	Major	0.12	0.32
	Minor	0.29	0.45
	Labor market regulations		
	Major	0.30	0.46
	Minor	0.33	0.47
	Business regulations and taxation		
	Major	0.32	0.47
	Minor	0.33	0.47
	Availability of adequate transport infrastructure		
	Major	0.17	0.37
	Minor	0.28	0.45
	Availability of finance		
	Major	0.21	0.41
	Minor	0.25	0.43
	Uncertainty about future		
	Major	0.38	0.49
	Minor	0.35	0.48
Adjustment	Investment, log(1 + investment)	12.44	3.96
	Percent change in employment in the last three years	0.13	0.44
	Investment over the last three years		
	too much	0.04	0.18
	about the right amount	0.79	0.41
	too little	0.17	0.37
	company did not exist three years ago	0.00	0.03
	Investment priority in the next three years		
	replacing capacity	0.35	0.48
	capacity expansion for existing products or services	0.28	0.45
	developing new products, processes or services	0.26	0.44
	no investment planned	0.09	0.29
Source of funds	internal funds or retained earnings	0.66	0.37
	external finance	0.31	0.35
	intra-group funding	0.02	0.12
	Finance constrained	0.07	0.25
Sample size		27,816	27,816

Note: All statistics are computed using sampling weights.

Table 4. Predictors of the dispersion of the marginal revenue products of capital and labor.

Regressor	Dependent variable	
	log(MRPK)	log(MRPL)
Demographics		
Firm age (omitted category: less than 5 years)		
5-9 years	-0.009	0.041**
	(0.039)	(0.019)
10-19 years	-0.218***	0.062***
	(0.035)	(0.018)
20+ years	-0.339***	0.097***
	(0.034)	(0.017)
log(employment)	0.031***	-0.024***
	(0.005)	(0.003)
Subsidiary	0.351***	0.139***
	(0.019)	(0.010)
Exporter	0.135***	0.241***
	(0.016)	(0.009)
Quality of capital and other inptus		
Share of state-of-the art machinery and equipment, including ICT	-0.123***	0.140***
	(0.023)	(0.012)
Share of high energy efficiency commercial building stock	-0.260***	0.055***
	(0.020)	(0.010)
Capacity utilization (omitted category: somewhat below full capacity)		
above maximum capacity	0.255***	0.099***
	(0.026)	(0.014)
at maximum capacity	0.134***	0.045***
	(0.014)	(0.007)
substantially below full capacity	-0.301***	-0.125***
	(0.022)	(0.012)
Obstacles to investment (omitted category: not an obstacle at all)		
Demand for products or services		
Major	0.078***	0.001
	(0.018)	(0.009)
Minor	0.058***	0.002
	(0.016)	(0.008)
Availability of staff with the right skills		
Major	0.065***	-0.054***
	(0.017)	(0.009)
Minor	0.036**	0.003
	(0.017)	(0.009)
Energy costs		
Major	-0.132***	-0.087***
	(0.018)	(0.009)
Minor	-0.092***	-0.052***
	(0.015)	(0.008)
Access to digital infrastructure		
Major	0.040*	0.021*
	(0.023)	(0.011)
Minor	0.003	0.036***
	(0.016)	(0.008)
Labor market regulations		
Major	0.003	-0.068***
	(0.018)	(0.009)
Minor	-0.018	-0.045***
	(0.016)	(0.008)
Business regulations and taxation		
Major	-0.041**	0.024***
	(0.019)	(0.009)
Minor	0.023	0.029***

	(0.016)	(0.009)
Availability of adequate transport infrastructure		
Major	-0.030	0.077***
	(0.019)	(0.010)
Minor	0.013	0.050***
	(0.015)	(0.008)
Availability of finance		
Major	-0.058***	-0.084***
	(0.018)	(0.010)
Minor	-0.006	-0.068***
	(0.015)	(0.008)
Uncertainty about future		
Major	0.052***	0.043***
	(0.018)	(0.010)
Minor	0.042**	0.035***
	(0.017)	(0.009)
Adjustment		
Investment, log(1 + investment)	-0.057***	0.049***
	(0.003)	(0.001)
Percent change in employment in the last three years	0.090***	-0.109***
	(0.015)	(0.008)
Investment over the last three years (omitted category: about the right amount)		
too much	-0.232***	-0.098***
	(0.030)	(0.016)
too little	-0.063***	-0.067***
	(0.016)	(0.008)
company did not exist three years ago	-0.269	-0.084
	(0.179)	(0.090)
Investment priority in the next three years (omitted category: no investment planned)		
replacing capacity	-0.054**	-0.028**
	(0.023)	(0.012)
capacity expansion for existing products or services	-0.068***	-0.006
	(0.024)	(0.012)
developing new products, processes or services	-0.056**	-0.007
	(0.025)	(0.012)
Source of funds (omitted category: external finance)		
internal funds or retained earnings	0.189***	0.066***
	(0.020)	(0.010)
intra-group funding	-0.124**	0.183***
	(0.062)	(0.031)
Credit constrained	-0.099***	-0.087***
	(0.024)	(0.012)
Sample size	27,816	27,663
R^2	0.527	0.776
Memorandum		
R^2 with country × industry × year fixed effects and no ***X***	0.492	0.736
R^2 with ***X*** and no fixed effects	0.138	0.289
R^2 with ***X*** and country fixed effects	0.170	0.526
R^2 with ***X*** and industry fixed effects	0.302	0.461
R^2 with ***X*** and year fixed effects	0.139	0.288
R^2 with ***X*** and country fixed effects, industry fixed effects and year fixed effects	0.329	0.676
R^2 with ***X*** and slopes varying by country	0.221	0.459
R^2 with ***X*** and slopes varying by industry	0.301	0.445
R^2 with ***X*** and slopes varying by year	0.141	0.292

Note: The table reports estimates of equation (4) with country × industry × year fixed effects. Industries are defined at 2-digit NACE level. All estimates are based on Huber robust regression. Observations are weighted so that the sample represents the population in terms of employment. Standard errors are clustered by industry and country. ***, **, * denote statistical significance at 1, 5 and 10 percent levels.

Table 5. Marginal R^2 of adding a group of variables to a specification with fixed effects.

Row	Group of variables	List of fixed effects: No fixed effects	Country	Industry	Year	Country + Industry + Year	Country × Industry × Year
		(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: MRPK						
(1)	Demographics	0.049	0.039	0.022	0.049	0.018	0.013
(2)	Quality of capital	0.016	0.009	0.015	0.016	0.008	0.007
(3)	Capacity utilization	0.016	0.016	0.009	0.016	0.010	0.008
(4)	Obstacles to investment	0.024	0.017	0.011	0.024	0.005	0.004
(5)	Adjustment	0.063	0.064	0.017	0.062	0.017	0.013
(6)	Source of funds	0.027	0.022	0.015	0.027	0.011	0.008
(7)	"Compensating differentials"	0.094	0.088	0.042	0.094	0.036	0.028
(8)	"Distortions"	0.085	0.070	0.037	0.085	0.029	0.021
(9)	All variables	0.138	0.124	0.070	0.138	0.058	0.045
	Panel B: MRPL						
(10)	Demographics	0.139	0.039	0.101	0.139	0.039	0.028
(11)	Quality of capital	0.030	0.008	0.020	0.030	0.007	0.005
(12)	Capacity utilization	0.014	0.009	0.011	0.014	0.006	0.004
(13)	Obstacles to investment	0.076	0.026	0.047	0.076	0.012	0.009
(14)	Adjustment	0.101	0.039	0.094	0.101	0.036	0.025
(15)	Source of funds	0.043	0.008	0.031	0.043	0.005	0.004
(16)	"Compensating differentials"	0.131	0.057	0.115	0.131	0.054	0.038
(17)	"Distortions"	0.186	0.051	0.132	0.185	0.036	0.026
(18)	All variables	0.289	0.092	0.218	0.288	0.072	0.052
	Panel C: MRPL - MRPK						
(19)	Demographics	0.100	0.061	0.054	0.100	0.027	0.020
(20)	Quality of capital	0.042	0.019	0.037	0.042	0.016	0.014
(21)	Capacity utilization	0.003	0.005	0.001	0.003	0.001	0.001
(22)	Obstacles to investment	0.029	0.016	0.014	0.029	0.005	0.005
(23)	Adjustment	0.188	0.122	0.119	0.187	0.067	0.048
(24)	Source of funds	0.022	0.012	0.017	0.022	0.009	0.007
(25)	"Compensating differentials"	0.218	0.140	0.145	0.218	0.080	0.061
(26)	"Distortions"	0.134	0.076	0.079	0.133	0.036	0.027
(27)	All variables	0.289	0.173	0.197	0.288	0.100	0.078
	Panel D: Productivity gain						
(28)	Demographics	0.331	0.105	0.234	0.330	0.093	0.066
(29)	Quality of capital	0.079	0.023	0.056	0.080	0.020	0.015
(30)	Capacity utilization	0.031	0.020	0.024	0.031	0.014	0.008
(31)	Obstacles to investment	0.171	0.060	0.104	0.171	0.027	0.020
(32)	Adjustment	0.281	0.125	0.242	0.280	0.101	0.071
(33)	Source of funds	0.099	0.021	0.072	0.099	0.013	0.011
(34)	"Compensating differentials"	0.356	0.170	0.295	0.356	0.143	0.102
(35)	"Distortions"	0.442	0.134	0.309	0.441	0.089	0.065
(36)	All variables	0.715	0.256	0.532	0.714	0.189	0.139

