Banc Ceannais na hÉireann
Central Bank of Ireland

Eurosystem

Research Technical Paper

The Sources of Capital Misallocation in Europe

Stephen Byrne and Robert Goodhead

Vol. 2025, No. 11

The Sources of Capital Misallocation in Europe*

Stephen Byrne† Robert Goodhead‡

September 9, 2025

Abstract

This paper decomposes the sources of capital misallocation at the country and industry level in Europe. Using a comprehensive dataset of European firms from 19 countries, we find that the majority of the observed misallocation stems from persistent firm-specific distortions, with a smaller role for adjustment costs and uncertainty. We document substantial differences in the sources of misallocation across industries. Our analysis reveals strong correlations between these permanent distortions and industry-level variation in both financial factors, and factors relating to productivity. Understanding the factors driving capital misallocation is important for policymakers seeking to address productivity constraints and stimulate growth in the long run.

JEL Classification: E0, O11, O4

*The opinions expressed in this paper are those of the authors and do not necessarily represent the views of the Central Bank of Ireland, or the European System of Central Banks. For helpful comments and feedback, we thank Peter Klenow, Bruno Morando, Cian Ruane, Laszlo Tetenyi, as well as participants at the ENRI Conference 2023 and the Central Bank of Ireland internal seminar series.

†Corresponding author. Irish Economic Analysis Division, Central Bank of Ireland & Trinity College Dublin; stephen.byrne@centralbank.ie.

‡Monetary Policy Division, Central Bank of Ireland; robert.goodhead@centralbank.ie.

Non-Technical Summary

One reason living standards differ between countries is productivity: how efficiently economies turn factors of production (capital and labour) into outputs. Productivity growth in Europe has been weak in recent decades as compared to the United States. This matters because weak productivity growth undermines the ability of governments to achieve social objectives, as well as reducing long-term real wage growth. However, the causes of this weak productivity growth are not well understood.

One potential explanation lies in how capital is allocated across firms. Some firms may grow too large because they are subsidised excessively. Others remain too small because they face frictions that impede their growth. This misallocation of resources drags down aggregate productivity in an economy.

In this study we focus on the misallocation of capital across manufacturing firms in 19 European economies, and decompose its sources into four determinants: (1) adjustment costs; (2) uncertainty; (3) firm-specific subsidies (or taxes) that can change over time; (4) firm-specific subsidies (or taxes) that are permanent.

We find that firm-specific distortions play an important role for explaining regional differences in productivity across Europe, for the example the lower productivity in the southern economies relative to western ones. However, we document substantial variation in our estimates across industries, which suggests that the sources of misallocation differ according to industry-specific factors. In order to gain understanding of the structural drivers that explain our findings, we map our estimates to data on the characteristics of industries. We find that industry-level characteristics relating to financial variables and productivity have an important role in explaining permanent distortions.

1 Introduction

Relative to the United States, European productivity performance has been weak, for over two decades at the time of writing. As reported by Draghi (2024), the EU-US gap in the level of GDP at 2015 prices has increased from around 15% in 2002 to 30% in 2023, with 70% of this difference down to lower total factor productivity (TFP) growth. This poor relative performance occurred during a period in which US growth itself fell, alongside global growth generally (IMF, 2024). Productivity growth is a major long-term determinant of living standards (Solow, 1957). Weak productivity growth reduces the ability of governments to achieve desirable social objectives. The need for reforms to increase productivity has recently been emphasised by numerous European policymakers (Draghi, 2024; European Commission, 2025a,b). An important question is where to focus such reforms. Prospects for fruitful policy action can vary not just across different countries, but also to a varying extent within industries.

The literature has offered numerous explanations for weak European productivity performance: the failure of European economies to reap gains from the information technology boom (Bloom et al., 2012; Pellegrino and Zingales, 2017; Van Ark et al., 2008); low rates of innovation and low business R&D (Teichgraeber and Van Reenen, 2022), coupled with a tendency to focus on "mid-technology" sectors such as automobiles (Fuest et al., 2024); financial frictions and undeveloped venture capital markets (Draghi, 2024; European Commission, 2025a); the presence of "zombie" firms (Schivardi et al., 2021); and a broad lack of business dynamism and weak "up-or-out" dynamics (Adilbish et al., 2025).[1]

A number of studies have indicated an important role of resource misallocation for explaining low European productivity (ECB, 2021; Gamberoni et al., 2016; Gopinath et al., 2017; Gorodnichenko et al., 2021).[2] Misallocation occurs when resources are directed away from

[1]For recent summaries, see Schnabel (2024) and Bergeaud (2024).

[2]Country-level investigations have been conducted for Italy (Calligaris, 2015; Calligaris et al., 2018; Lenzu and Manaresi, 2018), for the Netherlands (Bun and de Winter, 2022), for Portugal (Dias et al., 2016; Reis, 2013), and for Spain (García-Santana et al., 2020).

their most productive uses, on account of distortions (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2008). A large literature has documented the *causes* of misallocation, attributing distortions to taxes and subsidies, size-dependent regulations (Garicano et al., 2016), financial frictions (Midrigan and Xu, 2014), uncertainty (David et al., 2016), and markup dispersion (Peters, 2020). Researchers have also explained misallocation by factors such as adjustment costs (Asker et al., 2014), or simply mismeasurement (Bils et al., 2021).

We seek to understand the causes of misallocation in Europe, accounting for a number of alternative factors simultaneously. Our work builds most closely on David and Venkateswaran (2019) and David et al. (2021), applying their dynamic misallocation framework to a set of European countries. By decomposing the sources of misallocation at the country and industry level, we quantify the role of adjustment costs, firm-level uncertainty, and structural distortions in productivity differences across Europe. This approach provides new insights into the mechanisms driving misallocation and its impact on economic performance in the region.

This paper makes two main contributions to the literature. Firstly, we use the method introduced in David and Venkateswaran (2019) to decompose the sources of capital misallocation at the country level in 19 European economies using a comprehensive dataset of the balance sheets of firms in the manufacturing sector. This method explicitly accounts for adjustment costs and firm-level uncertainty. Secondly, we extend the analysis over previous studies by decomposing the sources of capital misallocation at the the two-digit NACE industry level.[3]

We document several important findings: First, most of the observed dispersion in marginal revenue products at the country level stems from firm-specific factors, with a smaller but still important role for adjustment costs and transitory distortions. Across the 19 European countries, we find that the variance of permanent firm-specific distortions is nearly twice as large as that of transitory distortions across countries. In some countries, such as Germany and Austria, the

[3]Kilumelume et al. (2025) also study industry-level estimates in their study of the implications of tariff policies for South African firms. Morando and Newman (2021) apply the approach of David and Venkateswaran (2019) to European farms.

permanent component is more than three times larger. Second, we find that the permanent firm-specific factors explain a larger fraction of the observed variation in marginal revenue products of capital in southern Europe relative to northern and western Europe. This suggests that permanent distortions are important for explaining lacklustre TFP outcomes in Southern Europe across our sample period. Third, we document substantial variation in these estimates at the *industry* level, suggesting that the sources of misallocation differ according to industry-specific (as opposed to manufacturing sector-specific) characteristics. In some industries, revenue product dispersion is more than twice as large as as the country median. Permanent distortions can vary by a factor of almost three across industries within countries, and in some sectors, adjustment costs are more then twice the country median. Fourth, to show that industry-level reforms could lead to an increases in aggregate productivity we conduct an exercise in which we mechanically reduce the level of Spanish adjustment costs to German levels. We show that this would lead to non-negligible total factor productivity increases in Spain. Conducting this exercise for each 2-digit Spanish manufacturing industry reduces the contribution of adjustment costs to revenue product dispersion from roughly 15 percentage points to 12 percentage points. Finally, to investigate which features of the macroeconomic environment are correlated with these sources of misallocation, we match our country-industry misallocation estimates to many potential predictors using micro-aggregated cross-country panel data from CompNet. Using a penalised regression approach to account for the high level of dimensionality in our data, we find that industry-level characteristics relating to financial variables and productivity have an important role in explaining these permanent distortions.

The remainder of the paper is organised as follows: Section 2 describes our dataset and the steps taken to generate a representative sample of European firms. Section 3 documents the methodology used to disentangle the sources of misallocation. Section 4 reports the results at the country and country industry level. Section 5 documents our investigations into how sources of misallocation conditionally correlate with economic features of the European economies we study. Section 6 concludes.

2 Data and Descriptive Statistics

2.1 Firm-Level Data – ORBIS

We construct nationally representative firm-level data for each year from 2008 to 2018 for a number of European countries. Our data come from the ORBIS/AMADEUS database, compiled by Bureau Van Dijk. The dataset contains firm-level balance sheet and ownership data originally derived from business registers and other regulatory sources for approximately 14 million European firms. Firms are divided across industries according to a 4-digit industry classification (NACE Rev.2). There are significant benefits to using ORBIS/AMADEUS. Much of the literature following HK calculate misallocation measures using only a subset of firms in a country's manufacturing sector. For example, David and Venkateswaran (2019) use data from Compustat for the US case, which contains data on large publicly listed firms. By contrast, ORBIS/AMADEUS contains data on small and private firms, with listed firms representing only around 1% of the sample. The data have further advantages over data derived from firm censuses, insofar that information on output and employment is contained alongside that relating to balance sheets and profit and loss accounts. Importantly for our analysis, in most European countries these filings are required by legislation for most firms. As such, the dataset is a rich source of information on the balance sheets of small and medium sized private enterprises, as well as large publicly listed firms.

While ORBIS/AMADEUS is the foremost cross-country comparable source of data on firms in Europe, the data do have well documented limitations. We closely follow the comprehensive data collection and cleaning process developed in Kalemli-Özcan et al. (2015), which documents how best to generate serviceable micro-data for analysis from this source. We discuss our approach to cleaning in more detail in Appendix Section A.

The ORBIS data we use provide us with relatively good coverage of employment and value added in the countries to be analysed. Appendix Table A.1 shows the coverage of operating revenue and employment in the sectors we analyse in comparison to Eurostat's Structural Business Statistics (SBS). The SBS data represent census measures, i.e. the universe of firms in each country. With respect to revenue and employment, coverage is in line with other papers that have used the ORBIS dataset for Europe. The ratios presented in Table A.1 represents the coverage of the data after completing the comprehensive cleaning steps outlined in Kalemli-Özcan et al. (2015). In Appendix Tables A.2 and A.3, we show that our sample is broadly representative in terms of the contribution of small and medium sized firms that account for much of the manufacturing activity across Europe. This is a significant advantage of our data.

2.2 Industry-Level Data – CompNet

The second part of our analysis uses data from the CompNet research network. The CompNet dataset contains aggregated firm-level information. The data are collected using firm-level datasets from national statistical institutes and central banks in 19 countries of the European Union. The data contain harmonised information on firm characteristics across a range of six categories: competitiveness, productivity, labour, trade, finance and other variables. The data are aggregated, but importantly for our matching to the results of our estimations using ORBIS, the data can be split by country and NACE two-digit industry. In order to ensure representativeness and harmonisation across countries, the variables are weighted by firm population weights.

There are ten vintages of the CompNet dataset, with each subsequent vintage incrementally adding new variables and countries. There are also two versions of the dataset available; the "20E" sample, which includes only firms who have twenty or more employees, and the "full" sample. We make use of the full sample version of the 8th vintage of the data, which covers the years 1999 to 2019 and contains all of the variables required for our analysis.

As well as the aggregated statistics for each country and industry, CompNet reports a number of moments of given industry-level variables (e.g. mean, standard deviation and percentiles of the distribution). The dataset also provides joint distributions from the data, i.e. summary statistics of variables by percentiles of the distribution of other variables. For example, productivity by size-deciles. The dataset also contains "discrete conditional variables", i.e. calculations of variables for firms with certain characteristics. For example "zombie-firms", in which case the data file would contain all distributions of a summarised variable conditional on whether or not the firm was classified as a "zombie". Details on the classification of variables by certain firm characteristics are based on the literature and a comprehensive summary can be found in CompNet (2018, 2021).

