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

First Draft
March 14, 2018

Abstract: Using a new survey, we show that the dispersion of marginal products across firms in the European Union is about twice as large as that in the United States. Reducing it to the US level would increase EU GDP by more than 30 percent. Alternatively, removing barriers between industries *and* countries would raise EU GDP by at least 25 percent. Firm characteristics, such as demographics, quality of inputs, utilization of resources, and dynamic adjustment of inputs, are predictors of the marginal 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 also show that cross-country differences in the dispersion of marginal products are more due to differences in how the business, institutional and policy environment translates firm characteristics into outcomes than to the differences in firm characteristics *per se*.

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

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

Acknowledgement: We thank Albert Park and participants at the 2018 ASSA meeting for helpful comments. In preparing this paper, Yuriy Gorodnichenko and Jan Svejnar worked under a grant from 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, European Union 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)). Indeed, in Figures 1 and 2 we show that the dispersion of MRPK within individual European countries and within individual economic sectors has been trending up since the mid-1990s and the dispersion of MRPL has also been rising, though to a smaller extent. These increases in dispersion are consistent with decreased productivity growth at the aggregate level.

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 a given economy and measures misallocation as dispersion of MRPK or MRPL across firms. But census-type data usually contain only income statements, balance sheet information about capital, and basic data on employment. As a result, researchers usually cannot tell why a given dispersion of MRPK or MRPL exists.

In this paper, we shed more light on this issue by using survey data. In particular, we develop a simple dynamic theoretical framework and use a new survey of firms to assess how much dispersion in marginal products exists and what drives these differentials. The EIB Investment Survey (henceforth EIBIS) was administered in 2016 to a stratified random sample of firms in each of the 28 countries of the European Union (EU). 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. In addition, firms in EIBIS are anonymously matched to the Orbis administrative database that permits us to validate and supplement the information furnished by EIBIS. 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.

With the above information we are able to estimate the extent to which various factors can account for cross-sectional dispersion in marginal products. We first estimate the extent of dispersion in MRPK and MRPL within the EU and within individual countries, an exercise that is important for understanding the extent of integration of the EU common market and the markets of individual countries. We next assess the quantitative contribution of each factor by building on Mincer (1958) and estimating dynamic equilibrium relationships (optimality conditions) in a regression framework. We assess how the explanatory variables that reflect the behavior and constraints of firms predict their MRPK and MRPL. Third, we examine the extent to which the dispersion of marginal products is related to firm-level characteristics as opposed to country-level or sector-level effects. Fourth, 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.

Although we do not have an unambiguously exogenous variation in our regressors, the predictive power of our regressors sheds light on what factors are likely to be quantitatively important and where future work should concentrate efforts to estimate causal effects. Given the nature of EIBIS data, we have more information about capital-related than labor-related issues. Our analysis of the determinants of MRPK is hence richer than that of MRPL. We also note that while the existing literature treats the dispersion of marginal products as reflecting barriers and distortions, this may not always be the case. If the dispersion is brought about by barriers and distortions, it is indeed economically undesirable. However, 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 may be optimal even from the standpoint of social welfare. Which of these phenomena is consistent with the data is therefore a key question that we start addressing.

We find that there is a sizable dispersion of marginal products measured across all 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. A less demanding proposition would be to say that after decades of economic integration the dispersion of marginal products in the EU ought to start approaching the dispersion found in the US. Comparing our findings on the dispersion of MRPK across firms in the 28 EU countries to the dispersion found by other researchers for the US indicates that the dispersion in the EU is about twice as large, a finding consistent with the view that the EU economy is far from being highly integrated. Our calculations suggest that by reducing the EU dispersion of MRPK to the level of the US would increase EU productivity (GDP) by more than 30 percent.

When we allow the firm-level characteristics alone to explain the dispersion in marginal products in EIBIS, we find that these variables account for 11.2 percent of the total variation in log(MRPK) and 27.1 percent of the variation in log(MRPL) – a sizable part of total dispersion. When we allow firm-level characteristics to have different effects by sector or country, we show that these variables explain most of the variation in marginal products within the EU. Overall, we find that in terms of dispersion of MRPK and MRPL the EU is far from being a single market and that firm-level characteristics coupled with fixed sector and country features explain most of the observed dispersion. Interestingly, 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.

When we exploit in detail the 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.

Our work is related to several strands of previous research. First, we contribute to the rapidly growing literature measuring misallocation of resources (e.g., Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Bartelsman et al. (2013); also see Restuccia and Rogerson (2013, 2017) for surveys). In particular, we document new facts about the allocation of capital and labor across 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 force 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), 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.

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 are 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 studying 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.

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 the 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

$$\pi_{it} = G_{it}\left[(U_{it}K_{it})^{\alpha_i}(E_{it}L_{it})^{\beta_i}\right]^{1-\frac{1}{\sigma_i}} - R_{it}(U_{it})K_{it} - W_{it}(E_{it})L_{it}$$
$$- \frac{\phi_K}{2}\left(\frac{K_{it}}{K_{i,t-1}} - 1\right)^2 R_{it}(U_{it})K_{i,t-1} - \frac{\phi_L}{2}\left(\frac{L_{it}}{L_{i,t-1}} - 1\right)^2 W_{it}(E_{it})L_{i,t-1}$$

where $\gamma_i = \alpha_i + \beta_i$ reflects returns to scale in production, K_{it} is capital, L_{it} is labor, 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), $R_{it}(U_{it})$ 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), 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_i is the elasticity of demand. For example, the price schedules could be $R_{it}(U_{it}) = R_{it}^{base} \times U_{it}^{\psi_K}/\psi_K$ and $W_{it}(E_{it}) = W_{it}^{base} \times E_{it}^{\psi_L}/\psi_L$, where ψ_K and ψ_L are slopes of the respective supply schedules. 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

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

where r_s 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 $Y_{it} \equiv G_{it}\left[(U_{it}K_{it})^{\alpha_i}(E_{it}L_{it})^{\beta_i}\right]^{1-\frac{1}{\sigma_i}}$ be the firm revenue. Then the optimality conditions for capital and labor are, respectively

$$MRPK_{it} \equiv (1-\sigma_i^{-1})\alpha_i \frac{Y_{it}}{K_{it}} \approx R_{it}(U_{it})\left\{1+\phi_K\left(\frac{K_{it}}{K_{i,t-1}}-1\right)-\frac{\phi_K}{1+r_{t+1}}\times\left(\frac{K_{i,t+1}}{K_{it}}-1\right)\right\} \quad (1')$$

$$MRPL_{it} \equiv (1-\sigma_i^{-1})\beta_i \frac{Y_{it}}{L_{it}} \approx W_{it}(E_{it})\left\{1+\phi_L\left(\frac{L_{it}}{L_{i,t-1}}-1\right)-\frac{\phi_L}{1+r_{t+1}}\times\left(\frac{L_{i,t+1}}{L_{it}}-1\right)\right\} \quad (1'')$$

Note that the marginal revenue products are defined in terms of physical units of capital and labor. Their variation across firms may reflect differences in adjustment costs, as well as input quality, utilization rates, and taxes or regulation (proxied by R_{it}^{base} and W_{it}^{base} as well as parameters ψ_K and ψ_L). However, if one adjusted inputs for quality and/or account for adjustment costs and if the price schedules were the same across firms (that is, R_{it}^{base} and W_{it}^{base} were the same across firms), then marginal revenue products for effective units of capital ($K_{it}U_{it}$) and labor ($E_{it}L_{it}$) should be equalized across firms and the cross-sectional dispersion ought to be zero. These optimality conditions show marginal revenue products (the left-hand side of equation (1') and (1'')) are functions of distortions and compensating differentials (the right-hand side).

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

$$R_i(U_i)K_i = \left((1-\sigma_i^{-1})\alpha_i \frac{Y_i}{K_i}\right)K_i = (1-\sigma_i^{-1})\alpha_i Y_i$$

$$W_i(E_i)L_i = \left((1-\sigma_i^{-1})\beta_i \frac{Y_i}{L_i}\right)L_i = (1-\sigma_i^{-1})\beta_i Y_i$$

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

$$s_i^K = \frac{R_i(U_i)K_i}{R_i(U_i)K_i + W_i(E_i)L_i} = \frac{\alpha_i}{\alpha_i+\beta_i} = \frac{\alpha_i}{\gamma_i} \Leftrightarrow \alpha_i = \gamma_i s_i^K$$

$$s_i^L = \frac{W_i(E_i)L_i}{R_i(U_i)K_i + W_i(E_i)L_i} = \frac{\beta_i}{\alpha_i+\beta_i} = \frac{\beta_i}{\gamma_i} \Leftrightarrow \beta_i = \gamma_i s_i^L$$

We use these expressions to replace β_i and α_i in the expressions for $MRPL_{it}$ and $MRPK_{it}$ to obtain

$$MRPK_{it} = (1 - \sigma_i^{-1})\gamma_i s_i^K \frac{Y_{it}}{K_{it}}$$

$$MRPL_{it} = (1 - \sigma_i^{-1})\gamma_i s_i^L \frac{Y_{it}}{L_{it}}$$

Note that since markup $\mu_i = (\sigma_i - 1)/\sigma_i$,

$$(1 - \sigma_i^{-1})\gamma_i = \frac{1}{\mu_i}\gamma_i = \left(1 - s_{\pi,i}\right) \approx 1$$

given that the share of pure economic profit $s_\pi \approx 0$ (e.g., Basu and Fernald 1997). Hence, we can further simplify the expressions for $MRPL_{it}$ and $MRPK_{it}$ to obtain

$$MRPK_{it} \approx s_i^K \frac{Y_{it}}{K_{it}}$$

$$MRPL_{it} \approx s_i^L \frac{Y_{it}}{L_{it}}$$

Although it is conventional to define marginal products for physical units of capital and labor (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

$$\widetilde{MRP}K_{it} \equiv (1 - \sigma_i^{-1})\alpha_i \frac{Y_{it}}{\tilde{R}_{it}K_{it}} \approx \frac{R_{it}(U_{it})}{\tilde{R}_{it}}\left\{1 + \phi_K\left(\frac{K_{it}}{K_{i,t-1}} - 1\right) - \frac{\phi_K}{1 + r_{t+1}} \times \left(\frac{K_{i,t+1}}{K_{it}} - 1\right)\right\}$$

where $\tilde{R}_{it}$ is a measure of 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).[1] We are

[1] For example, suppose that capital is bought at time τ 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. $p_\tau K_\tau$ is also the balance-sheet value of fixed assets. The market price of capital at time t is given by $p_t = p_\tau \left(\frac{\Pi}{A}\right)^{t-\tau}$ where Π and A are gross rates of inflation and technical change. Hence,

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

If $\Pi > A$, a large share of state-of-the-art capital means a lower $\widetilde{MRP}K_{it}$ 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.

