A tale of two sectors: why is misallocation higher in services than in manufacturing?*

Daniel A. Dias[†], Carlos Robalo Marques[‡] and Christine Richmond[§]

June 17, 2016

Abstract

Recent empirical studies documented that the level of resource misallocation in the service sector is significantly higher than in the manufacturing sector. In this paper, we quantify the importance of this difference and study its sources. Conservative estimates for Portugal for 2008 show that closing this gap, by reducing misallocation in the service sector to the manufacturing levels, would boost aggregate gross output by around 12 percent and aggregate value added by around 31 percent. Differences in the effect and size of productivity shocks explain most of the gap in misallocation between manufacturing and services, while the remainder is explained by differences in the firm productivity and age distributions. We interpret these results as stemming mainly from higher output-price rigidity, higher labor adjustment costs and higher informality in the service sector.

JEL classification: D24, O11, O41, O47

Key words: Misallocation, productivity, firm-level data, structural transformation, Gelbach decomposition.

1   Introduction

Recent empirical studies documented that the level of resource misallocation in the service sector is significantly higher than in the manufacturing sector. In this paper, we quantify the implications for aggregate productivity and aggregate GDP of the "excess misallocation" in the service sector, and investigate to what extent this misallocation gap reflects structural differences between the two sectors.

A now well accepted result in the growth literature is that differences in the degree of allocative efficiency are one of the reasons why countries differ in terms of aggregate total factor productivity (TFP). Most of the empirical studies linking misallocation of resources to differences in TFP have been based on data from the manufacturing or agriculture sectors (see, for instance, Hsieh and Klenow (2009), Camacho and Conover (2010), Machicado and Birbuet (2012), Ziebarth (2013), Bellone and Mallen-Pisano (2013), Adamopoulos and Restuccia (2014), Chen and Irarrazabal (2015), Gopinath et al. (2015), Calligaris (2015)). Despite services being the largest sector for most countries, either in terms of value added or in terms of total employment, it was only recently that estimates of misallocation for the service sector were made available (see for instance Dias et al. (2016), Garcia-Santana et al. (2015), and Benkovskis (2015) for Portugal, Spain, and Latvia, respectively).

One result that is new and common to these economy-wide studies is that the level of estimated efficiency gains in the service sector is significantly higher than in the manufacturing sector.[1] Estimates obtained in Dias et al. (2016) for Portugal, for the 2004-2011 period, show that resource misallocation is, on average, 24 percentage points (p.p) higher in services than in manufacturing, when evaluated in terms of gross output, or around 40 p.p. higher, when evaluated in terms of value added (see Table 4 in Dias et al. (2016)). In turn, estimates in Garcia-Santana et al. (2015) for Spain, for the 2001-2007 period, suggest that efficiency gains are around 22 p.p. higher in services

---

[1]Although not strictly correct, but because there is a one-to-one mapping between efficiency (or TFP) gains from reallocation of resources and misallocation, we refer to both interchangeably throughout the paper.

than in manufacturing (see Table 2 in Garcia-Santana et al. (2015)), while estimates in Benkovskis (2015) for Latvia, for the 2007-2013 period, allow us to compute an average misallocation gap of around 32 p.p. between the two sectors. Despite not strictly comparable (methodologies and sectoral definitions vary across papers), these numbers show that the level of resource misallocation in the service sector is significantly higher than in the manufacturing sector.

In order to address the empirical questions of this paper, concerning the sources of the "excess misallocation" in the service sector and its aggregate implications, we resort to the theoretical framework developed in Hsieh and Klenow (2009), with the three-factor extension presented in Dias et al. (2016).

Using firm-level data for the Portuguese economy, we show that the differences of misallocation between the manufacturing and service sectors, have important implications for aggregate TFP and aggregate GDP. We estimate that if the misallocation gap between services and manufacturing were closed (by making the level of misallocation in services be the same as in manufacturing), aggregate gross output (or aggregate TFP) would increase by about 12 percent, while aggregate value added (GDP) would increase by around 31 p.p..

We also document that the significantly higher level of resource misallocation in the service sector is not the result of a small number of industries with abnormal levels of misallocation, but the outcome of a strong regularity: the great majority of the industries belonging to the manufacturing sector rank among the industries with the lowest misallocation.

Based on regression analysis, we show that the higher levels of allocative inefficiency in the service sector can be fully explained by structural differences between the two sectors. Idiosyncratic productivity shocks, which impact allocative efficiency in the presence of (capital/labor) adjustment costs and/or output-price rigidity, is the most important factor contributing to the difference of misallocation between the two sectors. However, the contribution of productivity shocks stems more from the difference of

the impacts, than from the difference in the magnitude of the shocks between the two sectors. In particular, the impact of productivity shocks in the service sector is significantly higher than in the manufacturing sector. Further analysis of the different distortions suggests that higher output-price rigidity and higher labor adjustment costs may explain why the effect of productivity shocks on misallocation is larger in the service than in the manufacturing sector.

The sectoral structure, proxied by the skewness of the productivity distribution, emerges as the second most important factor to explain the difference of misallocation between the two sectors. Again, the bulk of the contribution of this regressor stems from its higher impact in the service sector. A higher proportion of low productivity firms in the service sector makes the productivity distribution more right-skewed, contributing to a higher level of misallocation in the service sector stemming from size dependent distortions. At the same time, the service sector seems to have a higher level of informality which may explain why the sectoral structure has more impact on the level of misallocation in the service sector.

Finally, our empirical model suggests that the proportion of young firms also has a bearing on the difference of misallocation between the two sectors. Young firms emerge as facing higher capital costs than older firms, so that we link this regressor to the presence of credit constraints imposed by financial institutions on young firms, due to a lack of credit history or because of insufficient guarantees. This regressor contributes with two opposite effects to the difference of misallocation between the two sectors. The higher proportion of young firms in the service sector contributes to increasing the difference in misallocation between the two sectors, but the lower impact in the service sector has the opposite effect, so that the total contribution of this regressor is negative. This means that in the absence of this effect the difference in misallocation between the service and the manufacturing sectors would be even higher.

Our findings have important consequences for developing countries and economies undergoing structural transformation. Duarte and Restuccia (2010) demonstrate that

differences in productivity in the service and agriculture sectors across countries are one of the main factors behind productivity differences between countries and that, in particular, low productivity in the service sector and lack of catch-up will explain the experiences of slowdown, stagnation, and decline observed across economies. Hsieh and Klenow (2009) show that differences in misallocation in the manufacturing sector are important to understanding the differences of total factor productivity between developed and developing countries. Using data for the manufacturing sector in China and India, the authors conclude that reducing the level of misallocation in these economies to the levels observed in the U.S. economy would increase productivity by 30-50 percent in China and 40-60 percent in India. However, if a significant difference of allocative efficiency between manufacturing and the service sector, similar to that documented for Portugal, Spain or Latvia, are present in other countries, the importance of resource misallocation to explaining productivity differences between developed and developing countries may even be higher than what the empirical evidence based on data from the manufacturing sector alone would suggest.