3 Methodology

The environment is identical to that of David and Venkateswaran (2019), which is an extension of the HK framework, and explicitly accounts for dynamic considerations of the firms investment decision. The model allows for adjustment costs, firm-level uncertainty, as well as the components of the HK distortion τ, broken into permanent, transitory, and correlated firm-specific distortions. A full description of the model is provided in Appendix Section B. We discuss the key elements here.

3.1 The Model

The model assumes a discrete time, infinite horizon economy with a representative household. Production is carried out by a continuum of firms who produce intermediate goods using capital and labour. Firm i produces according to a Cobb-Douglas production function, $Y_{it} = K_{it}^{\hat{\alpha}_1} L_{it}^{\hat{\alpha}_2}$, where $\hat{\alpha}_1 + \hat{\alpha}_2 \leq 1$. Intermediate goods are bundled to produce a single final good using a standard constant elasticity of substitution (CES) aggregator. The final good is produced under perfect competition, frictionlessly, by a representative firm.

Intermediate goods firms are heterogeneous, and their productivity is modelled as a stochastic process. The functional form of their productivity draws is AR(1) in logs, with autocorrelation parameter ρ, and a time-varying idiosyncratic white noise shock with variance σ_μ^2. Adjustment frictions are modelled using a quadratic adjustment cost function, similar to Asker et al. (2014). Firm-level uncertainty is modelled using the framework of David et al. (2016) by assuming that the firm, at the time of making its investment choice for period t, observes a noisy signal of next period's productivity a_{t+1}. DV show that a sufficient statistic for this uncertainty is the posterior variance of the firm, denoted $\mathcal{V}$. Because productivity is assumed to follow an AR(1) process, $\mathcal{V}$ ranges between 0 and σ_μ^2. If $\mathcal{V} = 0$, the firm is perfectly informed about next period's productivity. If $\mathcal{V} = \sigma_\mu^2$, the firm only has information about current productivity but has no signal regarding future realisations.

To model the other factors influencing investment decisions, DV follow HK by assuming that they take the form of an implicit proportional tax on the cost of capital, referred to throughout the literature as a "wedge". The model yields the following log-linearised law of motion for capital:

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

where ξ is a composite parameter that summarises the severity of adjustment costs, α is the curvature parameter for operating profits, $\alpha \equiv \alpha_1/(1-\alpha_2)$, and β is the discount rate.[4] The productivity term a_{t+1} takes the expectation operator $\mathbb{E}$, reflecting the fact that the firm may have imperfect information about future productivity innovations. Here τ_{it+1} is a distortion or wedge that acts on the firm's investment decisions.

We assume that τ_{it} has three components: $\tau_{it} = \gamma a_{it} + \chi_i + \epsilon_{it}$, where γa_{it} represents distortions

[4] Here $\alpha_j = (1-\theta^{-1})\hat{\alpha}_j,\ j = 1, 2$.

that are correlated with productivity, χ_i represents permanent firm-specific distortions, and ϵ_{it} represents idiosyncratic firm-specific distortions and is i.i.d over time. Misallocation sources other than adjustment costs and uncertainty are therefore parameterised by γ, and the variance of the permanent and transitory components, respectively denoted by σ_χ^2 and σ_ϵ^2.

3.2 Identification – Disentangling the Misallocation Sources

Section 3.1 shows that five separate forces all contribute to dispersion in $arpk$: adjustment costs, uncertainty and the three components of the wedge (τ), parameterised by $\xi, V, \gamma, \sigma_\chi^2$ and σ_ϵ^2. In a detailed derivation for the special case where productivity follows a random walk, DV prove that these parameters are uniquely identified by the following moments of the data: the variance and serial correlation of investment, the correlation of investment with lagged changes in productivity, a_{it}, the correlation of $arpk$ with productivity, and the cross sectional dispersion in $arpk$.

The intuition for this result provided in DV rests on the idea that while each of these moments is a function of multiple factors, making any single one insufficient for identification, analysing the moments in pairs is sufficient to pin down each of the parameters. For example, disentangling adjustment costs from other idiosyncratic factors that the response of investment to productivity rests on the insight that while both forces depress the variability of investment, they have opposing effects on its autocorrelation. Convex adjustment costs create incentives to smooth investment over time and so increase its serial correlation. A distortion that reduces the responsiveness to productivity, will lower the serial correlation of investment by raising the relative weight of other more transitory factors.

DV also show in their numerical analysis that that this result holds when productivity is assumed to follow an AR(1) process with $\rho \in (0, 1)$. They estimate the model using simulated method of moments (McFadden, 1989). We conduct the same exercise for European firms in this paper, but extending the analysis to both countries and industries, this allows us to quantify the severity of various forces and their impact on $arpk$ dispersion at a more granular level. Owing to identification issues we experienced with γ, we make the assumption that $\gamma = 0$ and focus on the permanent and transitory distortions.

3.3 Parameterisaton

We calibrate certain parameters on the basis of typical values in the literature. The discount rate is set to $\beta = 0.95$. The annual depreciation rate is assumed to be $\delta = 0.1$. We keep the elasticity of substitution θ common across countries and industries and set its value to 6, following David and Venkateswaran (2019). We assume constant returns to scale in production, but allow the parameters $\hat{\alpha}_1$ and $\hat{\alpha}_2$ to vary across country and industry. For each country, we assume that α is 0.62.

Firm-level productivity is given by $a_{it} = va_{it} - \alpha k_{it}$ where va_{it} denotes the log of value-added. Controlling for industry-year fixed effects to isolate the firm-specific component, we use a standard autoregression to estimate the persistence of the of the productivity process, ρ, and the variance of the innovation, σ_μ^2. To estimate adjustment costs, ξ, the quality of information, $\mathcal{V}$, and the variance of the idiosyncratic distortions, σ_ϵ^2, we follow DV and target a set of of moments as outlined in the previous subsection. Specifically, we target the correlation of investment growth with the lagged innovations in productivity ($\rho_{\iota,\Delta a_{-1}}$), the autocorrelation of investment growth ($\rho_{\iota,\iota_{-1}}$), and the variance of investment growth (σ_ι^2). Finally, to infer σ_χ^2, the variance of the fixed component, we match the overall dispersion in the average product of capital, σ_{arpk}^2. By construction, our estimation with the parameters outlined above, will match the observed $arpk$ dispersion in the data. This allows for a decomposition of the contribution of each factor.

3.4 Target Moments

In Figure 1 we display our measures of the variance of (log) average revenue products to factors. These represent summary measures of misallocation in our dataset, and ultimately the key feature of the data we seek to explain. While our study focusses on the variance of the average revenue product of capital, displayed in panel (a), we also display the variance of the average revenue product of labour for reference (panel b). We subdivide the countries in our sample into four broad geographic regions.[5] Our findings on the level of $arpk$ dispersion are broadly close to those found in the literature, including Gorodnichenko et al. (2021) and those estimated using census data in CompNet. Looking across countries, $arpk$ dispersion is highest in Romania, Latvia, and Sweden, but is also higher than average in the Netherlands, Spain, Slovenia, and Slovakia. We observe a clear distinction in between the countries sometimes labelled as the "core" (West) and the "periphery" (South). Specifically, the level of $arpk$ dispersion is lower than average in France, Germany, and Austria. The same stylised fact is apparent for the variance of the average product of labour, as can be observed in panel (b).

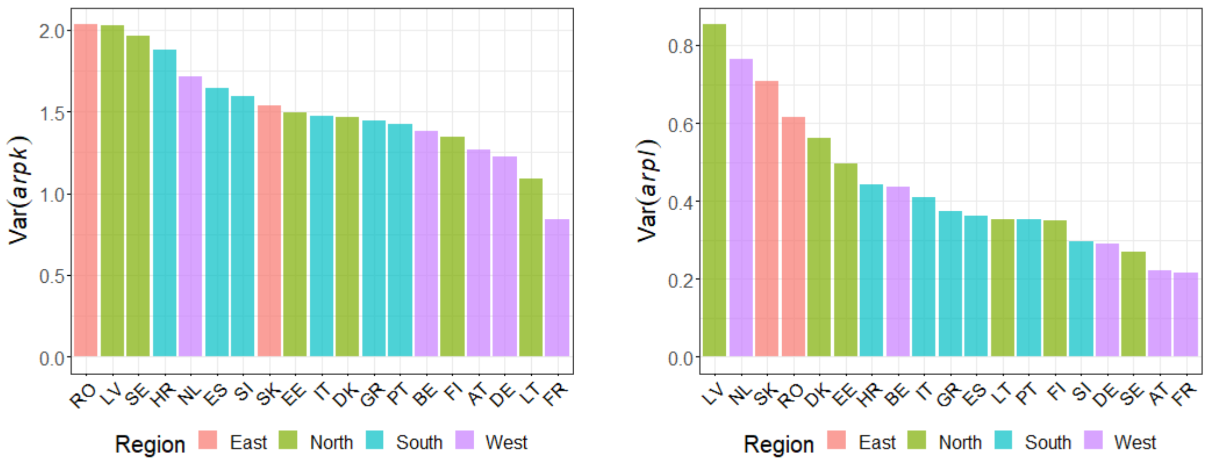
While we have established that the ORBIS/AMADEUS dataset broadly covers the manufacturing sector, there is still some scope for our misallocation measures to deviate from the underlying aggregates. To investigate this, we check whether our estimated misallocation measures correlate with country-level TFP. We find this indeed to be the case, as can be seen in Figure 2. We observe a negative correlation of -0.39 between the variance of $arpk$ and aggregate TFP of borderline statistical significance. The correlation patterns we see are reassuring, particularly as the TFP measure is based on national statistics, and therefore includes sectors outside of manufacturing. This suggests our measures can be generally useful for explaining broad patterns of cross-country performance.

Figure 3 displays the other target moments across the countries in our sample. Panel (a)

[5] North includes Latvia, Sweden, Estonia, Denmark, Finland, and Lithuania. West includes the Netherlands, Belgium, Austria, Germany, and France. South include Croatia, Spain, Slovenia, Italy, Greece, and Portugal. East includes Romania and Slovakia.

