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

Yuriy Gorodnichenko
University of California, Berkeley and NBER

Debora Revoltella
European Investment Bank

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

Christoph T. Weiss
European Investment Bank

This draft: June 14, 2020

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

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

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

Acknowledgement: We thank Russell Cooper, Peter Klenow, Albert Park and seminar participants in Berkeley, Bristol, Padova, at the 2018 ASSA meeting (Philadelphia), the CompNet 7th Annual Conference (Halle), the conference “Investment in the new monetary and financial environment” (Paris), the 2019 EEA Congress (Manchester), for helpful discussion and comments. In preparing this paper, Yuriy Gorodnichenko and Jan Svejnar worked under a grant from the European Investment Bank. Jan Svejnar was also supported by the Czech Science Foundation (grant No. 15-24642S). The views expressed in this paper are those of the authors and do not necessarily reflect those of the European Investment Bank.

I. Introduction

Although the large cross-country differentials in income per capita have been the subject of much research, accounting for sources of this dispersion has proven to be difficult. The most important factor appears to be differences in "productivity", which Moses Abramovitz famously 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 (EU) and other developed economies has generated a sense of urgency among policymakers and academics to identify impediments to productivity increases and to find ways to spur economic growth. Although a number of explanations has been put forth, rising misallocation of resources could be one of the culprits (see, e.g., Gopinath et al., 2017 for an analysis of European countries).[1]

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

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

In this paper we aim to make progress along three key dimensions. First, we continue and extend the long tradition of employing survey data for analyses of misallocation (e.g., Hsieh and Klenow, 2009) and use a new, large cross-country survey of firms: the European Investment Bank (EIB) Investment Survey (EIBIS).[2] This survey was administered annually since 2016 to a stratified random sample of firms in each of the 28 countries of the EU and is designed to be representative of the business population for each EU country for different sectors and firm sizes. EIBIS contains information about the investment behavior and constraints of firms – e.g., recent investment activities and future priorities, how firms obtain capital and whether the quantity is sufficient, whether their capital stock is state-of-the-art, and information about capacity utilization, rates of innovation, access to infrastructure, and foreign presence in management. Survey responses are also matched to administrative data of the firm (e.g., balance sheet information). Importantly, the design and implementation of the survey is consistent across countries and sectors, which is critical for understanding cross-country and cross-industry variation in the dispersion of marginal revenue products. In addition, the survey does not only cover firms in the manufacturing sector but also firms in services, utilities and construction.

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

Third, we examine the extent to which the dispersion of marginal products is related to firm-level characteristics (as opposed to the usual country-level or sector-level effects) and we compute the associated productivity gains. We note that while the existing literature treats the dispersion of marginal products as reflecting barriers and distortions, this may not always be the case. Some dispersion may reflects optimizing behavior of firms (e.g., paying compensating differentials in the labor market), in which case it is economically rational from the standpoint of the firms and

[2] Hsieh and Klenow (2009) use plant-level data from the Indian Annual Survey of Industries (ASI; 1987–1994) and firm-level data from the Chinese Annual Survey of Industrial Production (1998–2005).

should be interpreted as a "cost" to aggregate productivity. While we cannot always establish which of these phenomena is consistent with the data, we present a range of estimates consistent with various interpretations. Relatedly, we perform a Machado-Mata decomposition to construct counterfactual distributions of MRPK and MRPL for each country on the assumption that it has estimated coefficients or values of explanatory variables from another country (e.g., Greece and Germany). This decomposition exercise allows us to understand better whether observed dispersion in MRPK and MRPL is brought about by cross-country differences in firm characteristics or cross-country differences in how the business, institutional and policy environment guides the allocation of resources across heterogeneous firms.

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

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; see Restuccia and Rogerson, 2013, 2017 for literature 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), tight theoretical restrictions (e.g., David and Venkateswaran, 2017) or specific reforms (e.g., Rotemberg 2019), we draw on the richness of our survey and utilize *firm-level* information about various constraints and characteristics to account for cross-sectional variation in marginal revenue products with minimal restrictions. As a result, we can move beyond comparison of raw dispersion across countries, estimate the contribution of specific factors to misallocation, compute the associated productivity losses, and maintain high external validity.

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 et al., 2017). In particular, we document high consistency of responses in the survey data of EIBIS and the (census-type) administrative data of Orbis, thus showing that surveys can be a useful source of information so that applied work does not necessarily have to use only data with census-like coverage. Furthermore, data from administrative sources may contain manipulations (e.g., via imputation) that materially influence dispersion and hence implied productivity gains (Rotemberg and White 2020), and these manipulations vary across statistical agencies. As a result, we likely reduce systematic errors in our data by applying the same EIBIS survey instrument across countries.

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 the dispersion of worker earnings can be employed to understand the dispersion of marginal products across firms.

From the policy point of view, our estimates relate to the debate on the need to remove distortions in the EU single market. Launched in 1993, the single market allows for free mobility of labor, capital, goods, and services within the EU. Persistency of frictions and distortions have prevented the full exploitation of the associated benefits, with measurable costs of mis-allocation, or costs of "non-Europe".

The remainder of the paper is structured as follows. In Section II we present a dynamic model of a profit-maximizing firm that yields steady state conditions for MRPK and MRPL. We use these conditions in Section III to formulate our estimating equations. In Section IV we describe EIBIS and Orbis data sets and 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-Douglas production function, isoelastic demand function, and additively separable quadratic adjustment costs. Firm i's profit at time t is given by

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

where $\gamma = \alpha + \beta + \omega$ reflects returns to scale in production, K_{it} is capital, L_{it} is labor, X_{it} is an intermediate input, U_{it} is a measure of capital utilization (or quality), E_{it} is a measure of labor effort (this can also capture efficiency wages or labor quality), Δ_{it} is a measure of intermediate input quality (or utilization), $R_{it}(U_{it})$ is the price schedule for the price of capital as a function of capital utilization (or quality), $W_{it}(E_{it})$ is the price schedule for the price of labor as a function labor effort (or quality), $P_{it}^{X}(\Delta_{it})$ is the price schedule for the price of intermediate input as a function of its quality (or utilization), ϕ_K and ϕ_L capture the size of adjustment costs (these could be stochastic and firm specific), G_{it} is a combination of productivity and demand shifters, and σ is the elasticity of demand. The price schedules could be modelled as $R_{it}(U_{it}) = R_t^{base} \times U_{it}^{\psi_K}/\psi_K \times \xi_{it}^{R}$, $W_{it}(E_{it}) = W_t^{base} \times E_{it}^{\psi_L}/\psi_L \times \xi_{it}^{W}$, and $P_{it}^{X}(\Delta_{it}) = P_t^{X,base} \times \Delta_{it}^{\psi_X}/\psi_X \times \xi_{it}^{X}$, where ψ_K, ψ_L, and ψ_X are slopes of the respective supply schedules, $R_t^{base}, W_t^{base}, P_t^{X,base}$ are market prices for the base quality/utilization of capital, labor and intermediate input, and $\xi_{it}^{R}, \xi_{it}^{W}, \xi_{it}^{X}$ are random shocks (structural distortions) to the schedule.[3]

[3] We follow Hall (2004) and others in assuming that firms rent capital. Similar expressions can be derived for the case when firms own capital.

Firms are assumed to maximize the present value of their profits

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

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

Let $S_{it} \equiv G_{it}\left[(U_{it}K_{it})^{\alpha}(E_{it}L_{it})^{\beta}(\Delta_{it}X_{it})^{\omega}\right]^{1-\frac{1}{\sigma}}$ be the firm revenue (sales). The optimality conditions imply that marginal revenue products should be equal to (shadow) costs of inputs:

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

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

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

We treat shocks to G_{it} (productivity and demand), R_t^{base}, W_t^{base}, $P_t^{X,base}$ (base input prices that are common across firms), ξ_{it}^R, ξ_{it}^W, and ξ_{it}^X (idiosyncratic volatility in taxes, regulation, input quality, firm age, corporate structure, access to credit markets, etc.) as primitive sources of variation in marginal revenue across firms. Depending on the primitives, firms will choose different mixes of quantity and quality of inputs and, perhaps, different technologies (Jones 2005).

Given that the share of pure economic profits is small (e.g., Basu and Fernald, 1997), we can further simplify the expressions for marginal revenue products (see Appendix C, Section A) to obtain the corresponding expressions using observable cost shares s^K, s^L, s^X:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

These expressions lead us to conclude that the variation in marginal revenue products across firms that we attribute to distortions τ may also reflect differences in adjustment costs, as well as input quality, utilization rates, and taxes or regulation.

If one adjusted inputs for quality and/or account for adjustment costs and if the price schedules were the same across firms, then marginal revenue products for *effective* units of capital ($K_{it}U_{it}$), labor ($E_{it}L_{it}$), and intermediate inputs ($\Delta_{it}X_{it}$) should be equalized across firms and the cross-sectional dispersion ought to be zero. While appealing in theory, adjusting for quality differences is fraught with a number of issues. For example, Hsieh and Klenow (2009) adjust for quality differences in labor by dividing output by the wage bill (e.g., more educated workers receive higher wages and so the wage bill should be higher) rather than by the number of employees so that $(1-\sigma^{-1})\beta\frac{PY}{WL}=\frac{\tau^L}{\tau^Y}$ provides a measure of distortions. One can make a similar argument for using the ratio of output to the value of capital (rather than the number of machines). Apart from assuming that observed input prices are allocative[4], this approach implicitly assumes that distortions do not accrue to the owners of inputs and wages can reflect only differences in quality (more generally, market segmentation created by distortions is implicitly assumed to not influence factor prices in a given firm or industry). For example, a distortion in the labor market is not captured by workers and the distortion in revenue $(\tau_L-1)WL$ goes to a third party or is wasted. We view this assumption as potentially problematic. For example, distortions created by trade unions are likely to be captured at least partially by increased wages for union-covered workers, that is, the wage received by workers is $\tau^L W$ rather than W. Likewise, constraints on immigration almost certainly translate into higher wages for incumbents. As a result, normalizing output by the wage bill not only adjusts for quality differences but also eliminates some distortions that are absorbed into wages, i.e., this adjustment can yield $(1-\sigma^{-1})\beta\frac{PY}{\tau^L WL}=\frac{1}{\tau^Y}$ which does not include labor market distortions. Because we have direct measures of quality and distortions, we can account for the importance of these directly in the regression contexts by using these measures as controls.

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

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

where $t.i.d.$ captures terms independent of distortions.[5]

[4] This is unlikely to be true for balance sheet value of capital with long lives, see Section B of Appendix C.
[5] The Hsieh-Klenow framework does not include input-output structure of production and equation (3) may understate the cost of distortions (e.g., Jones 2011).

III. The Econometric Framework and Identification

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

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

where ψ_c is the set of country fixed effects, κ_j is the set of industry fixed effects, λ_t is the set of year fixed effects, $\boldsymbol{X}_{ijct}$ is the vector of explanatory variables (defined below), and ϵ_{ijct} is the disturbance term that captures unexplained variation in $MRPK$. By combining equation (4) with the empirical measurement of $MRPK$, we obtain an empirical "Mincerian-type" specification.[6] An analogous specification and approach is used for other inputs. Data permitting, one can also estimate a more flexible specification with country × sector × year fixed effects η_{jct}:

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

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

In estimating equation (4) and similar specifications, we generate several important "outputs". First, we can use ϵ_{ijct} to compute a "residual" measure of dispersion in MRPL and MRPK across countries to assess whether some cross-country variation can be explained by differences in observable firm characteristics and to quantify the contribution of various distortions and compensating differentials to the observed dispersion of marginal revenue products. Second, we obtain estimates of $\boldsymbol{b}$ and hence can evaluate how the explanatory variables $\boldsymbol{X}_{ijct}$ *predict* MRPK and MRPL. 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. We will cover this last point in Section V.D.

[6] Mincer (1958) developed a similar econometric specification linking wages to education to rationalize observed dispersion of earnings on the worker side. His insight was extended in subsequent work showing that wage dispersion can be accounted for by other compensating differentials (e.g., work experience) and distortions (e.g., discrimination). In a similar fashion to Mincer (1958) and subsequent work in labor economics, we aim to account for dispersion of marginal revenue products on the firm side.

A. Quantifying the contribution of observable characteristics

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

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

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

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

Because the variance of distortions is directly mapped into the dispersion of marginal revenue products, there is a simple way to quantify a productivity gain from "removing" a friction. Consider specification (4) with marginal revenue product of capital as the dependent variable. We are interested in how much variation in marginal revenue products is explained by a given regressor or a set of regressors. R^2 provides a convenient estimate for this object of interest.

We can quantify the contribution of a given friction to the variation in $MRPK$ across firms with the marginal R^2 associated with the friction, that is, the increase in R^2 when a regressor measuring the friction is added to some baseline regression. Because $V_{\tau Y} = var\ (\log(MRPL_i))$ under our assumptions, it follows that the change in $V_{\tau Y}$ after removing the friction is $var(\log(MRPL_i)) \times (marginal\ \ R^2)$.[9] Likewise, we can compute the change in $V_{\tau K}$ as $var(\log(MRPK_i) - \log(MRPL_i)) \times (marginal\ \ R^2)$ where $(\log(MRPK_i) - \log(MRPL_i))$ is the dependent variable in the regression. Thus, we measure productivity gains from removing a friction with

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

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

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

[9] The marginal R^2 is symmetric in the following sense. First, we can compute the marginal R^2 as a change in R^2 when we add regressor X_2 to the specification $Y = b_1X_1 + error$. Second, we can compute the marginal R^2 as a change in R^2 when we remove regressor X_2 in the specification $Y = b_1X_1 + b_2X_2 + error$.

B. Endogeneity of regressors

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

Second and more importantly, we are interested in how much variation in marginal revenue products is explained by a given regressor or a set of regressors, i.e., R^2 rather than $\boldsymbol{b}$. We know from basic econometrics that estimating specification (4) and similar specifications with OLS will (weakly) overstate R^2. That is, one can show (all derivations may be found in Appendix D) that for e.g., log $(MRPK)$

$$R_{OLS}^2 \equiv 1 - \frac{var(\hat{\epsilon}_{OLS})}{var(\log(MRPK))} \geq R_{IV}^2 \equiv 1 - \frac{var(\hat{\epsilon}_{IV})}{var(\log(MRPK))} \tag{5}$$

where $\hat{\epsilon}_{OLS}$ is the estimated error term in the OLS regression and $\hat{\epsilon}_{IV}$ is the *structural* error term in the instrumental variable (IV) regression. Intuitively, some of the attributed variation in OLS estimates could be due to, e.g., simultaneity or omitted variables that may confound an OLS estimated relationship between marginal revenue products and regressors.[10] Thus, (marginal) R^2 in an OLS estimate of specification (4) is a biased estimate but it provides an *upper bound* for how much variation in marginal revenue products can be due to a given friction or a given compensating differential measured in $\boldsymbol{X}$. As we discussed in the previous section, because larger R^2s are *ceteris paribus* associated with greater productivity gains and OLS yields an upper bound for R^2, we likely provide an *upper bound* for productivity gains from a better allocation of resources across firms. While having an upper bound may be only partially informative, our analysis does not rely on a structural interpretation of $\boldsymbol{b}$ and thus opens a number of opportunities. For example, one does not have to restrict the analysis only to variables with well-identified, exogenous variation. Likewise, one does not have to impose tight theoretical restrictions to achieve identification.

There are also a few practical concerns with using the insight of equation (5). First, R_{IV}^2 is not guaranteed to be between 0 and 1 and indeed R_{IV}^2 is often negative in empirical work, which

[10] To address this issue of potentially confounding factors, we include many control variables in specification (4).

makes model comparisons a challenge. Using $\frac{var(\widehat{\boldsymbol{b}}_{IV}\boldsymbol{X})}{var(\log(MRPK))}$ could be equally problematic for measuring the contribution of exogenous variation in $\boldsymbol{X}$ to variation in $\log(MRPK)$ because $\boldsymbol{X}$ is potentially correlated with the structural shock ϵ and some of the variation in $\boldsymbol{X}$ is due to structural error ϵ (as a result, $\frac{var(\widehat{\boldsymbol{b}}_{IV}\boldsymbol{X})}{var(\log(MRPK))}$ may be greater than 1). To address this issue, Pesaran and Smith (1994) propose a generalized R^2 which uses the *predictive* error in an IV regression: $GR_{IV}^2 \equiv 1 - \frac{var(\tilde{\epsilon}_{IV})}{var(\log(MRPK))} \in [0,1]$, where $\tilde{\epsilon}_{IV} \equiv \log(MRPK) - \widehat{\boldsymbol{b}}_{IV}\widehat{\boldsymbol{X}}_{IV}$ and $\widehat{\boldsymbol{X}}_{IV}$ is the predicted value of $\boldsymbol{X}$ from the first-stage regression (for OLS, $GR_{OLS}^2 = R_{OLS}^2$). We show in Appendix D that R_{OLS}^2 is guaranteed to be greater than GR_{IV}^2 when $\boldsymbol{X}$ is positively (negatively) correlated with ϵ and $\boldsymbol{X}$ is positively (negatively) correlated with $\log(MRPK)$. This condition is, for instance, satisfied when an omitted variable (e.g., the entrepreneurial talent of a manager) raises $\log(MRPK)$ and makes a distortion more binding because (e.g., a growing firm is more likely to run into red tape). In turn, this entails a higher $\log(MRPK)$ because the firm has too little capital due to red tape. This setting appears plausible to us and we take R_{OLS}^2 to provide an upper bound.