Note: The table reports change in R^2 in equation (4) when a group of variables is added to a specification with a given combination of industry and/or country and/or year fixed effects. Industries are defined at 2-digit NACE level. All estimates are based on Huber robust regression. Observations are weighted so that the sample represents the population in terms of employment. Standard errors are clustered by industry and country. The group "compensating differentials" includes "quality of capital", "capacity utilization" and "adjustment". The group "distortions" includes "demographics,", "obstacles for investment" and "source of funds". Productivity gain is computed according to equation (3').

Table 6. R^2 for various sets of fixed effects.

		List of fixed effects				
		Country	Industry	Year	Country + Industry + Year	Country × Industry × Year
		(2)	(3)	(4)	(5)	(6)
	Dispersion					
(1)	MRPK	0.170	0.239	0.000	0.275	0.492
(2)	MRPL	0.445	0.268	0.001	0.611	0.736
(3)	MRPL - MRPK	0.219	0.174	0.000	0.354	0.555
(4)	**Productivity gain**	1.023	0.631	0.002	1.423	1.760

Note: The table reports R^2 in equation (4) when a group of fixed effects is added to a specification with no other controls. Industries are defined at 2-digit NACE level. All estimates are based on Huber robust regression. Observations are weighted so that the sample represents the population in terms of employment. Productivity gain is computed according to equation (3').

Table 7. Machado-Mata decomposition of the marginal revenue products of capital and labor.

	$\sigma(MRPK)$					$\sigma(MRPL)$				
Country b	Own	Germany	Own	Greece	Own	Own	Germany	Own	Greece	Own
Country X	Own	Own	Germany	Own	Greece	Own	Own	Germany	Own	Greece
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Austria	1.38	0.91	1.43	1.62	1.62	0.75	0.62	0.76	0.91	0.83
Belgium	1.43	0.92	1.56	1.65	1.63	0.79	0.63	0.78	0.91	0.88
Bulgaria	1.34	0.98	1.36	1.70	1.56	1.02	0.67	1.01	0.90	1.08
Cyprus	1.72	0.94	1.77	1.75	1.83	1.15	0.67	1.36	0.94	1.38
Czech Rep.	1.15	0.86	1.37	1.68	1.47	0.73	0.62	0.77	0.82	0.79
Germany	0.92	0.92	0.92	1.66	0.94	0.61	0.61	0.61	0.91	0.69
Denmark	1.40	0.88	1.43	1.74	1.55	0.75	0.64	0.75	0.89	0.81
Estonia	1.56	0.98	1.53	1.75	1.70	1.03	0.65	0.97	0.90	1.08
Greece	1.64	0.94	1.66	1.64	1.64	0.91	0.69	0.91	0.91	0.91
Spain	1.04	0.88	1.14	1.59	1.14	0.71	0.64	0.71	0.90	0.78
Finland	1.26	0.93	1.33	1.75	1.56	0.75	0.61	0.80	0.88	0.95
France	1.09	0.91	1.08	1.73	1.07	0.60	0.65	0.59	0.91	0.62
Croatia	1.41	0.94	1.44	1.60	1.61	0.83	0.63	0.82	0.89	0.97
Hungary	1.29	0.93	1.32	1.71	1.41	0.85	0.64	0.84	0.91	0.94
Ireland	1.44	0.90	1.42	1.64	1.45	0.95	0.64	1.00	0.88	1.03
Italy	1.11	0.92	1.06	1.59	1.20	0.67	0.66	0.69	0.91	0.72
Lithuania	1.56	0.96	1.56	1.72	1.61	0.91	0.68	0.91	0.93	1.01
Luxembourg	1.88	0.91	2.39	1.73	2.47	0.88	0.66	1.08	0.93	1.08
Latvia	1.57	0.97	1.62	1.63	1.69	1.05	0.66	0.99	0.89	1.14
Malta	1.84	0.93	1.98	1.71	1.99	0.97	0.64	1.17	0.88	1.27
Netherlands	1.31	0.88	1.43	1.71	1.61	0.71	0.64	0.68	0.94	0.75
Poland	1.10	0.93	1.11	1.56	1.21	0.67	0.61	0.76	0.85	0.81
Portugal	1.41	0.93	1.52	1.57	1.63	0.82	0.65	0.92	0.90	0.90
Romania	1.31	0.96	1.33	1.63	1.43	0.88	0.65	0.92	0.88	0.96
Sweden	1.31	0.92	1.34	1.75	1.45	0.71	0.65	0.62	0.96	0.67
Slovenia	1.27	0.92	1.36	1.62	1.52	0.78	0.62	0.80	0.87	0.89
Slovakia	1.33	0.93	1.38	1.62	1.46	0.98	0.62	1.05	0.87	1.17
UK	1.12	0.89	1.17	1.63	1.20	0.63	0.63	0.63	0.92	0.70

Note: The table reports actual and counterfactual dispersion of marginal revenue products. See section V.D for more details.

Appendix A: Additional Tables

Appendix Table 1. Predictors of the dispersion of the marginal revenue products of capital and labor averaged across waves

Regressor	Dependent variable	
	log(MRPK)	log(MRPL)
Demographics		
Firm age (omitted category: less than 5 years)		
5-9 years	-0.216**	0.578***
	(0.088)	(0.085)
10-19 years	-0.208***	0.623***
	(0.072)	(0.078)
20+ years	-0.380***	0.665***
	(0.073)	(0.077)
log(employment)	0.079***	-0.085***
	(0.011)	(0.009)
Subsidiary	0.544***	0.092***
	(0.041)	(0.030)
Exporter	0.187***	0.324***
	(0.034)	(0.026)
Quality of capital and other inptus		
Share of state-of-the art machinery and equipment, including ICT	-0.028	0.228***
	(0.050)	(0.037)
Share of high energy efficiency commercial building stock	-0.331***	0.103***
	(0.046)	(0.029)
Capacity utilization (omitted category: somewhat below full capacity)		
above maximum capacity	0.312***	0.177***
	(0.072)	(0.054)
at maximum capacity	0.132***	0.108***
	(0.034)	(0.027)
substantially below full capacity	-0.411***	-0.127***
	(0.053)	(0.035)
Obstacles to investment (omitted category: not an obstacle at all)		
Demand for products or services		
Major	0.065*	0.003
	(0.034)	(0.024)
Minor	-0.005	0.043**
	(0.027)	(0.021)
Availability of staff with the right skills		
Major	0.067**	-0.041*
	(0.032)	(0.023)
Minor	0.106***	0.023
	(0.033)	(0.024)
Energy costs		
Major	-0.115***	-0.180***
	(0.032)	(0.025)
Minor	-0.060**	-0.066***
	(0.029)	(0.021)
Access to digital infrastructure		
Major	0.067	0.093***
	(0.043)	(0.031)
Minor	-0.001	-0.007
	(0.031)	(0.022)
Labor market regulations		
Major	0.110***	-0.086***
	(0.051)	(0.023)
Minor	0.010	-0.066***
	(0.032)	(0.024)
Business regulations and taxation		
Major	-0.063*	0.007
	(0.034)	(0.025)