Figure 1: Average Revenue Product Dispersion by Country

(a) Var($arpk$) (b) Var($arpl$)

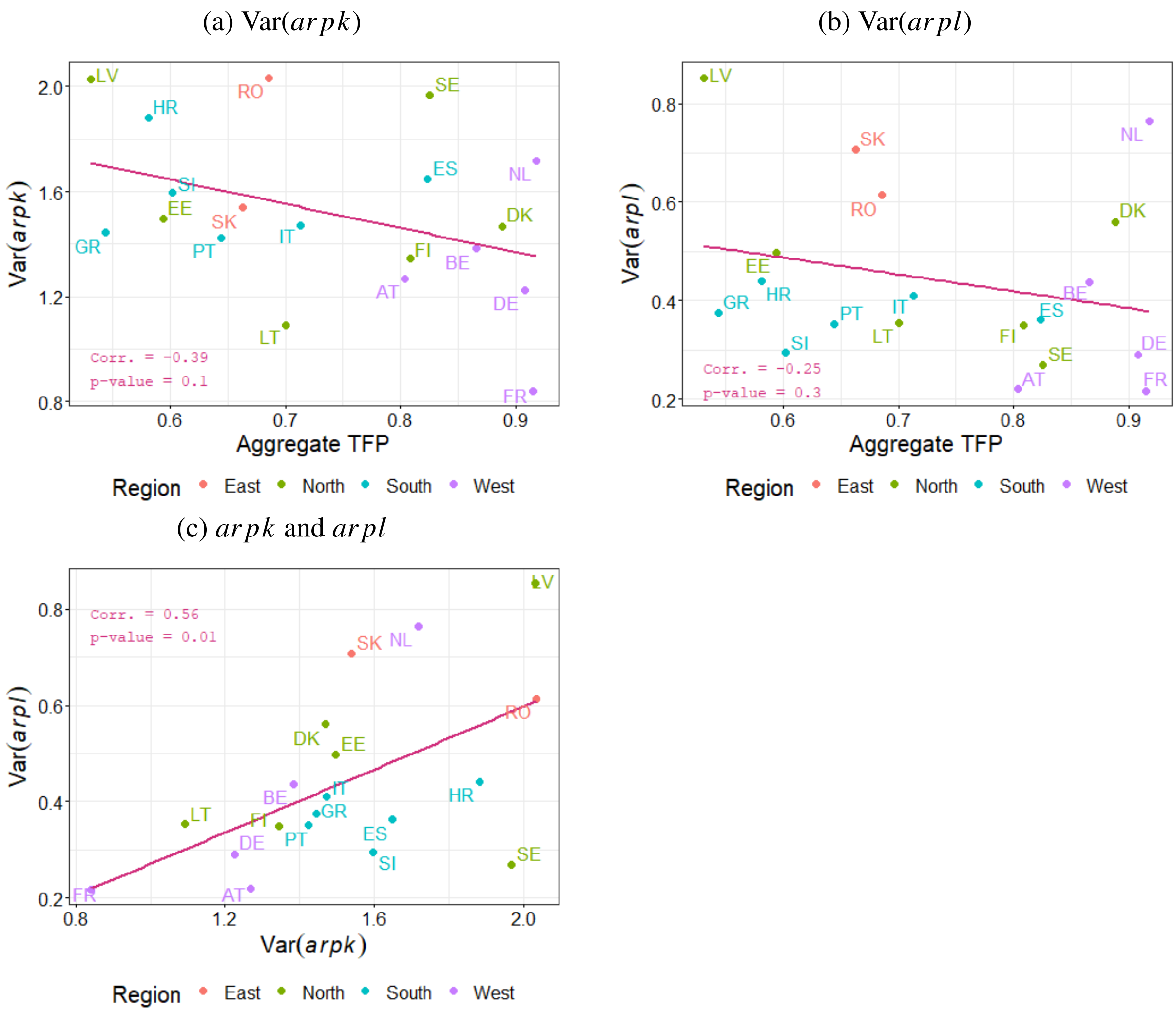


Notes: Panels (a) and (b) of the figure display the variance of (log) ARPK and (log) ARPL, respectively.

displays the estimated auto-correlation parameter for the (log) productivity process, which is broadly comparable across countries. Panel (b) demonstrates that the variance of productivity shocks is notably higher in the eastern European economies, as well as certain members of the Baltics. Investment growth rates show negative serial correlation and low variability across all countries, albeit with significant heterogeneity in the absolute levels of each (panels c. and d. respectively). For example, investment growth rates are significantly more variable in certain Northern economies, and the Baltics. The variance of investment growth is lowest in Western Europe, as well as a number of Southern European economies. Appendix Table C.1 reports the exact estimates for the target moments for each country, which have been displayed in Figures 1 and 3.

In the next section, we use these moments to shed light on the contributions to dispersion in $arpk$, our summary statistic for capital misallocation.

Figure 2: The Relation Between Aggregate TFP and Revenue Product Dispersion



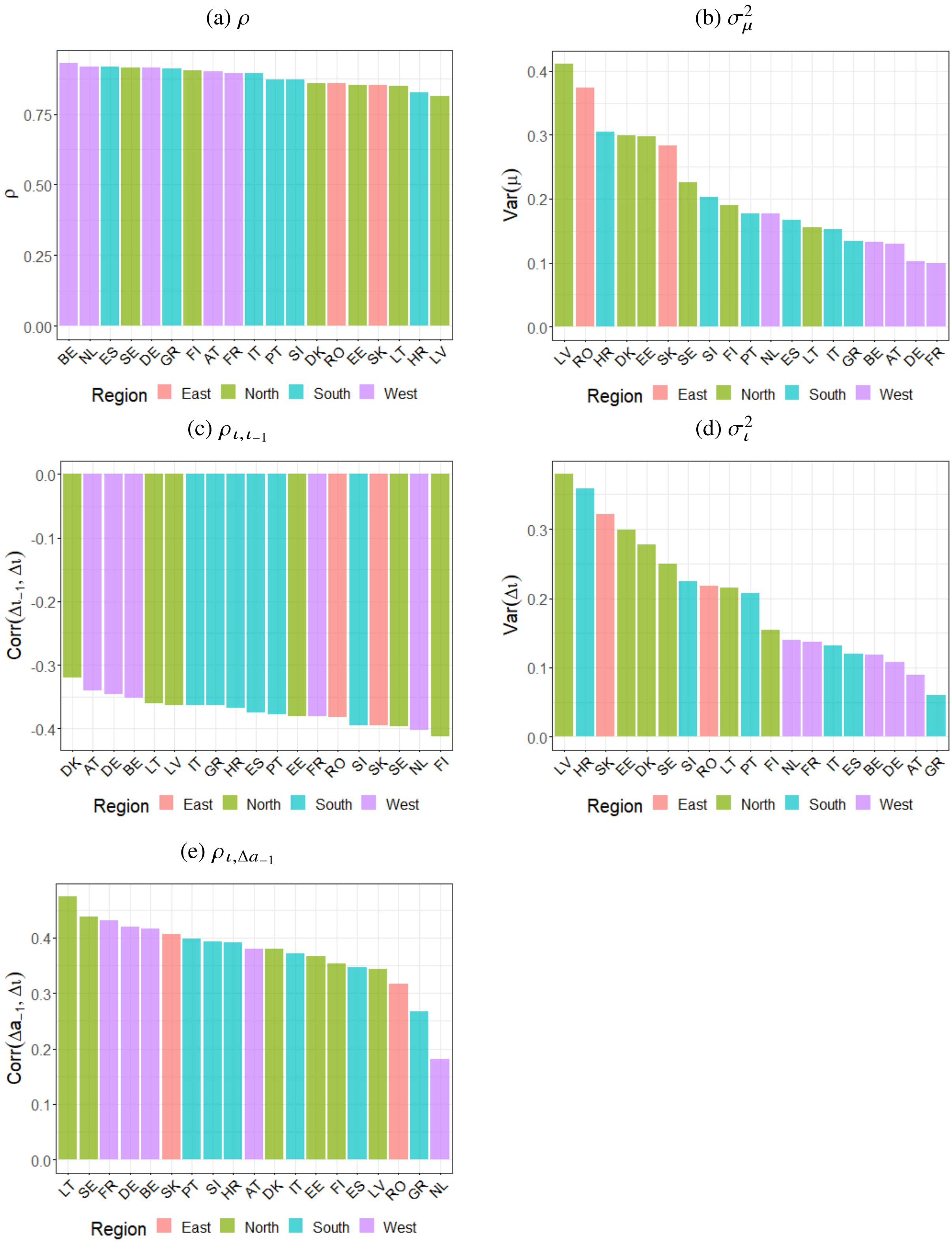
Notes: Panels (a) and (b) of the figure display the relationship between country-level TFP and the variance of (log) ARPK and (log) ARPL, respectively. Country-level TFP estimates are extracted from the Penn World Tables. Panel (c) displays the relationship between country-level (log) ARPK and (log) ARPL variance.

4 Results

4.1 Country-Level Estimates

Figure 4 displays estimated parameters, across the countries in our sample. These estimates are reported also in Appendix Table C.2 for reference. There are a number of interesting findings. Firstly, we can observe that the adjustment costs estimates do not obviously correlate by region. Each region appears to contain cases with both higher and lower adjustment cost estimates.

Figure 3: Moments by Country



Notes: Figure displays computed target moments for the countries in our sample.

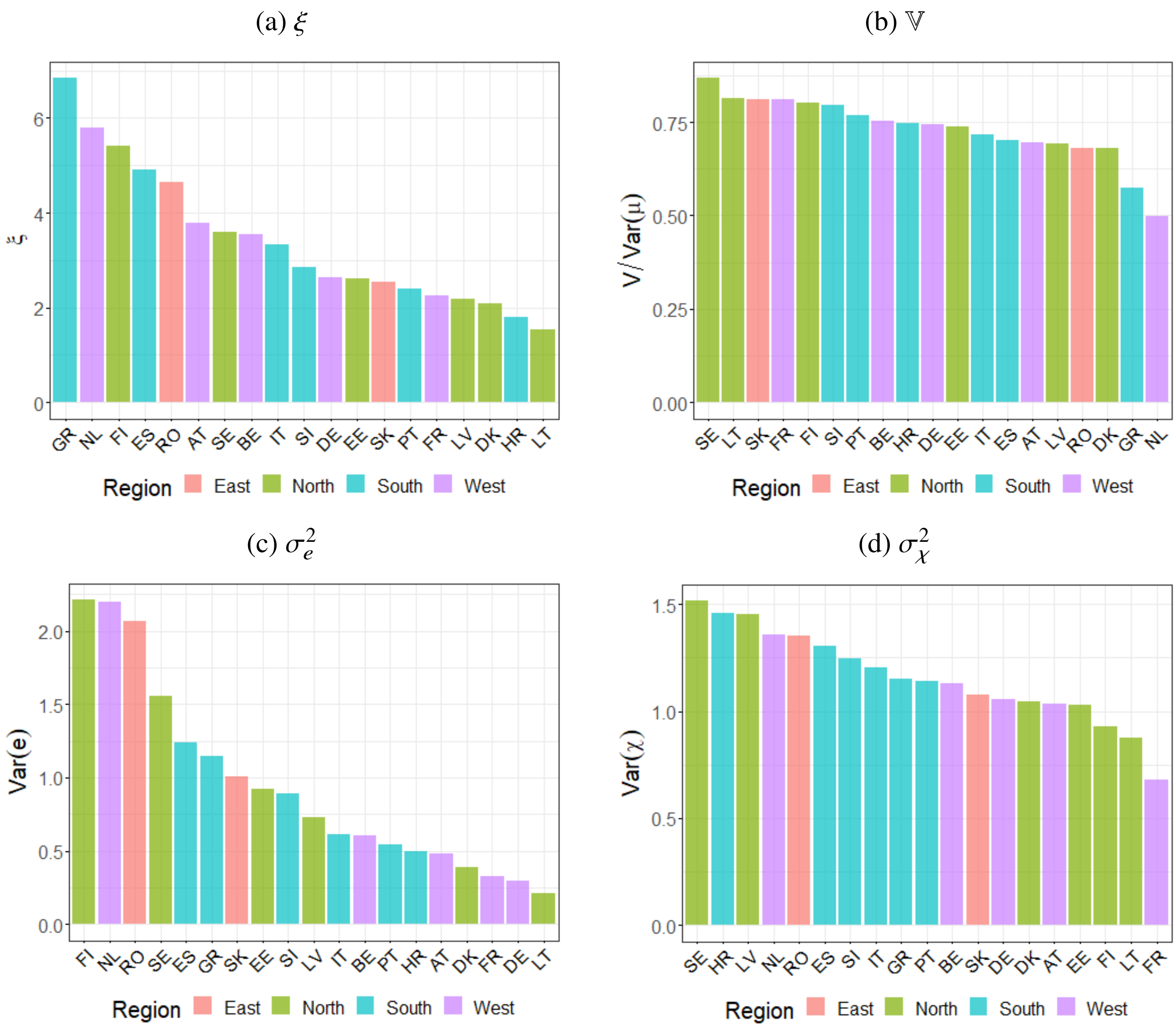
Within the four largest euro area economies, we can observe that Spain and Italy have larger adjustment costs according to our estimates, while Germany and France have lower values. We find that adjustment costs are highest in Greece, the Netherlands, Finland, and Spain. Each of these economies was at the lower end of estimates for the variance of investment (see panel d. of Figure 3). Adjustment costs act to reduce the variance of investment in the modelling framework we apply.