Second, measurement error in $\boldsymbol{X}$ will attenuate R_{OLS}^2 toward zero. In particular, we show in Appendix D that $R_{OLS}^2 = R_{true}^2 \frac{var(x)}{var(x)+var(\eta)}$ where $var(x)$ is the variance of a correctly measured regressor x and $var(\eta)$ is the variance of classical measurement error η. If one has N measurements of x with uncorrelated measurement errors, one can reduce the bias by running the regression on averaged values of x: $R_{OLS}^2 = R_{true}^2 \frac{var(x)}{var(x)+\frac{1}{N}var(\eta)}$. We will use this insight to assess how/if measurement errors can materially alter our conclusions. We also show that, with multiple measurements of x, one can use one measurement as an instrument for another measurement (e.g., Griliches 1986). In this case, we will generate $R_{IV}^2 \leq R_{OLS}^2$ and $GR_{IV}^2 = R_{OLS}^2$. Thus, even in this case R_{OLS}^2 is an upper bound.[11]

IV. Data

The main data source for our analysis is the EIB Investment Survey (EIBIS). In this section we provide information on the design and implementation of the survey. We also compare EIBIS

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

responses to the administrative data of the surveyed firms, as collected in the Orbis database. Once we establish consistency across the survey and administrative data, we describe survey questions that we use in the empirical analysis to account for the variation in MRPK and MRPL across firms.

A. The EIB Investment Survey (EIBIS)

EIBIS is an annual firm-level survey conducted by the market research company Ipsos MORI on behalf of the European Investment Bank (see Ipsos, 2019, for a detailed review of the survey). The first wave of EIBIS was administrated in 2016, targeting firms in the 28 EU member states, with the objective of being representative in each country for different size classes and sectors. The sampling targeted head offices.[12] Eligible respondents were senior persons with responsibility for investment decisions and how investments were financed. This person could be the owner, the finance manager, finance director or head of accounts, Chief Financial Officer (CFO), or Chief Executive Officer (CEO).

The sample was stratified disproportionally by country, industry group (sector) and size-class, and stratified proportionally by region within the country. The minimum number of employees of all enterprises is 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. Brutscher et al. (2020) provide evidence on representativeness of the data for the business population of interest (namely enterprises above 5 employees) by comparing distributions in EIBIS with the population of firm-level data available in Eurostat's Structural Business Statistics (SBS).

The fieldwork for the first wave 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, which is typical for surveys of executives (Cycyota and Harrison, 2006). The resulting sample of the first wave consists of 12,483 non-financial enterprises in the 28 EU member states in NACE categories C to J (industrial firms). The sample size varies across countries and ranges from 150 enterprises in Cyprus and Luxembourg to 600 in France, Germany, Italy and the UK. A total

[12] 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.

sample of 12,300 firms was targeted, with 150, 400, 475, or 600 interviews per country depending on the size of the population. Because the sampling frame as well as the resulting samples may not be fully representative, Ipsos MORI constructed weights to correct for possible imbalances. In particular, firms are weighted to make them representative of the EU economy based on country, sector and firm size (employment) where the population distribution is reported by SBS.

The second (2017) and third (2018) waves have similar properties and were conducted between April and August 2017 and 2018. In the end of each wave, firms are invited to participate in the next wave of the survey so that EIBIS has a panel component. Approximately 2,000 firms participated in all three waves and approximately 4,500 firms participated in two waves. We have on average 24 firms per country/year/industry (2-digit NACE classification) cell.

EIBIS is a rich source of information with a number of unique characteristics. First, EIBIS collects basic information on firms (e.g., number of employees, value of fixed assets, sales) which is matched to administrative data.[13] This feature of the survey allows us to cross-check survey responses against data from administrative sources and hence to assess the quality of survey data. Second, EIBIS gathers data on expectations and perceptions of firms' management (e.g., perceived obstacles to investment, plans for future investment) as well as statistics that are often not available in standard official sources (e.g., quality of capital, capacity utilization, sources of financing). These variables can inform us directly about sources of variation in marginal revenue products across firms and we can thus make progress relative to studies based only on income statements and balance sheets. Third, EIBIS data are collected in a *consistent* manner for a large number of firms across *many* countries and industries, thus permitting us to carry out a comparative analysis of resource allocation in various institutional settings. Using these unique data, we explore the relationship between MRPK and MRPL and a large number of explanatory variables at the *firm* level. To this end, we use questions on firm demographics, capacity utilization, quality of the capital stock, obstacles to long-term investment, investment plans, investment rate, employment growth, and sources of finance.

[13] 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 EIBIS does not have the name, the address, the contact details or any additional individual information that could identify the firms in the final sample. Note that not every firm in EIBIS has complete information in Orbis (e.g., Orbis may have missing information on employment while EIBIS does not).

B. Comparison of survey and Administrative Data

The Orbis database is a popular source of administrative data for cross-country analyses at the firm level.[14] We use these data to cross-check EIBIS responses. In particular, we match firms' EIBIS responses with their administrative (Orbis) data and compare cross-firm dispersion of the logarithm of sales, fixed assets and employment in the EIBIS and Orbis for those firms by country (Table 1) and by industry (Table 2). In columns (4), (7) and (10) of Tables 1 and 2, we also report correlations between the responses in EIBIS and the administrative data in Orbis.

We observe high consistency across the two sources of data. For example, the correlation between log employment in EIBIS and administrative data in Orbis is 0.91. The dispersion of the survey responses across firms is on average slightly larger than the dispersion in the administrative data, which is consistent with small noise (measurement error) in survey responses. Note that relative to the data on fixed assets, data on employment are available for fewer firms in the Orbis database.

Because the EIBIS uses a random sample of firms, we also evaluate whether dispersion of the sales-to-capital and sales-to-labor ratios estimated on the sample is consistent with the dispersion calculated for the population of firms. For this exercise, we use data for countries that have excellent (i.e., close to population) coverage in Orbis for each industry and firm size category (Brutscher et al. 2020): Italy, Portugal, and Romania. We find that mean absolute log difference between EIBIS and Orbis measures of standard deviations computed at the country level is approximately 2 percent, that is, the two sources of data produce very similar measures of dispersion. We conclude that EIBIS provides a satisfactory quality of firm-level data and that the survey responses are therefore suitable for our analysis.

C. Dispersion of marginal revenue products

We report descriptive statistics for the EIBIS sample in Table 3. The key statistic for our analysis is the dispersion of marginal revenue products. We observe a sizable dispersion across firms in the EU: the standard deviation is 1.43 for $\log(MRPK)$, 1.19 for $\log(MRPL)$, and 1.63 for $(\log(MRPL) - \log(MRPK))$.[15] For comparison, the dispersion of marginal value-added product across

---

[14] See Kalemli-Ozcan et al. (2015) and Bajgar et al. (2020) for a detailed analysis of the (dis)advantages of using this dataset.

[15] We find similar magnitudes when we use robust methods to estimate standard deviation. For example, when we employ median absolute deviation (MAD) to estimate $st.dev. = 1.48 \times MAD$ where $MAD = median|x_i - \tilde{x}|$, $\tilde{x} = median(x_i)$ and x_i is a random variable, we find that the standard deviation is1.49 for $\log MRPK$, 1.14 for $\log(MRPL)$ and 1.68 for $(\log(MRPL) - \log(MRPK))$.

establishments ("plants") for the U.S. is 0.98 for capital (Table 2 in Asker et al., 2014) and 0.58 for labor (Table 1 in Bartelsman et al., 2013).

Note that there are three potentially confounding sources of differences between our statistics and statistics reported for the U.S. First, EIBIS does not collect information on the cost of intermediate inputs and we therefore use sales to compute marginal revenue products for capital and labor, while previous studies use value added. Using EIBIS firms matched to the Orbis database (which has information on sales and value added), we find that the standard deviation of $\log(MRPK)$ based on sales is approximately 0.16 log points higher than the standard deviation of $\log(MRPK)$ based on value added. On the other hand, the standard deviation of $\log(MRPL)$ based on sales is approximately 0.21 log points lower than the standard deviation of $\log(MRPL)$ based on value added. Thus, using revenue rather than value added does not appear to explain the difference between the EU and US.

Second, our analysis is based on survey responses while US studies typically rely on administrative data. We might observe a US-EU difference in dispersion because survey data are more likely to have measurement error. As we discuss above, however, survey responses in EIBIS are broadly consistent with administrative data in Orbis.[16] To explore further the quantitative significance of measurement error, we exploit the panel component of EIBIS and compute the average $\log(MRPK)$ and $\log(MRPL)$ across years for a given firm, as well as cross-sectional dispersion of these averages. The disadvantage of taking averages across years (using repeated measurements) is that one attenuates not only measurement errors but also transitory factors (e.g., adjustment costs or high-frequency variation in demand) and the reduction in dispersion is likely to overstate the role of measurement errors. With this caveat in mind, we find that, in a consistent sample of firms, the standard deviation of average $\log MRPK$ is 8 percent lower for firms participating in two waves of the survey and 12 percent lower for firms participating in three waves. The corresponding figures for $\log MRPL$ are 4 percent for two-wave firms and 1 percent for three-wave firms. These results suggest that measurement error can account for only a portion of the EU-US difference in the dispersion of marginal revenue products. Note that while

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

measurement error can influence the *level* of dispersion, our objective is to study what *part* of dispersion in marginal revenue products can be rationalized by various measures of distortions and compensating differentials. As we show below, using three-year averages does not alter our calculations of variance contributions due to these forces.

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

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

D. EIBIS Variables Used in the Regression Analysis

We consider several blocks of variables available in the survey to construct vector $\boldsymbol{X}$ in equation (4). We next discuss possible relationships between the variables and marginal revenue products. The choice of variables is motivated by previous work and is partially constrained by data availability.[17] Some of the variables in the survey are qualitative, which creates some distance between an ideal theoretical metric and its survey measure. However, even these imperfect measures — which are often used for policy (e.g., World Bank uses similar measures to construct its *Ease of Doing Business* index) and research (e.g., Djankov et al. 2002) — allow us to go well beyond what one can achieve

[17] Lack of data prevents us from exploring the importance of firm-level variation in market power (mark-ups). To the extent other variables are correlated with markups, we can overstate the quantitative significance of these other variables in accounting for observed variation in marginal revenue products.

using only administrative data. Descriptive statistics are presented in Table 3 (and by survey wave in Appendix Table A.1). While we include many regressors to fully exploit the richness of the EIBIS, one should not be concerned with overfitting in specification (4) because the sample size is very large relative to the number of regressors.

Firm demographics

Employment size: Garicano et al. (2016), Bento and Restuccia (2017), Eslava et al. (2019) and other studies document that various size-based policies can introduce distortions to the scale of operation and, hence, allocation of resources (e.g., firms in low-productivity countries are systematically smaller than firms in high-productivity countries).

Firm age: Hsieh and Klenow (2014) and others argue that productivity of firms may have an important life-cycle component. For example, firms may accumulate more organizational capital as they age. On the other hand, one may expect that firms that have been longer in existence have older and probably lower quality capital so that older firms could have lower measured MRPK.

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

Exporter status: Being more exposed to competition, exporters are relatively more likely to employ high-quality, and hence more expensive inputs (see, e.g., Verhoogen, 2008).

Utilization and quality of inputs

Quality of capital: OECD (2001) and other statistical agencies emphasize that proper measurement of productivity requires adjustment for the quality of inputs (e.g., new vintages of capital goods). In our context, 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.

Capacity utilization: Fernald (2001) and others document that variable utilization of resources can materially influence measured productivity. For example, firms operating at (or even above) maximum capacity are 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. Utilization can also help us control for firm-level variation in the composition of transitory vs. persistent shocks. Intuitively, Abel and Eberly (1998) show that, when facing adjustment costs, firms should use utilization in response to transitory shocks and adjust inputs (buy more machines or hire more workers) in response to persistent shocks.

<u>Obstacles to investment</u>

The variables included in this cluster are answers of firms' top management to questions about constraints on investment. The World Bank and other institutions use similar questions to construct measures for barriers to entry, which have been related to reduced productivity (see e.g., Gorodnichenko et al. 2010; Commander and Svejnar 2011; World Bank 2018).[18] When asked about a specific potential constraint, the respondent in EIBIS reports whether he/she considers it to be a "major obstacle", "minor obstacle", or "not an obstacle at all". The list of constraints includes: *Demand for products or services*, *Availability of staff with the right skills*, *Energy costs*, *Access to digital infrastructure*, *Availability of adequate transport infrastructure*, *Labor market regulation*, *Business regulations and taxation*, *Availability of finance*, *Uncertainty about future*. For each obstacle, approximately 20-40 percent of firms report it to be a major obstacle and another 30 percent regard it as a minor obstacle.

<u>Dynamic adjustment</u>

Firms are exposed to a variety of shocks and with adjustment costs it may take time and resources for firms to reoptimize factor allocation. Although EIBIS data does not have a large panel component yet, the survey asks questions about firms' current and previous investment choices – an aspect that enables us to examine the dynamics of inflows and outflows of capital and labor.[19] These variables can help

[18] While these indicators may be imperfect measures of various barriers, they are relevant variables from the viewpoint of policymakers. Indeed, many governments focus on addressing poor scores on these variables so as to move up in various rankings.

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

understand misallocation because, as argued by Asker et al. (2014), the variability of input growth can inform us about the importance of adjustment costs. Indeed, the optimality conditions in equation (1) highlight that, with adjustment costs, *current* marginal revenue products depend on *past* and *future* choices of capital, labor and other inputs. The variables included in this cluster are:

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

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.

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.

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 activities. 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 investment priority for the next three years, which contrasts with studies such as Asker et al (2014) that have data only on realized investment. Specifically, firms can report whether investment will focus on "replacing capacity (existing buildings, machine, equipment and IT)", "expanding capacity for existing products and services", "developing or introducing new products, processes or services", or "do not have investment planned." There is no *a priori* expectation as to which types of investment (replacement versus capacity expansion) 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 investment priority is "replacing capacity" (35 percent).

Source of funding

Share of investment funded by internal and external finance: Firms may have different cost of capital depending for instance on how old they are or how connected they are to capital markets. 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: Midrigan and Xu (2014) and others document that credit frictions can generate misallocation of capital. To measure the importance of this channel, we use an indicator variable that 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.

Data filters and additional data

To minimize potentially adverse effects of extreme observations, we winsorize continuous variables at the top and bottom one percent. For firms with missing information for a given variable, we impute the average value of that variable in the industry-country cell.[20] For each variable, we create a corresponding indicator variable taking value one if the values were imputed. We include these indicator variables as additional regressors but do not report their estimated coefficients in the regression tables. We estimate the cost shares for labor and materials using data from the Industrial Analysis section of the OECD's Structural Analysis Database (STAN) or from Eurostat national accounts that are available at the level of the country, year and industry (2-digit NACE classification).[21]

[20] We impute 3.3 percent of observations. The results are qualitatively similar if we constrain the sample to observations with non-missing values.

[21] We use cost shares to make industries/countries more comparable so that we do not need to rely on fixed effects to study dispersion of marginal revenue products. We find similar results when we use industry-specific cost shares instead of firm-specific cost shares.

V. Empirical Analysis

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

A. Regression results

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

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

We report estimated coefficients of equation (4) in Appendix Table A.4.[22] Because the coefficients do not have a structural interpretation, we provide only a brief overview of selected correlations. Older, exporting, high-utilization firms are predicted to have higher MRPK and MRPL. Obstacles for investment generally have mixed associations with marginal revenue products. In the "adjustment" block, investment has a strong negative association with MRPK and a positive association with MRPL.[23] 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.[24] The "credit constrained" status is negatively correlated with MRPL and MRPK. Note that while variables in $\boldsymbol{X}$ may be correlated (the correlation matrix is reported in Appendix Table A.3), thus making the interpretation of individual coefficients challenging, we focus on R^2 which does not depend on how correlated regressors are. Furthermore, we also report results when we include blocks of variables one at a time, thus showing also the contribution of a given set of variables.

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, as may be seen in Table A.4, variables measuring firm demographics, dynamic adjustment of inputs, and source of funds appear to have robust predictive power. On the other hand, the

[22] The results are qualitatively similar when we restrict the sample to firms that participated in all three waves of the survey (Appendix Table A.5). The results are also similar when we estimate equation (4) for each wave separately.

[23] The results are similar when we include an indicator variable equal to one if a firm reports positive investment, and zero otherwise (the baseline specification uses log(1+investment)). See Appendix Table A.7 (analogue of Table A.4).

[24] One may be concerned that we use employment and capital (investment) measured with error to compute the dependent variables and use the same employment and capital as a regressor because measurement error can mechanically create a negative correlation between a regressor and a regressand. To address this concern, we use employment and investment from Orbis as regressors and we find nearly identical results (compare Appendix Tables A.4 and A.17).

contribution of the "constraint" variables — a direct, but partial measure of distortions — to the variation in MRPK and MRPL is modest, with some coefficients not being statistically significant. In the memorandum section of Table A.4 we also see that the R^2 for the specification without (with) fixed effects is 0.14 (0.49) for MRPK and 0.29 (0.74) for MRPL, which may appear somewhat small but in fact is in line with R^2s estimated for Mincerian wage regressions and for regressions with detailed firm/worker fixed effects (see, e.g., Card et al., 2013).

In order to quantify these observations, in Table 4 (panels A and B) we present marginal R^2s for blocks of variables, that is, by how much R^2 increases after a given block of variables is added to various fixed effects specifications. We observe for both MRPK and MRPL that the marginal R^2s are the largest for variables in the "adjustment" and "demographics" blocks, and that they are relatively low for variables in the "obstacles to investment" block.[25]

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. Our reasoning relates to Mincer (1958) emphasizing that more educated workers have higher productivity and hence demand a compensating premium for accumulating higher human capital. In a similar spirit, higher-quality capital may demand a compensating differential for the cost of inputs necessary to ensure higher quality. The second group of variables 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 4 that in terms of marginal R^2 the predictive power is similar for the two groups of variables. Conditional on accepting this classification of variables, one can reach two important conclusions. First, the "raw" dispersion in marginal products is likely to overstate the extent of misallocation since some variation is likely to be brought about by heterogeneity in the "quality" of inputs. Second, "distortions" are likely to be substantial and removing them may lead to significant gains in productivity.

In Appendix Tables A.6, A.7, A.8 we find similar results when we estimate equation (4') using a "between" regression — a regression that is estimated on average (across years) values of

[25] As we saturate the regression with fixed effects, the variance contribution in Y due to $\boldsymbol{X}$ shrinks. The remaining explanatory power of $\boldsymbol{X}$ in this case is comparable to 4-5 percent reported for time-varying worker characteristics in other studies (e.g., Card et al., 2013).

the regressors and regressands. This specification likely reduces the importance of transitory factors such as measurement errors and adjustment costs. These results suggest that measurement error is unlikely to overturn our conclusions.