Minor	-0.002	0.043*
	(0.029)	(0.025)
Availability of adequate transport infrastructure		
Major	-0.087**	0.053**
	(0.037)	(0.027)
Minor	-0.012	0.041*
	(0.030)	(0.023)
Availability of finance		
Major	-0.059*	-0.052**
	(0.034)	(0.025)
Minor	-0.021	-0.046**
	(0.028)	(0.022)
Uncertainty about future		
Major	0.005	0.127***
	(0.038)	(0.024)
Minor	0.002	0.088***
	(0.033)	(0.022)
Adjustment		
Investment, log(1 + investment)	-0.096***	0.092***
	(0.007)	(0.005)
Percent change in employment in the last three years	0.094**	-0.041
	(0.039)	(0.034)
Investment over the last three years (omitted category: about the right amount)		
too much	-0.464***	-0.336***
	(0.084)	(0.062)
too little	-0.121***	-0.074**
	(0.041)	(0.031)
company did not exist three years ago	0.761	0.116
	(0.507)	(0.404)
Investment priority in the next three years (omitted category: no investment planned)		
replacing capacity	-0.123**	0.120***
	(0.061)	(0.041)
capacity expansion for existing products or services	-0.149**	0.047
	(0.061)	(0.044)
developing new products, processes or services	-0.137**	0.113***
	(0.065)	(0.042)
Source of funds (omitted category: external finance)		
internal funds or retained earnings	0.257***	0.121***
	(0.049)	(0.036)
intra-group funding	-0.165	0.300***
	(0.178)	(0.111)
Credit constrained	-0.018	-0.145***
	(0.062)	(0.044)
Sample size	6,672	6,628
R^2	0.580	0.749
Memorandum		
R^2 with country × industry fixed effects and no $\boldsymbol{X}$	0.526	0.676
R^2 with $\boldsymbol{X}$ and no fixed effects	0.185	0.283
R^2 with $\boldsymbol{X}$ and country fixed effects	0.244	0.528
R^2 with $\boldsymbol{X}$ and industry fixed effects	0.359	0.506
R^2 with $\boldsymbol{X}$ and country fixed effects and industry fixed effects	0.391	0.671
R^2 with $\boldsymbol{X}$ and slopes varying by country	0.394	0.607
R2 with X and slopes varying by industry	0.511	0.631

Note: The table reports "*between*" estimates of equation (4) with country × industry × year fixed effects. Industries are defined at 2-digit NACE level. All estimates are based on Huber robust regression. Observations are weighted so that the sample represents the population in terms of employment. All RHS and LHS variables are averaged across waves. Standard errors are clustered by industry and country. ***, **, * denote statistical significance at 1, 5 and 10 percent levels.

Table A.2. Predictors of the dispersion of the marginal revenue products of capital and labor, balanced panel

Regressor	Dependent variable	
	log(MRPK)	log(MRPL)
Demographics		
Firm age (omitted category: less than 5 years)		
5-9 years	0.131	0.100**
	(0.095)	(0.051)
10-19 years	0.098	0.097**
	(0.089)	(0.048)
20+ years	-0.116	0.032
	(0.086)	(0.046)
log(employment)	0.026**	-0.040***
	(0.011)	(0.006)
Subsidiary	0.345***	0.125***
	(0.043)	(0.022)
Exporter	0.258***	0.295***
	(0.031)	(0.017)
Quality of capital and other inptus		
Share of state-of-the art machinery and equipment, including ICT	0.049	0.262***
	(0.050)	(0.024)
Share of high energy efficiency commercial building stock	-0.184***	0.136***
	(0.043)	(0.020)
Capacity utilization (omitted category: somewhat below full capacity)		
above maximum capacity	0.364***	0.270***
	(0.058)	(0.029)
at maximum capacity	0.135***	0.038***
	(0.030)	(0.014)
substantially below full capacity	-0.242***	-0.110***
	(0.050)	(0.024)
Obstacles to investment (omitted category: not an obstacle at all)		
Demand for products or services		
Major	0.111***	0.043**
	(0.038)	(0.019)
Minor	0.094***	0.113***
	(0.034)	(0.018)
Availability of staff with the right skills		
Major	0.026	-0.088***
	(0.036)	(0.019)
Minor	0.031	0.004
	(0.035)	(0.018)
Energy costs		
Major	-0.107***	-0.102***
	(0.038)	(0.021)
Minor	0.055	-0.041**
	(0.034)	(0.018)
Access to digital infrastructure		
Major	-0.134***	0.104***
	(0.049)	(0.024)
Minor	-0.117***	0.048***
	(0.034)	(0.016)
Labor market regulations		
Major	0.028	-0.098***
	(0.038)	(0.020)
Minor	0.099***	-0.043**
	(0.035)	(0.018)
Business regulations and taxation		
Major	-0.150***	0.026
	(0.040)	(0.020)

Minor	-0.047	0.058***
	(0.036)	(0.017)
Availability of adequate transport infrastructure		
Major	0.056	-0.009
	(0.042)	(0.021)
Minor	-0.134***	0.023
	(0.033)	(0.017)
Availability of finance		
Major	0.026	-0.169***
	(0.041)	(0.021)
Minor	0.137***	-0.080***
	(0.036)	(0.018)
Uncertainty about future		
Major	-0.064	0.063***
	(0.043)	(0.021)
Minor	0.079**	0.038**
	(0.039)	(0.019)
Adjustment		
Investment, log(1 + investment)	-0.051***	0.048***
	(0.006)	(0.003)
Percent change in employment in the last three years	0.187***	-0.126***
	(0.032)	(0.016)
Investment over the last three years (omitted category: about the right amount)		
too much	-0.294***	-0.071**
	(0.069)	(0.036)
too little	-0.104***	-0.043**
	(0.035)	(0.017)
company did not exist three years ago	-3.549***	-0.312
	(0.471)	(0.298)
Investment priority in the next three years (omitted category: no investment planned)		
replacing capacity	0.070	-0.029
	(0.049)	(0.023)
capacity expansion for existing products or services	0.007	-0.072***
	(0.052)	(0.025)
developing new products, processes or services	0.083	-0.043*
	(0.053)	(0.025)
Source of funds (omitted category: external finance)		
internal funds or retained earnings	0.111***	0.015
	(0.041)	(0.021)
intra-group funding	-0.200	0.155**
	(0.149)	(0.075)
Credit constrained	-0.161***	-0.157***
	(0.048)	(0.024)
Sample size	5,406	5,370
R^2	0.682	0.870
Memorandum		
R^2 with country × industry × year fixed effects and no ***X***	0.642	0.832
R^2 with ***X*** and no fixed effects	0.148	0.283
R^2 with ***X*** and country fixed effects	0.185	0.550
R^2 with ***X*** and industry fixed effects	0.356	0.472
R^2 with ***X*** and year fixed effects	0.149	0.286
R^2 with ***X*** and country fixed effects, industry fixed effects and year fixed effects	0.389	0.703
R^2 with ***X*** and slopes varying by country	0.386	0.583
R^2 with ***X*** and slopes varying by industry	0.475	0.585
R^2 with ***X*** and slopes varying by year	0.160	0.293

Note: The table reports estimates of equation (4) with country × industry × year fixed effects. Industries are defined at 2-digit NACE level. All estimates are based on Huber robust regression. Observations are weighted so that the sample represents the population in terms of employment. The sample is restricted to firms that participated in all three waves of EIBIS. Standard errors are clustered by industry and country. ***, **, * denote statistical significance at 1, 5 and 10 percent levels.

Table A.3. Predictors of the dispersion of the marginal revenue products of capital and labor, alternative measure of investment

Regressor	Dependent variable	
	log(MRPK)	log(MRPL)
Demographics		
Firm age (omitted category: less than 5 years)		
5-9 years	-0.002	0.025
	(0.039)	(0.020)
10-19 years	-0.212***	0.045**
	(0.036)	(0.018)
20+ years	-0.341***	0.074***
	(0.035)	(0.018)
log(employment)	-0.031***	0.025***
	(0.005)	(0.002)
Subsidiary	0.220***	0.198***
	(0.020)	(0.011)
Exporter	0.104***	0.261***
	(0.016)	(0.009)
Quality of capital and other inptus		
Share of state-of-the art machinery and equipment, including ICT	-0.160***	0.166***
	(0.023)	(0.012)
Share of high energy efficiency commercial building stock	-0.270***	0.067***
	(0.020)	(0.010)
Capacity utilization (omitted category: somewhat below full capacity)		
above maximum capacity	0.232***	0.097***
	(0.027)	(0.014)
at maximum capacity	0.126***	0.046***
	(0.014)	(0.007)
substantially below full capacity	-0.304***	-0.136***
	(0.022)	(0.012)
Obstacles to investment (omitted category: not an obstacle at all)		
Demand for products or services		
Major	0.060***	-0.001
	(0.018)	(0.009)
Minor	0.049***	0.003
	(0.016)	(0.008)
Availability of staff with the right skills		
Major	0.047***	-0.044***
	(0.017)	(0.009)
Minor	0.031*	0.005
	(0.017)	(0.009)
Energy costs		
Major	-0.129***	-0.078***
	(0.018)	(0.010)
Minor	-0.096***	-0.049***
	(0.015)	(0.008)
Access to digital infrastructure		
Major	0.034	0.028**
	(0.023)	(0.012)
Minor	-0.005	0.042***
	(0.016)	(0.008)
Labor market regulations		
Major	0.017	-0.071***
	(0.018)	(0.009)
Minor	-0.009	-0.043***
	(0.016)	(0.008)
Business regulations and taxation		
Major	-0.063***	0.024***
	(0.019)	(0.009)