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 $R_{it}(U_{it}) \approx \tilde{R}_{it}$.

III. The Econometric Framework and Definition of Variables

Given our derivations in the previous section, we can express the marginal revenue product of capital (the left-hand side of equation (1')) as

$$\log MRPK_{isc} = \log(s_{isc}^{K}) + \log\left(\frac{Y_{isc}}{K_{isc}}\right) \tag{2}$$

The discussion in Section II also makes it clear that $\log MRPK_{isc}$ is a function of the quality of inputs, intensity of input utilization, and other variables (the right-hand side of equation (1')), which after further linearization may be summarized as

$$\log MRPK_{isc} = \psi_c + \kappa_s + \boldsymbol{X}_{isc}\boldsymbol{b} + \epsilon_{isc} \tag{3}$$

where subscripts *i*, *s*, and *c* index firms, sectors and countries, ψ_c is the set of country fixed effects, κ_s is the set of industry fixed effects, and $\boldsymbol{X}_{isc}$ is the vector of explanatory variables of interest defined below. By combining (2) and (3) we obtain an empirical "Mincerian-type" specification. An analogous specification and approach is used for labor.

If one follows Hsieh and Klenow (2009) and assumes that cost shares s_{isc}^{K} are fixed within each country *or* industry (that is, $s_{isc}^{K} = s_{s}^{K}$ or $s_{isc}^{K} = s_{c}^{K}$), the first term in equation (2) is absorbed into country or industry fixed effects.[2] We can go a step further and allow, more realistically, the cost shares s_{isc}^{K} to be fixed within each country *and* industry, $s_{isc}^{K} = s_{sc}^{K}$. This permits us to estimate the following regression with country × sector fixed effects ω_{sc}

$$\log MRPK_{isc} = \omega_{sc} + \boldsymbol{X}_{isc}\boldsymbol{b} + \epsilon_{isc}. \tag{4}$$

One has to expect that a significant part of the overall variation in firm-specific MRPK and MRPL will be absorbed by these country × sector fixed effects ω_{sc} and that a smaller share of total variation will be explained by the vector $\boldsymbol{X}_{isc}$ of explanatory variables of interest. The explanatory power of the $\boldsymbol{X}_{isc}$ vector will thus be greatly enhanced in the less parsimonious specification where it is interacted with country, sector or country × sector fixed effects.

[2] 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_{isc} and returns to scale in production γ_{isc} are constant across countries, industry or country/industry cells so that fixed effects absorb variation in σ_{isc} and γ_{isc}.

In estimating equation (4) and similar specifications, we generate several important "outputs". First, we obtain estimates of $\boldsymbol{b}$ and can evaluate how the explanatory variables $\boldsymbol{X}_{isc}$ predict MRPK and MRPL. Second, we can use ϵ_{isc} 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 the expected effect of explanatory variables depends on whether they represent a movement along the marginal product curve of capital (labor) or a shift in it, *ceteris paribus*. Moreover, with capital, labor, energy, material inputs, and (digital and transport) infrastructure all being relevant factors of production, capital and (different categories of) labor may be substitutes or complements. This means, for instance, that the effects of labor regulation and shocks may have positive or negative effects on the marginal product of capital.

IV. Data

The main data source for our analysis is the EIB Investment Survey (EIBIS). We next provide information on the design and implementation of the survey. We also compare EIBIS responses to the official data available in the ORBIS database. Once we establish consistency across the data sources, we describe survey questions that we use in the empirical analysis to account for 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 to senior managers or financial directors in all 28 EU Member States in 2016. The sampling targeted head offices.[3] Eligible respondents were senior persons with responsibility for investment decisions and how these were financed. This person could be the owner, a finance manager, the finance director or head of accounts, the Chief Financial Officer (CFO) or the Chief Executive Officer (CEO).

[3] 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.

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 5, with full-time and part-time employees being counted as one employee and employees working less than 12 hours per week being excluded. The Orbis dataset of Bureau van Dijk was used as the sampling frame in all countries.[4]

The fieldwork started in July 2016 and continued until November 2016. The vast majority of the interviews were conducted in the months of August and September 2016. The interview was administrated by telephone using computer-assisting telephone interviewing (CATI). The responses refer to the fiscal year 2015. The response rate was approximately 13 percent. The resulting sample consists of 12,483 non-financial enterprises in the 28 EU member states in NACE categories C to J. The sample size varies across countries and ranges from 150 enterprises in Cyprus and Luxembourg to 622 in Italy. A total sample of 12,300 firms was targeted, with 150, 400, 475 or 600 interviews per country depending on the size of the population. Firms are weighted to make them representative of the EU economy based on the number of employees. The weights compare the number of employees covered by the firms in EIBIS to the number of employees in the same country, industry, and size group according to Eurostat.

We use data from EIBIS to explore the relationship between MRPK and MRPL and a large number of explanatory variables at the firm level. We use questions on basic demographics, capacity utilization, quality of the capital stock, obstacles to long-term investment, investment plans, investment rate, employment growth, and sources of finance. This information is typically difficult to observe at the firm level and with data comparable across all 28 EU countries. When the MRPK is the dependent variable, the sample has 8,164 observations. With MRPL, it has 9,202 observations.[5]

B. Orbis database

The Orbis database is a popular source of administrate data for cross-country analyses at the firm level.[6] We use these data to check the validity of EIBIS responses. Specifically, we compare cross-

[4] A potential drawback of using Orbis was the level of coverage of the target population, in particular for small and young firms. The data on each firm from EIBIS was merged with Orbis and the match was done by Ipsos-Mori, which provided anonymized data to the EIB. This means that the EIB does not have the name, the address, the contact details or any additional individual information that could identify the firms in the final sample.
[5] We exclude a handful of outlying observations with implausible values of MRPK or MRPL.
[6] See Kalemli-Ozcan et al. (2015) for a detailed analysis of the advantages and disadvantages of using this dataset.

firm dispersion of the logarithm of sales, fixed assets and employment in the EIBIS and Orbis databases by country (Table 1) and by 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. In this exercise, we use only Orbis firms that participated in the EIBIS.

We observe high consistency across the two sources of data. For example, the correlation between log employment in EIBIS and Orbis is 0.83. The dispersion of the survey responses across firms is on average a bit larger than the dispersion in the administrative data, which is consistent with small noise (measurement error) in survey responses. Note that relative to the data on fixed assets, data on employment are available for fewer firms in the Orbis database.

In Figure 3 we present scatterplots of the dispersion of MRPK and MRPL across firms within countries, computed using the survey and administrative data. Although the dispersion of MRPK and MRPL is somewhat larger in the survey than in the administrative data (consistent with some measurement error in survey responses), the correlation of dispersion across the sources is reasonably high at 0.64 for MRPK and 0.78 for MRPL. We conclude that EIBIS provides satisfactory quality of firm-level data and therefore that the survey responses are suitable for our analysis.

C. EIBIS Variables Used in the Regression Analysis

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

<u>Demographics of the firm</u>

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/selection effects), both MRPK and MRPL may be higher for older firms.

Employment size – 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. The average log employment in our sample is 4.84 (approximately 130 employees).

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 to 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 30 percent 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.

Utilization and quality of inputs:

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 arte state-of-the-art" and "energy efficient buildings" are 42 and 37 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 5 percent above maximum capacity.

Obstacles to investment

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 "major obstacle", "minor obstacle", and "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 20-30 percent of firms report it to be a major obstacle and another 20-30 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 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.[7] 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 (3) and (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.[8]

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 14 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 (78 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 EIBIS data is that the survey asks firms to report their expected investment for the next three years. Specifically, firms can report whether investment will be for "replacing capacity (existing

---

[7] Since EIBIS does not have information about material costs, we assume implicitly that materials may be adjusted quickly.

[8] The results are similar when we also include an indicator variable equal to one if a firm has zero investment.

buildings, machine, equipment and IT)", "expanding capacity for existing products and services", "developing or introducing new products, processes or services", or "do not have 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" (41 percent).

<u>Source of funding</u>

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.

<u>Data filters and additional data</u>

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.

Since the survey does not collect information about cost shares, we use the Orbis database as well as national statistics to construct cost shares. Specifically, the labor share s_{it}^{L} is estimated for each firm in EIBIS using data from Orbis on the wage bill and costs of goods sold as $s_{it}^{L} = \frac{costs\ of\ employees_{\ it}}{costs\ of\ goods\ sold_{\ it}}$, where *i* and *t* index firms and time. The capital share s_{it}^{K} is estimated using data on the employment share and material costs as $s_{it}^{K} = 1 - \frac{costs\ of\ employees_{\ it}}{costs\ of\ goods\ sold_{\ it}} - \frac{material\ costs_{\ it}}{costs\ of\ goods\ sold_{\ it}}$.[9] To minimize measurement error, we use an average of the cost shares over the 2000-2015 period or a subset of this period if information is available for fewer years.[10] When data on the labor share or material costs are not available at the firm level in Orbis, 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 three sets of results. First, we explore the extent to which firm characteristics predict log(MRPK) and log(MRPL). Second, 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. Third, 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).