B. Productivity Gains

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

We carry out calculations for several policy scenarios. First, assume that the policymakers would eliminate the dispersion in marginal revenue products brought about by the "distortions" group of variables. To have a level of variation in marginal revenue products, we use the dispersion from Table 3: $var(\log(MRPL)) = 1.19^2 \approx 1.42$ and $var(\log(MRPL) - \log(MRPK)) = 1.63^2 \approx 2.66$. Table 4 reports the marginal R^2 for various specifications without and with fixed effects. Clearly, the marginal R^2 decreases as we add a richer set of fixed effects. To obtain an upper bound on productivity gains, we consider the marginal R^2 without fixed effects (column (1)). The marginal R^2s of the "distortions group" are 0.186 for $\log(MRPL)$ (row 17, Table 4) and 0.134 for $(\log(MRPL) - \log(MRPK))$ (row 26, Table 4). It follows that the gain in productivity is $1.5 \times 1.42 \times 0.186 + 0.13 \times 2.66 \times 0.134 = 0.442$, which is reported in Panel D (row 35, column 1) of Table 4. In other words, removing distortions can raise aggregate productivity by more than 40 percent.

Second, consider the possibility that *all* variables in $\boldsymbol{X}$ capture distortions. In this case, the marginal R^2s are 0.289 for $\log(MRPL)$ (row 18, Table 4) and 0.289 for $(\log(MRPL) - \log(MRPK))$ (row 27, Table 4). With these marginal R^2s, the gain is $1.5 \times 1.42 \times 0.289 + 0.13 \times 2.66 \times 0.289 = 0.715$, which is reported in Panel D (row 36, column 1) of Table 4. Interestingly, variables in the "adjustment" block contribute to productivity gains (row 32) as much as variables in the "demographics" block (row 28). In short, treating all variables in $\boldsymbol{X}$ as distortions increases the magnitude of potential gains by over one-half relative to the first scenario. Consistent with our prediction that measurement errors can attenuate estimated productivity gains, we find higher gains when we use "between" regressions (Appendix Table A.8), but the difference from the baseline estimates in Table 4 is small.

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

C. Cross-country and cross-industry differences

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

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

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

The quantitative importance of country, industry and year fixed effects or the interaction terms country × industry × year raises an important identification challenge. In particular, fixed effects can absorb not only cross-country/industry compensating differentials related to the quality of inputs, but also barriers to 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.

In Table 5 we report R^2 for regression (4) with various sets of fixed effects and no controls $\boldsymbol{X}$.[26] Country fixed effects alone account for R^2 = 0.052 for MRPK and R^2 = 0.445 for MRPL. Industry (2-digit NACE level) fixed effects alone account for R^2 = 0.239 for MRPK and R^2 = 0.268 for MRPL. Time fixed effects have little explanatory power. The combined contribution of country, industry and year fixed effects in column 5 is R^2 = 0.275 for MRPK and R^2 = 0.611 for MRPL. To the extent that fixed effects embody distortions or compensating differentials common to countries or industries, these patterns suggest (for MRPL) either that moving a worker from one country to another is "costlier" than moving the worker from one industry to another – that is, countries are more segmented than industries and therefore differences in levels of MRPL are higher across countries than across industries and these differences are reflected in fixed effects – or that quality differences across workers are larger between countries than between industries. Indeed, the R^2 in the regression with country, industry and year fixed effects is similar to the R^2 with country fixed effects only, which suggests that industry is not likely to be the main driver of MRPL dispersion

[26] Appendix Table A.9 reports results based on "between" regressions that are qualitatively similar to our main specification.

across countries. This is also consistent with empirical evidence that labor supply to an industry is more elastic than to a country.

On the other hand, for capital the increment in the R^2 with country, industry and year fixed effects relative to the regression with no fixed effect is approximately equal to the sum of R^2 increments in the regression with country fixed effects and in regression with industry fixed effects relative to the regression with no fixed effects. Since the increment is somewhat larger for the regression with the industry fixed effect than country fixed effect, the interpretation is that moving a unit of capital from one country to another is "cheaper/easier" than moving it from one industry to another, or that quality differences in capital are smaller between countries than between industries.

Finally, there is a large increase in the R^2 when we introduce country × industry × year fixed effects: R^2 is 0.492 for MRPK and 0.736 for MRPL (column 6 of Table 5). Again, these results are consistent with two explanations. The first one is that 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. The alternative explanation is that there is an additional quality difference when workers or capital are compared across industries *and* countries. Irrespective of which view is taken, it is clear that there are quantitatively important complementarities in industry and country attributes.

If one interprets country and/or industry fixed effects as capturing barriers and distortions, then the EU is rather fragmented economically. This interpretation suggests that the EU can achieve considerable gains in productivity. For example, removing inequality in average marginal revenue products across countries (i.e., making the country fixed effects be all identical) could raise productivity by 102 percentage log points (row 4, column 1 of Table 5) according to the Hsieh-Klenow framework (equation (3')). Removing barriers between industries *and* countries could raise productivity by at least 143 percentage log points (row 4, column 5 of Table 5).

What set of estimates should one use to compute the cost of inadequately integrated Europe? In an ideal case, marginal revenue products should be equalized across countries *and* industries. However, achieving this objective may be infeasible. For example, industries may have a variety of idiosyncratic factors that could prevent equalization. As a result, one may propose a weaker goal: equalize marginal returns across countries. In our context, this means that one should focus on the case where we control only for industry fixed effects, i.e., within-industry marginal revenue products should be similar in "Germany" and "Greece". With this logic, the productivity cost of inadequately integrated Europe is approximately 31 percent if we use only "distortions" variables (row 35, column 3 in Table 4) and approximately 53 percent if we interpret all variables as measuring frictions (row 36, column 3, in Table 4).

D. Machado-Mata Decomposition

While our analysis so far is helpful for understanding what factors can predict MRPK and MRPL, it is also useful to understand whether the cross-country differences in dispersion are brought about by differences in firm characteristics or by the way how these characteristics are translated into differences in marginal revenue products. To address this question, we carry out a Machado and Mata (2005) decomposition of the variance in MRPK and MRPL.[27] We start by using Germany and Greece as two polar cases: $\sigma(\log(MRPK))$ is 0.92 for Germany and 1.64 for Greece, while $\sigma(\log(MRPL))$ is 0.61 for Germany and 0.91 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 *predictive* (not causal) 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 4, we depict the distribution of Greek MRPK in Panel A and Greek MRPL in Panel B. In each panel, we show the actual distribution using Greek $\boldsymbol{X}$ 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 in a less dispersed distribution of both MRPK and MRPL, suggesting that German business,

[27] 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 NACE industry fixed effects rather than 2-digit industry fixed effects as in Table 3. Note that in these decompositions we only use the variation due to observable characteristics, that is, we do not use unexplained ("residual") variation. Thus, when we report dispersion for country c, we report $\{X_{bc}\gamma_{cb}\}_{b=1}^{B}$. Because we have many countries in the sample, we report estimates of specification (4) for three blocs of EU countries (North/West, South, Center/East) in Appendix Tables A.10, A.11, A.12. Marginal R^2 and welfare calculations for the blocs are reported in Appendix Tables A.13-A.16.

institutional and policy environment would increase the efficiency of Greek firms by reducing the dispersion of marginal products of capital and labor across firms. In other words, German "prices" help increase the equalization of returns across firms. Indeed, the standard deviation of this counterfactual distribution is much closer to the actual distribution of marginal revenue products in Germany (e.g., for MRPK the counterfactual standard deviation for Greece is 0.94 rather than 1.66).

When we use German ***X*** and Greek ***b***, the distribution of MRPK is more dispersed and shifts to the right. The latter is consistent with German firms having characteristics associated with high levels of productivity. The former suggests that the dispersion of firm characteristics in Germany is greater than the corresponding dispersion in Greece which, when combined with the Greek business, institutional and policy environment ("prices"), results in a wider dispersion of marginal products than is actually observed in Greece. Interestingly, using German ***X*** and Greek ***b*** does not generate large differences in the mean or dispersion of MRPL. This pattern suggests that differences in firm characteristics are not likely to be a key determinant of the differences in the dispersion of MRPL between Germany and Greece. 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 or 3 for MRPK and columns 7 or 8 for MRPL) and (b) Greek ***X*** or ***b*** (columns 4 or 5 for MRPK and columns 9 or 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.[28]

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. In addition to presenting new cross-country evidence on allocation of resources, we propose a novel approach to quantify potential productivity gains from better allocations. This approach does not rely on exogenous variation in measured frictions (or compensating differentials) so that researchers can apply it in a wide range of settings.

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

Much of the overall dispersion in marginal products could be attributed to fixed differences among countries or sectors/industries. For example, if one would remove the dispersion in marginal products across countries (i.e., make the country fixed effects in the regression analysis all identical), EU productivity could rise by 102 percentage log points. We also find that the significant association between marginal products and firm characteristics is predominantly driven by variables measuring

[28] Consistent with this interpretation and evidence in Figure 1, we find that gains from reduced dispersion of marginal revenue products (measured as $\sigma(\boldsymbol{b}^{Germany}\boldsymbol{X}^{own})/\sigma(\boldsymbol{b}^{own}\boldsymbol{X}^{own})$) are correlated with popular measures of institutional quality such as World Bank's Governance Indicators and the International Country Risk Guide: countries with lower quality of institutions having greater gains from "importing" German institutions.

firm demographics, quality of inputs, utilization of resources, and dynamic adjustment of inputs. In contrast, the contribution of direct measures of "barriers and constraints" to cross-sectional variation in MRPK and MRPL seems to be modest. Finally, we show that cross-country variation in the within-country dispersion of marginal revenue products is largely rationalized by differences in how a country's business, institutional and policy environment translates firm characteristics into outcomes than by differences in firm characteristics *per se*.

Our work contributes to the growing literature measuring misallocation of resources, provides new insights into the nascent literature on the *sources* of the 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 the 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

Abel, Andrew, and Janice Eberly (1998). "The mix and scale of factors with irreversibility and fixed costs of investment." *Carnegie-Rochester Conference Series on Public Policy* 48(1): 101-135.

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

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo and Jonathan Timmis (2020). "Coverage and representativeness of Orbis data." OECD Science, Technology and Industry Working Paper No. 2020/06.

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

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

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

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

Brutscher, Philipp-Bastian, Andrea Coali, Julie Delanote, and Peter Harasztosi (2020). "EIB Group Survey on Investment and Investment Finance – A Technical Note on Data Quality." EIB Working Paper 2020/09.

Card, David, Jörg Heining, and Patrick Kline (2013). "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics* 128(3): 967-1015.

Commander, Simon, and Jan Svejnar (2011). "Business Environment, Exports, Ownership, and Firm Performance." *Review of Economics and Statistics* 93(1): 309-337.

Cycyota, Cynthia, and David A. Harrison (2006). "What (Not) to Expect When Surveying Executives: A Meta-Analysis of Top Manager Response Rates and Techniques Over Time." *Organizational Research Methods* 9(2): 133–160.

David, Joel M., and Venky Venkateswaran (2017). "The Sources of Capital Misallocation." NBER Working Paper No. 23129.

Desai, Mihir A., C. Fritz Foley, and James R. Hines (2004). "A Multinational Perspective on Capital Structure Choice and Internal Capital Markets." *Journal of Finance* 59(6): 2451-2487.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer (2002). "The Regulation of Entry." *Quarterly Journal of Economics* 117(1): 1-37.

Eslava, Marcela, John C. Haltiwanger, and Alvaro Pinzón (2019). "Job Creation in Colombia vs the U.S.: "Up or Out Dynamics" Meets "The Life Cycle of Plants"." NBER Working Paper No. 25550.

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.

Fernald, John G. (2012) "A Quarterly, Utilization-adjusted Series on Total Factor Productivity." Working Paper Series 2012-19, Federal Reserve Bank of San Francisco.

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

Garicano, Luis, Claire Lelarge, and John Van Reenen (2016). "Firm Size Distortions and the Productivity Distribution: Evidence from France." *American Economic Review* 106(11): 3439-79.

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.

Gorodnichenko, Yuriy, Jan Svejnar, and Katherine Terrell (2010). “Globalization and Innovation in Emerging Markets.” *American Economic Journal: Macroeconomics* 2(2): 194-226.

Griliches, Zvi (1986). “Data Issues in Econometrics.” In Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics*, Vol. III, North-Holland, Amsterdam, The Netherlands..

Hall, Robert E. (2004). “Measuring Factor Adjustment Costs.” *Quarterly Journal of Economics* 119(3): 899-927.

Heckman James J., Lance J. Lochner, and Petra E. Todd (2006). “Fifty Years of Mincer Earnings Regressions.” In E. Hanushek, and F. Welch (eds.), *Handbook of the Economics of Education*, Vol. 1, Elsevier, Amsterdam, 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 (2019). *EIB Group Survey of Investment and Investment Finance*. Technical Report. Available at https://www.eib.org/attachments/eibis-methodology-report-2019-en.pdf

Jones, Charles I. (2005). “The Shape of Production Functions and the Direction of Technical Change.” *Quarterly Journal of Economics* 120(2): 517-549.

Jones, Charles I. (2011). “Intermediate Goods and Weak Links in the Theory of Economic Development.” *American Economic Journal: Macroeconomics* 3(2): 1-28.

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 No. 21558.

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

Lipsey, Robert E. (2003) “Foreign Direct Investment and the Operations of Multinational Firms: Concepts, History, and Data.” 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, Nikolaj, James R. Markusen, and Bertel Schjerning (2013). "Foreign Firms, Domestic Wages." *Scandinavian Journal of Economics*, 115(2): 292–325.

Midrigan, Virgiliu, and Daniel Yi Xu (2014). "Finance and Misallocation: Evidence from Plant-Level Data." *American Economic Review* 104(2): 422-458.

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

OECD (2001). *Measuring Productivity: OECD Manual*. OECD Publications, Paris, France. Available at http://www.oecd.org/sdd/productivity-stats/2352458.pdf

Pesaran, M. Hashem and Richard J. Smith (1994). "A Generalized R^2 Criterion for Regression Models Estimated by the Instrumental Variables Method." *Econometrica*, 62(3), 705-710.

Restuccia, Diego, and Richard. Rogerson (2008). "Policy Distortions and Aggregate Productivity with Heterogeneous 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.

Rotemberg, Martin (2019). "Equilibrium Effects of Firm Subsidies." *American Economic Review* 109 (10): 3475-3513.

Rotemberg, Martin, and T. Kirk White (2020) "Measuring Cross-Country Differences in Misallocation." Unpublished manuscript.

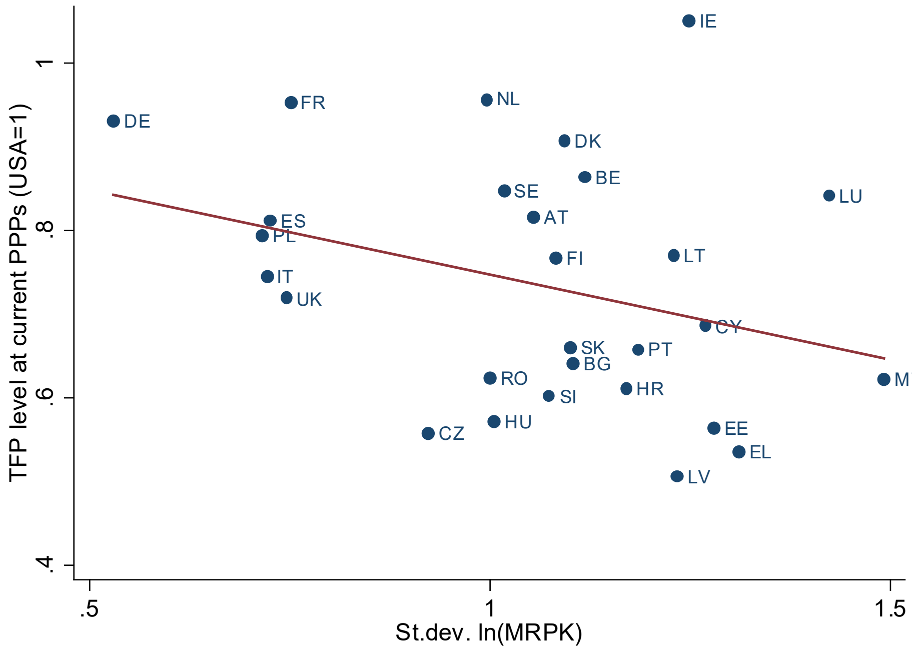
van Ark, Bart, Mary O'Mahoney, and Marcel P. Timmer (2008). "The Productivity Gap between Europe and the United States: Trends and Causes." *Journal of Economic Perspectives* 22(1): 25-44.

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

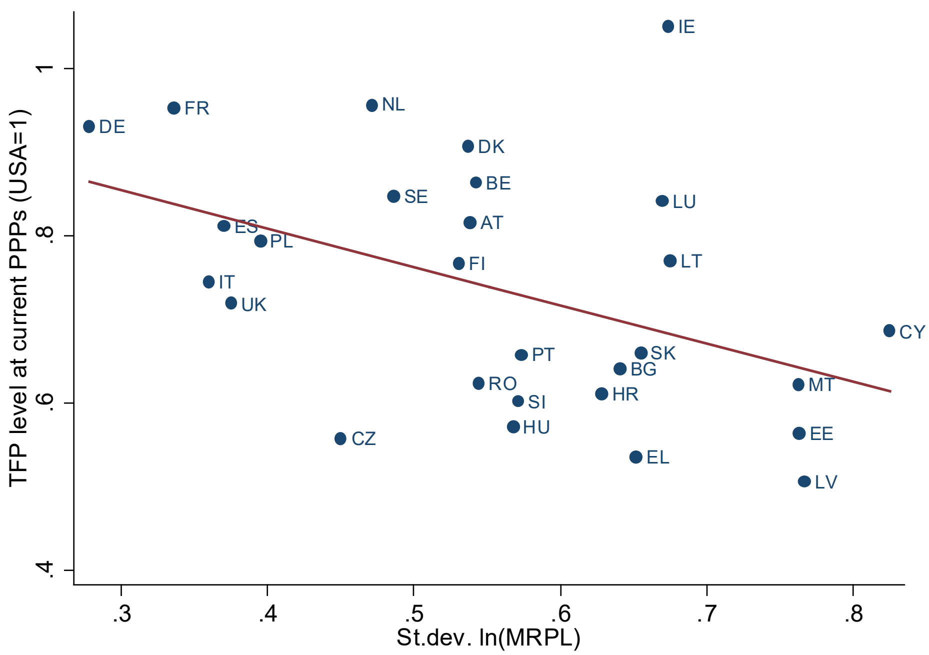
World Bank (2018). *Doing Business 2019: Training for Business*. 16th edition, The World Bank, Washington, DC.