Minor	0.017	0.030***
	(0.017)	(0.009)
Availability of adequate transport infrastructure		
Major	-0.050***	0.077***
	(0.019)	(0.011)
Minor	-0.006	0.057***
	(0.016)	(0.008)
Availability of finance		
Major	-0.023	-0.098***
	(0.018)	(0.010)
Minor	0.013	-0.077***
	(0.015)	(0.009)
Uncertainty about future		
Major	0.079***	0.047***
	(0.018)	(0.010)
Minor	0.056***	0.036***
	(0.017)	(0.009)
Adjustment		
Indicator variable for positive investment, **1**{investment > 0}	0.120***	0.215***
	(0.036)	(0.018)
Percent change in employment in the last three years	0.071***	-0.109***
	(0.015)	(0.008)
Investment over the last three years (omitted category: about the right amount)		
too much	-0.274***	-0.079***
	(0.030)	(0.016)
too little	-0.039**	-0.091***
	(0.016)	(0.008)
company did not exist three years ago	-0.189	-0.118
	(0.185)	(0.091)
Investment priority in the next three years (omitted category: no investment planned)		
replacing capacity	-0.095***	0.006
	(0.024)	(0.012)
capacity expansion for existing products or services	-0.116***	0.031**
	(0.024)	(0.012)
developing new products, processes or services	-0.106***	0.038***
	(0.025)	(0.012)
Source of funds (omitted category: external finance)		
internal funds or retained earnings	0.220***	0.025**
	(0.020)	(0.011)
intra-group funding	0.075	0.103***
	(0.064)	(0.031)
Credit constrained	-0.094***	-0.093***
	(0.024)	(0.012)
Sample size	27,815	27,648
R^2	0.518	0.77
Memorandum		
R^2 with country × industry × year fixed effects and no ***X***	0.492	0.736
R^2 with ***X*** and no fixed effects	0.119	0.247
R^2 with ***X*** and country fixed effects	0.146	0.513
R^2 with ***X*** and industry fixed effects	0.294	0.426
R^2 with ***X*** and year fixed effects	0.119	0.247
R^2 with ***X*** and country fixed effects, industry fixed effects and year fixed effects	0.321	0.666
R^2 with ***X*** and slopes varying by country	0.196	0.437
R^2 with ***X*** and slopes varying by industry	0.286	0.409
R^2 with ***X*** and slopes varying by year	0.122	0.251

Note: The table reports estimates of equation (4) with country × industry × year fixed effects. Industries are defined at 2-digit NACE level. All estimates are based on Huber robust regression. Observations are weighted so that the sample represents the population in terms of employment. Investment is measured as an indicator variable equal to one if a firm reports positive investment, and zero otherwise (the baseline specification uses log(1+investment)). Standard errors are clustered by industry and country. ***, **, * denote statistical significance at 1, 5 and 10 percent levels.

Appendix Table 4. Descriptive statistics

Group of variables	Variable	2015 Mean	2015 St. dev.	2016 Mean	2016 St. dev.	2017 Mean	2017 St. dev.
Outcome variables	log(sales)	16.56	2.32	16.53	2.39	16.65	2.38
	log(fixed assets)	15.29	2.75	15.27	2.79	15.37	2.71
	log(employment)	4.81	1.98	4.80	2.02	4.90	2.02
	log(MRPK)	-0.47	1.48	-0.41	1.43	-0.40	1.39
	log(MRPL)	10.21	1.22	10.20	1.17	10.25	1.17
	log(MRPL) - log(MRPK)	10.68	1.69	10.61	1.66	10.65	1.54
Demographics	Firm age						
	less than 5 years	0.03	0.18	0.04	0.20	0.03	0.17
	5-9 years	0.08	0.28	0.08	0.27	0.07	0.25
	10-19 years	0.20	0.40	0.21	0.41	0.20	0.40
	20+ years	0.68	0.47	0.67	0.47	0.71	0.46
	Subsidiary	0.33	0.47	0.34	0.47	0.35	0.48
	Exporter	0.51	0.50	0.52	0.50	0.51	0.50
Quality of capital and other inputs	Share of state-of-the art machinery and equipment	0.42	0.32	0.41	0.32	0.40	0.32
	Share of high energy efficiency commercial building stock	0.37	0.34	0.35	0.34	0.35	0.34
Capacity utilization	above maximum capacity	0.05	0.22	0.05	0.23	0.08	0.26
	at maximum capacity	0.44	0.50	0.46	0.50	0.42	0.49
	somewhat below full capacity	0.40	0.49	0.38	0.49	0.42	0.49
	substantially below full capacity	0.09	0.28	0.09	0.28	0.07	0.26
Obstacles to investment	Demand for products or services						
	Major	0.26	0.44	0.23	0.42	0.23	0.42
	Minor	0.24	0.43	0.27	0.44	0.26	0.44
	Availability of staff with the right skills						
	Major	0.38	0.49	0.43	0.49	0.46	0.50
	Minor	0.30	0.46	0.31	0.46	0.31	0.46
	Energy costs						
	Major	0.22	0.41	0.23	0.42	0.25	0.43
	Minor	0.32	0.47	0.35	0.48	0.35	0.48
	Access to digital infrastructure						
	Major	0.10	0.30	0.11	0.32	0.14	0.35
	Minor	0.26	0.44	0.31	0.46	0.29	0.46
	Labor market regulations						
	Major	0.28	0.45	0.31	0.46	0.30	0.46
	Minor	0.29	0.46	0.33	0.47	0.36	0.48
	Business regulations and taxation						

	Major	0.32	0.47	0.32	0.47	0.31	0.46
	Minor	0.28	0.45	0.33	0.47	0.37	0.48
	Availability of adequate transport infrastructure						
	Major	0.16	0.36	0.16	0.37	0.18	0.38
	Minor	0.24	0.43	0.28	0.45	0.31	0.46
	Availability of finance						
	Major	0.24	0.43	0.19	0.40	0.19	0.39
	Minor	0.22	0.41	0.26	0.44	0.26	0.44
	Uncertainty about future						
	Major	0.41	0.49	0.38	0.48	0.36	0.48
	Minor	0.32	0.47	0.36	0.48	0.37	0.48
Adjustment	Investment, log(1 + investment)	12.41	3.99	12.29	4.04	12.62	3.84
	Percent change in employment in the last three years	0.12	0.47	0.14	0.44	0.14	0.40
	Investment over the last three years						
	too much	0.04	0.19	0.03	0.18	0.04	0.19
	about the right amount	0.78	0.41	0.79	0.40	0.78	0.41
	too little	0.16	0.37	0.16	0.37	0.17	0.38
	company did not exist three years ago	0.00	0.03	0.00	0.04	0.00	0.02
	Investment priority in the next three years						
	replacing capacity	0.41	0.49	0.32	0.47	0.33	0.47
	capacity expansion for existing products or services	0.25	0.43	0.28	0.45	0.31	0.46
	developing new products, processes or services	0.24	0.43	0.28	0.45	0.26	0.44
	no investment planned	0.09	0.28	0.11	0.31	0.09	0.28
Source of funds	internal funds or retained earnings	0.66	0.37	0.66	0.37	0.67	0.37
	external finance	0.31	0.35	0.31	0.36	0.30	0.35
	intra-group funding	0.02	0.13	0.01	0.10	0.02	0.12
	Finance constrained	0.07	0.25	0.07	0.26	0.06	0.23
Sample size		8,926	8,926	9,447	9,447	9,443	9,443

Note: All statistics are computed using sampling weights.