By shedding light on the reasons behind the higher level of misallocation in the service sector relative to the manufacturing sector this paper also contributes to the understanding of the policies that may contribute to increasing productivity growth. Boosting competition so as to reduce output price rigidity in the service sector, avoiding size-contingent laws that may contribute to the survival of unproductive firms and reducing barriers to growth by eliminating credit constraints imposed by financial institutions on young firms, are measures that will contribute to reduce within-industry misallocation, especially in the service sector, and thus to increase aggregate TFP and aggregate value added (GDP).

The rest of the paper is structured as follows. Section 2 provides a brief description of the theoretical framework. Section 3 describes the dataset used in the analysis. Section 4 computes misallocation in the manufacturing and service sectors and discusses the aggregate implications of excess misallocation in the service sector. Section 5

reviews the main potential sources of misallocation suggested in the literature, presents the empirical results and discusses their interpretation. Finally, section 6 summarizes the main findings.

2 Theoretical framework

This section summarizes the methodology used to identify the linkage between aggregate productivity and resource misallocation that results from the existence of distortions and frictions affecting the optimal allocation of factors of production at the firm-level. We adopt the framework developed in Hsieh and Klenow (2009, 2011), but extend their model to consider a production function with intermediate inputs, as a third factor of production. The model with three factors of production, as well as the derivation of the full set of results was presented elsewhere (see, Dias et al. (2016)), so that here we just briefly review the model and summarize the main results needed for the purposes of the present paper.

The model assumes an economy with a single final good Y produced by a representative firm in a perfectly competitive market. This firm combines the output Y_s of S industries in the economy using a Cobb-Douglas production technology:

$$Y = \prod_{s=1}^{S} (Y_s)^{\theta_s} \tag{1}$$

with $\sum_{s=1}^{S} \theta_s = 1$ and $\theta_s = (P_s Y_s)/(PY)$, where P_s is the price of industry gross output, Y_s, and P is the price of the final good. At the industry level, gross output Y_s is a CES aggregate of M_s differentiated products:

$$Y_s = \left[\sum_{i=1}^{M_s} (Y_{si})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{2}$$

where Y_{si} stands for the gross output of firm i and parameter σ measures the elasticity of substitution between varieties of differentiated goods. At the firm level, the gross

output for each differentiated product is given by a Cobb-Douglas production function:

$$Y_{si} = A_{si} K_{si}^{\alpha_s} H_{si}^{\beta_s} Q_{si}^{1-\alpha_s-\beta_s} \tag{3}$$

where A_{si}, K_{si}, H_{si} and Q_{si}, stand for firm $i's$ total factor productivity (TFP), capital stock, labor and intermediate inputs, respectively.

With three factors of production it is possible to separately identify distortions that affect capital, labor and intermediate input prices simultaneously from distortions that affect the marginal product of one of the factors relative to the others. Thus, we introduce three types of distortions, or wedges, in the model: an output distortion, denoted $\tau_{y_{si}}$, a capital distortion, $\tau_{k_{si}}$, and a labor distortion, $\tau_{h_{si}}$, that take the form of a tax (or subsidy) on revenues, on capital services and on labor costs, respectively. Given these assumptions, profits are given by:

$$\pi_{si} = (1 - \tau_{y_{si}}) P_{si} Y_{si} - (1 + \tau_{k_{si}}) R_s K_{si} - (1 + \tau_{h_{si}}) W_s H_{si} - Z_s Q_{si} \tag{4}$$

where R_s, W_s and Z_s stand for the user cost of capital, labor wage and intermediate inputs price, respectively.[2]

Profit maximization yields the standard conditions that the firm's output price is a fixed markup over marginal cost:

$$P_{si} = \frac{\sigma}{\sigma - 1} \Psi_s \frac{(1 + \tau_{k_{si}})^{\alpha_s} (1 + \tau_{h_{si}})^{\beta_s}}{A_{si}(1 - \tau_{y_{si}})} \tag{5}$$

where

$$\Psi_s = \left[\left(\frac{R_s}{\alpha_s}\right)^{\alpha_s} \left(\frac{W_s}{\beta_s}\right)^{\beta_s} \left(\frac{Z_s}{1 - \alpha_s - \beta_s}\right)^{1-\alpha_s-\beta_s} \right] \tag{6}$$

---

[2]Equation (4) expresses the distortions in terms of output, capital and labor relative to the intermediate inputs distortion. Thus, in the model, an intermediate input distortion will show up as a higher output distortion and as lower capital and labor market distortions. An observationally equivalent characterization would be in terms of distortions to the absolute levels of capital, labor and intermediate input prices (and no output distortion).

In turn, from the first order conditions for profit maximization, we get:

$$(1 + \tau_{k_{si}}) = \frac{\alpha_s}{(1 - \alpha_s - \beta_s)} \frac{Z_s Q_{si}}{R_s K_{si}}$$

$$(1 + \tau_{h_{si}}) = \frac{\beta_s}{(1 - \alpha_s - \beta_s)} \frac{Z_s Q_{si}}{W_s H_{si}} \qquad (7)$$

$$(1 - \tau_{y_{si}}) = \frac{\sigma}{\sigma - 1} \frac{1}{(1 - \alpha_s - \beta_s)} \frac{Z_s Q_{si}}{P_{si} Y_{si}}$$

Equation (7) allows us to estimate the three wedges from information on gross output, input costs and the elasticities σ, α_s and β_s.

Next, defining total factor revenue productivity ($TFPR$) as $TFPR_{si} = P_{si} A_{si}$, we get from equation (5):

$$TFPR_{si} = \frac{\sigma}{\sigma - 1} \Psi_s \frac{(1 + \tau_{k_{si}})^{\alpha_s} (1 + \tau_{h_{si}})^{\beta_s}}{(1 - \tau_{y_{si}})} \qquad (8)$$

Equation (8) shows that TFPR does not vary across firms within the same industry unless firms face some type of distortion. Intuitively, in the absence of distortions, more capital, labor and intermediate inputs would be allocated to firms with higher TFP(A_{si}) to the point where their higher output results in a lower price and the exact same TFPR as in firms with lower TFP. In contrast, in the presence of distortions, a high (low) TFPR is a sign that the firm confronts barriers (benefits from subsidies) that raise (reduce) the firm's marginal products of the different factors of production, rendering the firm smaller (larger) than optimal.