Figure 1. Productivity and dispersion of marginal revenue products

Panel A. Marginal revenue product of capital (MRPK)



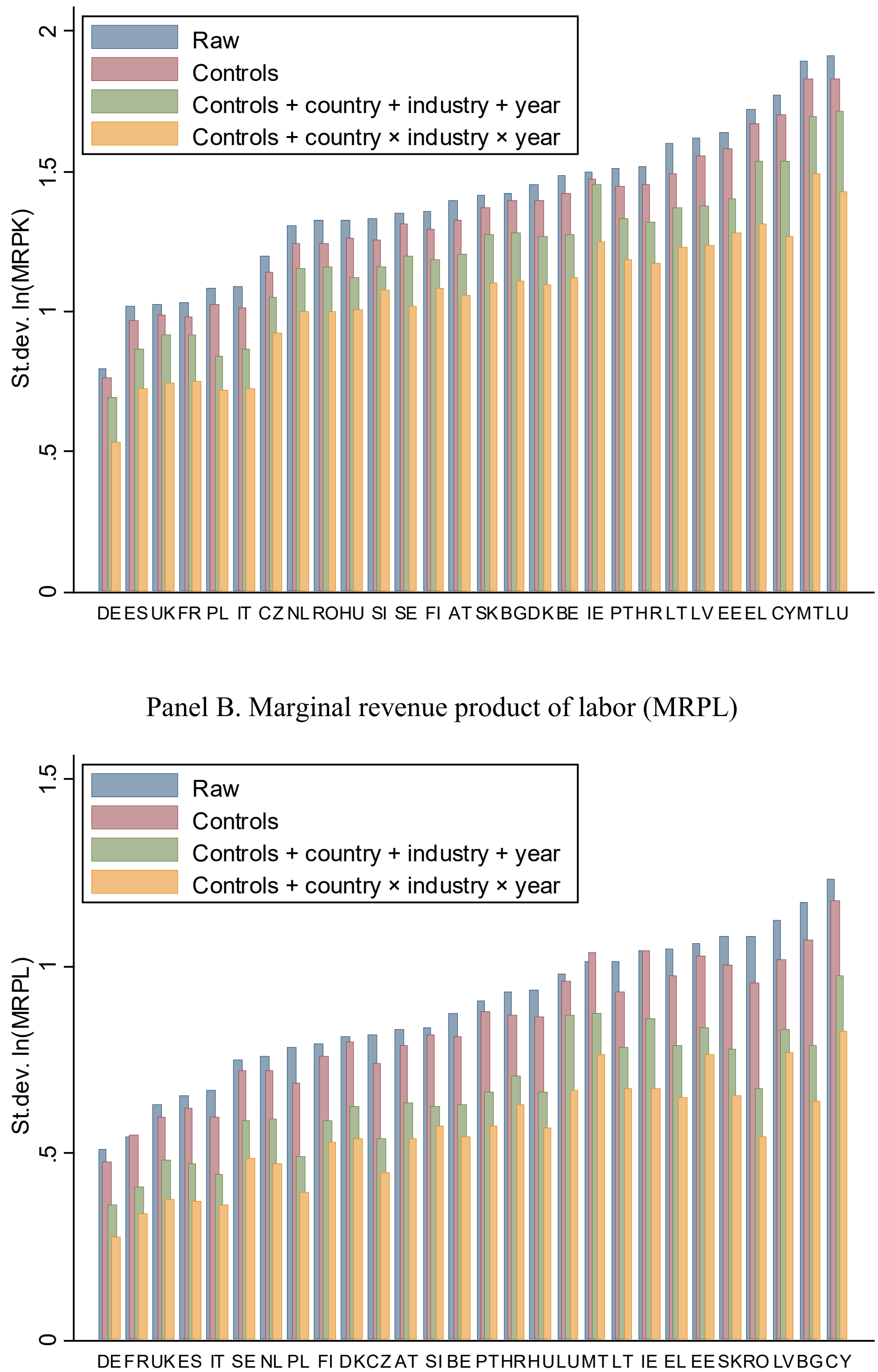
Panel B. Marginal revenue product of labor (MRPL)



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

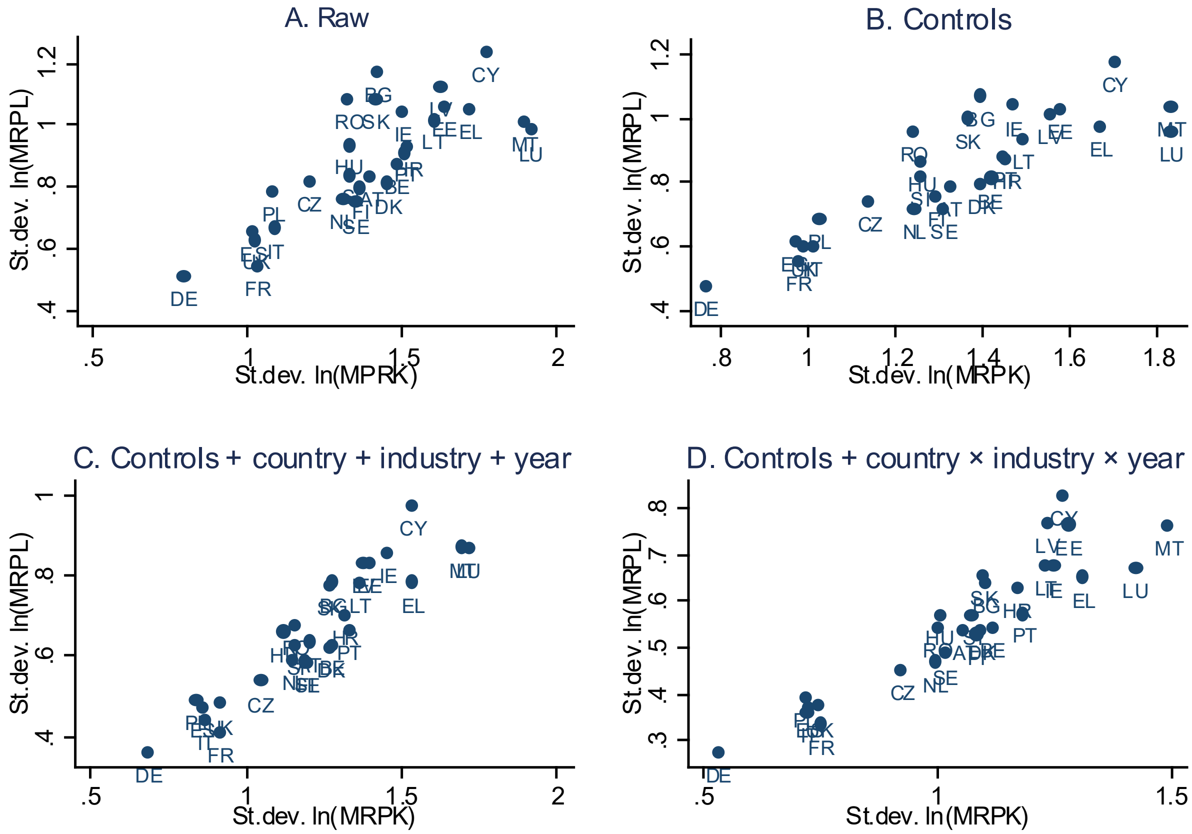
Figure 2. Raw and residual dispersion of the marginal revenue products of capital and labor

Panel A. Marginal revenue product of capital (MRPK)



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

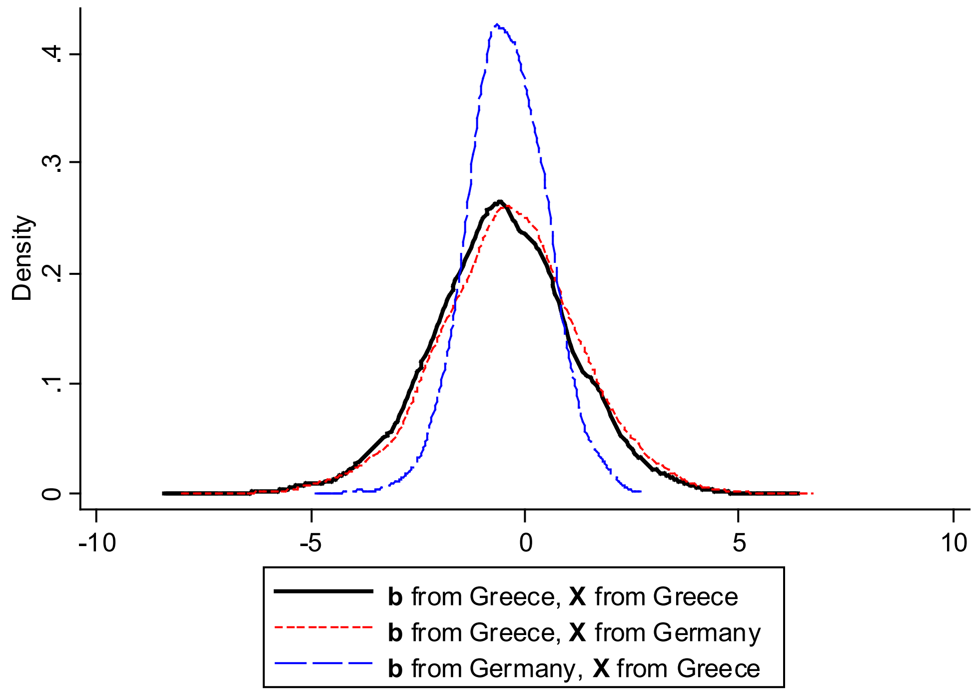
Figure 3. Association of the dispersion of the marginal revenue products of capital and labor



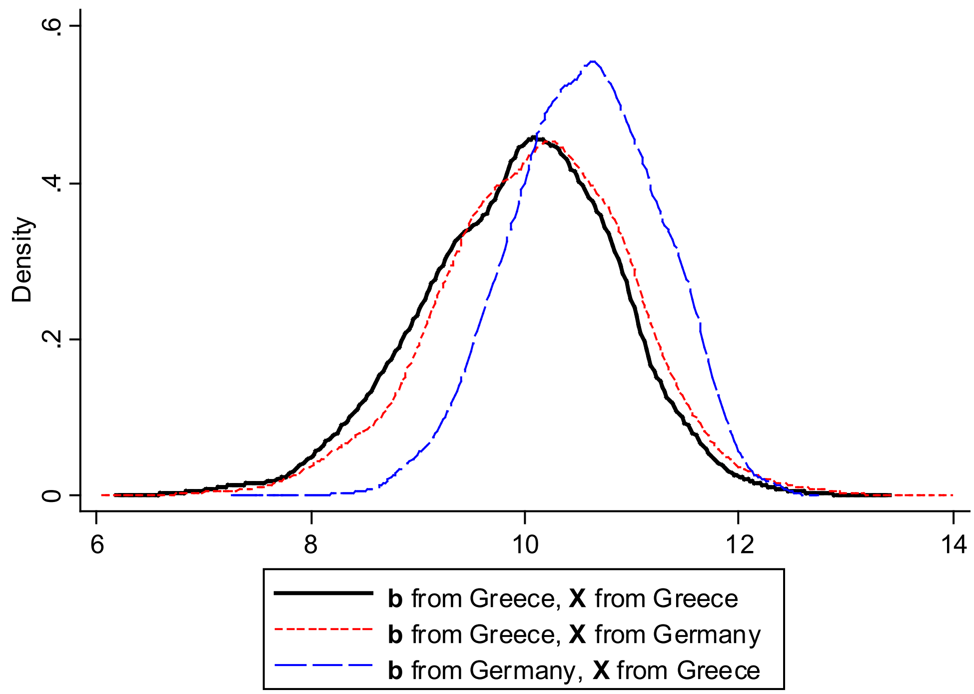
Note: The figures show the association between the dispersion of MRPK and MRPL across countries. "Raw" means no controls. "Controls" include the firm-level characteristics described in section IV. "Controls + country + industry" add fixed effects for countries, industries and years to firm-level characteristics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

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

Panel A. Marginal revenue product of capital (MRPK)



Panel B. Marginal revenue product of labor (MRPL)



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

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

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

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

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

Table 3. Descriptive statistics

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

Note: All statistics are computed using sampling weights.

Table 4. Marginal R^2 of adding a group of variables to a specification with fixed effects

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

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

Table 5. R^2 for various sets of fixed effects

		List of fixed effects				
		Country	Industry	Year	Country + Industry + Year	Country × Industry × Year
		(2)	(3)	(4)	(5)	(6)
	Dispersion					
(1)	MRPK	0.052	0.239	0.000	0.275	0.492
(2)	MRPL	0.445	0.268	0.001	0.611	0.736
(3)	MRPL - MRPK	0.219	0.174	0.000	0.354	0.555
(4)	**Productivity gain**	1.027	0.634	0.002	1.430	1.770

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

Table 6. Machado-Mata decomposition of the marginal revenue products of capital and labor

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

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

Appendix A: Additional Tables

Table A.1. Descriptive statistics, by year

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

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

Note: All statistics are computed using sampling weights.

Table A.2. Correlation matrix: variable index

Index	Group of variables	Variable	
1	**Outcome variables**	log(sales)	
2		log(fixed assets)	
3		log(employment)	
4		log(MRPK)	
5		log(MRPL)	
6		log(MRPL) - log(MRPK)	
7	**Demographics**	Firm age	less than 5 years
8			5-9 years
9			10-19 years
10			20+ years
11		Subsidiary	
12		Exporter	
13	**Quality of capital and other inputs**	Share of state-of-the art machinery and equipment	
14		Share of high energy efficiency commercial building stock	
15	**Capacity utilization**	above maximum capacity	
16		at maximum capacity	
17		somewhat below full capacity	
18		substantially below full capacity	
19	**Obstacles to investment**	Demand for products or services	Major
20			Minor
21		Availability of staff with the right skills	Major
22			Minor
23		Energy costs	Major
24			Minor
25		Access to digital infrastructure	Major
26			Minor
27		Labor market regulations	Major
28			Minor
29		Business regulations and taxation	Major
30			Minor
31		Availability of adequate transport infrastructure	Major
32			Minor
33		Availability of finance	Major
34			Minor
35		Uncertainty about future	Major
36			Minor
37		Investment, log(1 + investment)	
38		Percent change in employment in the last three years	
39	**Adjustment**	Investment over the last three years	too much
40			about the right amount
41			too little
42			company did not exist three years ago
43		Investment priority in the next three years	replacing capacity
44			capacity expansion for existing products or services
45			developing new products, processes or services
46			no investment planned
47			internal funds or retained earnings
48	**Source of funds**	external finance	
49		intra-group funding	
50		Finance constrained	