When we study levels of uncertainty across our sample, we again find little evidence for strong regional effects. For example, certain economies from Southern Europe have higher levels of uncertainty (Slovenia and Portugal), while Greece has one of the lowest levels of uncertainty. We estimate high levels of uncertainty in France, while our lowest estimate is from the Netherlands, suggesting that Western Europe also displays heterogeneous uncertainty levels. The low uncertainty estimates for Greece and the Netherlands result from the low values of the correlation of investment with lagged productivity, as can be observed in panel (c) of Figure 3.

Finally, turning to the two components of the distortion, τ, the estimates of the transitory and permanent components (panels c. and d. respectively) are reasonably heterogeneous across the countries in our sample. However, we do see greater evidence for variation by European region. With the exception of the Netherlands, estimates of the variance of idiosyncratic distortions tends to be lower in Western Europe. Southern European economies are congregated in the middle of the range of estimates, though Spain and Greece display estimates in the higher range. Turning to the permanent distortions, we also see greater evidence for a congregation of estimates by region. Western and Northern economies display lower estimates (with a couple of exceptions). Economies in the South of Europe show evidence for higher levels of fixed distortions.

Certain features of the estimates reported in Figure 4 are predictable given the nature of the target moments used to estimate the model. Firstly, some of the countries in our sample have quite a low variance of investment over the sample period. This suggests a role for adjustment costs. Quadratic adjustment costs tend to induce strong serial correlation in investment (Asker

Figure 4: Estimated Parameters by Country



Notes: Figure displays parameter estimates for the countries in our sample.

et al., 2014), and this is borne out by the countries with lower autocorrelation of investment having lower estimates for adjustment costs. Secondly, a high correlation between investment and lagged productivity suggests that firms do not respond immediately to shocks. This results in larger parameter estimates for the uncertainty component.

One general conclusion that emerges is that estimates for adjustment costs and uncertainty are heterogeneous within geographic regions, and that neighbouring economies with comparable industrial compositions can generate different estimates. For the case of transitory distortions, and in particular permanent distortions, we can see evidence for regional comparability of estimates. For example, Northern and Western Europe show lower levels for permanent distortions, relative to Southern Europe.

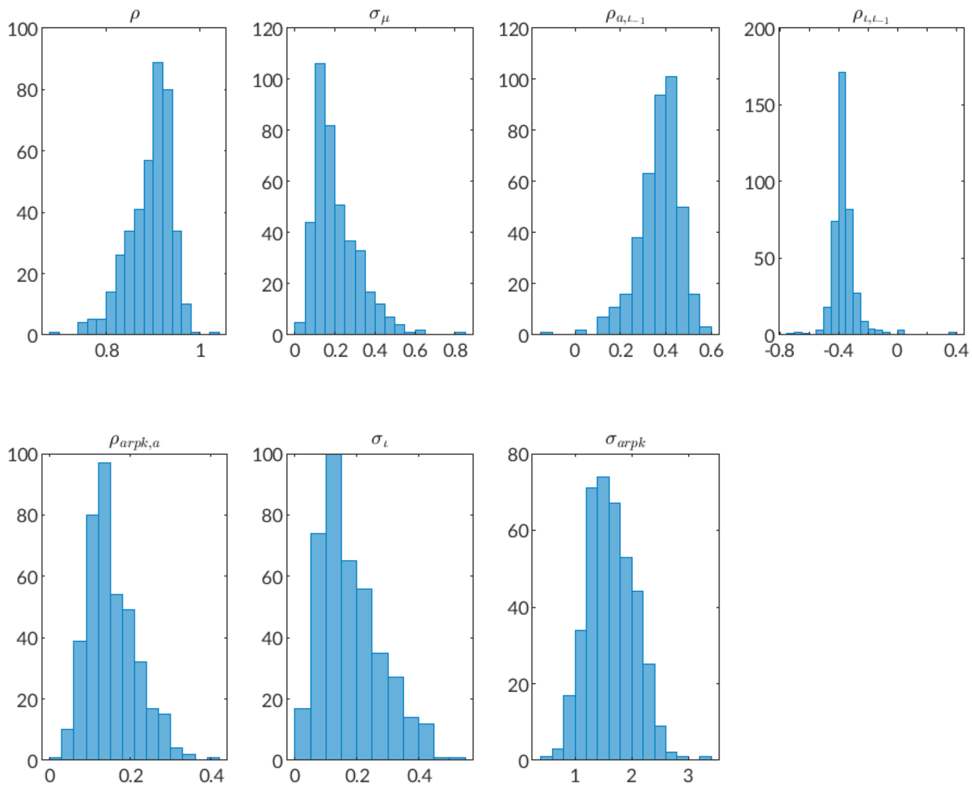
4.2 Industry-Level Estimates

While our country-level estimates are suggestive regarding the broad features giving rise to factor product dispersion across European economies, they treat the manufacturing sector in each case as an aggregate. An open question is the extent to which our estimates vary across the more narrow industries within the respective manufacturing sectors. We therefore estimate the factors contributing to $arpk$ dispersion at the industry level, by computing target moment for each NACE 2-digit industry within the manufacturing sector, before re-estimating the model. This avenue was not pursued in the study of DV. We document novel estimates of industry-level heterogeneity in misallocation source.

Figure 5 shows density plots of each of the target moments across countries and industries. The distribution of $arpk$ dispersion across countries and industries is broadly normally distributed, but it is clear that some industries display higher levels of misallocation than others. By applying the framework of DV across industries, we aim to shed light on the sources of such heterogeneity.

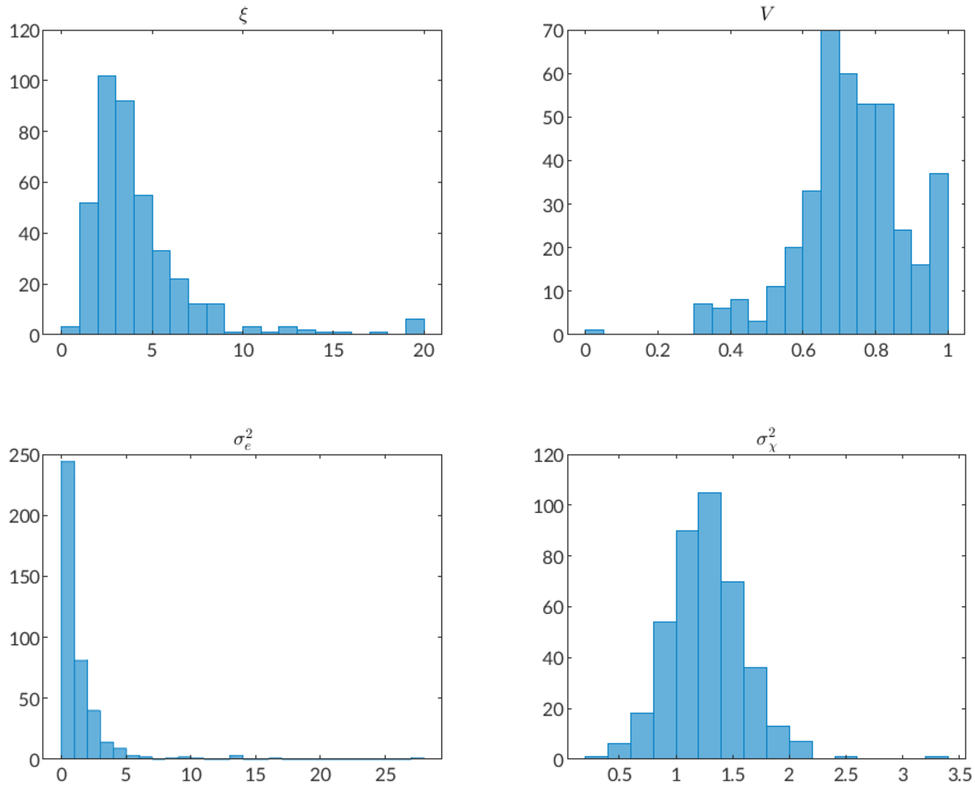
Figure 6 shows the distribution of each estimated parameter at the country industry level. Estimates of the adjustment costs parameter are positively skewed with a long right tail, indicating that the estimates for average adjustment costs are driven by a few country-industry pairs. The estimates for the uncertainty parameter at the country industry level are somewhat negatively skewed. Turning to the components of the distortion, τ, it is clear that for the majority of the industries examined idiosyncratic firm-specific distortions play a negligible role, with most of the mass of the distribution hovering around zero. The parameter for permanent distortions however is more symmetrically distributed at the country industry level.

Figure 5: Target Moments Across Industries and Countries



Notes: Figure shows histograms representing target moments computed from firm-level data from the manufacturing sector across 19 European economies. Moments are computed at the country-industry level.

Figure 6: Parameter Estimates Across Industries and Countries

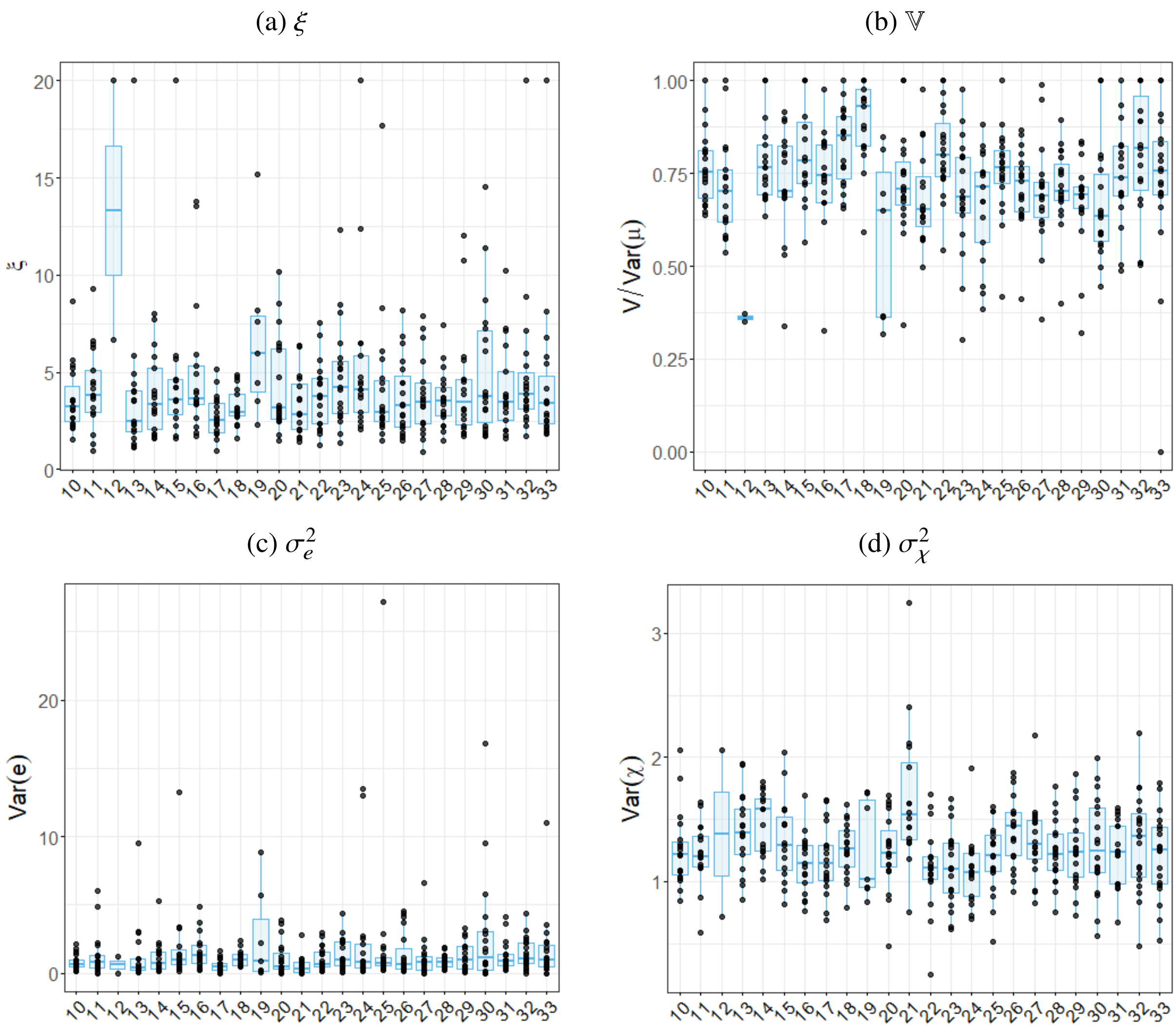


Notes: Figure shows histograms representing estimated parameters across individual country-industries.