Before reporting the regression estimates, we note that there is a sizable dispersion of marginal products measured across all the firms in our sample – the standard deviation of log(MRPK) is 1.44 and that of log(MRPL) is 0.93. Hence, if one took the 28 EU countries as a single market where marginal products would ideally be equalized across all firms, then the current state of Europe is very

[9] The cost shares for labor and capital need to be between 0 and 1. When data on *costs of employees* in the Orbis database are missing, we use data on *value added – ebitda.* Similarly, when data on the *cost of goods sold* in Orbis are missing, we use data on *turnover* or *sales* instead. When data on *material costs* are missing, we use data on *turnover – value added* instead.

[10] For example, if the firm started operating in 2013, the cost shares will refer to the averages over 2013, 2014 and 2015. But if the firms started operating before 2000, the cost shares will be an average over all years from 2000 to 2015.

far from that. With these baseline indicators in mind, we explore in the regression framework how much variation in MRPK and MRPL can be accounted for by the observed firm characteristics.

A. Regression results

Given our derivation of the estimating equation, our preferred specification for the regression analysis is equation (4) in which we enter as regressors variables $\boldsymbol{X}$ together with country × industry fixed effects. The estimated coefficients for this specification are reported in Table 4.

Turning to the coefficient estimates 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 welfare 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 welfare 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 (instrumental variable estimators have lower R^2 than OLS).

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 20 percent (log points) in firms that are 10-19 years old and 35 percent (log points) in firms that are 20+ years old. The effect of firm age on MRPL is estimated to increase by 10 percent after the first four years of existence and slightly increases over time thereafter, reaching 13 percent for firms that are 20+ years old. The higher MRPL in older firms could be related to a greater accumulation of firms-specific human capital over time. The estimated coefficient on log of employment in Table 4 is economically small and not statistically significant. On the other hand, employment size is positively associated with MRPK. Subsidiary status is positively related to MRPK and MRPL; the estimated coefficients are 45 percent (log points) for MRPK and 11 percent for MRPL, respectively. Export status is associated with higher MRPK (coefficient is 12 log points) and MRPL (coefficient is 18 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.

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 for state-of-the-art machinery and equipment, and insignificant for 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.

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 MRPL and MRPK. With capacity utilization already being controlled for, deficient demand may signal that firms are cautious in augmenting input use and prefer to keep marginal products high. Having availability of staff with the right skills as an obstacle to investment has no robust association with MRPK and a weak, 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 a major obstacle is generally associated with a positive coefficient on MRPL, suggesting that these types of infrastructure are complements to labor. The availability of adequate transport infrastructure as a constraint does not appear to have an association with MRPK, but when access to digital infrastructure is a major constraint on investment it does have a positive relationship with MRPK.

Labor market regulation as an obstacle to investment has a negative association with MRPL and no material correlation with MRPK. Somewhat surprisingly, business regulations and taxation has no discernible association with MRPL and has some negative correlation with MRPK, suggesting that while it may have other effects (e.g., on scale), it does not appear to have a significant effect on the relative use of capital and labor.

Unavailability of finance does not have a significant correlation with MRPK, but it is negatively related to MRPL when it is a major obstacle. Uncertainty about future has no statistically significant association with MRPK but it has positive correlation with MRPL when uncertainty represents a major constraint.

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.

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 most coefficients not being statistically significant. To quantify this observation, in Table 5 we present incremental 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 can observe for MRPK and MRPL that incremental R^2s are the largest for variables in the "adjustment" block and generally 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 "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 incremental R^2s 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.

To quantify the magnitudes of gains that one could obtain by eliminating constraints and distortions (if one interprets a given part of $\boldsymbol{X}$ as genuine distortions), we use the Hsieh-Klenow insight that (log) productivity losses due to misallocation of resources may be approximated with

$$\frac{\sigma}{2} var(\alpha_i \log(MRPK_i) + \beta_i \log(MRPL_i))$$

where σ is the elasticity of demand.[11] Assuming constant elasticities in the productions function (i.e., $\alpha_i = \alpha_j$ and $\beta_i = \beta_j$ for any firm i and firm j) and using $s^K \propto \alpha$ and $s^L \propto \beta$, we can obtain

$$\frac{\sigma}{2} var(\alpha \log(MRPK_i) + \beta \log(MRPL_i)) \approx$$

$$\frac{\sigma}{2}\{(s^K)^2 var(\log(MRPK_i)) + (s^L)^2 var(\log(MRPL_i))\}$$

$$+\frac{\sigma}{2} 2 s^K s^L \rho_{LK} \sqrt{var(\log(MRPK_i)) var(\log(MRPL_i))}$$

In the data, $\bar{s}^K = 0.28$ and $\bar{s}^L = 0.24$ and $\rho_{LK} \equiv \rho(MRPK_i, MRPL_i) \approx 0.21$. Since we do not have an estimate of the elasticity of demand (σ), we follow Hsieh and Klenow (2009) and set $\sigma = 3$ which likely yields a conservative estimate of welfare losses due to misallocation.

With these results, we can carry out calculations for several policy scenarios. First, assume that the policy makers would eliminate the dispersion in MRPK brought about by the "distortions" group, while holding the dispersion of MRPL fixed. Then, if we use the estimated dispersion of log(MRPK) from Table 3 and the incremental R^2 for the "distortions group" in the first column of Table 5, the gain in productivity is

$$\frac{\sigma}{2} \times (s^K)^2 \times var(\log(MRPK_i)) \times (incremental\ R^2) = 1.5 \times 0.28^2 \times 1.44^2 \times 0.063 = 0.015.$$

That is, removing distortions captured by ***X*** can generate productivity gains up to 2 log percentage points. A similar calculation for the marginal revenue product of labor yields

$$\frac{\sigma}{2} \times (s^L)^2 \times var(\log(MRPL_i)) \times (incremental\ R^2) = 1.5 \times 0.24^2 \times 0.93^2 \times 0.106 = 0.008.$$

Given that capital and labor distortions can reinforce each other (this corresponds to the term $\sigma s^K s^L \rho_{KL} \sqrt{var(\log(MRPK_i)) var(\log(MRPL_i))}$), the additional gain is 0.057, so the total productivity gain can be approximately 8 log percentage points.

Second, consider the possibility that all variables in ***X*** capture distortions (and hence we can use R^2 from regressions with ***X*** only, reported in Table 4). In this case, the gains from removing distortions in the allocation of capital are $1.5 \times 0.28^2 \times 1.44^2 \times 0.112 = 0.027$. The gains for labor are $1.5 \times 0.24^2 \times 0.93^2 \times 0.271 = 0.02$. The corresponding additional effect derived from removing distortions in labor *and* capital is 0.057. As a result, the total gain is approximately 10.4 log percentage points. Note that although significant, this estimated gain is smaller than the one that Hsieh and Klenow (2009) provide when they compare the U.S. to developing countries.

[11] Hsieh and Klenow (2009) assume that there are no distortions in inputs other than capital. We allow for distortions in the utilization of labor, but we assume that there are not distortions in the allocation of material inputs.

Obviously, our differences are smaller because we compare countries with more similar levels of development. More importantly, in the above calculations we compute gains from removing specific, measurable distortions (collected in $\boldsymbol{X}$) and therefore we do not treat the entire residual dispersion as a source of possible productivity gains. In contrast, Hsieh and Klenow (2009) compute gains using the raw dispersion, which treats all sources of variation as potential distortions – an estimate that we report in the next section.

B. Cross-country and cross-industry differences

Although a key objective of the EU has been to improve allocation of resources across member countries, the dispersion of marginal revenue products across member countries remains relatively large. For example, in Table 3 we show that the "raw" standard deviation of log (MRPK) across firms in the 28 European countries covered in EIBIS is 1.44, while this dispersion for the U.S. is 0.98 (Table 2 in Asker et al., 2014). Similarly, the "raw" standard deviation of log (MRPL) is 0.93 in EIBIS and it is found to be 0.58 for the U.S. (Table 1 in Bartelsman et al., 2013).

In our data, there is also considerable cross-country variation in the average marginal revenue products – 0.49 for log (MRPK) and 0.78 for log (MRPL) – but this variation is small relative to the within-country variation in MRPK and MRPL. In Figure 4, 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 and industry fixed effects, and variables $\boldsymbol{X}$ plus country×industry fixed effects) in a cross-country regression given by equations (3) and (4). As may be seen in Figure 5, 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 5 the dispersion of raw marginal products is particularly high in smaller countries such as Malta (MT), Luxembourg (LU), Lithuania (LT), and Ireland (IE). 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 5 (MRPK and MRPL adjusted for observed firm characteristics $\boldsymbol{X}$, with no fixed effects included). Although using ***X*** reduces the cross-sectional dispersion, it generally preserves the ranking of the countries. Adding industry and country fixed effects further reduces the levels of dispersion and the ranking of countries is generally preserved, although the ranking for some countries jumps (e.g., Romania (RO) is similar to Austria (AT) in terms of "raw" MRPK dispersion, but after this adjustment Romania becomes more similar to the Netherlands (NL)). Introducing country × industry fixed effects not only reduces the level of dispersion, it also attenuates differences across countries. For example, France, Italy, the UK, and Poland 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 or country × industry fixed effects 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.