Table A.3. Correlation matrix: coefficients

Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	1																							
2	0.86	1																						
3	0.86	0.80	1																					
4	-0.02	-0.51	-0.13	1																				
5	0.48	0.29	0.02	0.24	1																			
6	0.37	0.66	0.13	-0.71	0.52	1																		
7	-0.12	-0.13	-0.11	0.04	-0.05	-0.08	1																	
8	-0.16	-0.17	-0.14	0.08	-0.07	-0.12	-0.05	1																
9	-0.14	-0.13	-0.13	0.04	-0.07	-0.09	-0.10	-0.14	1															
10	0.26	0.27	0.24	-0.10	0.13	0.18	-0.28	-0.42	-0.75	1														
11	0.36	0.26	0.30	0.07	0.21	0.08	-0.01	-0.03	-0.02	0.04	1													
12	0.28	0.24	0.24	-0.05	0.11	0.13	-0.04	-0.06	-0.07	0.11	0.15	1												
13	0.06	0.08	-0.01	-0.07	0.14	0.16	0.01	0.01	0.00	-0.01	-0.02	0.09	1											
14	0.07	0.11	0.04	-0.10	0.07	0.14	-0.03	-0.02	0.02	0.01	-0.03	0.05	0.41	1										
15	-0.01	-0.04	-0.03	0.06	0.03	-0.03	0.03	0.01	0.01	-0.03	0.01	0.01	0.05	0.02	1									
16	0.02	-0.01	0.00	0.08	0.05	-0.03	-0.02	0.02	0.03	-0.03	-0.01	-0.07	0.06	0.07	-0.23	1								
17	0.01	0.04	0.03	-0.06	-0.03	0.03	0.01	-0.02	-0.04	0.04	0.01	0.06	-0.06	-0.06	-0.21	-0.72	1							
18	-0.10	-0.04	-0.09	-0.09	-0.07	0.02	0.01	0.01	-0.01	0.00	-0.03	0.00	-0.04	-0.04	-0.08	-0.26	-0.24	1						
19	0.01	0.03	0.01	-0.05	-0.03	0.02	-0.02	-0.03	-0.01	0.03	-0.03	0.03	-0.01	0.02	-0.04	-0.06	0.02	0.11	1					
20	-0.01	-0.02	-0.01	0.02	-0.03	-0.04	-0.01	0.00	0.00	0.01	0.02	0.02	-0.05	-0.03	-0.01	-0.04	0.06	0.00	-0.33	1				
21	-0.12	-0.12	-0.09	0.03	-0.10	-0.10	0.01	0.02	0.01	-0.03	-0.07	0.00	0.03	0.02	0.07	0.01	-0.04	0.00	0.17	0.03	1			
22	0.12	0.12	0.11	-0.03	0.04	0.06	-0.03	-0.03	0.01	0.02	0.04	0.06	-0.08	-0.06	-0.05	-0.02	0.05	-0.01	-0.05	0.09	-0.57	1		
23	-0.06	0.01	-0.03	-0.11	-0.08	0.04	0.00	-0.02	-0.01	0.02	-0.10	-0.04	0.02	0.02	-0.01	-0.07	0.07	0.03	0.15	0.02	0.17	-0.05	1	
24	0.01	0.01	0.03	0.00	-0.03	-0.02	-0.01	-0.02	0.01	0.01	-0.01	0.04	-0.04	-0.01	0.00	-0.01	0.02	0.00	0.06	0.12	0.07	0.11	-0.40	1
25	-0.01	0.00	-0.03	-0.02	0.04	0.05	0.00	-0.01	-0.01	0.01	-0.04	-0.02	0.06	0.04	0.00	-0.01	0.02	0.00	0.18	-0.04	0.17	-0.06	0.21	0.00
26	0.01	0.03	0.02	-0.02	-0.01	0.01	-0.02	-0.03	-0.01	0.04	-0.04	0.01	-0.05	-0.01	0.01	0.00	0.00	0.00	0.09	0.17	0.09	0.10	0.07	0.27
27	-0.11	-0.09	-0.09	-0.01	-0.09	-0.05	-0.01	0.00	0.01	0.00	-0.10	-0.05	0.01	0.00	0.02	-0.07	0.04	0.04	0.15	0.04	0.23	-0.08	0.26	0.01
28	0.06	0.05	0.06	-0.01	0.02	0.02	-0.01	-0.02	0.01	0.00	0.02	0.06	-0.04	-0.03	0.01	0.00	0.02	-0.02	0.01	0.11	0.01	0.14	-0.04	0.19
29	-0.10	-0.05	-0.09	-0.06	-0.05	0.01	-0.01	0.00	0.01	-0.01	-0.11	-0.06	0.04	0.05	0.01	-0.04	0.03	0.03	0.16	0.01	0.15	-0.06	0.26	-0.01
30	0.05	0.04	0.06	0.02	0.00	-0.01	-0.01	-0.01	-0.01	0.02	0.01	0.03	-0.07	-0.04	-0.01	0.00	0.02	-0.01	0.01	0.11	0.03	0.09	-0.06	0.22
31	0.01	0.05	0.02	-0.06	-0.03	0.03	-0.02	-0.02	-0.01	0.03	-0.05	-0.02	0.02	0.04	0.01	-0.01	0.00	0.02	0.18	-0.01	0.15	-0.07	0.22	-0.02
32	0.02	0.03	0.03	-0.01	-0.02	-0.01	0.00	-0.01	0.00	0.00	-0.01	0.03	-0.06	-0.04	0.00	-0.03	0.03	0.00	0.09	0.16	0.07	0.09	0.07	0.21
33	-0.14	-0.09	-0.12	-0.06	-0.10	-0.02	0.03	0.03	0.01	-0.04	-0.11	-0.06	-0.03	0.00	0.00	-0.06	0.01	0.12	0.27	-0.03	0.10	-0.06	0.20	-0.02
34	-0.04	-0.03	-0.02	0.00	-0.06	-0.04	0.00	0.01	0.02	-0.02	-0.02	0.00	-0.05	-0.03	-0.02	-0.01	0.03	-0.02	0.06	0.21	0.06	0.06	0.03	0.16
35	-0.09	-0.05	-0.08	-0.04	-0.08	-0.02	-0.01	-0.01	0.00	0.01	-0.12	-0.04	-0.02	0.00	-0.04	-0.08	0.04	0.12	0.31	-0.01	0.13	-0.05	0.26	0.00
36	0.10	0.08	0.09	0.01	0.04	0.02	-0.02	-0.02	0.00	0.02	0.07	0.05	-0.07	-0.04	0.00	0.01	0.02	-0.06	-0.12	0.18	-0.05	0.12	-0.11	0.14
37	0.67	0.67	0.65	-0.21	0.21	0.33	-0.07	-0.10	-0.11	0.18	0.26	0.23	0.09	0.07	0.01	0.00	0.03	-0.09	0.00	0.00	-0.05	0.10	-0.02	0.02
38	-0.02	-0.04	-0.01	0.05	-0.04	-0.07	0.11	0.08	0.03	-0.12	-0.03	0.01	0.07	0.03	0.09	0.06	-0.06	-0.08	-0.05	0.02	0.07	-0.02	0.00	0.01
39	-0.04	-0.01	-0.02	-0.05	-0.03	0.02	0.01	0.02	-0.01	0.00	-0.01	0.00	0.02	0.02	0.00	-0.03	0.01	0.06	0.04	0.00	0.02	-0.02	0.01	-0.01
40	0.09	0.07	0.05	0.00	0.09	0.06	-0.01	-0.01	0.02	-0.01	-0.01	0.00	0.13	0.10	0.01	0.10	-0.05	-0.12	-0.07	0.00	-0.04	0.01	-0.03	0.02
41	-0.06	-0.06	-0.03	0.01	-0.08	-0.06	0.00	0.00	-0.02	0.01	0.01	0.01	-0.14	-0.11	-0.01	-0.09	0.05	0.09	0.06	0.01	0.04	0.01	0.03	-0.01
42	-0.04	-0.03	-0.03	-0.01	-0.01	0.00	0.13	0.00	-0.01	-0.04	0.02	-0.01	0.02	-0.02	0.01	-0.01	-0.01	0.03	-0.01	0.01	0.00	0.01	0.01	0.00
43	-0.01	0.00	0.01	-0.02	-0.04	-0.01	0.01	-0.03	0.00	0.02	0.02	-0.09	-0.04	-0.03	-0.02	0.01	0.01	-0.02	-0.03	0.00	-0.02	-0.01	0.03	0.01
44	0.05	0.06	0.05	-0.03	0.02	0.04	0.02	0.02	0.02	-0.03	-0.02	0.03	0.04	0.03	0.02	0.03	-0.03	-0.04	0.00	-0.01	0.03	0.00	-0.03	0.01
45	0.10	0.08	0.07	0.00	0.06	0.04	-0.02	-0.01	-0.04	0.05	0.07	0.15	0.02	0.01	0.00	-0.06	0.05	0.03	0.03	0.04	0.01	0.04	0.00	-0.01
46	-0.21	-0.21	-0.21	0.07	-0.06	-0.10	0.00	0.04	0.03	-0.05	-0.10	-0.13	-0.03	-0.01	-0.01	0.01	-0.03	0.04	0.01	-0.02	-0.01	-0.05	0.00	-0.01
47	-0.08	-0.12	-0.06	0.11	-0.04	-0.13	0.01	0.01	0.02	-0.03	-0.06	-0.03	-0.03	0.00	-0.02	0.03	-0.03	0.01	-0.02	-0.02	-0.03	0.02	-0.08	0.00
48	0.03	0.08	0.02	-0.12	0.01	0.11	-0.02	-0.01	-0.03	0.04	-0.05	0.01	0.03	0.00	0.02	-0.03	0.04	-0.02	0.02	0.02	0.04	-0.03	0.09	0.00
49	0.10	0.09	0.08	-0.01	0.05	0.05	0.01	0.01	0.03	-0.03	0.22	0.03	0.02	0.02	-0.01	0.01	-0.01	0.02	0.01	0.02	-0.01	0.01	-0.02	0.01
50	-0.08	-0.04	-0.05	-0.04	-0.08	-0.02	0.03	0.02	0.01	-0.04	-0.10	-0.02	-0.03	-0.01	0.00	-0.04	0.02	0.05	0.01	0.00	0.01	0.01	0.03	-0.01

Index	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49
25	1																								
26	-0.23	1																							
27	0.22	0.13	1																						
28	-0.03	0.16	-0.45	1																					
29	0.19	0.07	0.41	-0.09	1																				
30	-0.02	0.19	-0.10	0.30	-0.48	1																			
31	0.24	0.05	0.23	-0.04	0.23	-0.02	1																		
32	0.04	0.27	0.09	0.18	0.08	0.19	-0.28	1																	
33	0.19	0.05	0.19	-0.04	0.21	-0.04	0.19	0.02	1																
34	0.01	0.24	0.08	0.11	0.04	0.16	0.02	0.27	-0.29	1															
35	0.16	0.10	0.27	-0.05	0.29	-0.05	0.19	0.08	0.28	0.07	1														
36	-0.07	0.09	-0.10	0.16	-0.12	0.17	-0.08	0.09	-0.12	0.10	-0.58	1													
37	0.02	0.01	-0.07	0.05	-0.07	0.06	0.01	0.04	-0.13	0.00	-0.11	0.09	1												
38	0.03	0.00	0.02	0.00	0.02	-0.01	0.02	0.04	-0.03	0.03	-0.06	0.02	0.05	1											
39	0.02	0.00	0.01	0.00	0.03	-0.01	0.00	0.02	0.06	0.00	0.04	-0.01	0.02	0.01	1										
40	-0.01	-0.03	-0.03	-0.01	-0.02	-0.01	-0.01	0.00	-0.16	-0.01	-0.09	0.04	0.11	0.06	-0.37	1									
41	0.00	0.03	0.03	0.01	0.01	0.02	0.02	0.00	0.14	0.02	0.08	-0.03	-0.09	-0.06	-0.09	-0.85	1								
42	0.00	-0.01	0.00	-0.01	0.00	-0.01	-0.01	0.00	0.03	-0.01	-0.01	0.00	-0.02	-0.01	-0.01	-0.06	-0.01	1							
43	-0.02	-0.02	0.02	-0.05	0.00	-0.02	0.00	-0.03	-0.01	0.01	-0.01	0.00	0.01	-0.04	-0.02	-0.03	0.05	-0.01	1						
44	0.00	0.02	-0.03	0.04	-0.02	0.05	0.02	0.02	-0.01	0.01	-0.03	0.02	0.08	0.07	-0.02	0.04	-0.04	0.01	-0.46	1					
45	0.02	0.02	0.00	0.04	0.01	0.00	-0.01	0.02	0.01	0.00	0.01	0.02	0.10	0.00	0.03	-0.02	0.02	-0.01	-0.44	-0.37	1				
46	0.01	-0.01	0.02	-0.03	0.03	-0.04	-0.01	-0.01	0.03	-0.02	0.06	-0.06	-0.29	-0.03	0.02	0.01	-0.04	0.01	-0.24	-0.20	-0.19	1			
47	-0.07	-0.02	-0.09	0.01	-0.06	-0.02	-0.07	-0.02	-0.11	-0.04	-0.06	0.01	-0.18	-0.03	-0.04	0.00	0.01	-0.01	-0.03	0.01	-0.01	0.03	1		
48	0.07	0.02	0.09	-0.01	0.06	0.02	0.07	0.02	0.11	0.05	0.07	-0.02	0.15	0.03	0.04	0.00	-0.01	-0.01	0.04	0.00	-0.01	-0.02	-0.94	1	
49	-0.01	0.01	0.01	-0.01	0.01	0.00	0.00	0.02	0.02	-0.01	0.00	-0.01	0.08	0.01	0.01	0.01	-0.02	0.06	-0.01	0.01	0.03	-0.04	-0.24	-0.08	1
50	-0.01	0.01	0.03	-0.01	0.05	-0.02	0.01	-0.02	0.19	-0.03	0.06	-0.01	-0.04	0.00	0.03	-0.10	0.08	0.05	-0.01	0.00	0.02	0.00	0.12	-0.13	0.03

Note: The table reports the correlation coefficients for the variables listed in Table A.2. All statistics are computed using sampling weights.

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

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

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

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

Table A.5. Predictors of the dispersion of the marginal revenue products of capital and labor, balanced panel

Regressor	Dependent variable	
	log(MRPK)	log(MRPL)
Demographics		
Firm age (omitted category: less than 5 years)		
5-9 years	0.131	0.100**
	(0.095)	(0.051)
10-19 years	0.098	0.097**
	(0.089)	(0.048)
20+ years	-0.116	0.032
	(0.086)	(0.046)
log(employment)	0.026**	-0.040***
	(0.011)	(0.006)
Subsidiary	0.345***	0.125***
	(0.043)	(0.022)
Exporter	0.258***	0.295***
	(0.031)	(0.017)
Quality of capital and other inptus		
Share of state-of-the art machinery and equipment, including ICT	0.049	0.262***
	(0.050)	(0.024)
Share of high energy efficiency commercial building stock	-0.184***	0.136***
	(0.043)	(0.020)
Capacity utilization (omitted category: somewhat below full capacity)		
above maximum capacity	0.364***	0.270***
	(0.058)	(0.029)
at maximum capacity	0.135***	0.038***
	(0.030)	(0.014)
substantially below full capacity	-0.242***	-0.110***
	(0.050)	(0.024)
Obstacles to investment (omitted category: not an obstacle at all)		
Demand for products or services		
Major	0.111***	0.043**
	(0.038)	(0.019)
Minor	0.094***	0.113***
	(0.034)	(0.018)
Availability of staff with the right skills		
Major	0.026	-0.088***
	(0.036)	(0.019)
Minor	0.031	0.004
	(0.035)	(0.018)
Energy costs		
Major	-0.107***	-0.102***
	(0.038)	(0.021)
Minor	0.055	-0.041**
	(0.034)	(0.018)
Access to digital infrastructure		
Major	-0.134***	0.104***
	(0.049)	(0.024)
Minor	-0.117***	0.048***
	(0.034)	(0.016)
Labor market regulations		
Major	0.028	-0.098***
	(0.038)	(0.020)
Minor	0.099***	-0.043**
	(0.035)	(0.018)
Business regulations and taxation		
Major	-0.150***	0.026
	(0.040)	(0.020)
Minor	-0.047	0.058***

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

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

Table A.6. Predictors of the dispersion of the marginal revenue products of capital and labor, sample averaged across waves

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

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

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

Table A.7. Predictors of the dispersion of the marginal revenue products of capital and labor, alternative measure of investment

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

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

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

Table A.8. Marginal R^2 of adding a group of variables to a specification with fixed effects, sample averaged across waves

Row	Group of variables	List of fixed effects: No fixed effects	Country	Industry	Country + Industry + Year	Country × Industry × Year
		(1)	(2)	(3)	(4)	(5)
	Panel A: average MRPK					
(1)	Demographics	0.048	0.036	0.027	0.021	0.017
(2)	Quality of capital	0.024	0.014	0.018	0.010	0.008
(3)	Capacity utilization	0.030	0.030	0.015	0.015	0.009
(4)	Obstacles to investment	0.038	0.027	0.019	0.011	0.008
(5)	Adjustment	0.078	0.080	0.028	0.023	0.020
(6)	Source of funds	0.034	0.026	0.015	0.011	0.010
(7)	"Compensating differentials"	0.128	0.119	0.064	0.051	0.038
(8)	"Distortions"	0.096	0.075	0.047	0.035	0.028
(9)	All variables	0.185	0.168	0.100	0.084	0.066
	Panel B: average MRPL					
(10)	Demographics	0.148	0.051	0.118	0.058	0.042
(11)	Quality of capital	0.049	0.018	0.036	0.014	0.013
(12)	Capacity utilization	0.021	0.013	0.015	0.007	0.006
(13)	Obstacles to investment	0.082	0.030	0.049	0.018	0.015
(14)	Adjustment	0.133	0.053	0.128	0.060	0.040
(15)	Source of funds	0.056	0.013	0.040	0.008	0.007
(16)	"Compensating differentials"	0.172	0.076	0.159	0.080	0.060
(17)	"Distortions"	0.203	0.067	0.147	0.057	0.041
(18)	All variables	0.350	0.130	0.281	0.112	0.091
	Panel C: average MRPL – average MRPK					
(19)	Demographics	0.109	0.077	0.060	0.039	0.026
(20)	Quality of capital	0.055	0.028	0.049	0.024	0.017
(21)	Capacity utilization	0.004	0.009	0.001	0.003	0.003
(22)	Obstacles to investment	0.034	0.022	0.014	0.011	0.009
(23)	Adjustment	0.227	0.164	0.150	0.100	0.076
(24)	Source of funds	0.020	0.012	0.013	0.009	0.008
(25)	"Compensating differentials"	0.265	0.190	0.184	0.121	0.092
(26)	"Distortions"	0.148	0.098	0.079	0.050	0.037
(27)	All variables	0.346	0.235	0.249	0.157	0.119
	Panel D: Productivity gain based on average marginal revenue products					
(28)	Demographics	0.283	0.106	0.219	0.110	0.080
(29)	Quality of capital	0.097	0.037	0.072	0.029	0.026
(30)	Capacity utilization	0.037	0.024	0.026	0.014	0.011
(31)	Obstacles to investment	0.150	0.058	0.089	0.033	0.028
(32)	Adjustment	0.281	0.127	0.256	0.125	0.086
(33)	Source of funds	0.103	0.025	0.072	0.015	0.014
(34)	"Compensating differentials"	0.358	0.173	0.317	0.165	0.125
(35)	"Distortions"	0.386	0.139	0.273	0.110	0.079
(36)	All variables	0.686	0.277	0.544	0.229	0.183

Note: The table replicates Table 4 in the paper for average marginal revenue products. Taking within-firm average marginal products should attenuate adverse effects of measurement errors. This table does not have a column with year fixed effects because panel data are collapsed to a cross-section. See notes for Table 4 for more details.

Table A.9. R^2 for various sets of fixed effects, sample averaged across waves

		List of fixed effects			
		Country	Industry	Country + Industry	Country × Industry
		(1)	(2)	(3)	(4)
	Dispersion				
(1)	MRPK	0.061	0.278	0.314	0.526
(2)	MRPL	0.408	0.243	0.563	0.676
(3)	MRPL - MRPK	0.193	0.190	0.346	0.554
(4)	**Productivity gain**	0.755	0.465	1.058	1.300

Note: The table replicates Table 5 in the paper for average marginal revenue products. Taking within-firm average marginal products should attenuate adverse effects of measurement errors. This table does not have a column with year fixed effects because panel data are collapsed to a cross-section. See notes for Table 5 for more details.