Appendix B: Derivations

Setup

The setup follows Hsieh and Klenow (2009). The objective function of the firm is

$$\max \tau_i^Y P_i Y_i - \tau_i^K r K_i - \tau_i^L w L_{it} - \tau_i^X P_X X_i$$

Subject to

Demand: $Y_i = Y\left(\frac{P_i}{P}\right)^{-\sigma}$

Production function: $Y_i = A_i K_i^{\alpha} L_i^{\beta} X_i^{1-\alpha-\beta}$

where i indexes firms (we skip time index to simplify notation), Y_i is output of firm i, Y is aggregate output,P_i is the price of firm i's output, P is the price index, K_i is capital, L_i is labor, X_i is materials (intermediate input), A_i is productivity, $\tau^Y, \tau^K, \tau^L, \tau^X$ are distortions in product and input market (no distortion corresponds to $\tau = 1$).

Aggregate demand is given by the Dixit-Stiglitz aggregator:

$$Y = \left(\int Y_i^{\frac{\sigma-1}{\sigma}} di\right)^{\frac{\sigma}{\sigma-1}}$$

We define TFP as $TFP_i \equiv \frac{Y_i}{K_i^{\alpha} L_i^{\beta} X_i^{1-\alpha-\beta}} = A_i$.

We define TFPR as $TFPR_i \equiv \frac{P_i Y_i}{K_i^{\alpha} L_i^{\beta} X_i^{1-\alpha-\beta}}$.

We define aggregate TFP as $A \equiv \frac{Y}{K^{\alpha} L^{\beta} X^{1-\alpha-\beta}}$ where aggregate capital, labor, and materials are $K = \int K_i di$, $L = \int L_i di$, $X = \int X_i di$.

We define marginal revenue product of capital as $MRPK_i \equiv \frac{\sigma-1}{\sigma}\frac{P_i Y_i}{K_i}$

We define marginal revenue product of labor as $MRPL_i \equiv \frac{\sigma-1}{\sigma}\frac{P_i Y_i}{L_i}$

We define marginal revenue product of intermediate inputs as $MRPX_i \equiv \frac{\sigma-1}{\sigma}\frac{P_i Y_i}{X_i}$

Note that using demand for firm i's output and the Dixit-Stiglitz aggregator we can find

$$Y = \left(\int Y_i^{\frac{\sigma-1}{\sigma}} di\right)^{\frac{\sigma}{\sigma-1}} = \left(\int \left[Y\left(\frac{P_i}{P}\right)^{-\sigma}\right]^{\frac{\sigma-1}{\sigma}} di\right)^{\frac{\sigma}{\sigma-1}} = YP^{\sigma}\left(\int P_i^{-(\sigma-1)} di\right)^{\frac{\sigma}{\sigma-1}}$$

which implies that

$$P = \left(\int P_i^{1-\sigma} di\right)^{\frac{1}{1-\sigma}}$$

Optimality conditions

The Lagrangian for the firm is

$$\mathcal{L} = \tau_i^Y P_i Y\left(\frac{P_i}{P}\right)^{-\sigma} - \tau_i^K r K_i - \tau_i^L w L_i - \tau_i^X P_X X_i - \lambda_i\left\{Y\left(\frac{P_i}{P}\right)^{-\sigma} - A_i K_i^{\alpha} L_i^{\beta} X_i^{1-\alpha-\beta}\right\}$$

Optimality conditions are:

$$\frac{\partial \mathcal{L}}{\partial K_i} = 0 \Rightarrow \tau_i^K r = \lambda_i \alpha \frac{Y_i}{K_i} \Rightarrow K_i = \lambda_i \alpha \frac{Y_i}{\tau_i^K r}$$

$$\frac{\partial \mathcal{L}}{\partial L_i} = 0 \Rightarrow \tau_i^L w = \lambda_i \beta \frac{Y_i}{L_i} \Rightarrow L_i = \lambda_i \beta \frac{Y_i}{\tau_i^L w}$$

$$\frac{\partial \mathcal{L}}{\partial X_i} = 0 \Rightarrow \tau_i^X P_X = \lambda_i (1-\alpha-\beta)\frac{Y_i}{X_i} \Rightarrow X_i = \lambda_i (1-\alpha-\beta)\frac{Y_i}{\tau_i^X P_X}$$

$$\frac{\partial \mathcal{L}}{\partial P_i} = 0 \Rightarrow P_i = \frac{\sigma}{\sigma-1}\frac{1}{\tau_i^Y}\lambda_i$$

Note that λ_i is the marginal cost for firm i. Using the production function and the optimality conditions for L_i, K_i, X_i, we can find

$$Y_i = A_i K_i^{\alpha} L_i^{\beta} X_i^{1-\alpha-\beta} = A_i \left(\lambda_i \alpha \frac{Y_i}{\tau_i^K r}\right)^{\alpha} \left(\lambda_i \beta \frac{Y_i}{\tau_i^L w}\right)^{\beta} \left(\lambda_i (1-\alpha-\beta) \frac{Y_i}{\tau_i^X P_X}\right)^{1-\alpha-\beta}$$

$$= A_i Y_i \lambda_i \left(\frac{\alpha}{\tau_i^K r}\right)^{\alpha} \left(\frac{\beta}{\tau_i^L w}\right)^{\beta} \left(\frac{(1-\alpha-\beta)}{\tau_i^X P_X}\right)^{1-\alpha-\beta}$$

which implies that

$$\lambda_i = \frac{1}{A_i}\left(\frac{\tau_i^K r}{\alpha}\right)^{\alpha} \left(\frac{\tau_i^L w}{\beta}\right)^{\beta} \left(\frac{\tau_i^X P_X}{(1-\alpha-\beta)}\right)^{1-\alpha-\beta} = \left[\left(\frac{r}{\alpha}\right)^{\alpha} \left(\frac{w}{\beta}\right)^{\beta} \left(\frac{P_X}{(1-\alpha-\beta)}\right)^{1-\alpha-\beta}\right] \frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i} =$$

$$= \boldsymbol{B} \frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i}$$

where $\boldsymbol{B} \equiv \left(\frac{r}{\alpha}\right)^{\alpha} \left(\frac{w}{\beta}\right)^{\beta} \left(\frac{P_X}{(1-\alpha-\beta)}\right)^{1-\alpha-\beta}$ does not depend on firm-specific distortions.

It follows that

$$P_i = \frac{\sigma}{\sigma-1} \frac{1}{\tau_i^Y} \frac{1}{A_i} \left(\frac{\tau_i^K r}{\alpha}\right)^{\alpha} \left(\frac{\tau_i^L w}{\beta}\right)^{\beta} \left(\frac{\tau_i^X P_X}{(1-\alpha-\beta)}\right)^{1-\alpha-\beta} = \frac{\sigma}{\sigma-1} \boldsymbol{B} \frac{1}{\tau_i^Y} \frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i}$$

We can also find expressions for marginal revenue products

$$MRPK_i \equiv \frac{\sigma-1}{\sigma} \alpha \frac{P_i Y_i}{K_i} = \frac{\tau_i^K}{\tau_i^Y} R$$

$$MRPL_i \equiv \frac{\sigma-1}{\sigma} \beta \frac{P_i Y_i}{L_i} = \frac{\tau_i^L}{\tau_i^Y} W$$

$$MRPX_i \equiv \frac{\sigma-1}{\sigma} (1-\alpha-\beta) \frac{P_i Y_i}{X_i} = \frac{\tau_i^X}{\tau_i^Y} P^X$$

Aggregation

Aggregate capital in the economy is given by

$$K = \int K_i di = \int \lambda_i \alpha \frac{Y_i}{\tau_i^K r} di = \int \lambda_i \alpha \frac{Y \left(\frac{P_i}{P}\right)^{-\sigma}}{\tau_i^K r} di = Y P^{\sigma} \left(\frac{\alpha}{r}\right) \int \frac{1}{\tau_i^K} \lambda_i P_i^{-\sigma} di$$

$$= Y P^{\sigma} \left(\frac{\alpha}{r}\right) \int \frac{1}{\tau_i^K} \lambda_i \left(\frac{\sigma}{\sigma-1} \frac{1}{\tau_i^Y} \lambda_i\right)^{-\sigma} di = Y P^{\sigma} \left(\frac{\alpha}{r}\right) \frac{\sigma}{\sigma-1} \int \frac{\left(\tau_i^Y\right)^{\sigma}}{\tau_i^K} \lambda_i^{1-\sigma} di$$

$$= Y P^{\sigma} \left(\frac{\alpha}{r}\right) \frac{\sigma}{\sigma-1} \int \frac{\left(\tau_i^Y\right)^{\sigma}}{\tau_i^K} \left(\boldsymbol{B} \frac{\left(\tau_i^K\right)^{\alpha} \left(\tau_i^L\right)^{\beta} \left(\tau_i^X\right)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma} di$$

$$= Y P^{\sigma} \left(\frac{\alpha}{r}\right) \frac{\sigma}{\sigma-1} \boldsymbol{B}^{1-\sigma} \int \frac{\left(\tau_i^Y\right)^{\sigma}}{\tau_i^K} \left(\frac{\left(\tau_i^K\right)^{\alpha} \left(\tau_i^L\right)^{\beta} \left(\tau_i^X\right)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma} di$$

Note that $Y P^{\sigma} \left(\frac{\alpha}{r}\right) \frac{\sigma}{\sigma-1} \boldsymbol{B}^{1-\sigma}$ does not depend on firm-specific outcomes.