As may be seen from the bottom rows of Table 4, vector ***X*** alone can account for $R^2 = 0.112$ share of variation in MRPK and $R^2 = 0.271$ of variation in MRPL when no fixed effects are included. Adding country fixed effects raises the R^2 to 0.117 for MRPK and 0.674 for MRPL. In contrast, adding (2-digit level) industry fixed effects raises R^2 to 0.209 for MRPK and 0.313 for MRPL. Including industry and country fixed effects raises R^2 to 0.272 for MRPK and 0.699 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 industry and country 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 and industry 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.[12] 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.

Moreover, there is a large increase in the R^2 when we introduce country × industry fixed effects: R^2 is 0.475 for MRPK and 0.777 for MRPL (bottom of Table 4). 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, if we use incremental R^2's relative to the specification with ***X***, removing inequality in average marginal revenue products across countries (i.e., making the country fixed effects be all identical) would raise productivity by at least 6 percentage log points using the Hsieh-Klenow framework. Removing barriers between industries *and* countries (i.e., making the country × industry fixed effects all identical) would raise productivity by at least 25 percentage log points.

Finally, we also perform an analogous exercise on the assumption that the effects of the ***X*** variables are industry or country specific. In such scenarios, also reported at the bottom of Table 4, the ***X*** variables account for more than 50 percent of total variation in MRPK and MRPL. This

[12] For MRPK, industry FE increment in R^2: 0.209-0.112 = 0.097; Country FE increment in R^2: 0.176-0.112 = 0.064; and Country + Industry FE increment in R^2: 0.275-0.112 = 0.163.

more flexible specification testifies to the potential importance of firm-specific variables in explaining the overall dispersion in MRPK and MRPL.[13]

C. 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.[14] We start by using Germany and Greece as two polar cases – $\sigma(\log(MRPK))$ is 0.98 for Germany and 1.81 for Greece, while $\sigma(\log(MRPL))$ is 0.46 for Germany and 0.77 for 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.

In Figure 6, 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 $\boldsymbol{X}$ and $\boldsymbol{b}$ (solid black line), as well as a counterfactual distribution using Greek $\boldsymbol{X}$ and German $\boldsymbol{b}$ (long-dash, blue line) and a counterfactual distribution using German $\boldsymbol{X}$ and Greek $\boldsymbol{b}$ (short-dash, red line). Using Greek $\boldsymbol{X}$ and German $\boldsymbol{b}$ results

[13] These analogous exercises are carried out on the assumption that the effects of the $\boldsymbol{X}$ variables are industry or country specific. As may be seen from the bottom of Table 4, when we interact firm characteristics $\boldsymbol{X}$ with industry fixed effects, hence allowing the effects of the firm characteristics to vary by industry (but not country), we find that these variables account for 54.5 percent of the total variation in log (MRPK) and 60 percent of the variation in log (MRPL). Similarly, when we interact firm characteristics with country fixed effects, allowing the effects to vary by country (but not industry), we find that these variables account for 53.8 percent of the total variation in log (MRPK) and 73.1 percent of the variation in log (MRPL).

[14] 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 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 1.08 rather than 1.81).

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 6 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 6, 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.

Using a simple dynamic theoretical framework as a guide, we find in our empirical work 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. Comparing our findings on the dispersion of MRPK across firms in the 28 EU countries to the dispersion found by other researchers for the US indicates that the dispersion in the EU is about twice as large. Our calculations suggest that by reducing the EU dispersion of MRPK to the level of the US would increase EU productivity (GDP) by more than 30 percent.

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 industries *and* countries (i.e., made the country × industry fixed effects in the regression analysis all identical), EU GDP (productivity) would rise by at least 25 percentage log points. When we allow the firm-level characteristics alone to explain the dispersion in marginal products in EIBIS, we find that these variables account for 11.2 percent of the total variation in log (MRPK) and 27.1 percent of the variation in log (MRPL). When we allow the effects of firm characteristics to be industry or country specific, they account for more than 50 percent of total variation in log (MRPK) and log (MRPL). This more flexible specification testifies to the potential importance of firm-specific variables in explaining the overall dispersion in MRPK and MRPL.

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 a Machado-Mata decomposition, we show that cross-country variation in the within-country dispersion of marginal revenue products is largely brought about 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.

Bils, Mark, Peter J. Klenow, and Cian Ruane (2017). “Misallocation or Mismeasurement?” manuscript.

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.

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.

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.

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

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

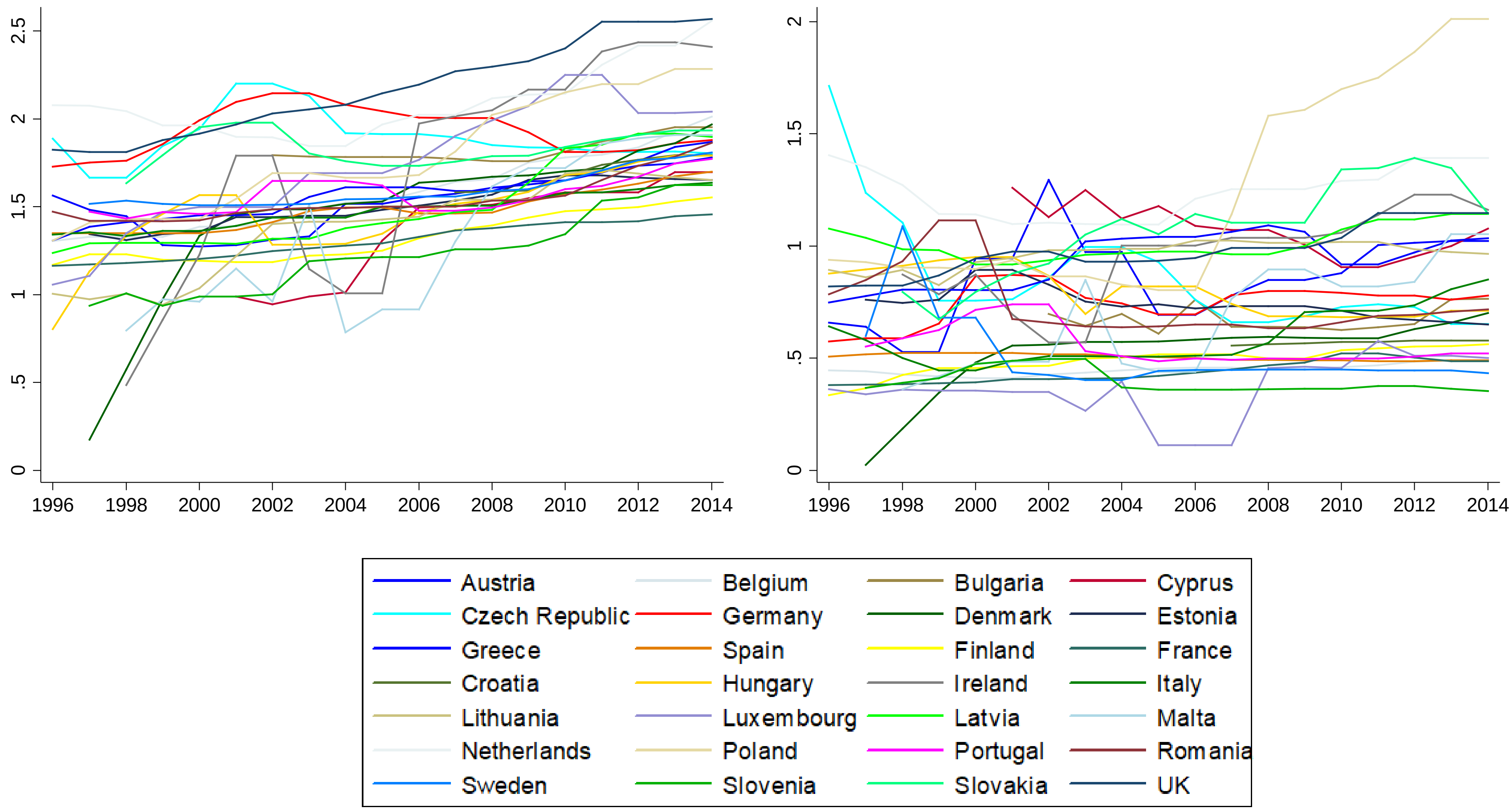
Restuccia, Diego, and Richard Rogerson (2017). “The Causes and Costs of Misallocation,” *Journal of Economic Perspectives* 31(3): 151-74.

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

Figure 1. Evolution of the dispersion of the marginal revenue products of capital and labor, by country

Panel A. Marginal revenue product of capital (MRPK)

Panel B. Marginal revenue product of labor (MRPL)