Importantly, we find that much of the variation across country-industry pairs stems from differences between European economies, even within narrow 2-digit NACE categories. In Figure 6 we display, within each 2-digit NACE category, the dispersion in estimates across countries. There are many cases for which individual countries display estimates that are markedly removed from the median across cases. We view this result as helpful from a policy perspective, since it is unclear, for example, why adjustment costs in the manufacture of wood products are higher in Spain, relative to Germany.

To demonstrate how our approach could inform potential policy choices, we perform a simulation exercise. Specifically, we take estimates of adjustment costs from Spain and Germany. As can be seen from panel (a) of Figure 8, Spanish estimates are higher than German ones across every manufacturing industry, not just for the aggregate sector. We then compute the contribution to $arpk$ dispersion for the German and Spanish cases, as displayed in panel (b). We also compute

Figure 7: Country-Industry Estimates by Country



(a) ξ (b) $\mathbb{V}$ (c) σ_e^2 (d) σ_χ^2

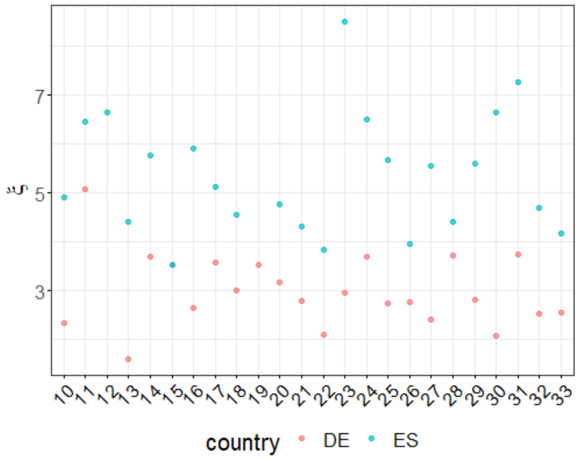
Notes: Figure displays the relationship between country-level parameter estimates and moments.

an estimated contribution for the case when Spanish adjustment costs are set to German levels, and find a sizeable reduction in misallocation costs for this case. Although adjustment costs are frequently treated as "efficient" distortions in the literature, in the sense that they are an irreducible part of any policy environment, our approach suggests that such determinants could potentially be used to reduce misallocation levels within Europe.

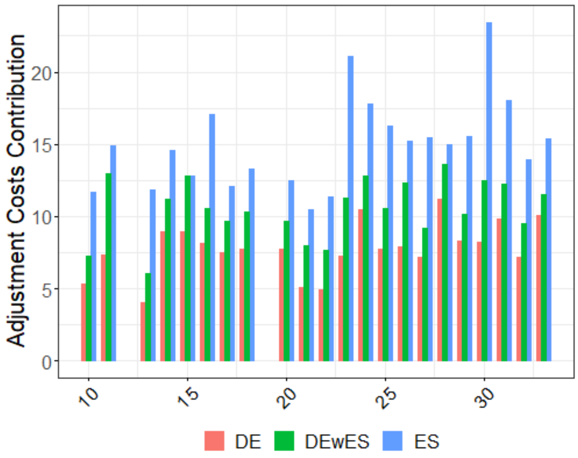
In order to understand this heterogeneity, in the next section, we investigate which factors at the country industry level are conditionally correlated with variation in these parameters.

Figure 8: Reduction of Spanish Adjustment Costs to German Levels

(a) DE vs. ES Adjustment Costs



(b) Adjustment Cost Contribution



Notes: Figure displays results from setting Spanish adjustment costs to German values, on a sub-industry by sub-industry basis.

5 Which Factors Predict Distortions?

At this stage our analysis points to the following conclusions: 1) there is meaningful dispersion in $arpk$ across a sample of European countries; 2) that "efficient" sources do not account for the whole picture, i.e. when explicitly accounting for adjustment costs and firm-level uncertainty, we still find meaningful contributions of the components of the wedge τ to overall dispersion in $arpk$ both across countries and across countries and industries.

But which macroeconomic features of these countries and industries in turn predict these results? To investigate this, we turn to the CompNet dataset. These data contain aggregated firm-level information at the country industry level, with a richer set of variables than is available from the ORBIS dataset. As discussed in Section 2.2, CompNet contains data sub-divided into six categories: competitiveness, productivity, labour, trade, finance, and a residual category termed "other". These data are presented as aggregate means at the country industry level, but the variance and higher order moments are also included.

This rich dataset allows us to investigate which factors lead to variation in the country-industry-level parameter estimates displayed in Figure 6. Note that we do not interpret these investigations as causal, and they are designed to chart conditional correlations of interest.

5.1 Estimation Strategy

Our estimation strategy is simple, we regress the parameter estimates for the components of misallocation (adjustment costs, uncertainty and the components of the wedge) for each country and industry on a set of potential predictors from CompNet.

The set of predictors are chosen based on the literature. We regress each of the parameters on country-industry measures of labour costs, investment ratios, firm demographic variables, a measure of firm financial constraints, and the wage share of industry i in country j. The results are presented in Table 1.

Table 1: OLS Regression Results for Country-Industry Misallocation Components

VARIABLES	(1) ξ	(2) $\mathbb{V}/\sigma_u^2$	(3) σ_e^2	(4) σ_χ^2
Growth rate (from t-1): Ratio: labour cost per employee	-2.573	1.548***	27.75**	-1.473
	(10.54)	(0.568)	(13.02)	(0.982)
Growth rate (from t-1): labor = number of employees	-48.53***	0.0224	-37.18**	2.137*
	(13.69)	(0.738)	(16.91)	(1.275)
Ratio: investment (change in nk + depr) as fraction of nk (t-1)	1.638**	0.0108	0.601	0.111
	(0.748)	(0.0403)	(0.924)	(0.0697)
Age of firm in years	-0.242*	0.0108	-0.165	0.0117
	(0.135)	(0.00729)	(0.167)	(0.0126)
D = 1, if firm is financially constrained	-4.623	-0.627	-7.920	1.855*
	(10.51)	(0.567)	(12.98)	(0.979)
Nominal labor costs	-2.63e-05	-2.07e-05**	-0.000229	1.05e-05
	(0.000178)	(9.61e-06)	(0.000220)	(1.66e-05)
Ratio: wageshare: nom. labor cost / nom. value-added	3.782	0.288	9.824*	-0.197
	(4.532)	(0.244)	(5.599)	(0.422)
Firm's labor market power: rev-based CD, sec, WD	0.0736	0.00700*	0.0582	-0.00884
	(0.0664)	(0.00358)	(0.0820)	(0.00619)
Ratio: Unit labor costs: nom. labor cost / real value-added	-4.898	0.147	-3.255	0.420
	(3.134)	(0.169)	(3.872)	(0.292)
Constant	8.909**	0.376**	-1.600	0.931***
	(3.517)	(0.190)	(4.345)	(0.328)
Industry FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Observations	155	155	155	155
R-squared	0.583	0.491	0.306	0.787

Notes: Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

In Column 1, we find preliminary evidence of a negative relationship between adjustment costs and firm demographics, particularly firm size and age. Column 2 shows that labour market power and the growth rate of labour costs help explain the variation in the uncertainty measure across countries and industries. In Column 3, industry-level wages appear to account for some of the variation in transitory distortions, though the negative coefficient on the growth rate of employees suggests that the effect is driven by wage costs per worker rather than total wage costs. Finally, Column 4 indicates a role for financial constraints and employment growth in explaining the variation in permanent distortions.

5.2 LASSO Estimation

In section 5.1 we choose a set of predictor variables based on theory. However, another approach would be to use the richness of the data in CompNet and search over variables that are useful for predicting distortions. However, CompNet contains many potential predictors (>300 when the estimations include the variance of each of the predictors) and our parameter estimates from ORBIS give comparatively fewer observations. To address this, we adopt the "elastic net" least absolute shrinkage and selection operator (LASSO) from Zou and Hastie (2005), and solve

$$\min_{\gamma} \sum_{c} \sum_{i} \left(y_{ci} - \gamma^T X_{ci} \right)^2 + \lambda \left[\alpha \left\| \gamma \right\|_1 + (1 - \alpha) \left\| \gamma \right\|_2^2 \right], \tag{2}$$

for some $\lambda > 0$ and $\alpha \in [0, 1]$. For country c and industry i, y_{ci} denotes a dependent variable of interest and X_{ci} is an $m \times 1$ vector of potential explanatory variables. In our context y_{ci} are estimated parameters at the country industry level, and X_{ci} represents country-industry information from CompNet. We set $\alpha = 0.99$. We estimate λ by cross-validation.

It is well known that the LASSO estimator can run into problems when estimating with highly correlated independent variables, since the algorithm can select parameters in a somewhat arbitrary manner (Taddy, 2017). In our CompNet dataset, we include both the mean and the standard deviation of all our potential explanatory variables. Not only this, the variables within each of the six categories are highly correlated in many cases.

To account for this, we summarise our the results from our LASSO estimation using a non-parametric bootstrap. We draw with replacement from our dataset 500 times, and store the estimated parameters. We compute the adjusted R^2 by re-estimating our prediction equation via OLS conditional on the subset of variables that were assigned non-zero coefficients by the LASSO algorithm.

Table 2 reports the bootstrapped median Adjusted R^2 values from post-LASSO OLS regressions of the country-industry parameter estimates for the factors affecting $arpk$ dispersion on the categories of variables selected by the LASSO estimation.

We have four separate elastic net regressions, for each of the four parameters summarising misallocation sources. For parsimony, we present the results in the following way. The values shown in Table 2 represent the percentage point deviations in the adjusted R^2 as a result of omitting the variables in that category. The more significant the deviation, the more explanatory power is lost by omitting those variables and hence the greater importance we can assign in terms of explaining the variation in the parameters contributing to $arpk$ dispersion. We use these results as a guide to investigate which variables from those categories have the highest selection probability in the baseline model (including all explanatory variables). We can then in turn report the coefficients on those variables and the relevant interpretation.[6].

With respect to the distortions, the main message is that the financial and productivity variables have the greatest explanatory power, though labour market variables have some explanatory power for the transitory distortions. For adjustment costs, productivity variables again generate the largest percentage point change in the adjusted R^2 when omitted from the model. For the case of uncertainty, each of the groups of independent variables play a role at explaining variation in our parameter estimates. However, we again find that factors relating to industry-level productivity play a role at explaining firm-level uncertainty.