Aggregate labor in the economy is given by

$$L = \int L_i di = \int \lambda_i \beta \frac{Y_i}{\tau_i^L w} di = \int \lambda_i \beta \frac{Y \left(\frac{P_i}{P}\right)^{-\sigma}}{\tau_i^L w} di = Y P^{\sigma} \left(\frac{\beta}{w}\right) \int \frac{1}{\tau_i^L} \lambda_i P_i^{-\sigma} di$$

$$= Y P^{\sigma} \left(\frac{\beta}{w}\right) \int \frac{1}{\tau_i^L} \lambda_i \left(\frac{\sigma}{\sigma-1} \frac{1}{\tau_i^Y} \lambda_i\right)^{-\sigma} di = Y P^{\sigma} \left(\frac{\beta}{w}\right) \frac{\sigma}{\sigma-1} \int \frac{\left(\tau_i^Y\right)^{\sigma}}{\tau_i^L} \lambda_i^{1-\sigma} di$$

$$= YP^{\sigma}\left(\frac{\beta}{w}\right)\frac{\sigma}{\sigma-1}\int\frac{(\tau_i^Y)^{\sigma}}{\tau_i^L}\left(\boldsymbol{B}\frac{(\tau_i^K)^{\alpha}(\tau_i^L)^{\beta}(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di$$

$$= YP^{\sigma}\left(\frac{\beta}{w}\right)\frac{\sigma}{\sigma-1}\boldsymbol{B}^{1-\sigma}\int\frac{(\tau_i^Y)^{\sigma}}{\tau_i^L}\left(\frac{(\tau_i^K)^{\alpha}(\tau_i^L)^{\beta}(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di$$

Aggregate intermediate input in the economy is given by

$$X = \int X_i di = \int \lambda_i(1-\alpha-\beta)\frac{Y_i}{\tau_i^X P_X}di = \int \lambda_i(1-\alpha-\beta)\frac{Y\left(\frac{P_i}{P}\right)^{-\sigma}}{\tau_i^X P_X}di$$

$$= YP^{\sigma}\left(\frac{1-\alpha-\beta}{P_X}\right)\int\frac{1}{\tau_i^X}\lambda_i P_i^{-\sigma}di$$

$$= YP^{\sigma}\left(\frac{1-\alpha-\beta}{P_X}\right)\int\frac{1}{\tau_i^X}\lambda_i\left(\frac{\sigma}{\sigma-1}\frac{1}{\tau_i^Y}\lambda_i\right)^{-\sigma}di = YP^{\sigma}\left(\frac{1-\alpha-\beta}{P_X}\right)\frac{\sigma}{\sigma-1}\int\frac{(\tau_i^Y)^{\sigma}}{\tau_i^X}\lambda_i^{1-\sigma}di$$

$$= YP^{\sigma}\left(\frac{1-\alpha-\beta}{P_X}\right)\frac{\sigma}{\sigma-1}\int\frac{(\tau_i^Y)^{\sigma}}{\tau_i^X}\left(\boldsymbol{B}\frac{(\tau_i^K)^{\alpha}(\tau_i^L)^{\beta}(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di$$

$$= YP^{\sigma}\left(\frac{1-\alpha-\beta}{P_X}\right)\frac{\sigma}{\sigma-1}\boldsymbol{B}^{1-\sigma}\int\frac{(\tau_i^Y)^{\sigma}}{\tau_i^X}\left(\frac{(\tau_i^K)^{\alpha}(\tau_i^L)^{\beta}(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di$$

Aggregate price index is given by

$$P = \left(\int P_i^{1-\sigma}di\right)^{\frac{1}{1-\sigma}} = \left(\int\left(\frac{\sigma}{\sigma-1}\boldsymbol{B}\frac{1}{\tau_i^Y}\frac{(\tau_i^K)^{\alpha}(\tau_i^L)^{\beta}(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di\right)^{\frac{1}{1-\sigma}}$$

$$= \frac{\sigma}{\sigma-1}\boldsymbol{B}\left(\int\left(\frac{1}{\tau_i^Y}\frac{(\tau_i^K)^{\alpha}(\tau_i^L)^{\beta}(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di\right)^{\frac{1}{1-\sigma}}$$

Aggregate TFP

Using our definition of aggregate TFP, we have

$$A \equiv \frac{Y}{K^{\alpha}L^{\beta}X^{1-\alpha-\beta}}$$

Let's compute the denominator of this expression:

$$K^{\alpha}L^{\beta}X^{1-\alpha-\beta} = \left(YP^{\sigma}\left(\frac{\alpha}{r}\right)\frac{\sigma}{\sigma-1}\boldsymbol{B}^{1-\sigma}\int\frac{(\tau_i^Y)^{\sigma}}{\tau_i^K}\left(\frac{(\tau_i^K)^{\alpha}(\tau_i^L)^{\beta}(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di\right)^{\alpha}$$

$$\times\left(YP^{\sigma}\left(\frac{\beta}{w}\right)\frac{\sigma}{\sigma-1}\boldsymbol{B}^{1-\sigma}\int\frac{(\tau_i^Y)^{\sigma}}{\tau_i^L}\left(\frac{(\tau_i^K)^{\alpha}(\tau_i^L)^{\beta}(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di\right)^{\beta}$$

$$\times\left(YP^{\sigma}\left(\frac{1-\alpha-\beta}{P_X}\right)\frac{\sigma}{\sigma-1}\boldsymbol{B}^{1-\sigma}\int\frac{(\tau_i^Y)^{\sigma}}{\tau_i^X}\left(\frac{(\tau_i^K)^{\alpha}(\tau_i^L)^{\beta}(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di\right)^{1-\alpha-\beta}$$

$$= \left[Y P^{\sigma} \frac{\sigma}{\sigma-1} \boldsymbol{B}^{1-\sigma} \left(\frac{\alpha}{r}\right)^{\alpha} \left(\frac{\beta}{w}\right)^{\beta} \left(\frac{1-\alpha-\beta}{P_X}\right)^{1-\alpha-\beta} \right] \times \left(\int \frac{(\tau_i^Y)^{\sigma}}{\tau_i^K} \left(\frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i} \right)^{1-\sigma} di \right)^{\alpha} \times$$

$$\left(\int \frac{(\tau_i^Y)^{\sigma}}{\tau_i^L} \left(\frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i} \right)^{1-\sigma} di \right)^{\beta} \times \left(\int \frac{(\tau_i^Y)^{\sigma}}{\tau_i^X} \left(\frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i} \right)^{1-\sigma} di \right)^{1-\alpha-\beta}$$

Note that we defined $\boldsymbol{B} \equiv \left(\frac{r}{\alpha}\right)^{\alpha} \left(\frac{w}{\beta}\right)^{\beta} \left(\frac{P_X}{(1-\alpha-\beta)}\right)^{1-\alpha-\beta}$ and so we can simplify this expression a bit more:

$$K^{\alpha} L^{\beta} X^{1-\alpha-\beta} = \left[Y P^{\sigma} \frac{\sigma}{\sigma-1} \boldsymbol{B}^{-\sigma} \right] \times \left(\int \frac{(\tau_i^Y)^{\sigma}}{\tau_i^K} \left(\frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i} \right)^{1-\sigma} di \right)^{\alpha}$$

$$\times \left(\int \frac{(\tau_i^Y)^{\sigma}}{\tau_i^L} \left(\frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i} \right)^{1-\sigma} di \right)^{\beta} \times \left(\int \frac{(\tau_i^Y)^{\sigma}}{\tau_i^X} \left(\frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i} \right)^{1-\sigma} di \right)^{1-\alpha-\beta}$$