Table A.10. Predictors of the dispersion of the marginal revenue products of capital and labor, EU North/West

Regressor	Dependent variable	
	log(MRPK)	log(MRPL)
Demographics		
Firm age (omitted category: less than 5 years)		
5-9 years	0.109*	0.001
	(0.064)	(0.029)
10-19 years	-0.103*	0.062**
	(0.056)	(0.026)
20+ years	-0.241***	0.079***
	(0.052)	(0.026)
log(employment)	0.075***	-0.028***
	(0.009)	(0.005)
Subsidiary	0.398***	0.101***
	(0.029)	(0.013)
Exporter	0.206***	0.139***
	(0.025)	(0.013)
Quality of capital and other inputs		
Share of state-of-the art machinery and equipment, including ICT	-0.149***	0.087***
	(0.040)	(0.017)
Share of high energy efficiency commercial building stock	-0.129***	0.006
	(0.035)	(0.015)
Capacity utilization (omitted category: somewhat below full capacity)		
above maximum capacity	0.284***	0.133***
	(0.039)	(0.019)
at maximum capacity	0.178***	0.065***
	(0.022)	(0.010)
substantially below full capacity	-0.223***	-0.096***
	(0.037)	(0.017)
Obstacles to investment (omitted category: not an obstacle at all)		
Demand for products or services		
Major	-0.047	-0.004
	(0.032)	(0.014)
Minor	-0.006	-0.011
	(0.026)	(0.011)
Availability of staff with the right skills		
Major	0.139***	-0.074***
	(0.029)	(0.012)
Minor	0.059**	-0.014
	(0.027)	(0.013)
Energy costs		
Major	-0.131***	-0.069***
	(0.033)	(0.015)
Minor	-0.108***	-0.054***
	(0.024)	(0.012)
Access to digital infrastructure		
Major	0.031	-0.006
	(0.038)	(0.017)
Minor	-0.028	0.015
	(0.026)	(0.012)
Labor market regulations		
Major	-0.021	-0.072***
	(0.031)	(0.014)
Minor	-0.062**	-0.046***
	(0.025)	(0.012)
Business regulations and taxation		
Major	-0.128***	0.015
	(0.033)	(0.014)
Minor	-0.017	0.013
	(0.026)	(0.012)
Availability of adequate transport infrastructure		
Major	-0.039	0.118***
	(0.035)	(0.017)
Minor	0.009	0.049***
	(0.026)	(0.012)

Availability of finance		
Major	-0.017	-0.090***
	(0.034)	(0.015)
Minor	-0.009	-0.072***
	(0.026)	(0.013)
Uncertainty about future		
Major	0.173***	0.043***
	(0.030)	(0.014)
Minor	0.133***	0.023**
	(0.025)	(0.012)
Adjustment		
Investment, log(1 + investment)	-0.090***	0.047***
	(0.005)	(0.002)
Percent change in employment in the last three years	-0.001	-0.116***
	(0.022)	(0.011)
Investment over the last three years (omitted category: about the right amount)		
too much	-0.230***	-0.044*
	(0.054)	(0.025)
too little	-0.041	0.000
	(0.027)	(0.012)
company did not exist three years ago	-0.620***	0.109
	(0.205)	(0.121)
Investment priority in the next three years (omitted category: no investment planned)		
replacing capacity	-0.070*	-0.010
	(0.039)	(0.017)
capacity expansion for existing products or services	-0.035	0.004
	(0.039)	(0.018)
developing new products, processes or services	-0.064	-0.004
	(0.043)	(0.018)
Source of funds (omitted category: external finance)		
internal funds or retained earnings	0.285***	0.087***
	(0.034)	(0.015)
intra-group funding	0.268***	0.203***
	(0.085)	(0.040)
Credit constrained	-0.206***	-0.138***
	(0.048)	(0.020)
Sample size	11,172	11,016
R^2	0.509	0.649
Memorandum		
R^2 with country × industry × year fixed effects and no $\boldsymbol{X}$	0.456	0.591
R^2 with $\boldsymbol{X}$ and no fixed effects	0.166	0.180
R^2 with $\boldsymbol{X}$ and country fixed effects	0.174	0.210
R^2 with $\boldsymbol{X}$ and industry fixed effects	0.309	0.439
R^2 with $\boldsymbol{X}$ and year fixed effects	0.169	0.180
R^2 with $\boldsymbol{X}$ and country fixed effects, industry fixed effects and year fixed effects	0.318	0.456
R^2 with $\boldsymbol{X}$ and slopes varying by country	0.229	0.252
R^2 with $\boldsymbol{X}$ and slopes varying by industry	0.388	0.445
R^2 with $\boldsymbol{X}$ and slopes varying by year	0.176	0.189

Note: The sample is restricted to Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Luxembourg, the Netherlands, Sweden, and the UK. See notes for Appendix Table A.4 for more details.

Table A.11. Predictors of the dispersion of the marginal revenue products of capital and labor, EU South

Regressor	Dependent variable	
	log(MRPK)	log(MRPL)
Demographics		
Firm age (omitted category: less than 5 years)		
5-9 years	-0.357***	0.022
	(0.108)	(0.049)
10-19 years	-0.472***	-0.013
	(0.092)	(0.046)
20+ years	0.664***	0.087*
	(0.094)	(0.044)
log(employment)	0.025**	-0.043***
	(0.011)	(0.007)
Subsidiary	0.253***	0.078***
	(0.050)	(0.023)
Exporter	0.159***	0.316***
	(0.035)	(0.019)
Quality of capital and other inptus		
Share of state-of-the art machinery and equipment, including ICT	-0.108**	0.080***
	(0.052)	(0.027)
Share of high energy efficiency commercial building stock	-0.272***	0.074***
	(0.046)	(0.025)
Capacity utilization (omitted category: somewhat below full capacity)		
above maximum capacity	0.347***	0.013
	(0.066)	(0.037)
at maximum capacity	0.082***	0.037**
	(0.031)	(0.017)
substantially below full capacity	-0.471***	-0.158***
	(0.052)	(0.026)
Obstacles to investment (omitted category: not an obstacle at all)		
Demand for products or services		
Major	0.173***	0.009
	(0.039)	(0.021)
Minor	0.179***	0.046**
	(0.041)	(0.022)
Availability of staff with the right skills		
Major	0.120***	-0.033*
	(0.039)	(0.019)
Minor	0.150***	0.004
	(0.039)	(0.020)
Energy costs		
Major	-0.173***	-0.026
	(0.037)	(0.021)
Minor	-0.129***	0.004
	(0.039)	(0.022)
Access to digital infrastructure		
Major	0.042	0.044*
	(0.048)	(0.023)
Minor	-0.013	0.074***
	(0.038)	(0.019)
Labor market regulations		
Major	0.143***	-0.105***
	(0.043)	(0.022)
Minor	0.017	-0.056**
	(0.044)	(0.023)
Business regulations and taxation		
Major	-0.06	0.064***
	(0.044)	(0.021)
Minor	-0.041	0.032
	(0.045)	(0.023)
Availability of adequate transport infrastructure		

Major	0.018	0.044**
	(0.037)	(0.021)
Minor	0.017	0.018
	(0.036)	(0.019)
Availability of finance		
Major	-0.165***	-0.099***
	(0.036)	(0.021)
Minor	-0.068*	-0.074***
	(0.040)	(0.023)
Uncertainty about future		
Major	-0.168***	0.004
	(0.049)	(0.028)
Minor	-0.118**	-0.045
	(0.051)	(0.028)
Adjustment		
Investment, log(1 + investment)	-0.051***	0.048***
	(0.006)	(0.003)
Percent change in employment in the last three years	0.160***	-0.128***
	(0.036)	(0.020)
Investment over the last three years (omitted category: about the right amount)		
too much	-0.231***	-0.124***
	(0.069)	(0.039)
too little	-0.184***	-0.105***
	(0.043)	(0.020)
company did not exist three years ago	-0.701	-0.418*
	(0.491)	(0.238)
Investment priority in the next three years (omitted category: no investment planned)		
replacing capacity	-0.040	0.017
	(0.052)	(0.024)
capacity expansion for existing products or services	-0.043	0.015
	(0.051)	(0.025)
developing new products, processes or services	-0.010	0.054**
	(0.053)	(0.024)
Source of funds (omitted category: external finance)		
internal funds or retained earnings	0.098**	0.044*
	(0.043)	(0.024)
intra-group funding	-0.613***	0.314***
	(0.147)	(0.084)
Credit constrained	-0.075	-0.075***
	(0.056)	(0.026)
Sample size	5,657	5,602
R^2	0.483	0.627
Memorandum		
R^2 with country × industry × year fixed effects and no $\boldsymbol{X}$	0.432	0.574
R^2 with $\boldsymbol{X}$ and no fixed effects	0.136	0.185
R^2 with $\boldsymbol{X}$ and country fixed effects	0.147	0.208
R^2 with $\boldsymbol{X}$ and industry fixed effects	0.327	0.471
R^2 with $\boldsymbol{X}$ and year fixed effects	0.136	0.185
R^2 with $\boldsymbol{X}$ and country fixed effects, industry fixed effects and year fixed effects	0.341	0.525
R^2 with $\boldsymbol{X}$ and slopes varying by country	0.193	0.248
R^2 with $\boldsymbol{X}$ and slopes varying by industry	0.489	0.568
R^2 with $\boldsymbol{X}$ and slopes varying by year	0.153	0.200

Note: The sample is restricted to Cyprus, Greece, Italy, Malta, Portugal, and Spain. See notes for Appendix Table A.4 for more details.

Table A.12. Predictors of the dispersion of the marginal revenue products of capital and labor, EU Center/East

Regressor	Dependent variable	
	log(MRPK)	log(MRPL)
Demographics		
Firm age (omitted category: less than 5 years)		
5-9 years	0.006	0.066**
	(0.051)	(0.030)
10-19 years	-0.271***	0.093***
	(0.051)	(0.028)
20+ years	-0.374***	0.097***
	(0.050)	(0.028)
log(employment)	0.007	-0.017***
	(0.009)	(0.005)
Subsidiary	0.293***	0.227***
	(0.033)	(0.021)
Exporter	0.096***	0.368***
	(0.027)	(0.016)
Quality of capital and other inptus		
Share of state-of-the art machinery and equipment, including ICT	-0.080**	0.262***
	(0.034)	(0.020)
Share of high energy efficiency commercial building stock	-0.298***	0.101***
	(0.029)	(0.017)
Capacity utilization (omitted category: somewhat below full capacity)		
above maximum capacity	0.184***	0.113***
	(0.045)	(0.027)
at maximum capacity	0.058***	0.020*
	(0.021)	(0.012)
substantially below full capacity	-0.260***	-0.132***
	(0.034)	(0.019)
Obstacles to investment (omitted category: not an obstacle at all)		
Demand for products or services		
Major	0.068***	0.042***
	(0.026)	(0.016)
Minor	0.025	0.023*
	(0.022)	(0.013)
Availability of staff with the right skills		
Major	-0.023	-0.028*
	(0.027)	(0.015)
Minor	-0.045	0.016
	(0.028)	(0.015)
Energy costs		
Major	-0.114***	-0.131***
	(0.027)	(0.016)
Minor	-0.062***	-0.065***
	(0.023)	(0.014)
Access to digital infrastructure		
Major	0.034	0.019
	(0.044)	(0.026)
Minor	0.035	0.032**
	(0.024)	(0.013)
Labor market regulations		
Major	-0.070**	-0.058***
	(0.028)	(0.015)
Minor	0.019	-0.052***
	(0.024)	(0.014)
Business regulations and taxation		
Major	0.012	-0.007
	(0.028)	(0.016)
Minor	0.093***	0.063***
	(0.025)	(0.014)
Availability of adequate transport infrastructure		
Major	0.039	0.060***
	(0.027)	(0.017)
Minor	0.026	0.051***
	(0.023)	(0.013)

Availability of finance		
Major	-0.074***	-0.069***
	(0.026)	(0.016)
Minor	-0.032	-0.050***
	(0.021)	(0.013)
Uncertainty about future		
Major	0.032	0.050***
	(0.028)	(0.017)
Minor	-0.003	0.057***
	(0.027)	(0.015)
Adjustment		
Investment, log(1 + investment)	-0.050***	0.058***
	(0.004)	(0.003)
Percent change in employment in the last three years	0.172***	-0.077***
	(0.026)	(0.014)
Investment over the last three years (omitted category: about the right amount)		
too much	-0.208***	-0.162***
	(0.041)	(0.026)
too little	-0.006	-0.094***
	(0.023)	(0.013)
company did not exist three years ago	1.142***	-0.115
	(0.393)	(0.195)
Investment priority in the next three years (omitted category: no investment planned)		
replacing capacity	-0.06	-0.052**
	(0.037)	(0.021)
capacity expansion for existing products or services	-0.115***	-0.036
	(0.036)	(0.023)
developing new products, processes or services	-0.089**	-0.050**
	(0.038)	(0.023)
Source of funds (omitted category: external finance)		
internal funds or retained earnings	0.178***	0.043**
	(0.030)	(0.019)
intra-group funding	-0.294**	0.143**
	(0.123)	(0.061)
Credit constrained	-0.040	-0.052***
	(0.032)	(0.018)
Sample size	11,184	11,101
R^2	0.527	0.720
Memorandum		
R^2 with country × industry × year fixed effects and no $\boldsymbol{X}$	0.496	0.643
R^2 with $\boldsymbol{X}$ and no fixed effects	0.155	0.256
R^2 with $\boldsymbol{X}$ and country fixed effects	0.185	0.389
R^2 with $\boldsymbol{X}$ and industry fixed effects	0.342	0.519
R^2 with $\boldsymbol{X}$ and year fixed effects	0.155	0.257
R^2 with $\boldsymbol{X}$ and country fixed effects, industry fixed effects and year fixed effects	0.373	0.619
R^2 with $\boldsymbol{X}$ and slopes varying by country	0.227	0.366
R^2 with $\boldsymbol{X}$ and slopes varying by industry	0.407	0.523
R^2 with $\boldsymbol{X}$ and slopes varying by year	0.162	0.265

Note: The sample is restricted to Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, and Slovenia. See notes for Appendix Table A.4 for more details.

Table A.13. Marginal R^2 of adding a group of variables to a specification with fixed effects, EU North/West

Row	Group of variables	No fixed effects	Country	Industry	Year	Country + Industry + Year	Country × Industry × Year
		List of fixed effects					
		(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: MRPK						
(1)	Demographics	0.041	0.035	0.027	0.042	0.024	0.019
(2)	Quality of capital	0.020	0.011	0.015	0.020	0.010	0.007
(3)	Capacity utilization	0.012	0.017	0.005	0.012	0.008	0.008
(4)	Obstacles to investment	0.028	0.025	0.010	0.028	0.009	0.008
(5)	Adjustment	0.073	0.070	0.030	0.073	0.029	0.021
(6)	Source of funds	0.042	0.041	0.028	0.040	0.024	0.022
(7)	"Compensating differentials"	0.103	0.096	0.052	0.103	0.046	0.038
(8)	"Distortions"	0.092	0.086	0.048	0.091	0.044	0.037
(9)	All variables	0.166	0.154	0.091	0.166	0.084	0.071
	Panel B: MRPL						
(10)	Demographics	0.079	0.063	0.059	0.079	0.050	0.034
(11)	Quality of capital	0.001	0.004	0.002	0.002	0.004	0.002
(12)	Capacity utilization	0.011	0.011	0.006	0.011	0.006	0.004
(13)	Obstacles to investment	0.057	0.044	0.023	0.057	0.018	0.016
(14)	Adjustment	0.056	0.048	0.050	0.056	0.044	0.030
(15)	Source of funds	0.050	0.040	0.023	0.050	0.019	0.013
(16)	"Compensating differentials"	0.076	0.068	0.064	0.076	0.058	0.045
(17)	"Distortions"	0.126	0.101	0.069	0.126	0.058	0.040
(18)	All variables	0.180	0.150	0.113	0.180	0.098	0.074
	Panel C: MRPL - MRPK						
(19)	Demographics	0.051	0.047	0.019	0.050	0.018	0.011
(20)	Quality of capital	0.017	0.015	0.016	0.017	0.014	0.010
(21)	Capacity utilization	0.006	0.008	0.001	0.006	0.003	0.004
(22)	Obstacles to investment	0.027	0.028	0.010	0.027	0.010	0.008
(23)	Adjustment	0.160	0.151	0.091	0.160	0.080	0.052
(24)	Source of funds	0.014	0.015	0.012	0.013	0.010	0.009
(25)	"Compensating differentials"	0.176	0.166	0.104	0.175	0.092	0.061
(26)	"Distortions"	0.084	0.080	0.037	0.082	0.034	0.026
(27)	All variables	0.229	0.222	0.138	0.229	0.129	0.093
	Panel D: Productivity gain						
(28)	Demographics	0.138	0.111	0.096	0.137	0.082	0.055
(29)	Quality of capital	0.008	0.011	0.008	0.008	0.010	0.006
(30)	Capacity utilization	0.018	0.019	0.010	0.018	0.011	0.008
(31)	Obstacles to investment	0.096	0.076	0.039	0.096	0.031	0.027
(32)	Adjustment	0.137	0.122	0.105	0.137	0.092	0.063
(33)	Source of funds	0.082	0.066	0.039	0.081	0.032	0.023
(34)	"Compensating differentials"	0.172	0.157	0.132	0.171	0.119	0.088
(35)	"Distortions"	0.220	0.180	0.117	0.219	0.099	0.069
(36)	All variables	0.348	0.301	0.217	0.348	0.191	0.143

Note: The sample is restricted to Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Luxembourg, the Netherlands, Sweden, and the UK. See notes for Table 4 for more details.

Table A.14. Marginal R^2 of adding a group of variables to a specification with fixed effects, EU South.