[6]The full set of results is available on request

Table 2: LASSO Category Analysis

	Permanent	Transitory	Uncertainty	Adjustment Costs
Baseline R^2	*93.9*	*65.7*	*78.6*	*64.3*
Competitiveness	-0.4	-1.0	-1.1	-0.2
Financial	-9.4	-3.6	-2.4	-1.6
Labour	-0.3	-3.5	-3.3	-0.9
Productivity	-2.7	-5.8	-4.3	-5.7

Notes: Figures are expressed as the percentage point deviation from the baseline R^2.

6 Conclusion

A large literature has established that misallocation of factor inputs, as measured by dispersion in the marginal revenue product of those inputs, are an important factor behind differences in cross-country TFP. In Europe, the literature has shown that misallocation of capital flows has been one of the main sources of the dispersion in the marginal revenue product of capital, particularly in southern countries (Gopinath et al., 2017).

Our paper sheds important light on both the *sources* of capital misallocation in Europe and macroeconomic features at the country industry level that are correlated with those sources. We find that a large proportion of the observed dispersion stems from permanent firm-specific factors. This is particularly true in the periphery countries. We also find an important role for adjustment costs and uncertainty at the country industry level, and a more limited role for idiosyncratic firm-specific distortions. We additionally find that industry-level characteristics relating to financial variables and productivity have an important role in explaining permanent distortions.

The findings in this paper shed light on the factors influencing capital misallocation in Europe. By better understanding these relationships, policymakers can identify the impediments to productivity growth and output expansion. This is of particular importance in the euro area, where divergences in output growth have been substantial in recent decades. The findings have implications for designing policies aimed at enhancing productivity and promoting output growth in the region.

References

Adilbish, Oyun Erdene, Diego A. Cerdeiro, Romain A. Duval, Gee Hee Hong, Luca Mazzone, Lorenzo Rotunno, Hasan H. Toprak, and Maryam Vaziri (2025) "Europe's Productivity Weakness: Firm-Level Roots and Remedies," IMF Working Paper WP/25/40, International Monetary Fund.

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

Bergeaud, Antonin (2024) "The Past, Present and Future of European Productivity," July, Presented at the ECB Forum on Central Banking.

Bils, Mark, Peter J. Klenow, and Cian Ruane (2021) "Misallocation or Mismeasurement?," *Journal of Monetary Economics*, 124S, S39–S56.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2012) "Americans Do IT Better: US Multinationals and the Productivity Miracle," *American Economic Review*, 102 (1), 167–201.

Bun, Maurice J. G. and Jasper de Winter (2022) "Capital and Labor Misallocation in the Netherlands," *Journal of Productivity Analysis*, 57, 93–113.

Calligaris, Sara (2015) "Misallocation and Total Factor Productivity in Italy: Evidence from Firm Level Data," *LABOUR*, 29 (4), 367–393.

Calligaris, Sara, Massimo Del Gatto, Fadi Hassan, Gianmarco I. P. Ottaviano, and Fabiano Schivardi (2018) "The Productivity Puzzle and Misallocation: An Italian Perspective," *Economic Policy*, 33 (95), 635–684, 10.1093/epolic/eiy009.

CompNet (2018) "Assessing the Reliability of the CompNet Micro-Aggregated Dataset for Policy Analysis and Research: Coverage, Representativeness and Cross-EU Comparability,"Technical report, The Competitiveness Research Network (CompNet), Prepared by the Working Group on "Cross-country comparability" of the CompNet Dataset, chaired by Marc Melitz.

——— (2021) "User Guide for the 8$^{\text{th}}$ Vintage CompNet Dataset,"Technical report, CompNet: The Competitiveness Research Network.

David, Joel M., Hugo A. Hopenhayn, and Venky Venkateswaran (2016) "Information, Misallocation, and Aggregate Productivity," *The Quarterly Journal of Economics*, 131 (2), 943–1006.

David, Joel M. and Venky Venkateswaran (2019) "The Sources of Capital Misallocation," *American Economic Review*, 109 (7), 2531–2567.

David, Joel M, Venky Venkateswaran, Ana Paula Cusolito, and Tatiana Didier (2021) "Capital Allocation in Developing Countries," *The World Bank Economic Review*, 35 (4), 1102–1121.

Dias, Daniel A., Carlos Robalo Marques, and Christine Richmond (2016) "Misallocation and productivity in the lead up to the Eurozone crisis," *Journal of Macroeconomics*, 49, 46–70.

Draghi, Mario (2024) "The Future of European Competitiveness. Part A: A Competitiveness Strategy for Europe," working paper, European Commission.

ECB (2021) "Occasional Paper Series: Key Factors behind Productivity Trends in EU Countries,"Technical Report 268, European Central Bank, Revised December 2021.

European Commission (2025a) "Savings and Investments Union: A Strategy to Foster Citizens' Wealth and Economic Competitiveness in the EU," Communication from the Commission to the European Parliament, the European Council, the Council, the European Central Bank, the European Economic and Social Committee and the Committee of the Regions, March, COM(2025) 124 final.

——— (2025b) "A Competitiveness Compass for the EU," jan, Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions.

Fuest, Clemens, Daniel Gros, Pierre-Louis Mengel, Giovanni Presidente, and Jean Tirole (2024) "EU Innovation Policy: How to Escape the Middle Technology Trap,"Technical report, European Policy Analysis Group, Report.

Gamberoni, Elisa, Claire Giordano, and Paloma Lopez-Garcia (2016) "Capital and Labour (Mis)allocation in the Euro Area: Some Stylized Facts and Determinants,"Technical Report 1981, European Central Bank.

García-Santana, Manuel, Enrique Moral-Benito, Josep Pijoan-Mas, and Roberto Ramos (2020) "Growing Like Spain: 1995–2007," *International Economic Review*, 61 (1), 383–416.

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, Debora Revoltella, Jan Svejnar, and Christoph T. Weiss (2021) "Resource Misallocation in European Firms: The Role of Constraints, Firm Characteristics and Managerial Decisions," June, *mimeo*.

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

IMF (2024) *Slowdown in Global Medium-Term Growth: What Will It Take to Turn the Tide?*, Washington, DC: International Monetary Fund.

Kalemli-Özcan, Şebnem, Bent Sørensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yeşiltaş (2015) "How to Construct Nationally Representative Firm Level data from the ORBIS Global Database," Working Paper 21558, NBER.

Kilumelume, Michael, Bruno Morando, Carol Newman, and John Rand (2025) "Tariffs, Productivity and Resource Misallocation," *The World Bank Economic Review*, Forthcoming.

Lenzu, Simone and Francesco Manaresi (2018) "Sources and Implications of Resource Misallocation: New Evidence from Firm-Level Marginal Products and User Costs," November, *mimeo*.

McFadden, Daniel (1989) "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57 (5), 995–1026.

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

Morando, Bruno and Carol Newman (2021) "Capital Misallocation, Agricultural Subsidies and Productivity: A European Perspective,"Technical Report 0221, Trinity Economics Papers, Department of Economics.

Pellegrino, Bruno and Luigi Zingales (2017) "Diagnosing the Italian Disease," Working Paper 23964, National Bureau of Economic Research.

Peters, Michael (2020) "Heterogeneous Markups, Growth, and Endogenous Misallocation," *Econometrica*, 88 (5), 2037–2073.

Reis, Ricardo (2013) "The Portuguese Slump and Crash and the Euro Crisis," *Brookings Papers on Economic Activity*, Spring, 143–193.

Restuccia, Diego and Richard Rogerson (2008) "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, 11, 707–720.

Schivardi, Fabiano, Enrico Sette, and Guido Tabellini (2021) "Credit Misallocation During the European Financial Crisis," *The Economic Journal*, 132 (1), 391–423.

Schnabel, Isabel (2024) "From laggard to leader? Closing the euro area's technology gap," Speech, Inaugural lecture of the EMU Lab, European University Institute, Florence, 16 February, https://www.ecb.europa.eu/press/key/date/2024/html/ecb.sp240216~df6f8d9c31.en.html.

Solow, Robert M. (1957) "Technical Change and the Aggregate Production Function," *The Review of Economics and Statistics*, 39 (3), 312–320.

Taddy, Matt (2017) "One-Step Estimator Paths for Concave Regularization," *Journal of Computational and Graphical Statistics*, 26 (3), 525–536.

Teichgraeber, Andreas and John Van Reenen (2022) "A Policy Toolkit to Increase Research and Innovation in the European Union,"Technical Report 2022/02, European Commission, Directorate-General for Research and Innovation, 10.2777/553584.

Van Ark, Bart, Mary O'Mahony, 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.

Zou, Hui and Trevor Hastie (2005) "Regularization and Variable Selection via the Elastic Net," *Journal of the Royal Statistical Society B*, 67, 301–320, Part 2.

Appendix

A Cleaning the ORBIS Dataset

In this Appendix section we discuss our approach to cleaning the ORBIS data in more detail.

One of the primary issues affecting data coverage is that some versions of ORBIS drop data for firms who have not reported balance sheet information in the previous five years. This creates an artificial survivorship bias in the data. We accessed the dataset through Bureau Van Dijk's proprietary web platform. Fortunately this platform allows us to keep observations for firms who have not reported, but for whom Bureau Van Dijk have information from business registers that the firm is still active. These firms these enter our dataset as missing variables for that year. The platform also allows us to keep information on inactive companies who reported in previous years.

An additional issue faced by some researchers using ORBIS is that the web platform limits each data download to a maximum of one million cells (firms $\times$ variables). This is further complicated by the fact that the panel element of the raw dataset is "wide", i.e. each year is an additional variable. To avoid any missing observations, we downloaded our required set of variables for firms in each NACE two-digit industry by methodically downloading the data at NUTS3 administrative division. For example, downloading our dataset for firms in NACE industry 10 (manufacture of food products) located in a particular province in Italy, then moving on to food manufacturers in the next province, and so on until each NACE code is complete. This process ensures the best possible coverage and is similar to the process outlined in Kalemli-Özcan et al. (2015), who accessed the dataset using discs. In order to minimise challenges with the estimation of the production function, we restrict our analysis to the manufacturing industry (NACE industries ten to thirty-three).

Some further important cleaning steps are required. Firstly, to remove duplicates we keep only unconsolidated accounts. We also adopt the convention in the literature that the current year is assigned only if the account closing date is after the 31st of May and the previous year is assigned otherwise. Our next step is to construct the "TFP Sample" outlined in Kalemli-Özcan et al. (2015). This step involves keeping only firms with positive values for employment or the wage-bill, and positive values for tangible fixed assets, gross output and materials. We clean the data further by dropping firm-year observations that are missing information on each of total assets and operating revenues, sales and employment. We drop firms if total assets, sales or fixed assets are negative in any year. We also drop firms where employment in the firm is negative or greater than 2 million in any year.

Crucially for estimating the average revenue product of capital central to our analysis, we drop firm-year observations with missing, zero or negative values for materials, operating revenue, or total assets. We then check the internal consistency of the balance sheet data by comparing the sum of the variables belonging to some aggregate to their respective aggregate. This involves generating a set of ratios for each country in our analysis, as outlined in detail in Appendix A.2 of Gopinath et al. (2017). For example, we construct the sum of tangible fixed assets, intangible fixed assets, and other fixed assets as a ratio of total fixed assets. Following Gopinath et al. (2017), we then drop extreme values from the analysis by excluding observations that are below the 0.1 percentile or above the 99.9 percentile of the distribution of ratios.

We conduct some further checks to examine the quality of the data, again following the steps in Appendix A.2 of Gopinath et al. (2017). These checks mostly involve verifying the internal consistency of the data. For example ensuring that the implied values for firm age, liabilities, or the wage bill are not negative or missing. We also implement checks to ensure that those variables that are most important for our estimation are of good quality, for example we test whether the difference between total assets and total liabilities is equal to shareholders funds, and drop observations where this identity does not hold.