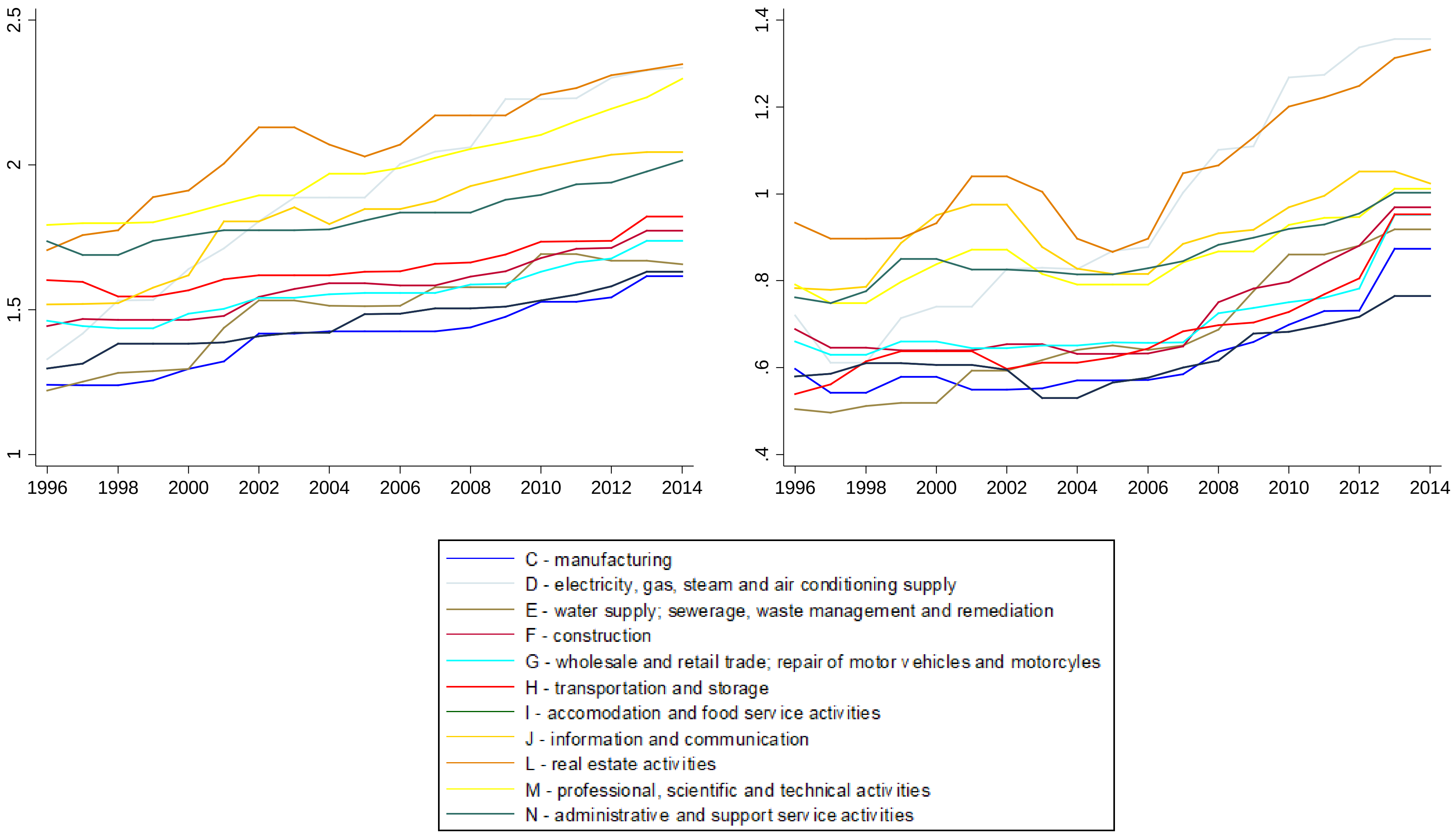


Note: The figure plots time series (3-year moving median) of the standard deviation of the logarithm of the marginal revenue product of capital (MRPK) and labor (MRPL) for each country using data on firms in Orbis. The dispersion is computed after projecting MRPK and MRPL on country × industry fixed effects (industry at 2-digit NACE).

Figure 2. Evolution of the dispersion of marginal revenue products, by industry

Panel A. Marginal revenue product of capital (MRPK)

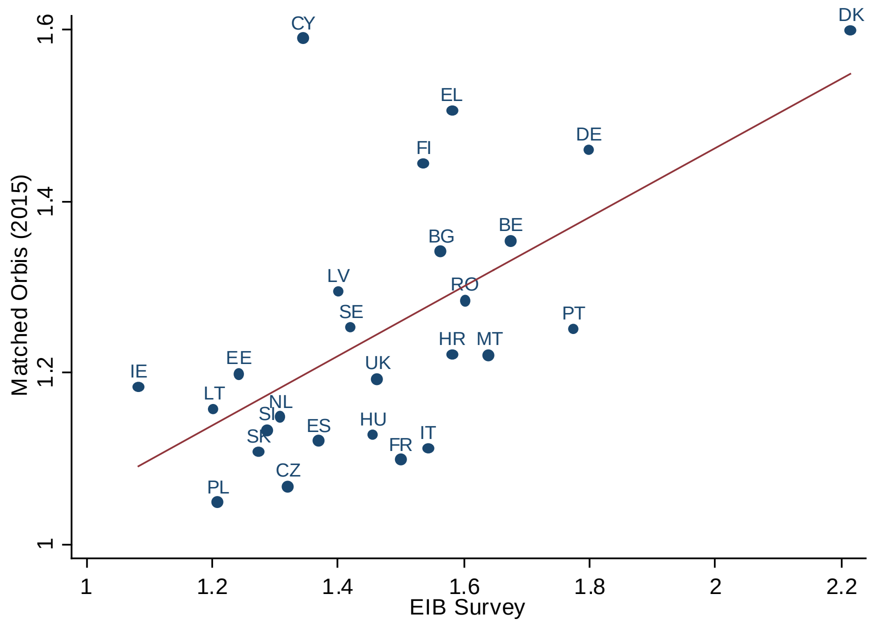
Panel B. Marginal revenue product of labor (MRPL)



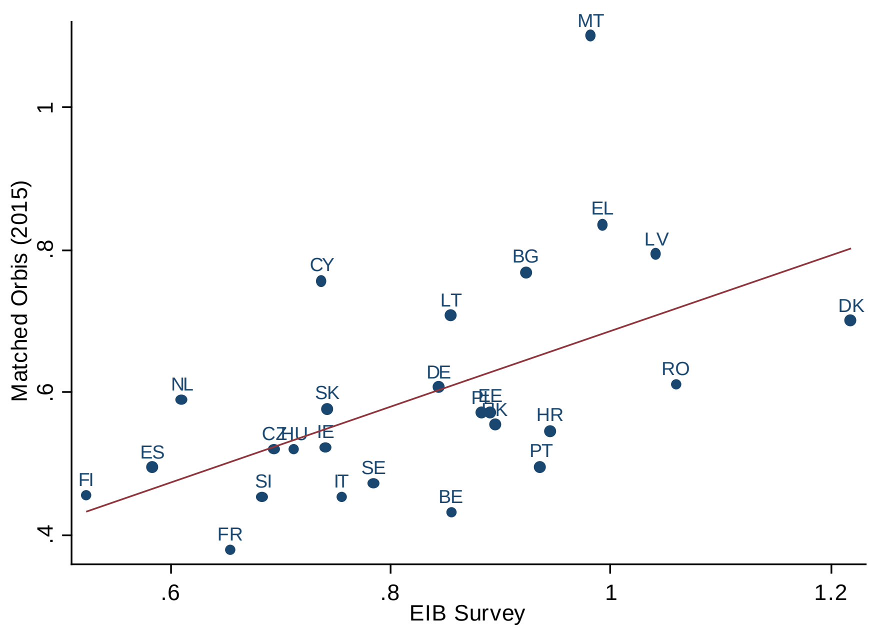
Notes: The figure plots time series (3-year moving median) of standard deviation of log marginal revenue product of capital (MRPK) and labor (MRPL) for each country. The dispersion is computed after projecting log marginal revenue products on country × industry fixed effects (industry at 2-digit NACE).

Figure 3. Dispersion of marginal revenue products of capital and labor in the EIB Investment Survey and Orbis

Panel A. Marginal revenue product of capital (MRPK)



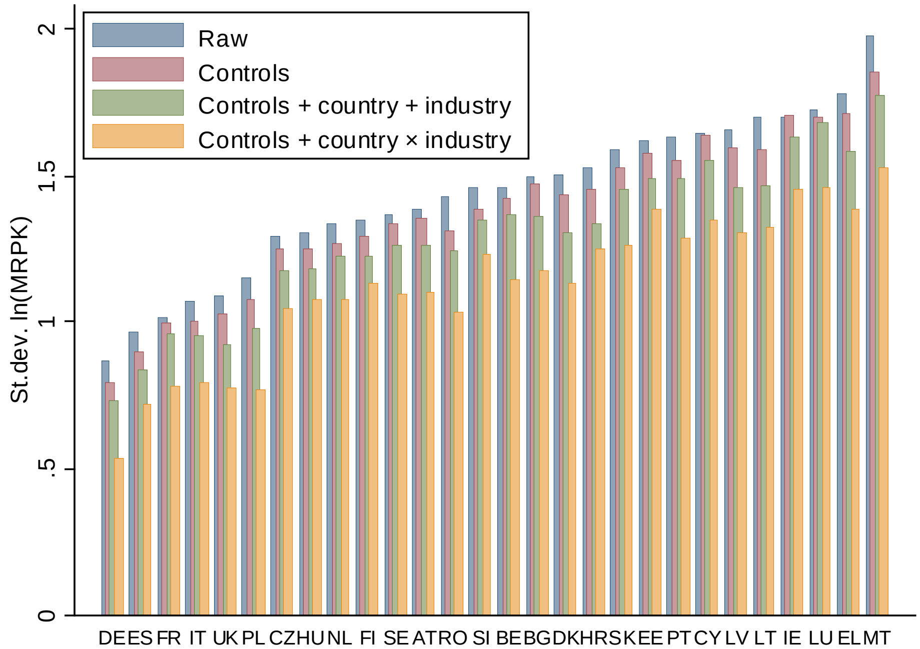
Panel B. Marginal revenue product of labor (MRPL)