Denoting the levels of efficient real and nominal output as Y_{si}^* and $(P_{si} Y_{si})^*$, it can be shown (see Dias et al. (2016)) that:

$$Y_{si}^* = \left(\frac{A_{si}}{TFPR_s^*} \right)^{\sigma} = Y_{si} \left(\frac{TFPR_{si}}{TFPR_s^*} \right)^{\sigma} \qquad (9)$$

$$(P_{si}Y_{si})^* = \left(\frac{A_{si}}{TFPR_s^*}\right)^{\sigma-1} = P_{si}Y_{si}\left(\frac{TFPR_{si}}{TFPR_s^*}\right)^{\sigma-1} \tag{10}$$

where $TFPR_s^*$ is the efficient level of total factor revenue productivity common to all firms in industry s that will prevail if distortions are eliminated from the industry. $TFPR_s^*$ is defined so that all firms face the same average wedges, and these are such that the demand for factors of production at the industry level is the same before and after the reallocation of resources.[3] The average wedges denoted as $(1 + \overline{\tau}_{k_s})$, $(1 + \overline{\tau}_{h_s})$ and $(1 - \overline{\tau}_{y_s})$ are given by the following expressions:

$$
\begin{aligned}
(1 + \overline{\tau}_{k_s}) &= \frac{\alpha_s}{(1 - \alpha_s - \beta_s)}\frac{Z_s Q_s}{R_s K_s}\\
(1 + \overline{\tau}_{h_s}) &= \frac{\beta_s}{(1 - \alpha_s - \beta_s)}\frac{Z_s Q_s}{W_s H_s}\\
(1 - \overline{\tau}_{y_s}) &= \frac{\sigma}{\sigma - 1}\frac{1}{(1 - \alpha_s - \beta_s)}\frac{Z_s Q_s}{(P_s Y_s)^*}
\end{aligned} \tag{11}
$$

where $(P_s Y_s)^* = \sum_{i=1}^{M_s}(P_{si}Y_{si})^*$ is the industry efficient nominal output, $K_s = \sum_{i=1}^{M_s}K_{si}$, $H_s = \sum_{i=1}^{M_s}H_{si}$ and $Q_s = \sum_{i=1}^{M_s}Q_{si}$ are the actual industry levels of the capital stock, labor and intermediate inputs, respectively.

Replacing the firm-specific wedges by the industry average wedges in (8) we have:

$$TFPR_s^* = \frac{\sigma}{\sigma - 1}\Psi_s\frac{(1 + \overline{\tau}_{k_s})^{\alpha_s}(1 + \overline{\tau}_{h_s})^{\beta_s}}{(1 - \overline{\tau}_{y_s})} \tag{12}$$

so that we can decompose the (log) scaled TFPR $(TFPR_{si}/TFPR_s^*)$ for each firm as a weighted sum of the (log) scaled capital, labor and output wedges:

$$ln\left(\frac{TFPR_{si}}{TFPR_s^*}\right) = \alpha_s ln\left(\frac{1 + \tau_{k_{si}}}{1 + \overline{\tau}_{k_s}}\right) + \beta_s ln\left(\frac{1 + \tau_{h_{si}}}{1 + \overline{\tau}_{h_s}}\right) - ln\left(\frac{1 - \tau_{y_{si}}}{1 - \overline{\tau}_{y_s}}\right) \tag{13}$$

---

[3] $TFPR_s^*$ should not be confused with $\overline{TFPR}_s$ used in Hsieh and Klenow (2009). $\overline{TFPR}_s$ in Hsieh and Klenow (2009) is the average TFPR in industry s in the inefficient, or observed, allocation of resources. In contrast, $TFPR_s^*$ is the average TFPR in industry s when all available factors of production are efficiently allocated.

This equation allows us to tell what happens to the output of the firm if distortions are eliminated from the economy, so that TFPR is equalized across firms in each industry. If scaled $TFPR_{si}$ is above one, the firm is being "taxed", so that it will increase production if distortions are eliminated: in the absence of distortions more resources are allocated to this firm to the point where its higher output results in lower price and its TFPR equalizes $TFPR_s^*$. By looking at the right-hand side of this equation we are able to tell where the increase in production comes from. If, for instance, the scaled capital wedge $(1 + \tau_{k_{si}})/(1 + \overline{\tau}_{k_s})$ is larger than one, the firm is facing a capital tax, so that it will increase the capital stock if the distortion is eliminated. Similarly, for the scaled labor wedge. In contrast, firms for which the scaled output wedge, $(1 - \tau_{y_{si}})/(1 - \overline{\tau}_{y_s})$, is above one, are benefitting from output subsidies, so that they would decrease production if those subsidies are eliminated.

Finally, by combining the various results presented above it is straightforward to write the expression for the real gross-output gains at the industry level:

$$
\frac{Y_s^*}{Y_s} = \frac{\left[\sum_{i=1}^{M_s} (Y_{si}^*)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}}{\left[\sum_{i=1}^{M_s} (Y_{si})^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}} = \left[\frac{\sum_{i=1}^{M_s} A_{si}^{\sigma-1}}{\sum_{i=1}^{M_s} \left(A_{si} \frac{TFPR_s^*}{TFPR_{si}}\right)^{\sigma-1}}\right]^{\frac{\sigma}{\sigma-1}}
$$

$$
= \left[\frac{1}{\sum_{i=1}^{M_s} \omega_{si} \cdot \left(\frac{1}{TFPR_{si}/TFPR_s^*}\right)^{\sigma-1}}\right]^{\frac{\sigma}{\sigma-1}} \tag{14}
$$

where

$$
\omega_{si} = \left(\frac{A_{si}}{\left(\sum_{i=1}^{M_s} A_{si}^{\sigma-1}\right)^{\frac{1}{\sigma-1}}}\right)^{\sigma-1} = \left(\frac{A_{si}}{TFPs^*}\right)^{\sigma-1} = \frac{A_{si}^{\sigma-1}}{\sum_{i=1}^{M_s} A_{si}^{\sigma-1}}. \tag{15}
$$

The interpretation of equation (14) is very intuitive as it is simply the ratio of efficient output to observed output in industry s. Note that $\sum_{i=1}^{M_s} \omega_{si} = 1$ and that $TFP_s^* = \left(\sum_{i=1}^{M_s} A_{si}^{\sigma-1}\right)^{(1/(\sigma-1))}$ is the industry level TFP in the absence of distortions (see Hsieh and Klenow (2009)). Thus, equation (14) shows that efficiency gains in industry s

are a weighted sum of the inverse scaled TFPR $(1/(TFPR_{si}/TFPR_s^*))$ across firms, where the weights are the contribution of each firm to the efficient industry TFP.[4] The smaller is this weighted sum, the larger are the efficiency gains obtained if distortions are eliminated from the industry. In particular, this sum will be small and, thus, efficiency gains will be large if there is a strong positive correlation between the weights ω_{si} and the scaled TFPR $(TFPR_{si}/TFPR_s^*)$. In other words, efficiency gains will be higher if, on average, more productive firms face higher distortions. From (14), we can also intuitively see that, everything else constant, efficiency gains will be higher the larger the dispersion of scaled TFPR.[5]