Table A.1: Coverage of the ORBIS Dataset

	AT	BE	DE	EE	ES	FI	FR	GR	LU	IE	IT	LT	LV	NL	SI	SK	PT
Operating Revenue																	
2010	36	75	24	46	65	15	67	32	48	19	57	1	32	18	79	39	79
2011	47	95	24	58	67	43	71	35	81	34	62	45	69	26	81	87	83
2012	50	98	24	59	70	42	72	35	71	37	62	46	68	28	83	89	84
2013	50	97	25	60	73	40	79	37	73	42	66	49	72	29	85	90	88
2014	55	98	25	62	75	45	83	37	72	50	70	55	77	29	86	90	89
2015	61	107	26	62	77	51	86	42	41	28	73	58	81	27	88	91	91
2016	63	103	27	64	79	48	83	50	38	51	75	62	83	30	90	94	92
2017	66	104	31	65	80	59	78	47	37	51	74	61	85	31	90	94	94
2018	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Employment																	
2010	0	51	16	46	56	14	38	.	32	22	37	1	27	9	67	32	66
2011	0	98	17	53	60	43	37	.	51	26	49	52	64	13	70	78	71
2012	1	72	18	54	62	44	33	.	50	29	52	48	67	14	73	77	76
2013	4	73	19	57	65	46	42	.	51	.	55	55	68	16	76	80	79
2014	30	75	17	59	66	50	52	.	51	32	59	58	74	17	73	83	80
2015	52	77	21	60	70	49	55	43	9	33	62	59	78	17	80	86	83
2016	52	78	23	62	70	61	59	44	2	50	64	63	80	17	82	85	85
2017	53	119	25	62	71	69	59	49	0	57	66	68	81	17	83	90	83
2018	53	79	24	63	69	57	54	40	0	51	68	71	81	16	86	88	84

Notes: Table displays coverage ratios for operating revenue and employment. We compare totals computed across our ORBIS/AMADEUS dataset, relative to totals from Eurostat Structural Business Statistics.

Table A.2: Comparison of the ORBIS Size-Distribution with Eurostat Data (2018) – Operating Revenue

Size Class	Dataset	AT	BE	DE	EE	ES	FI	FR	GR	LU	IE	IT	LT	LV	NL	SI	SK	PT
0–9	ORBIS	1	2	2	10	5	3	1	5	.	1	6	1	7	2	7	2	5
0–9	SBS	3	4	2	8	6	4	4	9	0	3	8	3	6	5	8	6	7
10–19	ORBIS	1	2	3	7	5	4	1	8	.	1	10	2	6	0	6	2	6
10–19	SBS	2	2	2	6	4	3	2	5	0	0	8	3	5	3	5	2	6
20–49	ORBIS	5	7	6	18	12	11	5	18	.	8	16	10	15	9	11	5	13
20–49	SBS	5	6	4	13	9	6	4	8	5	0	12	8	13	7	10	5	12
50–249	ORBIS	20	22	18	40	26	19	17	39	.	16	31	45	42	40	25	16	32
50–249	SBS	19	15	12	43	20	19	11	19	21	0	25	29	39	26	22	16	32
250+	ORBIS	73	67	72	25	51	63	76	31	.	74	36	42	30	50	51	75	44
250+	SBS	70	72	80	30	61	68	78	58	71	0	46	57	36	59	54	72	43

Table A.3: Comparison of the ORBIS Size-Distribution with Eurostat Data (2018) – Employment

Size Class	Dataset	AT	BE	DE	EE	ES	FI	FR	GR	LU	IE	IT	LT	LV	NL	SI	SK	PT
0–9	ORBIS	1	1	1	13	11	7	1	5	.	1	10	2	13	0	10	4	11
0–9	SBS	8	11	6	13	17	10	12	32	0	11	22	12	16	15	16	20	17
10–19	ORBIS	2	2	5	10	11	7	2	11	.	1	14	4	10	0	7	4	11
10–19	SBS	5	5	6	8	9	7	5	10	0	0	14	7	9	8	7	4	10
20–49	ORBIS	7	8	12	21	19	13	7	23	.	3	21	12	19	2	11	9	20
20–49	SBS	9	11	8	16	15	11	8	13	9	0	15	13	16	13	10	8	18
50–249	ORBIS	20	33	22	39	29	26	26	39	.	22	30	42	41	34	28	27	35
50–249	SBS	22	23	20	40	23	24	16	21	27	0	22	32	37	32	25	22	31
250+	ORBIS	70	56	61	17	31	48	63	22	.	73	26	40	18	64	44	55	24
250+	SBS	55	50	60	24	37	49	60	24	56	0	27	36	22	32	43	46	23

B The David and Venkateswaran (2019) Framework

We use the model of David and Venkateswaran (2019), which in turn is an extension of the model of monopolistic competition with heterogeneous firms used in Hsieh and Klenow (2009). We consider a discrete time, infinite horizon economy populated by a representative household. The household inelastically supplies a fixed quantity of labour N and has preferences over consumption of a final good. The household discounts time at a rate β.

Production in the economy is carried out by a continuum of firms of fixed measure one, indexed by i. The produce intermediate goods using capital and labour according to

$$Y_{it} = K_{it}^{\hat{\alpha}_1} N_{it}^{\hat{\alpha}_2}, \quad \hat{\alpha}_1 + \hat{\alpha}_2 \leq 1. \tag{3}$$

Intermediate goods are bundled to produce the single final good using a standard constant elasticity of substitution (CES) aggregator

$$Y_t = \left(\int \hat{A}_{it} Y^{\frac{\theta-1}{\theta}} di \right)^{\frac{\theta}{\theta-1}}, \tag{4}$$

where $\theta \in (1, \infty)$ is the elasticity of substitution between intermediate goods and $\hat{A}_{it}$ represents a firm specific idiosyncratic component in production/demand. This idiosyncratic productivity shock is the only source of risk in the economy.

The final good is produced under perfect competition, frictionlessly, by a representative firm. This yields a standard demand function for intermediate good i

$$Y_{it} = P_{it}^{-\theta} \hat{A}_{it}^{\theta} Y_t \implies P_{it} = \left(\frac{Y_{it}}{Y_t}\right)^{-\frac{1}{\theta}} \hat{A}_{it}, \tag{5}$$

where P_{it} denotes the relative price of good i in terms of the final good, which serves as the numeraire.

Revenues for firm i at time t are

$$P_{it} Y_{it} = Y_t^{\frac{1}{\theta}} \hat{A}_{it} K_{it}^{\alpha_1} N_{it}^{\alpha_2}, \tag{6}$$

where $\alpha_j = \left(1 - \frac{1}{\theta}\right) \hat{\alpha}_j, \quad j = 1, 2.$

As such, $\hat{A}_{it}$ can be interpreted as either a firm-specific shifter of quality/demand or productive efficiency.

Firms hire labour period by period under full information at a competitive wage W_{it}. At the end of each period, firms choose capital for the following period. Investment is subject to capital adjustment costs, given by

$$\Phi(K_{it+1}, K_{it}) = \frac{\hat{\xi}}{2} \left(\frac{K_{it+1}}{K_{it}} - (1 - \delta)\right)^2 K_{it}, \tag{7}$$

where $\hat{\xi}$ controls the severity of the adjustment costs and δ represents the rate of depreciation.

Following Hsieh-Klenow (2009), the model introduces other factors affecting investment decisions as firm-specific proportional "taxes" on the flow cost of capital. These taxes are denoted $T_{i,t+1}^{K}$. Using the terminology from Hsieh-Klenow, these are referred to as distortions and/or wedges in the analysis. The wedge affects the firms choice of capital in the next period

$K_{i,t+1}$.

The firms problem in a stationary equilibrium can be represented by

$$\begin{aligned}\mathcal{V}(K_{it}, \mathcal{I}_{it}) = \max_{N_{it}, K_{it+1}} \mathbb{E}_{it} \Big[& Y^{\frac{1}{\theta}} \hat{A}_{it} K_{it}^{\alpha_1} N_{it}^{\alpha_2} - W N_{it} \\ & - T_{it+1}^{K} K_{it+1} (1 - \beta(1 - \delta)) - \Phi(K_{it+1}, K_{it}) \Big] \\ & + \beta \mathbb{E}_{it} [\mathcal{V}(K_{it+1}, \mathcal{I}_{it+1})].\end{aligned} \tag{8}$$

where $\mathbb{E}_{it}[\cdot]$ denotes expectations conditional on $\mathcal{I}_{it}$, the firms information set a the time that it chooses investment for next period K_{it+1}. The wedge $T_{it+1}^{K} K_{it+1}$ affects both the capital decision and the capital-labour ratio.

After maximising over labour N_{it}, this becomes

$$\begin{aligned}\mathcal{V}(K_{it}, \mathcal{I}_{it}) = \max_{K_{it+1}} \mathbb{E}_{it} & \left[G A_{it} K_{it}^{\alpha} - T_{it+1}^{K} K_{it+1} (1 - \beta(1 - \delta)) - \Phi(K_{it+1}, K_{it}) \right] \\ & + \beta \mathbb{E}_{it} [\mathcal{V}(K_{it+1}, \mathcal{I}_{it+1})].\end{aligned} \tag{9}$$

where $\alpha \equiv \frac{\alpha_1}{1-\alpha_2}$ is the curvature of operating profits (value added less labour expenses) and $A_{it} \equiv \hat{A}_{it}^{\frac{1}{1-\alpha_2}}$ is the firm specific profitability of capital, which David and Venkateswaran (2019) call "productivity" and the term $G \equiv (1 - \alpha_2) \left(\frac{\alpha_2}{W}\right)^{\frac{\alpha_2}{1-\alpha_2}} Y^{\frac{1}{\theta} \frac{1}{1-\alpha_2}}$ is a constant that captures the effects of the aggregate variables.

The stationary equilibrium in the model is defined as a set of value and policy functions for the firm, $\mathcal{V}(K_{it}, \mathcal{I}_{it})$, $N_{it}(K_{it}, \mathcal{I}_{it})$ and $K_{it+1}(K_{it}, \mathcal{I}_{it})$; a wage W; and a joint distribution over $(K_{it}, \mathcal{I}_{it})$ such that taking wages as given and the law of motion for Investment $\mathcal{I}_{it}$. The value and policy functions solve the firm's optimisation problem, the labour market clears and the joint distribution remains constant over time.

The model is solved using perturbation methods. Following David and Venkateswaran (2019), the firms optimality condition and laws of motion are log-linearised around $A_{it} = \bar{A}$ (the unconditional average level of productivity) and setting $T_K = 1$ (i.e. the case with no distortions), the model yields a log-linearised Euler equation

$$k_{it+1}((1+\beta)\xi + 1 - \alpha) = \mathbb{E}_{it}[a_{it+1} + \tau_{it+1}] + \beta\xi\mathbb{E}_{it}[k_{it+2}] + \xi k_{it}, \tag{10}$$

where ξ and τ_{it+1} are re-scaled versions of the adjustment cost parameter and the distortion/wedge, respectively.

Productivity A_{it} is assumed to follow an AR(1) process in logs with normally distributed iid innovations, i.e.$a_{it} = \rho a_{it-1} + \mu_{it}, \quad \mu_{it} \sim \mathcal{N}(0, \sigma_\mu^2)$, where ρ is the persistence and σ_μ^2 is the variance of the innovations.

The distortion τ_{it} takes the form

$$\tau_{it} = \gamma a_{it} + \epsilon_{it} + \chi_i, \quad \epsilon_{it} \sim \mathcal{N}(0, \sigma_\epsilon^2), \quad \chi_i \sim \mathcal{N}(0, \sigma_\chi^2), \tag{11}$$

where the parameter γ controls the extent to which τ_{it} co-moves with productivity. If $\gamma < 0$ the distortion encourages investment by firms with higher productivity and discourages investment by firms with lower productivity. The opposite is true if $\gamma > 0$. The remaining components of τ_{it} are both uncorrelated with a_{it}. The term ξ_{it} is permanent while ϵ_{it} is i.i.d over time. Thus, the severity of factors is summarised by three parameters; γ, σ_ϵ^2 which is the volatility of the i.i.d shocks to τ_{it} and σ_χ^2, which is the cross-sectional variation in the fixed component.