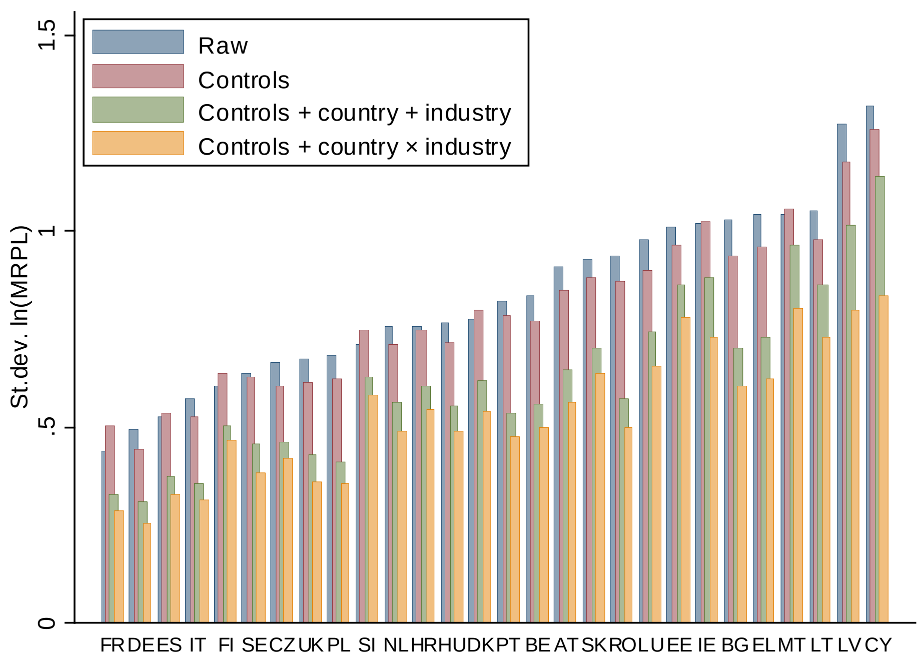
Notes: The figures plot standard deviations of the logarithm of the marginal revenue product of capital (MRPK) and marginal revenue product of labor (MRPL) across firms using data of the EIB Investment Survey (horizontal axis) and the Orbis database (vertical axis). Standard deviations are computed after controlling for country × industry fixed effects (industry at 2-digit NACE). The red, solid line shows the fit of a linear regression. For MRPK, the slope of the fitted line is 0.40 (s.e.=0.11, ρ=0.60). For MRPL, the slope of the fitted line is 0.53 (s.e.=0.16, ρ=0.55). Austria and Luxemburg are outliers and are excluded from the scatterplots. 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. Raw and residual dispersion of the marginal revenue products of capital and labor

Panel A. Marginal revenue product of capital (MRPK)

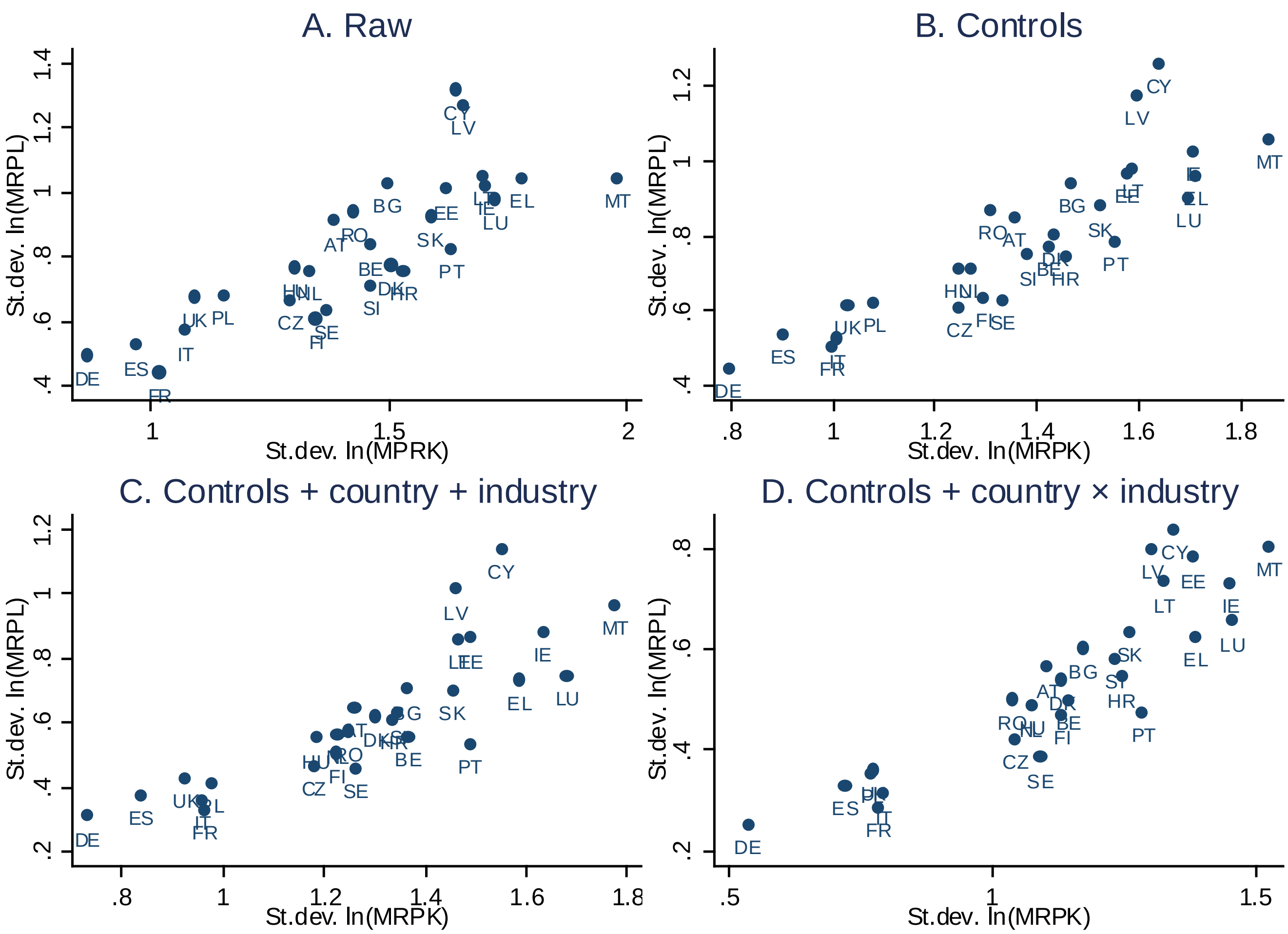


Panel B. Marginal revenue product of labor (MRPL)



Notes: 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” add fixed effects for industries and for countries to firm-level characteristics. “Controls + country × industry” add country × industry 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.

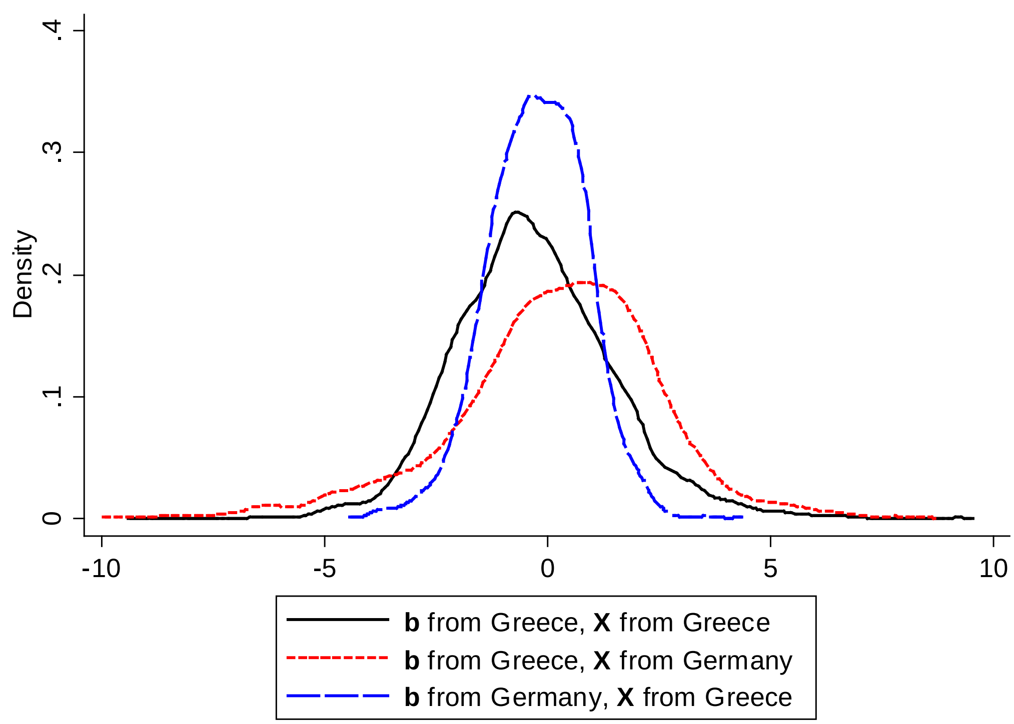
Figure 5. Association of the dispersion of the marginal revenue products of capital and labor