		List of fixed effects					
Row	Group of variables	No fixed effects	Country	Industry	Year	Country + Industry + Year	Country × Industry × Year
		(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: MRPK						
(1)	Demographics	0.040	0.036	0.019	0.040	0.015	0.013
(2)	Quality of capital	0.008	0.007	0.005	0.008	0.004	0.005
(3)	Capacity utilization	0.028	0.025	0.018	0.028	0.017	0.016
(4)	Obstacles to investment	0.030	0.029	0.015	0.030	0.014	0.014
(5)	Adjustment	0.057	0.053	0.018	0.056	0.016	0.013
(6)	Source of funds	0.014	0.011	0.005	0.013	0.005	0.003
(7)	"Compensating differentials"	0.092	0.084	0.043	0.092	0.036	0.034
(8)	"Distortions"	0.076	0.069	0.033	0.075	0.030	0.027
(9)	All variables	0.136	0.126	0.068	0.135	0.060	0.057
	Panel B: MRPL						
(10)	Demographics	0.081	0.086	0.041	0.080	0.049	0.044
(11)	Quality of capital	0.016	0.014	0.006	0.016	0.006	0.007
(12)	Capacity utilization	0.014	0.014	0.006	0.013	0.007	0.007
(13)	Obstacles to investment	0.055	0.048	0.021	0.055	0.015	0.013
(14)	Adjustment	0.084	0.070	0.060	0.083	0.042	0.039
(15)	Source of funds	0.011	0.008	0.010	0.010	0.005	0.005
(16)	"Compensating differentials"	0.125	0.117	0.079	0.124	0.070	0.065
(17)	"Distortions"	0.098	0.088	0.048	0.097	0.037	0.034
(18)	All variables	0.185	0.170	0.108	0.183	0.093	0.085
	Panel C: MRPL - MRPK						
(19)	Demographics	0.094	0.086	0.039	0.094	0.032	0.027
(20)	Quality of capital	0.015	0.012	0.010	0.015	0.009	0.009
(21)	Capacity utilization	0.013	0.009	0.005	0.013	0.004	0.006
(22)	Obstacles to investment	0.032	0.028	0.014	0.031	0.011	0.012
(23)	Adjustment	0.124	0.107	0.063	0.123	0.047	0.034
(24)	Source of funds	0.023	0.014	0.017	0.022	0.007	0.004
(25)	"Compensating differentials"	0.161	0.141	0.080	0.160	0.062	0.050
(26)	"Distortions"	0.122	0.102	0.062	0.121	0.043	0.037
(27)	All variables	0.227	0.200	0.120	0.226	0.095	0.083
	Panel D: Productivity gain						
(28)	Demographics	0.168	0.175	0.082	0.167	0.094	0.083
(29)	Quality of capital	0.032	0.027	0.014	0.033	0.014	0.015
(30)	Capacity utilization	0.027	0.027	0.012	0.027	0.013	0.014
(31)	Obstacles to investment	0.102	0.090	0.040	0.102	0.029	0.027
(32)	Adjustment	0.185	0.156	0.124	0.183	0.088	0.077
(33)	Source of funds	0.027	0.018	0.023	0.026	0.010	0.010
(34)	"Compensating differentials"	0.267	0.246	0.161	0.265	0.140	0.126
(35)	"Distortions"	0.208	0.183	0.103	0.206	0.077	0.070
(36)	All variables	0.391	0.356	0.223	0.387	0.190	0.172

Note: The sample is restricted to Cyprus, Greece, Italy, Malta, Portugal, and Spain. See notes for Table 4 for more details.

Table A.15. Marginal R^2 of adding a group of variables to a specification with fixed effects, EU Center/East

		List of fixed effects					
Row	Group of variables	No fixed effects	Country	Industry	Year	Country + Industry + Year	Country × Industry × Year
		(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: MRPK						
(1)	Demographics	0.063	0.060	0.019	0.063	0.020	0.016
(2)	Quality of capital	0.018	0.011	0.018	0.018	0.011	0.009
(3)	Capacity utilization	0.010	0.012	0.006	0.010	0.007	0.004
(4)	Obstacles to investment	0.020	0.017	0.009	0.020	0.007	0.004
(5)	Adjustment	0.087	0.080	0.027	0.087	0.023	0.018
(6)	Source of funds	0.019	0.016	0.008	0.019	0.008	0.006
(7)	"Compensating differentials"	0.111	0.100	0.049	0.111	0.041	0.029
(8)	"Distortions"	0.093	0.086	0.034	0.093	0.032	0.024
(9)	All variables	0.155	0.142	0.074	0.155	0.063	0.046
	Panel B: MRPL						
(10)	Demographics	0.107	0.059	0.106	0.106	0.073	0.059
(11)	Quality of capital	0.064	0.029	0.047	0.065	0.021	0.020
(12)	Capacity utilization	0.029	0.018	0.018	0.029	0.009	0.007
(13)	Obstacles to investment	0.059	0.040	0.031	0.059	0.020	0.016
(14)	Adjustment	0.098	0.071	0.087	0.097	0.060	0.048
(15)	Source of funds	0.006	0.005	0.006	0.006	0.005	0.005
(16)	"Compensating differentials"	0.175	0.107	0.155	0.175	0.094	0.079
(17)	"Distortions"	0.118	0.076	0.083	0.117	0.063	0.049
(18)	All variables	0.256	0.153	0.198	0.255	0.121	0.101
	Panel C: MRPL - MRPK						
(19)	Demographics	0.136	0.102	0.076	0.135	0.054	0.043
(20)	Quality of capital	0.067	0.039	0.055	0.068	0.030	0.026
(21)	Capacity utilization	0.002	0.002	0.000	0.002	0.000	0.000
(22)	Obstacles to investment	0.032	0.020	0.016	0.032	0.008	0.008
(23)	Adjustment	0.209	0.153	0.123	0.208	0.092	0.074
(24)	Source of funds	0.025	0.023	0.014	0.025	0.014	0.012
(25)	"Compensating differentials"	0.252	0.180	0.165	0.251	0.114	0.093
(26)	"Distortions"	0.154	0.121	0.082	0.153	0.063	0.051
(27)	All variables	0.309	0.220	0.202	0.308	0.136	0.111
	Panel D: Productivity gain						
(28)	Demographics	0.252	0.149	0.231	0.251	0.158	0.128
(29)	Quality of capital	0.146	0.070	0.110	0.149	0.051	0.047
(30)	Capacity utilization	0.057	0.035	0.035	0.057	0.018	0.013
(31)	Obstacles to investment	0.124	0.083	0.065	0.123	0.041	0.034
(32)	Adjustment	0.260	0.189	0.209	0.258	0.148	0.117
(33)	Source of funds	0.021	0.018	0.016	0.020	0.014	0.013
(34)	"Compensating differentials"	0.423	0.268	0.355	0.422	0.221	0.185
(35)	"Distortions"	0.280	0.188	0.187	0.278	0.142	0.112
(36)	All variables	0.599	0.370	0.449	0.596	0.279	0.232

Note: The sample is restricted to Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, and Slovenia. See notes for Table 4 for more details.

Table A.16. R^2 for various sets of fixed effects, by EU region

		List of fixed effects				
		Country	Industry	Year	Country + Industry + Year	Country × Industry × Year
		(2)	(3)	(4)	(5)	(6)
Panel A: EU North/West						
	Dispersion					
(1)	MRPK	0.019	0.226	0.003	0.240	0.456
(2)	MRPL	0.065	0.329	0.001	0.370	0.591
(3)	MRPL - MRPK	0.033	0.199	0.003	0.227	0.482
(4)	**Productivity gain**	0.110	0.567	0.002	0.639	1.058
Panel B: EU South						
	Dispersion					
(1)	MRPK	0.023	0.264	0.000	0.286	0.432
(2)	MRPL	0.040	0.371	0.002	0.444	0.574
(3)	MRPL - MRPK	0.074	0.246	0.002	0.297	0.459
(4)	Productivity gain	0.094	0.705	0.004	0.845	1.120
Panel C: EU Center/East						
	Dispersion					
(1)	MRPK	0.048	0.277	0.000	0.319	0.496
(2)	MRPL	0.246	0.349	0.002	0.518	0.643
(3)	MRPL - MRPK	0.187	0.216	0.002	0.355	0.541
(4)	Productivity gain	0.536	0.744	0.005	1.116	1.420

Note: The EU North/West sample is restricted to Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Luxembourg, the Netherlands, Sweden, and the UK. The EU South sample is restricted to Cyprus, Greece, Italy, Malta, Portugal, and Spain. The EU Center/East sample is restricted to Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, and Slovenia. See notes for Table 5 for more details.

Table A.17. Predictors of the dispersion of the marginal revenue products of capital and labor, capital and labor regressands are from Orbis

Regressor	Dependent variable	
	log(MRPK)	log(MRPL)
Demographics		
Firm age (omitted category: less than 5 years)		
5-9 years	-0.001	0.045**
	(0.038)	(0.019)
10-19 years	-0.212***	0.063***
	(0.035)	(0.018)
20+ years	-0.335***	0.094***
	(0.034)	(0.017)
log(employment)	0.048***	-0.016***
	(0.006)	(0.003)
Subsidiary	0.323***	0.171***
	(0.019)	(0.010)
Exporter	0.132***	0.244***
	(0.016)	(0.009)
Quality of capital and other inptus		
Share of state-of-the art machinery and equipment, including ICT	-0.124***	0.148***
	(0.023)	(0.012)
Share of high energy efficiency commercial building stock	-0.258***	0.059***
	(0.020)	(0.010)
Capacity utilization (omitted category: somewhat below full capacity)		
above maximum capacity	0.249***	0.094***
	(0.026)	(0.013)
at maximum capacity	0.128***	0.044***
	(0.014)	(0.007)
substantially below full capacity	-0.315***	-0.133***
	(0.022)	(0.011)
Obstacles to investment (omitted category: not an obstacle at all)		
Demand for products or services		
Major	0.070***	0.008
	(0.018)	(0.009)
Minor	0.052***	0.011
	(0.016)	(0.008)
Availability of staff with the right skills		
Major	0.064***	-0.050***
	(0.017)	(0.009)
Minor	0.032*	-0.005
	(0.017)	(0.009)
Energy costs		
Major	-0.127***	-0.086***
	(0.018)	(0.009)
Minor	-0.091***	-0.055***
	(0.015)	(0.008)
Access to digital infrastructure		
Major	0.036	0.015
	(0.023)	(0.012)
Minor	0.01	0.035***
	(0.016)	(0.008)
Labor market regulations		
Major	-0.009	-0.070***
	(0.018)	(0.009)
Minor	-0.02	-0.043***
	(0.016)	(0.008)
Business regulations and taxation		
Major	-0.046**	0.027***

	(0.019)	(0.009)
Minor	0.022	0.028***
	(0.017)	(0.009)
Availability of adequate transport infrastructure		
Major	-0.024	0.072***
	(0.019)	(0.010)
Minor	0.008	0.048***
	(0.015)	(0.008)
Availability of finance		
Major	-0.051***	-0.086***
	(0.018)	(0.010)
Minor	-0.003	-0.072***
	(0.015)	(0.008)
Uncertainty about future		
Major	0.059***	0.048***
	(0.018)	(0.010)
Minor	0.048***	0.035***
	(0.017)	(0.009)
Adjustment		
Investment, log(1 + investment)	-0.067***	0.048***
	(0.003)	(0.001)
Percent change in employment in the last three years	0.099***	-0.097***
	(0.015)	(0.008)
Investment over the last three years (omitted category: about the right amount)		
too much	-0.223***	-0.099***
	(0.030)	(0.016)
too little	-0.065***	-0.070***
	(0.016)	(0.008)
company did not exist three years ago	-0.255	-0.066
	(0.176)	(0.087)
Investment priority in the next three years (omitted category: no investment planned)		
replacing capacity	-0.069***	-0.021*
	(0.023)	(0.012)
capacity expansion for existing products or services	-0.087***	-0.002
	(0.023)	(0.012)
developing new products, processes or services	-0.080***	0.008
	(0.025)	(0.012)
Source of funds (omitted category: external finance)		
internal funds or retained earnings	0.183***	0.066***
	(0.020)	(0.010)
intra-group funding	-0.099	0.122***
	(0.063)	(0.030)
Credit constrained	-0.091***	-0.093***
	(0.024)	(0.012)
Sample size	27,832	27,627
R^2	0.527	0.777
Memorandum		
R^2 with country × industry × year fixed effects and no ***X***	0.492	0.736
R^2 with ***X*** and no fixed effects	0.144	0.287
R^2 with ***X*** and country fixed effects	0.175	0.53
R^2 with ***X*** and industry fixed effects	0.304	0.460
R^2 with ***X*** and year fixed effects	0.145	0.287
R^2 with ***X*** and country fixed effects, industry fixed effects and year fixed effects	0.332	0.677
R^2 with ***X*** and slopes varying by country	0.226	0.459
R^2 with ***X*** and slopes varying by industry	0.305	0.445
R^2 with ***X*** and slopes varying by year	0.147	0.290

Note: In this table, the regressors log(employment) and log(1+investment) use data from the Orbis database. Other variables are from EIBIS. See notes to Appendix Table A.4 for more details.

Appendix B: Derivations for the Hsieh-Klenow model

Setup

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

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

Subject to

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

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

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

Aggregate demand is given by the Dixit-Stiglitz aggregator:

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

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

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

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

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

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

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

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

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

which implies that

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

Optimality conditions

The Lagrangian for the firm is

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

Optimality conditions are:

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

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

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

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

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

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

which implies that

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

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

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

We can also find expressions for marginal revenue products

$$MRPK_i \equiv \frac{\sigma-1}{\sigma}\alpha\frac{P_i Y_i}{K_i} = \frac{\tau_i^K}{\tau_i^Y} R$$
$$MRPL_i \equiv \frac{\sigma-1}{\sigma}\beta\frac{P_i Y_i}{L_i} = \frac{\tau_i^L}{\tau_i^Y} W$$
$$MRPX_i \equiv \frac{\sigma-1}{\sigma}(1-\alpha-\beta)\frac{P_i Y_i}{X_i} = \frac{\tau_i^X}{\tau_i^Y} P^X$$

AGGREGATION

Aggregate capital in the economy is given by

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

Note that $Y P^\sigma \left(\frac{\alpha}{r}\right)\frac{\sigma}{\sigma-1}\boldsymbol{B}^{1-\sigma}$ does not depend on firm-specific outcomes.
Aggregate labor in the economy is given by

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

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

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

Aggregate intermediate input in the economy is given by

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

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

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

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

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

Aggregate price index is given by

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

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

AGGREGATE TFP

Using our definition of aggregate TFP, we have

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

Let's compute the denominator of this expression:

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

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

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

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

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

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

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

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

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

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

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

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

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

APPROXIMATION TO AGGREGATE PRODUCTIVITY

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

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

Consider the aggregate price level:

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

Then using $\int z_i di = E(z_i)$ (that is a cross-sectional average is equal to the mathematical expectation of random variable z_i) and the property of log-normal variable $E(z) = \exp\left(mean_z + \frac{1}{2} variance_z\right)$,[29] we have

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\log A = \log\left[P^{-\sigma}\frac{\sigma-1}{\sigma}\boldsymbol{B}^{\sigma}\right]+\log Green+\log Red+\log Blue$$
$$= constant - \sigma\log P+\log Green+\log Red+\log Blue$$
$$= constant+\left\{\frac{\sigma(\sigma-1)\alpha^2}{2}V_{\tau K}+\frac{\sigma(\sigma-1)\beta^2}{2}V_{\tau L}+\frac{\sigma(\sigma-1)(1-\alpha-\beta)^2}{2}V_{\tau X}+\frac{\sigma(\sigma-1)}{2}V_{\tau Y}+\sigma\mu_A+\frac{\sigma(\sigma-1)}{2}V_A\right\}$$
$$-\alpha\left(\frac{(\alpha(1-\sigma)-1)^2}{2}V_{\tau K}+\frac{(\beta(1-\sigma))^2}{2}V_{\tau L}+\frac{\left((1-\alpha-\beta)(1-\sigma)\right)^2}{2}V_{\tau X}+\frac{\sigma^2}{2}V_{\tau Y}+(\sigma-1)\mu_A+\frac{(\sigma-1)^2}{2}V_A\right)$$
$$-\beta\left(\frac{\left(\alpha(1-\sigma)\right)^2}{2}V_{\tau K}+\frac{(\beta(1-\sigma)-1)^2}{2}V_{\tau L}+\frac{\left((1-\alpha-\beta)(1-\sigma)\right)^2}{2}V_{\tau X}+\frac{\sigma^2}{2}V_{\tau Y}+(\sigma-1)\mu_A+\frac{(\sigma-1)^2}{2}V_A\right)$$
$$-(1-\alpha-\beta)\left(\frac{\left(\alpha(1-\sigma)\right)^2}{2}V_{\tau K}+\frac{\left(\beta(1-\sigma)\right)^2}{2}V_{\tau L}+\frac{\left((1-\alpha-\beta)(1-\sigma)-1\right)^2}{2}V_{\tau X}+\frac{\sigma^2}{2}V_{\tau Y}+(\sigma-1)\mu_A+\frac{(\sigma-1)^2}{2}V_A\right)$$
$$= constant-\left\{\frac{\alpha(1-\alpha)}{2}+\frac{\alpha^2\sigma}{2}\right\}V_{\tau K}-\left\{\frac{\beta(1-\beta)}{2}+\frac{\beta^2\sigma}{2}\right\}V_{\tau L}-\left\{\frac{(1-\alpha-\beta)(\alpha+\beta)}{2}+\frac{(1-\alpha-\beta)^2\sigma}{2}\right\}V_{\tau X}-\frac{\sigma}{2}V_{\tau Y}+\mu_A+\frac{(\sigma-1)}{2}V_A$$

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

IDENTIFICATION OF DISTORTIONS τ

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

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

Using the same logic, we have

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

We have four unknowns $\tau_i^Y, \tau_i^K, \tau_i^L, \tau_i^X$ and three moments $MPRK_i, MRPL_i, MRPX_i$. The system is not identified. We need to impose an identifying assumption.
The Hsieh-Klenow framework assumes that $\tau_i^L = 1$ for all i and hence, one can find

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

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

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

and hence

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

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

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

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

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

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

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

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

If we make the other assumption, then

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

Appendix C: Additional derivations

Section A. Cost shares

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

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

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

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

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

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

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

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

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

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

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

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

given that the share of pure economic profit in total revenue s_π is approximately zero in the data (e.g., Basu and Fernald, 1997).[30] Hence, we can further simplify the expressions for marginal revenue products to obtain

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

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

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

[30] Weaker assumptions may suffice. For example, to study a cross-sectional dispersion of marginal revenue products, it is enough to have γ/μ constant across industries and countries. If even this assumption is not satisfied, one ought to consider using country and/or industry fixed effects to control for the variation in market power and in returns to scale.