Using the Cobb-Douglas aggregator given by equation (1) we get the economy-wide potential gross-output (or TFP) gains from resource reallocation:

$$
\frac{Y^*}{Y} = \prod_{s=1}^{S} \left\{ \frac{Y_s^*}{Y_s} \right\}^{\theta_s} = \prod_{s=1}^{S} \left\{ \left[\frac{\sum_{i=1}^{M_s} A_{si}^{\sigma-1}}{\sum_{i=1}^{M_s} \left(A_{si} \frac{TFPR_s^*}{TFPR_{si}} \right)^{\sigma-1}} \right]^{\frac{\sigma}{\sigma-1}} \right\}^{\theta_s}
$$
$$
= \prod_{s=1}^{S} \left\{ \left[\frac{1}{\sum_{i=1}^{M_s} \omega_{si} \cdot \left(\frac{1}{TFPR_{si}^{**}} \right)^{\sigma-1}} \right]^{\frac{\sigma}{\sigma-1}} \right\}^{\theta_s}
$$

(16)

Equations (14) and (16) will be used to compute industry and economy aggregate gross output reallocation gains, respectively. As the exercise fixes the total amount of inputs and calculates how much gross output could be increased by reallocating resources between firms within each industry, it follows that potential gross-output gains coincide with potential TFP gains, so that (16) gives us the potential efficiency gains both in terms of gross output and TFP. In the empirical section we compute gross-output gains for different groupings of industries (manufacturing vs services) and

---

[4]Using equation (10), it is straightforward to show that these weights also correspond to the firm's gross-output market share, when all resources are efficiently allocated across firms, that is, $\omega_{si} = (P_{si}Y_{si})^* / \sum_{i=1}^{M_s} (P_{si}Y_{si})^* = A_{si}^{\sigma-1} / \sum_{i=1}^{M_s} A_{si}^{\sigma-1}$.

[5]Note that efficiency gains are zero if scaled TFPR is equal to one for all firms, i.e., if there are no distortions in the industry, which means that dispersion of $(TFPR_{si}/TFPR_s^*)$ is zero. Introducing distortions implies, in practice, making the dispersion of $(TFPR_{si}/TFPR_s^*)$ differ from zero.

for each case these formulas will be adjusted accordingly.

3 The data

In this paper we use firm-level balance-sheet data and industry-level factor shares. The firm-level data draws on annual information for Portuguese firms reported under the *Informação Empresarial Simplificada* (IES). IES data exist from 2006 onwards and covers virtually the universe of Portuguese non-financial firms. The almost universal coverage of IES emerges from the fact that it is the system through which firms report mandatory information to the tax administration and the statistical authorities like the *Instituto Nacional de Estatística* (INE) (the Portuguese Statistics Institute), and the *Banco de Portugal* (the Portuguese central bank). The data provide very detailed information on the firms' balance sheets and income statements. From this dataset we get information on firm's gross output, value added, consumption of intermediate inputs, labour costs (wages and benefits including social security contributions), employment (average number of employees), gross investment (or gross fixed capital formation), annual and accumulated depreciations, and the book values of gross and net capital stock.

For the purpose of this paper, even though we report results only for 2008 and 2010, we also use data for 2007 and 2009, because we need consecutive years for the construction of some ancillary variables such as productivity shocks. In the IES there are 375,783 observations (different firms) in 2008 and 370,326 observations in 2010. Before using the data, we clean the dataset by dropping firms that do not report strictly positive figures for gross output (production), labor costs, employment, capital stock, intermediate consumption and value added. After cleaning the data, we are left with 236,022 and 230,157 observations for 2008 and 2010, respectively.

Table 1 records the relative importance of agriculture, manufacturing and services in our dataset in terms of employment, gross output and value added. Note the small

contribution of agriculture for total employment and value added (around 2 percent), while the service sector contributes around 75 percent. Manufacturing, which has been the focus of most empirical studies, contributes only 22-24 percent to total value added.[6]

Table 1: Relative importance of each sector in the dataset
(Percentage)

	2008			2010		
	Agric.	Manuf.	Serv.	Agric.	Manuf.	Serv.
Employment	1.97	25.34	72.69	2.04	23.69	74.26
Gross output	2.42	34.46	63.12	1.92	32.71	65.36
Value added	2.35	23.57	74.08	1.76	22.24	76.00
Number of firms	6,069	34,257	195,696	6,351	32,096	191,710

Note: Agriculture also includes forestry, fishing, mining and quarrying; Services also include construction and energy.

The dataset also includes information on the firm's main industry of operation based on NACE classification (Rev. 2.1 and Rev. 3) both at 3- and 5-digit disaggregation level. However, the exercises in our paper are conducted with industries defined at the 3-digit NACE code (Rev. 2.1). In fact, with disaggregation at 5-digit level, the number of firms is very small in many industries implying noisy estimates of industry-level efficiency gains, as well as of some of the regressors used in our analysis. Overall, the 3-digit NACE code classification for 2008 implies 213 different industries (16 for agriculture (including forestry, fishing, mining and quarrying), 101 for manufacturing and 96 for services (including energy and construction)). For 2010 we have 215 industries (16 for agriculture, 101 for manufacturing and 98 for services).

For the industry-level factor shares, we use the average factor shares that are observed in the US during the period 1998 to 2010, which are published by the BEA (Bureau of Economic Analysis).[7] An important remark to be made is that, as with

---

[6]According to information from the National Accounts, in 2008, agriculture, manufacturing and services contribute 2.4, 14.1 and 83.5 percent for aggregate GDP, respectively. Thus, if anything, our dataset appears to be slightly skewed towards manufacturing and against the service sector.

[7]In our model, it is not possible to separately identify the average input distortions (average wedges) and the input elasticities in each industry. Thus, using factor shares from the U.S. economy is a simple

the BEA data, our firm-level data on worker compensation include the salaries and other labor costs such as pension contributions or fringe benefits. Because industry classification is different in the two countries we make an approximate concordance between the two classifications.[8]

4 Misallocation in the manufacturing and service sectors

In order to take the model to the data we must choose a value for the elasticity of substitution parameter (σ), decide how to treat outliers and choose the group of firms included in the analysis. It is known that these assumptions impact the estimated levels of misallocation (see Hsieh and Klenow (2009) or Dias et al. (2016)). In the Appendix, where we compute the efficiency gains under alternative assumptions, it is shown that changing these assumptions also has consequences for the estimated difference of misallocation between the manufacturing and service sectors. Table 2 shows the estimates of efficiency gains for 2008 and 2010, for two distinct sets of assumptions, which are thoroughly discussed in the Appendix. The "baseline model" assumes σ=3.0 (the number usually assumed in the literature), trims 1 percent tails of the scaled TFPR and TFP distributions (also the usual trimming used in the literature) and includes all the firms in the dataset. The "final model" assumes σ=4.5 (the average sigma estimated for Portugal), trims 2.5 percent tails of the scaled TFPR and TFP distributions and excludes very small firms (firms with 10 or less employees).

Under the "baseline model" assumptions, we see from Table 2 that, if distortions in the economy were eliminated (by equalizing TFPR across firms in each industry and

---

way to control for distortions that could affect the input shares in the Portuguese economy, while the U.S. is taken as a benchmark of a relatively undistorted economy.