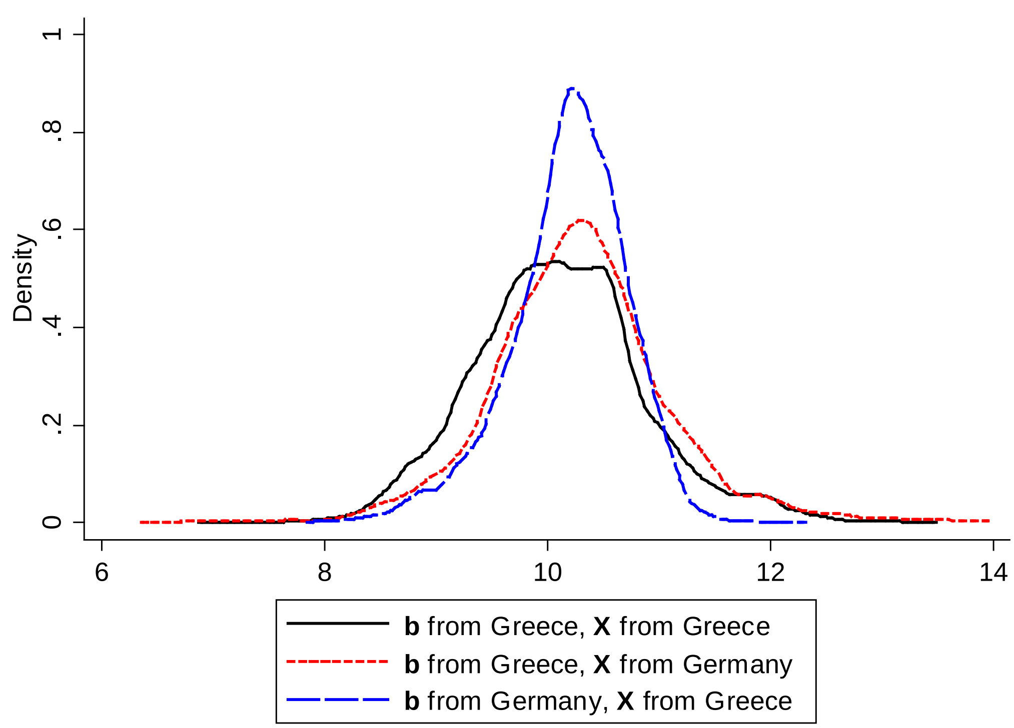
Note: The figures show the association between the dispersion of MRPK and MRPL across countries. "Raw" means no controls. "Controls" include the firm-level characteristics described in section IV. "Controls + country + industry" add fixed effects for industries and for countries to firm-level characteristics. "Controls + country × industry" add country × industry 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 6. 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)	(10)
Austria	475	2.53	2.56	0.60	2.89	3.02	0.74	1.74	2.18	0.83
Belgium	421	1.84	2.42	0.79	2.88	3.07	0.77	1.75	2.56	0.72
Bulgaria	474	2.46	2.54	0.92	2.86	2.97	0.88	1.84	1.78	0.97
Cyprus	147	0.92	2.49	0.86	1.85	2.55	0.67	0.82	1.57	0.92
Czech Rep.	467	2.12	2.36	0.81	2.44	2.66	0.80	1.73	1.76	0.94
Germany	590	2.24	2.83	0.70	2.99	2.90	0.82	1.73	2.30	0.62
Denmark	456	1.95	2.30	0.65	3.29	3.08	0.92	1.79	2.19	0.65
Estonia	383	2.06	2.04	0.89	2.56	2.27	0.83	1.70	1.89	0.95
Greece	396	2.04	2.29	0.85	2.69	2.59	0.84	1.71	1.67	0.91
Spain	452	2.29	2.58	0.87	2.81	2.96	0.87	2.01	2.09	0.94
Finland	428	2.60	2.55	0.82	3.02	2.90	0.89	2.32	2.39	0.83
France	570	2.27	2.28	0.81	2.72	2.75	0.77	1.66	2.06	0.96
Croatia	462	2.37	2.37	0.84	3.14	2.69	0.82	2.03	1.99	0.96
Hungary	424	2.59	2.56	0.92	3.21	3.10	0.90	2.26	2.24	0.99
Ireland	360	1.65	2.37	0.95	2.91	2.36	0.65	1.68	2.38	0.58
Italy	589	2.42	2.77	0.92	3.12	3.49	0.86	2.08	2.30	0.94
Lithuania	402	2.50	2.52	0.91	3.08	3.21	0.89	1.95	1.83	0.97
Luxembourg	148	1.74	2.36	0.56	2.64	2.60	0.35	1.35	1.85	0.86
Latvia	378	2.28	2.60	0.93	3.19	3.16	0.89	1.65	1.92	0.94
Malta	159	2.51	1.50	0.38	2.37	2.13	0.63	1.24	1.40	0.98
Netherlands	472	1.55	2.37	0.94	3.07	2.84	0.61	2.06	2.25	0.84
Poland	462	1.92	2.23	0.87	2.25	2.62	0.85	2.68	1.87	0.97
Portugal	468	2.07	2.29	0.81	2.72	2.74	0.67	1.72	1.80	0.97
Romania	457	2.05	2.16	0.80	3.01	2.86	0.83	1.86	1.74	0.96
Sweden	439	2.34	2.50	0.86	2.96	2.76	0.82	1.97	2.19	0.91
Slovenia	400	2.01	2.08	0.87	2.57	2.49	0.79	1.65	1.74	0.88
Slovakia	380	2.15	2.35	0.95	2.78	2.58	0.83	1.89	1.87	0.95
UK	520	1.63	2.54	0.77	2.88	2.94	0.85	1.38	2.19	0.78
All countries	11,179	2.13	2.54	0.80	2.87	2.91	0.82	1.83	2.14	0.83

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) St. dev. Orbis	log(sales) St. dev. EIBIS	log(sales) Correl. coeff.	log(fixed assets) St. dev. Orbis	log(fixed assets) St. dev. EIBIS	log(fixed assets) Correl. coeff.	log(employment) St. dev. Orbis	log(employment) St. dev. EIBIS	log(employment) Correl. coeff.
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
10-12	food; beverages; tobacco	499	1.98	2.34	0.93	2.07	2.33	0.73	1.70	1.76	0.94
13-15	textiles; apparel; leather and related products	288	2.09	2.11	0.90	2.59	2.30	0.83	1.47	1.58	0.95
16-18	wood; paper; printing and recorded media	336	2.00	2.60	0.82	2.40	2.26	0.75	1.87	1.81	0.88
19-20	coke and refined petroleum; chemicals	130	1.95	2.25	0.83	2.87	2.87	0.66	1.61	1.90	0.85
21	pharmaceutical products	45	2.08	2.47	0.81	1.80	2.60	0.82	1.20	1.09	0.89
22-23	rubber and plastic products; mineral products	376	1.89	2.19	0.87	2.22	2.59	0.75	1.51	1.78	0.89
24-25	basic and fabricated metal products	609	2.17	2.20	0.86	2.45	2.48	0.86	1.64	1.73	0.94
26	computer, electronic and optical products	137	1.90	2.40	0.94	2.21	2.76	0.82	1.46	2.00	0.95
27	electrical equipment	131	2.15	2.37	0.96	3.27	1.92	0.69	2.00	2.03	0.97
28	machinery and equipment	333	1.86	1.83	0.88	2.34	2.23	0.77	1.50	1.62	0.92
29-30	motor vehicles; other transport equipment	136	1.93	2.02	0.93	1.93	2.07	0.81	1.51	1.59	0.97
31-33	furniture; other manuf.; repair and installation	333	1.92	2.06	0.80	2.20	2.20	0.79	1.65	1.72	0.94
35	electricity, gas, steam and air conditioning	221	2.66	2.97	0.82	2.99	3.18	0.88	2.03	2.35	0.89
36-39	water supply; sewerage and waste management	482	2.12	2.36	0.95	2.76	3.01	0.89	1.80	1.85	0.95
41	construction of buildings	799	2.65	2.91	0.87	2.86	3.10	0.69	2.03	2.23	0.89
42	civil engineering	404	2.29	2.52	0.88	2.45	2.66	0.81	1.87	2.10	0.90
43	specialised construction activities	1,416	2.25	2.22	0.91	2.36	2.62	0.67	1.75	1.93	0.95
45	wholesale and retail trade	354	2.29	2.42	0.82	2.20	2.49	0.75	1.88	2.05	0.94
46	wholesale trade, except of motor vehicles	1,230	2.32	2.93	0.75	2.88	2.93	0.85	1.90	2.33	0.74
47	retail trade, except of motor vehicles	777	2.47	2.57	0.74	2.82	2.77	0.71	2.20	2.23	0.76
49-53	transportation and storage	1,796	2.54	2.49	0.81	3.26	3.14	0.83	2.07	2.18	0.90
55-56	accommodation and food service activities	508	2.12	2.09	0.87	2.87	3.03	0.72	1.78	1.98	0.91
58-63	information and communication	348	2.68	2.75	0.97	3.62	3.24	0.87	2.02	2.00	0.95
64-99	other services	91	2.07	3.17	0.84	3.65	3.48	0.94	1.14	3.48	0.42
10-99	all industries	11,779	2.25	2.45	0.84	2.69	2.74	0.78	1.83	2.03	0.87

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)	16.36	2.26
	log(fixed assets)*	15.16	2.74
	log(employment)	4.61	1.93
	log(MRPK)*	-0.14	1.44
	log(MRPL)	10.15	0.93
Demographics	Firm age		
	less than 5 years	0.03	0.18
	5-9 years	0.09	0.28
	10-19 years	0.22	0.41
	20+ years	0.67	0.47
	Subsidiary	0.30	0.46
	Exporter	0.49	0.50
Quality of capital and other inputs	Share of state-of-the art machinery and equipment	0.42	0.32
	Share of high energy efficiency commercial building stock	0.37	0.33
Capacity utilization	above maximum capacity	0.05	0.22
	at maximum capacity	0.44	0.50
	somewhat below full capacity	0.38	0.49
	substantially below full capacity	0.10	0.29
Obstacles to investment	Demand for products or services		
	Major	0.26	0.44
	Minor	0.23	0.42
	Availability of staff with the right skills		
	Major	0.40	0.49
	Minor	0.29	0.45
	Energy costs		
	Major	0.21	0.41
	Minor	0.32	0.47
	Access to digital infrastructure		
	Major	0.11	0.31
	Minor	0.25	0.43
	Labor market regulations		
	Major	0.29	0.45
	Minor	0.30	0.46
	Business regulations and taxation		
	Major	0.31	0.46
	Minor	0.28	0.45
	Availability of adequate transport infrastructure		
	Major	0.15	0.35
	Minor	0.24	0.43
	Availability of finance		
	Major	0.25	0.43
	Minor	0.21	0.41
	Uncertainty about future		
	Major	0.40	0.49
	Minor	0.31	0.46
Adjustment	Investment, log(1 + investment)	11.96	4.18
	Percent change in employment in the last three years	0.14	0.55
	Investment over the last three years		
	too much	0.03	0.18
	about the right amount	0.78	0.41
	too little	0.17	0.38
	company did not exist three years ago	0.00	0.02
	Investment priority in the next three years		
	replacing capacity	0.41	0.49
	capacity expansion for existing products or services	0.24	0.43
	developing new products, processes or services	0.24	0.43
	no investment planned	0.10	0.30
Source of funds	internal funds or retained earnings	0.66	0.37
	external finance	0.32	0.36
	intra-group funding	0.02	0.13
	Finance constrained	0.07	0.26
Sample size*		9,202	9,202

Note: All statistics are computed using sampling weights. * The sample size is 8,164 for the variables fixed assets and MRPK.