Section B. Quality adjustment for capital

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

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

where $\tilde{R}_{it}$ is a measure of the capital price used in constructing the replacement value or the balance sheet value of fixed assets. In the case of replacement value of capital, we may have $R_{it}(U_{it}) \approx \tilde{R}_{it}$. With the balance sheet value of fixed asset, $\tilde{R}_{it}$ likely reflects the historical price rather than the current market price. Given technical change and inflation, the difference between the market and historical prices can be large, especially for assets that were acquired a long time ago (e.g., buildings).

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

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

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

We are fortunate to have proxy information that enables us to try to correct for this effect. In particular, from EIBIS we know the share of capital (including machinery, equipment and ICT) that the management considers to be "state-of-the-art", which presumably means capital that has been obtained recently. Thus, for firms with a large share of state-of-the-art capital we can expect $R_{it}(U_{it}) \approx \tilde{R}_{it}$.

Appendix D: R^2 for OLS vs. Instrumental Variables

Appendix D derives and compares R^2 for ordinary least squares (OLS) and instrumental variable (IV) estimators under various assumptions.

A. Structural model

Consider the following system:

$$y = \alpha x + u \tag{D.1}$$

$$x = z + e \tag{D.2}$$

where z is an instrumental variable (without loss of generality we can assume that the coefficient on z is equal to one; we can always normalize z so that this is the case), y is an outcome variable, x is a (potentially endogenous) regressor, u and e are error terms which could be correlated (x is endogenous) or uncorrelated (x is exogenous). We assume that z is a valid instrument and is uncorrelated with e and u.

B. OLS

When we run an OLS regression for equation (D.1), we find that R^2 is given by

$$R^2_{OLS} \equiv 1 - \frac{var(\hat{u}_{OLS})}{var(y)} \tag{D.3}$$

where $\hat{u}_{OLS}$ is the residual of the regression. Note that by construction of OLS regression we have

$$var(y) = var(\hat{\alpha}_{OLS}x) + var(\hat{u}_{OLS}) \tag{D.4}$$

and so it does not matter whether we use equation (D.3) to compute R^2 or instead we use

$$R^2_{OLS} = \frac{var(\hat{\alpha}_{OLS}x)}{var(y)} = \frac{\hat{\alpha}^2_{OLS}var(x)}{var(y)}. \tag{D.5}$$

We can find that

$$\hat{u}_{OLS} \equiv y - \hat{\alpha}x = (\alpha x + u) - \frac{cov(y,x)}{var(x)}x = (\alpha x + u) - \frac{cov(\alpha x + u,x)}{var(x)}x$$

$$= (\alpha x + u) - \alpha x - \frac{cov(u,x)}{var(x)}x = u - \frac{cov(u,x)}{var(x)}x \tag{D.6}$$

It follows that

$$var(\hat{u}_{OLS}) = var\left(u - \frac{cov(u,x)}{var(x)}x\right)$$

$$= var(u) + \left[\frac{cov(u,x)}{var(x)}\right]^2 var(x) - 2cov(u,x)\frac{cov(u,x)}{var(x)}$$

$$= var(u) - \frac{[cov(u,x)]^2}{var(x)} < var(u) \tag{D.7}$$

Thus, OLS understates the variance of the error term and so OLS will inflate R^2: $var(\hat{u}_{OLS}) < var(u)$.

Also notice that $cov(x, u) = cov(z + e, u) = cov(e, u)$ and so we can alternatively write

$$var(\hat{u}_{OLS}) = var(u) - \frac{[cov(u,e)]^2}{var(x)} \tag{D.7'}$$

C. Instrumental variables (IV)

Now consider the IV regression where the first stage is given by equation (D.2). In this case, we can show that we can recover a consistent estimate of α. In this case, we can compute the *structural* error term in equation (1) as

$$\hat{u}_{IV} = y - \hat{\alpha}_{IV} x \tag{D.8}$$

Because the estimate is consistent, we have also the correct estimate of the variance of the error term u as the sample size increases to infinity:

$$var(\hat{u}_{IV}) = var(u). \tag{D.9}$$

But this means that $R_{IV}^2 \equiv 1 - \frac{var(\hat{u}_{IV})}{var(y)} \leq R_{OLS}^2$.

If $cov(u, x) = 0$ (or equivalently $cov(e, u) = 0$), we have $R_{IV}^2 = R_{OLS}^2$. However, notice that with the IV regression we can no longer use equations (3) and (5) equivalently because:

$$var(y) = var(\hat{\alpha}_{IV} x + u) = \hat{\alpha}_{IV}^2 var(x) + var(u) + 2\hat{\alpha}_{IV} cov(x, u) \tag{D.10}$$

and hence

$$\frac{\hat{\alpha}_{IV}^2 var(x)}{var(y)} \neq 1 - \frac{var(\hat{u}_{IV})}{var(y)}. \tag{D.11}$$

As a result, R_{IV}^2 can be negative.

While we can guarantee that $1 - \frac{var(\hat{u}_{IV})}{var(y)} \leq 1 - \frac{var(\hat{u}_{OLS})}{var(y)}$, we cannot guarantee that

$$\frac{\hat{\alpha}_{IV}^2 var(x)}{var(y)} \leq \frac{\hat{\alpha}_{OLS}^2 var(x)}{var(y)}. \tag{D.12}$$

D. Predictive error

To address this issue, Pesaran and Smith (1994) propose using *predictive* error in IV regressions to do model comparisons and model selection. In contrast to structural errors, these errors can be used for model selections because they have a well-defined measure of fit: generalized R^2 (or GR^2) which is calculated as follows

$$\tilde{u}_{IV} = y - \hat{\alpha}_{IV} \hat{x} \tag{D.13}$$

where $\hat{x}$ is the predicted value of x in the first-stage regression (D.2). Effectively, computing this residual amounts to regressing y on z, i.e., running a reduced-form regression. In this case, because z is uncorrelated with e or u, we have

$$var(y) = var(\hat{\alpha}_{IV} z + u) = \hat{\alpha}_{IV}^2 var(z) + var(u) \tag{D.14}$$

What is the relation of $var(\tilde{u}_{IV})$ to $var(\hat{u}_{OLS})$? Using the logic of equation (D.7) and the fact that $\hat{\alpha}_{IV} \rightarrow_p \alpha$, we can find

$$var(\tilde{u}_{IV}) = var(y - \hat{\alpha}_{IV}\hat{x}) = var(y - \alpha z) = var\big(y - \alpha(x - e)\big)$$

$$= var\big(\alpha x + u - \alpha(x - e)\big) = var(u + \alpha e)$$

$$= var(u) + \alpha^2 var(e) + 2\alpha cov(e, u) \qquad \text{(D.15)}$$

The second term in equation (D.15) is positive (and so we push R^2 down) but the third term is ambiguous: the sign depends how x influences y (the sign of α) and how the error terms are correlated.

Now using equation (D.7'), we compute

$$var(\hat{u}_{OLS}) - var(\tilde{u}_{IV}) = \left\{var(u) - \frac{[cov(u,e)]^2}{var(x)}\right\} - \{var(u) + \alpha^2 var(e) + 2\alpha cov(e, u)\}$$

$$= -\frac{[cov(u,e)]^2}{var(x)} - \alpha^2 var(e) - 2\alpha cov(e, u) \qquad \text{(D.16)}$$

The first two terms are negative (and so this helps us ensure that R^2_{OLS} is greater that GR^2_{IV}) but the last term is ambiguous because we need to know more about the nature of the omitted variable that gives a non-zero correlation between e and u as well as how x influences y.

One special case to guarantee $R^2_{OLS} \geq GR^2_{IV}$ is when the omitted variable moves y and x in the same direction (i.e., $cov(u, e) > 0$) as x moves y. For example, $cov(u, e) > 0$ and $\alpha > 0$. Alternatively, $cov(u, e) < 0$ and $\alpha < 0$. If $\alpha cov(e, u) \geq 0$, OLS overstates the contribution of x to variable of y and hence R^2_{OLS} continues to be an upper bound for the variation in y due to x.

One may expect condition $\alpha cov(e, u) = cov(x, u) \geq 0$ to be satisfied in our setting: more productive firms are more likely to run into distortionary constraints (e.g., red tape) and these constraints are likely to raise the marginal revenue product of an input (red tape makes firms employ too little of the input). Consider the following model as an illustration of this point. Suppose that a firm is maximizing profits subject to a size constraint:

$$\max_{E,L} A(EL)^\alpha - w(E)L \;\; s.t. \;\; L \leq \bar{L}$$

where L is labor input, $\bar{L}$ is the maximum firm size allowed by red tape, E is effort of workers, $w(E)$ is the wage function for effort, A is productivity (entrepreneurial talents of the firm's manager). We will assume that the wage function is increasing and convex: $w'(E) > 0, w''(E) > 0$. The parameter $\alpha < 1$ measures returns to scale. Firms are heterogenous in their draws of A.

The Lagrangian is $\mathcal{L} = A(EL)^\alpha - w(E)L - \lambda(L - \bar{L})$ where λ is the Lagrange multiplier on the size constraint. Optimality conditions for E and L are respectively:

$$\alpha A E^{\alpha-1} L^\alpha = w'(E)L,$$

$$MRPL \equiv \alpha A E^\alpha L^{\alpha-1} = w(E) + \lambda.$$

Note that when the size constraint is not binding, $\lambda = 0$ and hence the optimal level of effort E^* is given by

$$\frac{w'(E^*)E^*}{w(E^*)} = 1.$$

and so effort does not depend on productivity A. If the size constraint is not binding, the marginal revenue product of labor is

$$MRPL = w(E^*).$$

If the size constraint is binding, the optimal effort is given by

$$\alpha A(E^*)^{\alpha-1}\bar{L}^{\alpha} = w'(E^*)\bar{L} \Leftrightarrow A = \frac{\bar{L}^{1-\alpha}}{\alpha} w'(E^*)(E^*)^{1-\alpha}.$$

Because $w''(E) > 0$ and $\alpha < 1$, it follows that E^* is increasing in productivity A. To determine λ in this case, we note that

$$\alpha A(E^*)^{\alpha}\bar{L}^{\alpha-1} = w(E^*) + \lambda \Longrightarrow w'(E^*)E^* = w(E^*) + \lambda \Longrightarrow \lambda = w'(E^*)E^* - w(E^*)$$

Given our assumptions about the wage function, we have

$$\frac{\partial\lambda}{\partial E^*} = w''(E^*)E^* + w'(E^*) - w'(E^*) = w''(E^*)E^* > 0,$$

that is, the Lagrange multiplier on the size constraint is increasing in effort. Trivially, $MRPL$ is increasing in effort.

Now suppose we have a measure of how binding the size constraint is (e.g., in the survey, we ask firms to report how binding various constraints are: major issue, minor issue, not an issue) but we do not observe effort. We regress $MRPL$ on λ. The optimality condition for labor implies that the regression takes the following form:

$$MRPL = \lambda + w(E) = \lambda + error,$$

where the error term $error$ absorbs unobserved effort $w(E)$. Clearly the structural coefficient on λ is positive and $cov(\lambda, error) = cov(\lambda, w(E)) > 0$. Thus, the condition for $R^2_{OLS} \geq R^2_{IV}$ and $R^2_{OLS} \geq GR^2_{IV}$ is satisfied.

E. Measurement error

Suppose we do not have the right measure of x and instead have two imperfect measures of x: $x_1^* = x + \eta_1$ and $x_2^* = x + \eta_2$. To keep algebra simple, we assume that:

1) η is uncorrelated with u and e
2) η_1 is uncorrelated with η_2.
3) $var(\eta_1) = var(\eta_2) = var(\eta)$

Also, to focus on the measurement error, we assume that $cov(u, e) = 0$ and so we do not have endogeneity considered above.

Using x_1^* (or x_2^*) in estimating equation (D.1) yields

$$\hat{\alpha}^{OLS}_{ME} = \alpha\frac{var(x)}{var(x)+var(\eta)} \qquad \text{(D.17)}$$

$$\hat{u}_{ME,OLS} = y - \hat{\alpha}_{ME}^{OLS} x_1^* = (\alpha x + u) - \alpha \frac{var(x)}{var(x)+var(\eta)} (x + \eta_1)$$

$$= u + \alpha \frac{var(\eta)}{var(x)+var(\eta)} x - \alpha \frac{var(x)}{var(x)+var(\eta)} \eta_1 \tag{D.18}$$

Because u, x, η_1 are assumed to be uncorrelated, we have

$$var(\hat{u}_{ME,OLS}) = var(u) + \alpha^2 \left[\frac{var(\eta)}{var(x)+var(\eta)}\right]^2 var(x) + \alpha^2 \left[\frac{var(x)}{var(x)+var(\eta)}\right]^2 var(\eta)$$

$$= var(u) + \alpha^2 \frac{var(\eta)var(x)}{\left(var(x)+var(\eta)\right)^2} \{var(x) + var(\eta)\} =$$

$$= var(u) + \alpha^2 \frac{var(x)}{var(x)+var(\eta)} var(\eta) \tag{D.19}$$

Because $var(\hat{u}_{ME,OLS}) > var(u)$, it follows that

$$R_{ME,OLS}^2 = 1 - \frac{var(\hat{u}_{ME,OLS})}{var(y)} < R_{true}^2 = 1 - \frac{var(u)}{var(y)} \tag{D.20}$$

Furthermore, one can show that

$$R_{ME,OLS}^2 = 1 - \frac{var(\hat{u}_{ME,OLS})}{var(y)} = 1 - \frac{var(u)+\alpha^2 \frac{var(x)}{var(x)+var(\eta)} var(\eta)}{var(y)}$$

$$= \left(1 - \frac{var(u)}{var(y)}\right) - \frac{\alpha^2 var(x)}{var(y)} \times \frac{var(\eta)}{var(x)+var(\eta)} = R_{true}^2 - R_{true}^2 \frac{var(\eta)}{var(x)+var(\eta)}$$

$$= R_{true}^2 \frac{var(x)}{var(x)+var(\eta)} \tag{D.21}$$

where $R_{true}^2 \equiv 1 - \frac{var(u)}{var(y)} = \frac{\alpha^2 var(x)}{var(y)}$.

We can use x_2^* as an instrument to x_1^* because η_1 and η_2 are uncorrelated. This IV regression will give us $\hat{\alpha}_{IV} \to_p \alpha$.

However, having an IV estimate does not help here, because the structural error in this regression will still have measurement error

$$\hat{u}_{ME,IV} = y - \hat{\alpha}_{IV} x_1^* = y - \alpha x_1^* = \alpha x + u - \alpha(x + \eta_1) = u - \alpha \eta_1 \tag{D.22}$$

and hence

$$var(\hat{u}_{ME,IV}) = var(u) + \alpha^2 var(\eta). \tag{D.23}$$

Because $var(\hat{u}_{ME,IV}) > var(u)$, we have

$$R_{ME,IV}^2 = 1 - \frac{var(\hat{u}_{ME,IV})}{var(y)} < R_{true}^2$$

Because $var(\hat{u}_{ME,IV}) > var(\hat{u}_{ME,OLS})$, we also have $R^2_{ME,IV} < R^2_{ME,OLS}$.

Now consider using a predictive IV error term $\tilde{u}_{IV} = y - \hat{\alpha}_{IV}\hat{x}_1^*$ where $\hat{x}_1^* = \hat{\psi}x_2^*$ and $\hat{\psi} = \frac{cov(x_1^*,x_2^*)}{var(x_2^*)} = \frac{var(x)}{var(x)+var(\eta)}$. Because x_2^* also has measurement error, we obtain

$$\tilde{u}_{IV} = y - \hat{\alpha}_{IV}\hat{x}_1^* = y - \alpha\hat{\psi}x_2^*$$

$$= \alpha x + u - \alpha\hat{\psi}(x + \eta_2) = u + \alpha(1-\hat{\psi})x - \alpha\hat{\psi}\eta_2 \quad \text{(D.24)}$$

Because u, x, η_2 are uncorrelated, we have

$$var(\tilde{u}_{IV}) = var(u) + \alpha^2(1-\hat{\psi})^2 var(x) + \alpha^2\hat{\psi}^2 var(\eta) =$$

$$= var(u) + \alpha^2\left(\frac{var(\eta)}{var(x)+var(\eta)}\right)^2 var(x) + \alpha^2\left(\frac{var(x)}{var(x)+var(\eta)}\right)^2 var(\eta) =$$

$$= var(u) + \alpha^2\frac{var(\eta)var(x)}{[var(x)+var(\eta)]^2}\{var(x) + var(\eta)\} =$$

$$= var(u) + \alpha^2\frac{var(x)}{var(x)+var(\eta)}var(\eta) \quad \text{(D.25)}$$

Because $var(\tilde{u}_{IV}) > var(u)$, we know

$$GR^2_{ME,IV} \equiv 1 - \frac{var(\tilde{u}_{IV})}{var(y)} < R^2_{true}. \quad \text{(D.26)}$$

When we compare $GR^2_{ME,IV}$ (equation D.26) and $R^2_{ME,OLS}$ (equation D.19), we can see that $GR^2_{ME,IV} = R^2_{ME,OLS}$. Thus, OLS continues to provide an upper bound for predictive power of x in explaining y.

We can do a little better even when we have multiple measurements of x. Notice that we can use $\bar{x}^* = (x_1^* + x_2^*)/2$ as a regressor. In this case, the measurement error is reduced: $var(\bar{x}^*) = var\left(x + \frac{\eta_1}{2} + \frac{\eta_2}{2}\right) = var(x) + \frac{1}{4}var(\eta) + \frac{1}{4}var(\eta) = var(x) + \frac{1}{2}var(\eta)$. As a result,

$$R^2_{ME,OLS} = R^2_{true}\frac{var(x)}{var(x)+\frac{1}{2}var(\eta)} \quad \text{(D.27)}$$

and hence $R^2_{ME,OLS}$ is less biased. Thus, by running a regression with average values of x^*, we can reduce the size of measurement errors and try to improve the upper bound for R^2.