The firms information set at the time of choosing next periods capital is denoted by $\mathcal{I}_{it}$. This includes the entire history of productivity realisations through to period t. i.e. $\{a_{it-s}\}_{s=0}^\infty$. Since this is assumed to be AR(1), this can be summarised by the most recent observation, a_{it}. The firm also observes a noisy signal of next periods productivity innovation

$$s_{it+1} = \mu_{it+1} + e_{it+1}, \quad e_{it+1} \sim \mathcal{N}(0, \sigma_e^2), \tag{12}$$

where e_{it+1} is an i.i.d, mean zero and normally distributed noise term. The firms also perfectly observe the uncorrelated transitory component of distortions at the time of choosing period t investment. Firms also observe the fixed component of the distortion, χ_i. They do not see the correlated component, but are aware of its structure, i.e. they know γ. As such, the firms information set is given by

$$\mathcal{I}_{it} = (a_{it}, s_{it+1}, \epsilon_{it+1}, \chi_i). \tag{13}$$

A direct application of Bayes' rule yields the conditional expectation of productivity

$$\begin{aligned} a_{it+1} \mid \mathcal{I}_{it} &\sim \mathcal{N}(\mathbb{E}_{it}[a_{it+1}], \mathbb{V}) \\ \mathbb{E}_{it}[a_{it+1}] &\sim \rho a_{it} + \frac{\mathbb{V}}{\sigma_e^2} s_{it+1}, \quad \mathbb{V} = \left(\frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_e^2}\right)^{-1}. \end{aligned} \tag{14}$$

In the absence of news, i.e. $\sigma_e^2 = \infty$ we have that $\mathbb{V} = \sigma_\mu^2$. This corresponds to a standard one period time to build assumption. Alternatively $\sigma_e^2 = 0$ implies that $\mathbb{V} = \sigma_\mu^2$ so the firm is perfectly informed about next periods productivity a_{it+1}.

The Euler equation in 8 can be solved forward to obtain the law of motion for capital

$$k_{it+1} = \psi_1 k_{it} + \psi_2 (1+\gamma)\mathbb{E}_{it}[a_{it+1}] + \psi_3 \epsilon_{it+1} + \psi_4 \chi_i, \tag{15}$$

where

$$\begin{aligned} \xi(\beta\psi_1^2 + 1) &= \psi_1((1+\beta)\xi + 1 - \alpha) \\ \psi_2 = \frac{\psi_1}{\xi(1-\beta\rho\psi_1)}, &\quad \psi_3 = \frac{\psi_1}{\xi}, \quad \psi_4 = \frac{1-\psi_1}{1-\alpha}. \end{aligned} \tag{16}$$

The coefficients ψ_1-ψ_4 depend only on production (and preference) parameters, including the adjustment cost, and are independent of assumptions about information and distortions.The coefficient ψ_1 is increasing and ψ_2-ψ_4 are decreasing in the severity of adjustment costs. if there is no adjustment costs, (i.e. $\xi = 0$), $\psi_1 = 0$ and $\psi_2 = \psi_3 = \psi_4 = \frac{1}{1-\alpha}$. On the other hand, if

$\xi \to \infty$, $\psi_1 \to 1$ and$\psi_2 - \psi_4$ vanish. As adjustment costs become large, the firm's choice of capital becomes more autocorrelated and less responsive to productivity and distortions.

Aggregate output can be expressed as

$$\log Y \equiv y = a + \hat{\alpha}_1 k + \hat{\alpha}_2 n, \tag{17}$$

where k and n denote the logs of the aggregate capital stock and labour inputs, respectively. Aggregate TFP, denoted by a, is given by

$$a = a^* - \frac{(\theta\hat{\alpha}_1 + \hat{\alpha}_2)\hat{\alpha}_1}{2}\sigma^2_{arpk}, \quad \frac{da}{d\sigma^2_{arpk}} = -\frac{(\theta\hat{\alpha}_1 + \hat{\alpha}_2)\hat{\alpha}_1}{2}, \tag{18}$$

where a^* is aggregate TFP if static capital products ($arpk_{it}$) are equalised across firms and σ^2_{arpk} is the cross sectional dispersion in (the log of) the static average product of capital ($arpk_{it} = p_{it}y_{it} - k_{it}$).

C Additional Tables

Table C.1: Estimated Moments from ORBIS Data (Differences)

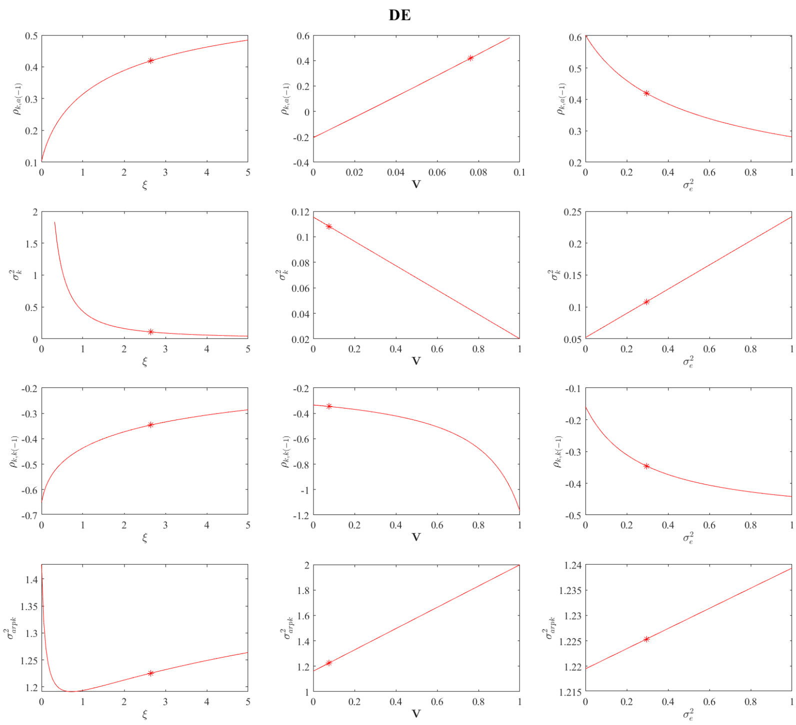
Country	α	ρ	σ_μ	$\rho_{a,\iota_{-1}}$	$\rho_{\iota,\iota_{-1}}$	$\rho_{MPK,a}$	σ_ι	σ_{APK}
AT	0.62	0.90	0.13	0.38	-0.34	0.71	0.09	1.27
BE	0.62	0.93	0.13	0.42	-0.35	0.74	0.12	1.38
DE	0.62	0.91	0.10	0.42	-0.35	0.75	0.11	1.23
EE	0.62	0.85	0.30	0.37	-0.38	0.75	0.30	1.50
ES	0.62	0.92	0.17	0.35	-0.37	0.76	0.12	1.65
FI	0.62	0.90	0.19	0.35	-0.41	0.72	0.15	1.34
FR	0.62	0.90	0.10	0.43	-0.38	0.70	0.14	0.84
GR	0.62	0.91	0.13	0.27	-0.36	0.82	0.06	1.44
HR	0.62	0.83	0.30	0.39	-0.37	0.76	0.36	1.88
IT	0.62	0.89	0.15	0.37	-0.36	0.81	0.13	1.47
LT	0.62	0.85	0.16	0.47	-0.36	0.70	0.22	1.09
LV	0.62	0.81	0.41	0.34	-0.36	0.78	0.38	2.03
NL	0.62	0.92	0.18	0.18	-0.40	0.79	0.14	1.72
PT	0.62	0.87	0.18	0.40	-0.38	0.75	0.21	1.42
RO	0.62	0.86	0.37	0.32	-0.38	0.77	0.22	2.03
SE	0.62	0.92	0.23	0.44	-0.40	0.74	0.25	1.97
SI	0.62	0.87	0.20	0.39	-0.39	0.73	0.22	1.60
SK	0.62	0.85	0.28	0.41	-0.40	0.75	0.32	1.54

Table C.2: Parameter Estimates

Country	ξ	V	σ_e^2	σ_χ^2	log Error
AT	3.789	0.696	0.478	1.033	-27.740
BE	3.558	0.753	0.603	1.133	-27.305
DE	2.643	0.746	0.296	1.054	-28.174
EE	2.609	0.739	0.919	1.030	-25.892
ES	4.905	0.703	1.236	1.307	-29.434
FI	5.404	0.800	2.212	0.931	-28.436
FR	2.250	0.810	0.325	0.678	-26.326
GR	6.842	0.575	1.146	1.154	-29.640
HR	1.801	0.747	0.499	1.459	-25.195
IT	3.337	0.716	0.613	1.203	-28.984
LT	1.538	0.815	0.214	0.874	-26.479
LV	2.172	0.693	0.731	1.455	-24.866
NL	5.785	0.498	2.196	1.355	-30.012
PT	2.393	0.770	0.541	1.140	-26.334
RO	4.655	0.681	2.068	1.352	-27.656
SE	3.596	0.868	1.557	1.515	-27.637
SI	2.859	0.796	0.889	1.249	-26.630
SK	2.533	0.812	1.005	1.078	-25.914

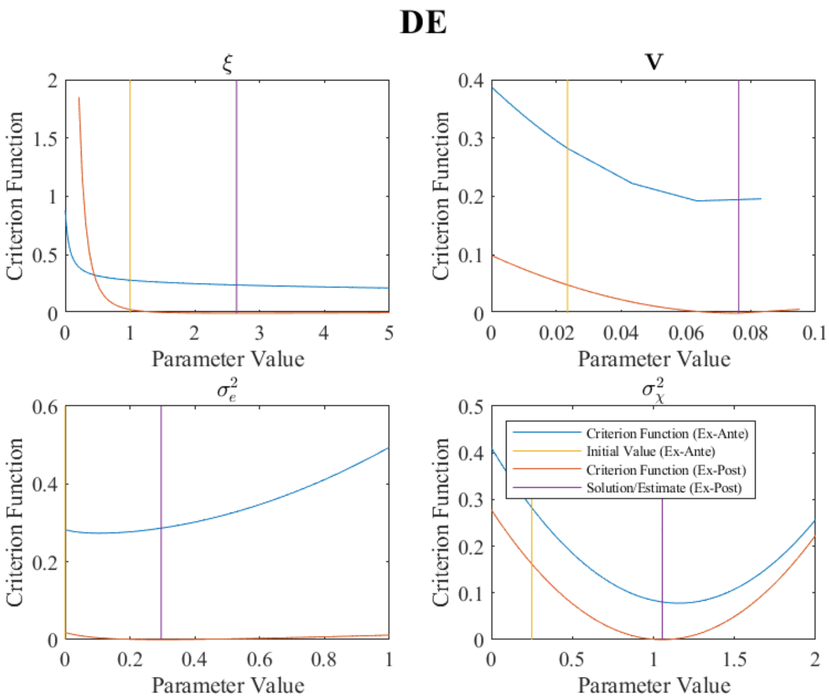
D Additional Figures

Figure D.1: Simulated Moments Across The Parameter Space



Notes: Figure shows predicted moments as individual parameters are varied.

Figure D.2: Parameter Estimates Across Industries and Countries



Notes: Figure criterion function around SMM estimates. The criterion function around the initial values is also displayed.

T: +353 (0)1 224 6000
www.centralbank.ie
publications@centralbank.ie

Bosca PO 559, Baile Átha Cliath 1, Éire
PO Box 559, Dublin 1, Ireland