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

Regressor	Dependent variable	
	log(MRPK)	log(MRPL)
Demographics		
Firm age (omitted category: less than 5 years)		
5-9 years	-0.001	0.103***
	(0.075)	(0.032)
10-19 years	-0.204***	0.117***
	(0.063)	(0.029)
20+ years	-0.356***	0.131***
	(0.063)	(0.027)
log(employment)	0.027**	0.004
	(0.010)	(0.005)
Subsidiary	0.448***	0.110***
	(0.037)	(0.016)
Exporter	0.115***	0.180***
	(0.033)	(0.014)
Quality of capital and other inputs		
Share of state-of-the art machinery and equipment, including ICT	-0.160***	0.096***
	(0.045)	(0.019)
Share of high energy efficiency commercial building stock	-0.182***	-0.005
	(0.041)	(0.016)
Capacity utilization (omitted category: somewhat below full capacity)		
above maximum capacity	0.319***	0.139***
	(0.055)	(0.025)
at maximum capacity	0.058**	0.026**
	(0.028)	(0.011)
substantially below full capacity	-0.280***	-0.093***
	(0.043)	(0.017)
Obstacles to investment (omitted category: not an obstacle at all)		
Demand for products or services		
Major	0.141***	0.041***
	(0.034)	(0.014)
Minor	0.117***	0.038***
	(0.029)	(0.013)
Availability of staff with the right skills		
Major	0.042	-0.036***
	(0.035)	(0.014)
Minor	0.028	0.018
	(0.034)	(0.014)
Energy costs		
Major	-0.115***	-0.062***
	(0.036)	(0.015)
Minor	-0.135***	-0.011
	(0.031)	(0.014)
Access to digital infrastructure		
Major	0.084**	0.035**
	(0.043)	(0.016)
Minor	0.001	0.006
	(0.032)	(0.014)
Labor market regulations		
Major	-0.053	-0.098***
	(0.036)	(0.014)
Minor	0.024	-0.050***
	(0.032)	(0.013)
Business regulations and taxation		
Major	-0.087**	0.005
	(0.037)	(0.014)

Minor	-0.017	-0.005
	(0.034)	(0.015)
Availability of adequate transport infrastructure		
Major	0.007	0.070***
	(0.037)	(0.016)
Minor	0.028	0.015
	(0.032)	(0.013)
Availability of finance		
Major	-0.008	-0.047***
	(0.035)	(0.015)
Minor	-0.034	-0.023
	(0.033)	(0.014)
Uncertainty about future		
Major	0.024	0.024*
	(0.035)	(0.014)
Minor	-0.004	0.004
	(0.034)	(0.015)
Adjustment		
Investment, log(1 + investment)	-0.058***	0.030***
	(0.005)	(0.002)
Percent change in employment in the last three years	0.072***	-0.075***
	(0.023)	(0.009)
Investment over the last three years (omitted category: about the right amount)		
too much	-0.283***	-0.083***
	(0.060)	(0.024)
too little	-0.111***	-0.024*
	(0.031)	(0.013)
company did not exist three years ago	-0.996**	-0.601***
	(0.451)	(0.159)
Investment priority in the next three years (omitted category: no investment planned)		
replacing capacity	-0.108**	0.007
	(0.047)	(0.019)
capacity expansion for existing products or services	-0.126***	0.001
	(0.049)	(0.021)
developing new products, processes or services	-0.030	0.038*
	(0.050)	(0.021)
Source of funds (omitted category: external finance)		
internal funds or retained earnings	0.152***	0.055***
	(0.038)	(0.016)
intra-group funding	-0.160	0.135***
	(0.113)	(0.051)
Credit constrained	-0.104**	-0.083***
	(0.045)	(0.019)
Sample size	8,164	9,202
R^2	0.477	0.777
Memorandum:		
R^2 with country × industry fixed effects and no $\boldsymbol{X}$	0.430	0.746
R^2 with $\boldsymbol{X}$ and no fixed effects	0.112	0.271
R^2 with $\boldsymbol{X}$ and country fixed effects	0.176	0.674
R^2 with $\boldsymbol{X}$ and industry fixed effects	0.209	0.313
R^2 with $\boldsymbol{X}$ and country fixed effects and industry fixed effects	0.275	0.699
R^2 with $\boldsymbol{X}$ and slopes varying by industry	0.545	0.600
R^2 with $\boldsymbol{X}$ and slopes varying by country	0.538	0.731

Note: The table reports estimates of equation (4) with industry × country 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. Change in R^2 when adding a group of variables to a specification with fixed effects

Group of variables	List of fixed effects: No fixed effects	Country	Industry	Country + industry	Country × industry
Panel A: MRPK					
Demographics	0.039	0.038	0.023	0.022	0.017
Quality of capital	0.014	0.007	0.011	0.006	0.007
Capacity utilization	0.011	0.012	0.006	0.008	0.006
Obstacles for investment	0.023	0.018	0.011	0.009	0.008
Adjustment	0.049	0.049	0.018	0.015	0.015
Source of funds	0.011	0.012	0.009	0.009	0.009
"Compensating differentials" group	0.073	0.067	0.035	0.029	0.028
"Distortions" group	0.063	0.060	0.035	0.035	0.027
Panel B: MRPL					
Demographics	0.078	0.023	0.062	0.016	0.013
Quality of capital	0.022	0.003	0.018	0.003	0.002
Capacity utilization	0.006	0.002	0.005	0.002	0.002
Obstacles for investment	0.038	0.012	0.029	0.008	0.007
Adjustment	0.073	0.029	0.062	0.023	0.017
Source of funds	0.015	0.003	0.010	0.002	0.002
"Compensating differentials" group	0.095	0.038	0.081	0.030	0.025
"Distortions" group	0.106	0.026	0.085	0.017	0.015

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 country 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".

Table 6. 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	1.35	1.01	1.71	2.15	1.72	0.66	0.47	0.75	0.91	0.74
Belgium	1.43	1.05	1.73	2.20	2.12	0.60	0.44	0.69	0.95	0.67
Bulgaria	1.47	1.06	1.56	2.19	1.56	0.79	0.51	0.94	0.92	0.92
Cyprus	1.93	1.03	3.34	1.95	2.70	1.19	0.55	2.50	0.82	2.17
Czech Rep.	1.26	1.01	1.62	2.74	1.58	0.56	0.52	0.63	0.87	0.76
Germany	0.98	0.98	0.98	2.37	1.08	0.46	0.46	0.46	0.86	0.52
Denmark	1.41	1.04	1.44	2.29	1.89	0.64	0.45	0.65	0.84	0.82
Estonia	1.55	1.04	2.13	2.02	1.97	0.90	0.51	1.12	0.86	1.09
Greece	1.81	1.08	2.37	1.81	1.81	0.77	0.52	0.86	0.77	0.77
Spain	0.91	1.07	1.12	2.04	1.02	0.46	0.49	0.49	0.84	0.56
Finland	1.32	1.04	1.48	2.19	1.65	0.51	0.47	0.59	0.87	0.65
France	1.08	1.01	1.24	2.94	1.22	0.36	0.49	0.36	0.77	0.38
Croatia	1.43	1.10	1.60	2.41	1.58	0.61	0.48	0.64	0.89	0.67
Hungary	1.17	1.04	1.55	2.11	1.36	0.59	0.50	0.67	0.89	0.69
Ireland	1.63	1.03	2.50	2.02	1.90	0.87	0.44	0.99	0.83	1.02
Italy	1.06	1.10	1.33	2.02	1.22	0.47	0.50	0.52	0.86	0.49
Lithuania	1.64	1.10	1.77	2.34	1.75	0.93	0.47	0.92	1.01	1.02
Luxembourg	1.74	1.01	3.84	2.21	3.35	0.85	0.46	2.14	0.93	1.63
Latvia	1.66	1.11	2.00	2.05	1.81	1.12	0.51	1.11	0.90	1.29
Malta	1.97	1.02	3.28	2.10	2.98	0.96	0.50	1.58	0.80	1.63
Netherlands	1.31	1.01	1.52	2.29	1.64	0.63	0.50	0.62	0.94	0.63
Poland	1.13	1.03	1.62	2.01	1.49	0.48	0.50	0.60	0.89	0.55
Portugal	1.56	1.07	1.94	2.04	1.88	0.55	0.47	0.63	0.87	0.71
Romania	1.45	1.06	1.80	2.06	1.69	0.60	0.47	0.78	1.00	0.76
Sweden	1.31	1.07	1.57	2.46	1.59	0.49	0.46	0.56	0.98	0.59
Slovenia	1.43	1.03	1.61	2.18	1.61	0.64	0.47	0.68	0.93	0.83
Slovakia	1.51	1.07	1.83	2.12	1.78	0.72	0.51	0.83	0.87	1.00
UK	1.16	1.00	1.49	2.17	1.53	0.56	0.49	0.65	0.95	0.58

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