[8]In the small number of cases for which we were not able to find a good match, we used the average for the whole economy in the U.S.. Between 1998 and 2010, gross output was composed of 46 percent consumption of intermediate inputs, 33 percent labor compensation and the remaining 21 percent were the compensation to capital owners.

Table 2: Efficiency gains under alternative assumptions

Assumptions	2008				2010			
	Total	M	S	S-M	Total	M	S	S-M
Baseline model	43.36	16.02	59.19	43.18	49.33	16.81	66.46	49.65
Final model	28.46	14.15	37.66	23.51	31.28	14.43	40.82	26.39

Note: Efficiency gains in the baseline model are computed taking all the firms in the dataset, assuming $\sigma=3.0$ and trimming 1 percent tails of $\log(TFPR_{si}/TFPR_s^*)$ and $log(A_{si}M^{\frac{1}{\sigma-1}}/TFP_s^*)$. Efficiency gains in the final model are computed taking only firms with more than 10 employees, assuming $\sigma=4.5$ and trimming 2.5 percent tails of the scaled TFPR and TFP distributions. M stands for manufacturing and S for services. S-M is the difference between the service and the manufacturing sectors. The total also includes firms from agriculture.

keeping industry level factor demand constant), gross output (or TFP) for the whole economy would be around 43 percent above actual gross output (or actual TFP) in 2008 and around 49 percent in 2010. Efficiency gains are also clearly higher in the service sector (around 59 percent in 2008 and 66 percent in 2010) than in the manufacturing sector (around 16 and 17 percent in 2008 and 2010, respectively). Thus, the service sector emerges as far more inefficient than the manufacturing sector in line with the results in Dias et al. (2016). Under the "final model" assumptions, the estimated efficiency gains for the whole economy become smaller: about 28 percent in 2008 and 31 percent in 2010.[9] The difference between the service and the manufacturing sectors also becomes smaller, but remains very high: about 24 pp in 2008 and 26 pp in 2010. The decreasing differences of misallocation between the two sectors reflect a higher presence of small firms and a higher frequency of outliers in the service sector, so that dropping small firms from the dataset and increasing the trimming of the scaled TFP and TFPR distributions, as we move from the baseline to the final model, affect estimated efficiency gains in the service sector disproportionately.

Figures 1 and 2 depict industries ordered by the level of efficiency gains for 2008 under the baseline and the final model, respectively. The striking message from this figures is that the significantly higher levels of efficiency gains in the service sector

---

[9]Note that these figures are for gross output and that the efficiency gains evaluated in terms of value added (the type of estimates usually available in the literature) are significantly higher. For 2008, the corresponding value-added efficiency gains, under the "final model" assumptions, are 71.75 percent for the whole economy, 47.64 percent for manufacturing and 84.05 percent for services.
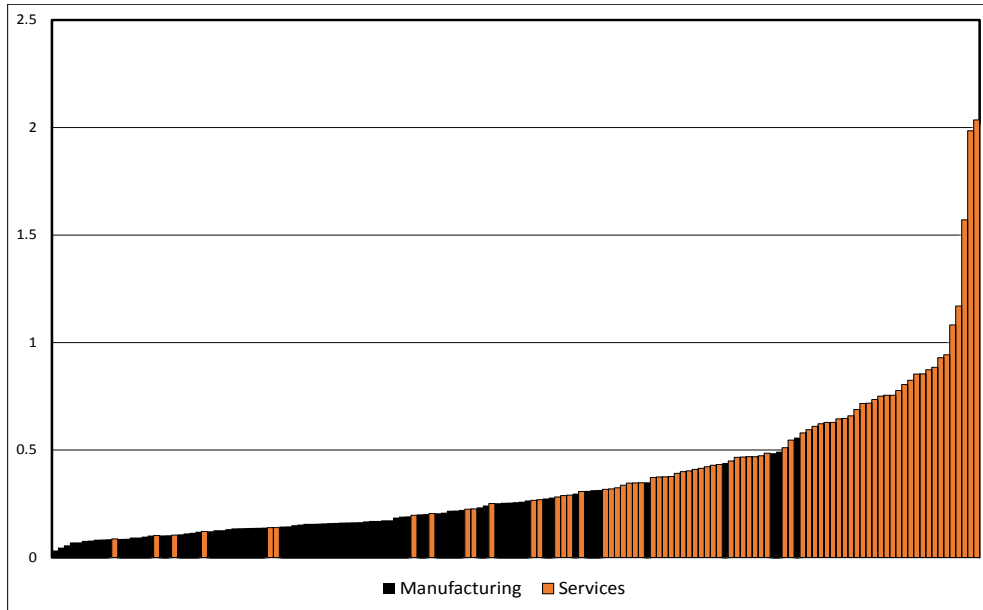
15

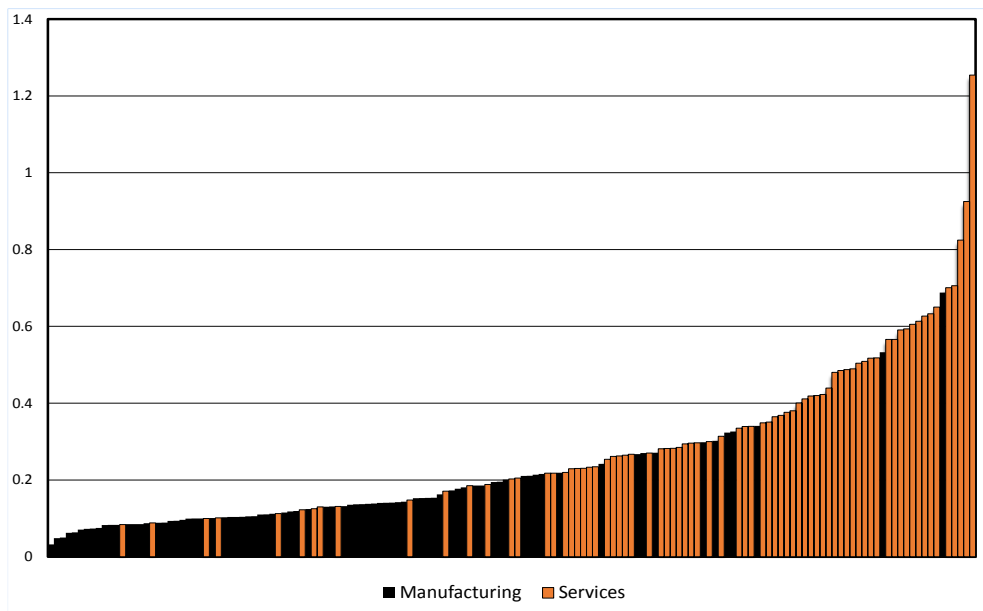Figure 1: TFP gains from reallocation in 2008 - Baseline model

Figure 2: TFP gains from reallocation in 2008 - Final model

are not the result of a small number of industries with abnormal levels of efficiency gains, but the result of a strong regularity: the bulk of the industries pertaining to the manufacturing sector ranks first, while the bulk of the industries of the service sector appears on the right-hand side of the chart. More specifically, among the 50 percent of the industries with the lowest TFP gains (77 industries) only 11 industries (14.7 percent) belong to the service sector in Figure 1, and only 13 industries (17.3 percent) belong to the service sector in Figure 2. This result shows that the presence of higher levels of inefficiency is a widespread phenomenon in the service sector. For reasons discussed in the Appendix, in what follows we focus on the "final model" assumptions, as these give rise to more conservative estimates of the differences of misallocation between the two sectors, which are the main concern of this paper.