Because Y appears in the numerator and denominator of A, it follows that

$$A \equiv \frac{Y}{K^{\alpha} L^{\beta} X^{1-\alpha-\beta}} =$$

$$= \left[P^{-\sigma} \frac{\sigma-1}{\sigma} \boldsymbol{B}^{\sigma} \right] \times \left(\int \frac{(\tau_i^Y)^{\sigma}}{\tau_i^K} \left(\frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i} \right)^{1-\sigma} di \right)^{-\alpha}$$

$$\times \left(\int \frac{(\tau_i^Y)^{\sigma}}{\tau_i^L} \left(\frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i} \right)^{1-\sigma} di \right)^{-\beta}$$

$$\times \left(\int \frac{(\tau_i^Y)^{\sigma}}{\tau_i^X} \left(\frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i} \right)^{1-\sigma} di \right)^{-(1-\alpha-\beta)}$$

Approximation to aggregate productivity

Hsieh and Klenow (2009) assume log-normal distribution of firm-specific variables $A_i, \tau_i^Y, \tau_i^K, \tau_i^L, \tau_i^X$. We can use this assumption to derive exact formulae for output lost due to frictions $\tau_i^Y, \tau_i^K, \tau_i^L, \tau_i^X$. Assume that each of these variables are distributed independently (zero covariance):

$$\log A_i \sim N(\mu_A, V_A)$$
$$\log \tau_i^Y \sim N(0, V_{\tau Y})$$
$$\log \tau_i^K \sim N(0, V_{\tau K})$$
$$\log \tau_i^L \sim N(0, V_{\tau L})$$
$$\log \tau_i^X \sim N(0, V_{\tau X})$$

Consider the aggregate price level:

$$P^{-\sigma} = \frac{\sigma}{\sigma-1} \boldsymbol{B} \left(\int \left(\frac{1}{\tau_i^Y} \frac{(\tau_i^K)^{\alpha} (\tau_i^L)^{\beta} (\tau_i^X)^{1-\alpha-\beta}}{A_i} \right)^{1-\sigma} di \right)^{\frac{\sigma}{\sigma-1}}$$

Then using $\int z_i di = E(z_i)$ (that is a cross-sectional average is equal to the mathematical expectation of random variable z_i) and the property of log-normal variable $E(z) = \exp\left(mean_z + \frac{1}{2} variance_z\right)$,[22] we have

[22] Note that $E(z^a) = \exp\left(a \times mean_z + \frac{a^2}{2} variance_z\right)$.

$$-\sigma\log P = \log\frac{\sigma}{\sigma-1}\boldsymbol{B} + \frac{\sigma}{1-\sigma}\log\left(\int\left(\frac{1}{\tau_i^Y}\frac{(\tau_i^K)^\alpha(\tau_i^L)^\beta(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di\right)$$

$$= constant + \frac{\sigma}{\sigma-1}\log\big(E\big[\exp\{(1-\sigma)\alpha\log\tau_i^K + (1-\sigma)\beta\log\tau_i^L + (1-\sigma)(1-\alpha-\beta)\log\tau_i^X - (1-\sigma)\log\tau_i^Y - (1-\sigma)\log A_i\}\big]\big)$$

$$= constant + \frac{\sigma}{\sigma-1}\log\big(E\big[\exp\{-(\sigma-1)\alpha\log\tau_i^K - (\sigma-1)\beta\log\tau_i^L - (\sigma-1)(1-\alpha-\beta)\log\tau_i^X + (\sigma-1)\log\tau_i^Y + (\sigma-1)\log A_i\}\big]\big)$$

$$= constant + \frac{\sigma}{\sigma-1}\log\left(\left[\exp\left\{\frac{[(\sigma-1)\alpha]^2}{2}V_{\tau K} + \frac{[(\sigma-1)\beta]^2}{2}V_{\tau L} + \frac{[(\sigma-1)(1-\alpha-\beta)]^2}{2}V_{\tau X} + \frac{[(\sigma-1)]^2}{2}V_{\tau Y} + (\sigma-1)\mu_A + \frac{(\sigma-1)^2}{2}V_A\right\}\right]\right)$$

$$= constant + \frac{\sigma}{\sigma-1}\left\{\frac{[(\sigma-1)\alpha]^2}{2}V_{\tau K} + \frac{[(\sigma-1)\beta]^2}{2}V_{\tau L} + \frac{[(\sigma-1)(1-\alpha-\beta)]^2}{2}V_{\tau X} + \frac{[(\sigma-1)]^2}{2}V_{\tau Y} + (\sigma-1)\mu_A + \frac{(\sigma-1)^2}{2}V_A\right\}$$

$$= constant + \left\{\frac{\sigma(\sigma-1)\alpha^2}{2}V_{\tau K} + \frac{\sigma(\sigma-1)\beta^2}{2}V_{\tau L} + \frac{\sigma(\sigma-1)(1-\alpha-\beta)^2}{2}V_{\tau X} + \frac{\sigma(\sigma-1)}{2}V_{\tau Y} + \sigma\mu_A + \frac{\sigma(\sigma-1)}{2}V_A\right\}$$

Now let's derive terms (in logs) that are highlighted in green, red and blue:

$$\log\left[\left(\int\frac{(\tau_i^Y)^\sigma}{\tau_i^K}\left(\frac{(\tau_i^K)^\alpha(\tau_i^L)^\beta(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di\right)^{-\alpha}\right] =$$

$$= -\alpha\log E\left(\frac{(\tau_i^Y)^\sigma}{\tau_i^K}\left(\frac{(\tau_i^K)^\alpha(\tau_i^L)^\beta(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}\right)$$

$$= -\alpha\log E\left((\tau_i^Y)^\sigma(\tau_i^K)^{\alpha(1-\sigma)-1}(\tau_i^L)^{\beta(1-\sigma)}(\tau_i^X)^{(1-\alpha-\beta)(1-\sigma)}A_i^{\sigma-1}\right)$$

$$= -\alpha\left(\frac{(\alpha(1-\sigma)-1)^2}{2}V_{\tau K} + \frac{(\beta(1-\sigma))^2}{2}V_{\tau L} + \frac{\big((1-\alpha-\beta)(1-\sigma)\big)^2}{2}V_{\tau X} + \frac{\sigma^2}{2}V_{\tau Y} + (\sigma-1)\mu_A + \frac{(\sigma-1)^2}{2}V_A\right)$$

$$\log\left[\left(\int\frac{(\tau_i^Y)^\sigma}{\tau_i^L}\left(\frac{(\tau_i^K)^\alpha(\tau_i^L)^\beta(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di\right)^{-\beta}\right] =$$

$$= -\beta\log E\left(\frac{(\tau_i^Y)^\sigma}{\tau_i^L}\left(\frac{(\tau_i^K)^\alpha(\tau_i^L)^\beta(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}\right)$$

$$= -\beta\log E\left((\tau_i^Y)^\sigma(\tau_i^K)^{\alpha(1-\sigma)}(\tau_i^L)^{\beta(1-\sigma)-1}(\tau_i^X)^{(1-\alpha-\beta)(1-\sigma)}A_i^{\sigma-1}\right)$$

$$= -\beta\left(\frac{\big(\alpha(1-\sigma)\big)^2}{2}V_{\tau K} + \frac{(\beta(1-\sigma)-1)^2}{2}V_{\tau L} + \frac{\big((1-\alpha-\beta)(1-\sigma)\big)^2}{2}V_{\tau X} + \frac{\sigma^2}{2}V_{\tau Y} + (\sigma-1)\mu_A + \frac{(\sigma-1)^2}{2}V_A\right)$$

$$\log\left[\left(\int\frac{(\tau_i^Y)^\sigma}{\tau_i^X}\left(\frac{(\tau_i^K)^\alpha(\tau_i^L)^\beta(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}di\right)^{-(1-\alpha-\beta)}\right] =$$

$$= -(1-\alpha-\beta)\log E\left(\frac{(\tau_i^Y)^\sigma}{\tau_i^X}\left(\frac{(\tau_i^K)^\alpha(\tau_i^L)^\beta(\tau_i^X)^{1-\alpha-\beta}}{A_i}\right)^{1-\sigma}\right)$$

$$= -(1-\alpha-\beta)\log E\left(\left(\tau_i^Y\right)^{\sigma}\left(\tau_i^K\right)^{\alpha(1-\sigma)}\left(\tau_i^L\right)^{\beta(1-\sigma)}\left(\tau_i^X\right)^{(1-\alpha-\beta)(1-\sigma)-1}A_i^{\sigma-1}\right)$$

$$= -(1-\alpha-\beta)\left(\frac{\left(\alpha(1-\sigma)\right)^2}{2}V_{\tau K}+\frac{\left(\beta(1-\sigma)\right)^2}{2}V_{\tau L}+\frac{\left((1-\alpha-\beta)(1-\sigma)-1\right)^2}{2}V_{\tau X}+\frac{\sigma^2}{2}V_{\tau Y}+(\sigma-1)\mu_A+\frac{(\sigma-1)^2}{2}V_A\right)$$