Using the conservative estimates recorded in Table 2, we conclude that closing the misallocation gap, that is, reducing misallocation in the service sector to the manufacturing levels in 2008 (from 37.66 percent to 14.15 percent) would lead to a 12 percent boost in aggregate gross output (or aggregate TFP) and a 31 percent boost in aggregate value added (GDP).[10] These are large numbers deserving a thorough investigation of the determinants underlying the misallocation gap between the two sectors. In the next section we resort to the literature on misallocation and regression analysis to identify the most relevant factors that might explain the difference in misallocation levels between the manufacturing and the service sectors.

---

[10]The efficiency gains in terms of gross output are computed as $(1.3766/1.1415)^{\theta_s}$, where θ_s is the share of the service sector in aggregate gross output. The value-added efficiency gains are computed using equation (24) in Dias et al. (2016). Notice that this exercise is similar to the one carried out in Hsieh and Klenow (2009) where the authors estimate the TFP losses in China and India stemming from the misallocation gap between these countries and the U.S..

5 Explaining differences in misallocation between the service and manufacturing sectors

In this section, we use regression analysis and the Gelbach decomposition to identify which factors are the most relevant to explaining the differences of misallocation between the service and manufacturing sectors. We start by reviewing the main theories of misallocation suggested in the literature for guidance on the identification of the regressors required to complete the model specification. Then, we briefly present the Gelbach decomposition (see Gelbach (2016)) for our model and discuss the empirical results.

5.1 Theories of misallocation

The list of potential sources of misallocation suggested in the literature is long and varied. For presentation purposes we group them in four different categories: a) adjustment costs; b) distortions to input prices; c) financial frictions; and d) firm-specific markups.

a) Adjustment costs

The Hsieh and Klenow (2009) model described above is static in the sense that it assumes all the adjustments are instantaneous, so that the firms should always be in their long-run static equilibrium relationship no matter the frequency, size or type of shocks that hit the firms. Thus, any short-term deviation from the static equilibrium relationship stemming from idiosyncratic shocks and the presence of adjustment costs will show up as misallocation in the model. In a recent paper, Asker et al. (2014) investigated the role of adjustment costs in shaping the dispersion of marginal product of inputs (see also Bartelsman et al. (2013) and Song and Wu (2013)). In their model firms acquire the inputs in a frictionless spot market, but are hit by idiosyncratic productivity shocks and face costs when adjusting the capital stock. In such a framework,

dispersion in the marginal revenue product of capital arises naturally and the resource allocation, while appearing as inefficient in a static model, may be efficient in a dynamic sense. An important result for our paper is that, in the presence of adjustment costs, as the volatility of TFP shocks increase, so does the dispersion of marginal product of the inputs. Thus, in the context of our model, we may expect volatility of TFP shocks to have a bearing on the dispersion of TFPR and thus, on misallocation, especially through higher dispersion of the capital and/or labor wedges (see equation (13)).[11]

Another potential source of misallocation, stemming from the static nature of the Hsieh and Klenow (2009) model, is the implicit assumption of instantaneous price adjustment following productivity shocks. According to the model, there is a one-to-one contemporaneous negative relationship between productivity (A_{si}) and prices (P_{si}) (see equation (5)), that is, the model assumes that a one percent increase in productivity implies an instantaneous one percent decrease in prices. It is widely known, however, that price stickiness is a pervasive phenomenon in the economy and that firms may react differently to shocks (demand or cost shocks) and asymmetrically to positive and negative shocks.[12] Thus, similarly to what happens with capital or labor adjustment costs, we may expect firm-level productivity shocks in the presence of price-rigidity to imply additional TFPR dispersion (through higher dispersion of the output wedge) and thus, to give rise to increased misallocation.[13]

Below, we account for the possibility that productivity shocks might help to explain the differences in misallocation between the manufacturing and the service sector. As we shall show, this will be the case if the average size of TFP shocks and/or the effect

---

[11]Imperfect information has also been suggested as a foundation for adjustment costs, giving rise to a sluggish response of inputs to fundamentals and thus to misallocation (see David et al. (2014)). In our model it is not possible, however, to distinguish between alternative sources of adjustment costs (technological versus informational frictions).

[12]Information costs and/or menu costs incurred by the firm to determine the optimal price and/or to change the price are usually suggested in the literature as the main sources of price rigidity. For empirical evidence on price rigidity at the firm level see, for instance, Fabiani et al. (2006) for the euro area and Dias et al. (2015) for Portugal.

[13]Our framework does not allow distinguishing productivity shocks from other type of shocks. So, in what follows, we use the term productivity shocks to designate a range of time-varying shocks to production that include TFP shocks, demand shocks, natural disasters, changes to informal barriers, etc.

of these shocks differ across the two sectors. If the importance of adjustment costs or the degree of price rigidity vary across industries, we should expect the impact of productivity shocks on misallocation to also differ across industries. If, for instance, on average, price rigidity is higher in the service than in the manufacturing sector, we might expect the impact of productivity shocks on misallocation to be higher in the former.

b) Distortions to input prices

A second potential source of misallocation is the presence of distortions on the prices of the production factors (for a discussion, see Hsieh and Klenow (2009), Guner et al. (2008) or Restuccia and Rogerson (2008), among others). For instance, non-competitive banking systems may offer favourable interest rates on loans to some producers based on non-economic factors, leading to misallocation of credit across firms. Governments may offer subsidies, special tax incentives or lucrative contracts to specific producers. Various product and labor-market regulations may drive up the cost of labor in the formal vis-à-vis the informal sector, or in large versus small firms, or drive down the cost of capital in small firms through special lines of credit. In the specific case of Portugal there are many regulations that tend to benefit small and medium-sized firms, by granting these firms access to labor and investment subsidies that are not easily accessible or not accessible at all to large firms. These include i) less bureaucratic processes for worker's dismissals; ii) less costly conditions regarding the provision of health care services; iii) smaller fines for labor law breach; iv) better accessibility conditions (including reduction of social contributions) to subsidized employment creation, workers training and professional internships programs; v) better accessibility conditions to investment subsidies, including access to specific investment programs; and vi) access to specific fiscal benefits (smaller taxes on profits).[14]

In practice, the impact of policies that favor smaller (or larger) firms through lower capital or labor costs is expected to vary across industries depending on the importance