Now we can put together to compute the log of aggregate TFP

$$\log A = \log\left[P^{-\sigma}\frac{\sigma-1}{\sigma}\boldsymbol{B}^{\sigma}\right]+\log Green+\log Red+\log Blue$$

$$= constant - \sigma\log P+\log Green+\log Red+\log Blue$$

$$= constant+\left\{\frac{\sigma(\sigma-1)\alpha^2}{2}V_{\tau K}+\frac{\sigma(\sigma-1)\beta^2}{2}V_{\tau L}+\frac{\sigma(\sigma-1)(1-\alpha-\beta)^2}{2}V_{\tau X}+\frac{\sigma(\sigma-1)}{2}V_{\tau Y}+\sigma\mu_A+\frac{\sigma(\sigma-1)}{2}V_A\right\}$$
$$-\alpha\left(\frac{(\alpha(1-\sigma)-1)^2}{2}V_{\tau K}+\frac{(\beta(1-\sigma))^2}{2}V_{\tau L}+\frac{\left((1-\alpha-\beta)(1-\sigma)\right)^2}{2}V_{\tau X}+\frac{\sigma^2}{2}V_{\tau Y}+(\sigma-1)\mu_A+\frac{(\sigma-1)^2}{2}V_A\right)$$
$$-\beta\left(\frac{\left(\alpha(1-\sigma)\right)^2}{2}V_{\tau K}+\frac{(\beta(1-\sigma)-1)^2}{2}V_{\tau L}+\frac{\left((1-\alpha-\beta)(1-\sigma)\right)^2}{2}V_{\tau X}+\frac{\sigma^2}{2}V_{\tau Y}+(\sigma-1)\mu_A+\frac{(\sigma-1)^2}{2}V_A\right)$$
$$-(1-\alpha-\beta)\left(\frac{\left(\alpha(1-\sigma)\right)^2}{2}V_{\tau K}+\frac{\left(\beta(1-\sigma)\right)^2}{2}V_{\tau L}+\frac{\left((1-\alpha-\beta)(1-\sigma)-1\right)^2}{2}V_{\tau X}+\frac{\sigma^2}{2}V_{\tau Y}+(\sigma-1)\mu_A+\frac{(\sigma-1)^2}{2}V_A\right)$$

$$= constant-\left\{\frac{\alpha(1-\alpha)}{2}+\frac{\alpha^2\sigma}{2}\right\}V_{\tau K}-\left\{\frac{\beta(1-\beta)}{2}+\frac{\beta^2\sigma}{2}\right\}V_{\tau L}-\left\{\frac{(1-\alpha-\beta)(\alpha+\beta)}{2}+\frac{(1-\alpha-\beta)^2\sigma}{2}\right\}V_{\tau X}-\frac{\sigma}{2}V_{\tau Y}+\mu_A+\frac{(\sigma-1)}{2}V_A$$

Note that TFP is increasing in the variance of productivity V_A and it is decreasing in the variable of distortions $V_{\tau K}, V_{\tau L}, V_{\tau X}, V_{\tau Y}$

Identification of distortions τ

Using the optimality condition for capital and the expression for the optimal price

$$MRPK_i \equiv \frac{P_iY_i}{K_i} = P_i\frac{Y_i}{K_i} = P_i\times\frac{\tau_i^K r}{\lambda_i\alpha} = \frac{\sigma}{\sigma-1}\frac{1}{\tau_i^Y}\lambda_i\times\frac{\tau_i^K r}{\lambda_i\alpha} = \frac{\sigma}{\sigma-1}\frac{\tau_i^K}{\tau_i^Y}\frac{r}{\alpha}$$

Using the same logic, we have

$$MRPL_i \equiv \frac{P_iY_i}{L_i} = P_i\frac{Y_i}{L_i} = P_i\times\frac{\tau_i^L w}{\lambda_i\beta} = \frac{\sigma}{\sigma-1}\frac{1}{\tau_i^Y}\lambda_i\times\frac{\tau_i^L w}{\lambda_i\beta} = \frac{\sigma}{\sigma-1}\frac{\tau_i^L}{\tau_i^Y}\frac{w}{\beta}$$

$$MRPX_i \equiv \frac{P_iY_i}{X_i} = P_i\frac{Y_i}{X_i} = P_i\times\frac{\tau_i^X P_X}{\lambda_i(1-\alpha-\beta)} = \frac{\sigma}{\sigma-1}\frac{1}{\tau_i^Y}\lambda_i\times\frac{\tau_i^X P_X}{\lambda_i(1-\alpha-\beta)} = \frac{\sigma}{\sigma-1}\frac{\tau_i^X}{\tau_i^Y}\frac{P_X}{(1-\alpha-\beta)}$$

We have four unknowns $\tau_i^Y, \tau_i^K, \tau_i^L, \tau_i^X$ and three moments $MPRK_i, MRPL_i, MRPX_i$. The system is not identified. We need to impose an identifying assumption.

The Hsieh-Klenow framework assumes that $\tau_i^L = 1$ for all i and hence, one can find

$$\log MRPL_i = constant - \log(\tau_i^Y)$$

so that one can estimate $V_{\tau Y} \equiv var\left(\log(\tau_i^Y)\right) = var\,(\log MRPL_i) = V_{MRPL}$. Then one can note that

$$\log MRPK_i = constant + \log(\tau_i^K) - \log(\tau_i^Y) = constant + \log(\tau_i^K) + \log MRPL_i$$

and hence

$$\log(\tau_i^K) = constant + \log MRPK_i - \log MRPL_i$$

$$\begin{aligned} V_{\tau K} \equiv var\left(\log(\tau_i^K)\right) &= var(\log MRPK_i - \log MRPL_i) \\ &= var\,(\log MRPL_i) + var\,(\log MRPK_i) - 2cov(\log MRPL_i\,, \log MRPK_i) \\ &= V_{MRPL} + V_{MRPK} - 2V_{MRPL,MRPK} \end{aligned}$$

One can alternatively assume that $\tau_i^K = 1$ for all i and τ_i^L is varying across firms. Then

$$V_{\tau Y} \equiv var\left(\log(\tau_i^Y)\right) = var\,(\log MRPK_i) = V_{MRPK}$$

$$\begin{aligned} V_{\tau L} \equiv var\left(\log(\tau_i^L)\right) &= var(\log MRPL_i - \log MRPK_i) \\ &= var\,(\log MRPL_i) + var\,(\log MRPK_i) - 2cov(\log MRPL_i\,, \log MRPK_i) = \\ &= V_{MRPL} + V_{MRPK} - 2V_{MRPL,MRPK} \end{aligned}$$

Note that we do not have materials in the EIBIS. Is this a problem? The answer is not necessarily. We know that $V_{\tau X} \geq 0$ and hence a distortion in the intermediate input market will lower aggregate TFP. If we do not observe $MRPX$, we likely understate the effect of the distortions and thus our estimate is conservative.

If we make the assumption as in Hsieh and Klenow, then the (conservative) loss in aggregate TFP (and hence aggregate output) is

$$loss = -\left\{\frac{\alpha(1-\alpha)}{2} + \frac{\alpha^2\sigma}{2}\right\} V_{\tau K} - \frac{\sigma}{2} V_{\tau Y} = -\left\{\frac{\alpha(1-\alpha)}{2} + \frac{\alpha^2\sigma}{2}\right\} \left[V_{MRPL} + V_{MRPK} - 2V_{MRPL,MRPK}\right] - \frac{\sigma}{2} V_{MRPL}$$

If we make the other assumption, then

$$loss = -\left\{\frac{\beta(1-\beta)}{2} + \frac{\beta^2\sigma}{2}\right\} V_{\tau L} - \frac{\sigma}{2} V_{\tau Y} = -\left\{\frac{\beta(1-\beta)}{2} + \frac{\beta^2\sigma}{2}\right\} \left[V_{MRPL} + V_{MRPK} - 2V_{MRPL,MRPK}\right] - \frac{\sigma}{2} V_{MRPK}$$