---

[14]Illustrations of size-dependent polices for Italy and France can be seen in Guner et al. (2008) and Gourio and Roys (2014) and Garicano et al. (2013), respectively.

of smaller (or larger) firms in each industry. Below, we capture this phenomenon by looking at the skewness of the industry TFP distribution. We use this statistic as a summary measure to characterize the structure of each industry in terms of productivity. By looking at TFP rather than at other measures of size (employment or gross output) we avoid some additional endogeneity problems, as employment and gross output are distortions dependent.[15] By definition, skewness is high in industries with a high proportion of firms with low productivity, and a few firms with very high productivity. Thus, we interpret the correlation between this statistic and misallocation, at the industry level, as measuring the importance of size-dependent distortions (if distortions were allocated purely randomly across firms the covariance between misallocation and skewness should be zero). Further, by linking productivity to the individual wedges we will be able to tell whether less productive firms are being taxed or subsidized and where the tax or the subsidy comes from.

c) Financial frictions

A third potential source of misallocation is the presence of financial frictions. For instance, financial institutions may be unable or unwilling to provide credit to firms that are highly productive but have no credit history or insufficient guarantees, preventing these firms from expanding their activities (for a discussion, see Midrigan and Xu (2014), Moll (2014), Gilchrist et al. (2013), Banerjee and Moll (2010) or Hosono and Takizawa (2012), among others). But, financial frictions may instead affect larger firms with larger scale of operation and larger financing needs (see Buera et al. (2011)).

Below, we try to accommodate the possibility that young firms (irrespective of their size) face financial constraints by investigating how misallocation varies with the age of the firms. In practice, our model allows us to distinguish financial frictions that operate as quantity restrictions from size-dependent policies that operate through the

---

[15]Bartelsman et al. (2013) show that misallocation stemming from policy induced distortions may affect the correlation between the distribution of productivity and the size of the firm. Recall also that, according to the discussion in Section 2, industry-level misallocation will be higher the stronger the (positive) firm-level correlation between TFP and scaled TFPR.

price of inputs. The presence of size-dependent policies that affect the price of the inputs may be detected by looking at the relationship between the size of the firm and the individual scaled wedges. For instance, if distortions are due to firm-size contingent policies that favour smaller firms by reducing the cost of capital (through special lines of credit) or the cost of labor (through especial labor regulations), returns to additional capital and labor would be expected to be lower in smaller firms, that is, we would expect a positive relationship between size and TFPR or size and the capital or labor wedge. In contrast, if misallocation is due to financial market failures that constrain young firms, we would expect the presence of many of these firms that do not grow because they could not secure access to credit. In other words, we would expect a negative relationship between the age of the firm and the capital wedge. Of course, the two situations may coexist in practice: there may be at the same time small firms that benefit from lower capital costs (special lines of credit) and young firms (eventually small) that are facing capital constraints.

d) Firm-specific markups

Imperfectly competitive product markets with firm-specific markups have also been suggested as a potential source of misallocation (see, among others, Peters (2013) and Syverson (2004*a,b*)). In an environment with imperfectly competitive output markets, misallocation, as identified by the static model, occurs because firms have monopoly power and set firm-specific markups. While distortions on the prices of inputs imply that firms with relatively high TFPR are constrained (or "taxed"), imperfectly competitive output markets predict that high TFPR is indicative of market power (captures higher markups).[16] According to this type of models, industry-level misallocation is expected to be negatively correlated with the level of competition. For instance, barriers to substitution across producers, stemming from various forms of product differentiation (spatial, physical or brand-driven), may explain different levels of misallocation

---

[16]In the model suggested by Peters (2013), $TFPR_{si}$ is proportional to the firm-specific markup, which replaces the right-hand side of equation (8) above.

among industries.

To capture the impact of competition on misallocation we need reliable measures of competition across industries. However, satisfactory measures of competition are very difficult to build, either because they are theoretically unsatisfactory or because appropriate data are not available. Statistics such as output concentration ratios, advertising intensity measures, price-cost margins, Herfindahl or the Ellison and Glaeser (1997) concentration indexes, measures of sunk entry costs or of international exposure, do not capture the degree of competition in the relevant markets for all industries. For instance, for Portugal, the concentration ratio (the importance of sales of the largest 4 or 8 firms in total industry sales) is higher in the manufacturing than in the service sector. Taken at face value that would mean that the manufacturing sector is less competitive than the service sector. However, the number of firms in the industry or its concentration ratio does not identify the degree competition in the product market if a significant part of the industry output is exported. What matters is competition in the destination market. Similar comments can be made about most of the other statistics.[17]

Thus, the estimated model below does not explicitly include any direct competition measure. However, it is important to note that competition affects the degree of price rigidity, as well as the characteristics of the productivity distribution. We may expect less price rigidity, and thus less misallocation, in industries where competition is higher. Similarly, in industries in which it is easy for customers to switch between competing suppliers, the productivity distributions should exhibit higher minima, less dispersion and lower skewness. Thus, in our empirical model, the skewness of the productivity distribution may also be interpreted as an indirect measure of competition: higher skewness signals lower competition.

---

[17]For a a thorough discussion on this issue see Holmes and Schmitz (2010).

23

5.2 The Gelbach decomposition

Let us denote the efficiency gains in industry s by $Z_s = Y_s^*/Y_s$, and let D be a dummy variable, which equals 1 if the industry belongs to the service sector and 0 if it belongs to the manufacturing sector.[18] In the simple cross-section regression

$$Z_s = a_0 + a_1 D_s + u_s \qquad (17)$$

the coefficient a_1 measures the difference between the efficiency gains in the service and manufacturing sectors. The D_s variable in the simple regression (17) may be thought of as proxying for differences of certain factors (characteristics) between the manufacturing and service sectors.

The theories of misallocation, surveyed above, suggest that we can expect industry-level efficiency gains to be correlated with productivity shocks, skewness of the TFP distribution and the proportion of young firms in each industry. Let us denote these three regressors by X_{1s}, X_{2s} and X_{3s}, respectively. If we account for the possibility of each regressor having a different impact on the service and manufacturing sectors, the general model may be written as:

$$Z_s = a_0 + a_1 D_s + b_1 X_{1s} + c_1 D_s X_{1s} + b_2 X_{2s} + c_2 D_s X_{2s} + b_3 X_{3s} + c_3 D_s X_{3s} + v_s \quad (18)$$

where the c_i coefficient measures the difference of the impact of the X_{is} regressor between the two sectors. From the estimates for the full model (18), we can tell whether the regressors are able to fully explain the difference between the misallocation in the two sectors. That will be the case if the coefficient a_1 is not statistically different from zero.

The Gelbach decomposition of omitted variable bias allows us to quantify the impact that each variable/factor has on the change in the estimate of a_1. Using the results in

---

[18]In the analysis that follows we drop the agriculture sector, as we are only interested in explaining the differences between misallocation in the manufacturing and service sectors.

(Gelbach (2016)), we can write:

$$\widehat{a}_1^{base} = \widehat{a}_1^{full} + \widehat{\theta}_1\widehat{b}_1 + \widehat{\mu}_1\widehat{c}_1 + \widehat{\theta}_2\widehat{b}_2 + \widehat{\mu}_2\widehat{c}_2 + \widehat{\theta}_3\widehat{b}_3 + \widehat{\mu}_3\widehat{c}_3 \qquad (19)$$

where:

a) $\widehat{a}_1^{base}$ and $\widehat{a}_1^{full}$ are the OLS estimate of a_1 in the base model (equation (17)) and the full model (18), respectively.

b) $\widehat{b}_i$ and $\widehat{c}_i$ (i=1,2,3) are the OLS estimates of the b_i and c_i parameters in the full model (18).

c) $\widehat{\theta}_i$ and $\widehat{\mu}_i$ are the OLS estimates of the θ_i and μ_i parameters in the simple regressions $X_{is} = b_{i0} + \theta_i D_s + \varepsilon_{is}$ and $D_s X_{is} = c_{i0} + \mu_i D_s + \varepsilon_{is}$, (i=1,2,3), respectively.

In our case, the OLS estimates of the parameters in equation (19) lend themselves to a very intuitive interpretation:

i) $\widehat{a}_1^{base}$ is the difference between the efficiency gains in the service and the manufacturing sectors;

ii) $\widehat{a}_1^{full}$ is the residual difference between the efficiency gains in the service and manufacturing sectors not accounted for by the model;

iii) $\widehat{\theta}_i$ is the difference between the mean of X_{is} in the service and manufacturing sectors, and

iv) $\widehat{\mu}_i$ is the mean of X_{is} in the service sector.

The total contribution of the X_{is} regressor is given by $\widehat{\theta}_i\widehat{b}_i + \widehat{\mu}_i\widehat{c}_i$. Thus, X_{is} accounts for the difference between the efficiency gains in the two sectors, if at least one of the two conditions is met: i) the mean of X_{is} differs across the two sectors $(\widehat{\theta}_i \neq 0)$, or ii) the impact of X_{is} differs between the two sectors $(\widehat{c}_i \neq 0)$.[19]

---

[19]We note that the Gelbach decomposition for our particular model is similar to the so-called Oaxaca-Blinder decomposition that has been extensively used in the literature to decompose mean wage differentials (see, for instance, Blinder (1973), Oaxaca (1973) and Jann (2008)). One important difference, however, is that the methodology developed in Gelbach (2016) allows for statistical inference regarding the decomposition, while the Oaxaca-Blinder method did not.

5.3 Regression analysis

The results of decomposition (19) are presented in Table 3 for 2008 and 2010, with robust standard errors in parentheses. The first row records the $\widehat{a}_1^{base}$ estimates, that is, the difference between efficiency gains in the service and the manufacturing sectors.[20] The second row reports the explained difference, that is the sum of contributions of the 3 regressors. Finally, the second row from bottom records the unexplained difference, that is, $\widehat{a}_1^{full}$. For each regressor X_{is}, the total contribution is divided into the two above mentioned components: one stemming from the difference between the mean of the regressors in the two sectors (characteristics effect), and one stemming from the difference of the regressors's impact in the two sectors (coefficients effect).

An important result is that the model fully accounts for the difference between the efficiency gains in the two sectors. The unexplained difference, that is, $\widehat{a}_1^{full}$, is not significantly different from zero both in 2008 and 2010. Similarly to Figures 1 and 2 above, if we now order industries by the level of unexplained efficiency gains in 2008, we conclude that among the 50 percent of the industries with the lowest TFP gains (77 industries), 43 industries (58.1 percent) belong to the service sector (compared to 17.3 percent in Figure 2). These numbers suggest that our model is able to explain the differences of misallocation between industries of the two sectors to the point where no systematic differences of TFP gains in the two sectors remain.

Productivity shocks emerge as the most important factor explaining misallocation differences between the two sectors.[21] Importantly, the contribution of productivity shocks stems mostly from the difference of the impacts between the two sectors (coefficients effect). In particular, the impact of productivity shocks in the service sector is

---

[20]Note that the difference in efficiency gains between the two sectors in Table 3 is a non-weighted average, which explains the difference vis-à-vis the figures reported in the last row of Table 2.

[21]To identify the productivity shocks, we assume that TFP follows the AR(1) process: $a_{si,t} = \mu_s + \rho_s a_{si,t-1} + \phi_s v_{si,t}$, where $a_{si,t}$ stands for the log of TFP of firm i, in industry s, in period t, and $v_{si,t} \sim N(0,1)$ is an independent and identically distributed (i.i.d.) standard normal random variable. The ϕ_s term measures productivity shocks in industry s. To estimate ϕ_s we use two years of consecutive data, but restrict the sample to firms that appear in both years. When estimated freely, ρ_s is close to unity for the great majority of industries. Thus, ultimately, we compute the productivity shocks as the industry-level standard deviation of $(a_{si,t} - a_{si,t-1})$.

significantly higher than in the manufacturing sector. This is an interesting result that warrants some explanations.

Table 3: Efficiency gains
(Difference between services and manufacturing)
(Gelbach Decompositions)

	2008	2010
Difference in efficiency gains	0.202 (8.31)	0.205 (7.90)
Explained difference:	0.225 (2.36)	0.207 (2.45)
a) Productivity shocks	0.175 (1.82)	0.123 (1.64)
a_1) characteristics effect	0.027 (1.77)	0.023 (2.17)
a_2) coefficients effect	0.148 (1.44)	0.099 (1.23)
b) Sectoral structure	0.086 (2.60)	0.113 (3.07)
b_1) characteristics effect	0.016 (2.15)	0.002 (0.33)
b_2) coefficients effect	0.069 (2.00)	0.112 (2.85)
c) Importance of young firms	−0.036 (−0.96)	−0.029 (−0.97)
c_1) characteristics effect	0.021 (1.83)	0.014 (1.70)
c_2) coefficients effect	−0.057 (−1.32)	−0.042 (−1.37)
Unexplained difference	−0.023 (−0.26)	−0.002 (−0.02)
Number of industries	154	154

Note: Efficiency gains are obtained assuming case (8) in Table 2. Difference in efficiency gains is given by the coefficient of the industry-dummy in regression (17), while the unexplained difference is given by the coefficient of the industry-dummy in regression (18). Robust t-statistics in parentheses.

For reasons discussed above, we expect industry-level efficiency gains to be positively correlated with productivity shocks. In the presence of adjustment costs, a firm can adjust capital or labor only with some time lags as it takes time to install capital or to hire new employees. A similar process takes place in the presence of output price rigidity. Thus, when hit by an idiosyncratic productivity shock, a firm responds with a lag and adjusts the input level or the output price sluggishly, which leads to variation of TFPR across firms. With this lagged response, greater idiosyncratic shocks lead to greater variation of TFPR across firms and thus, to greater misallocation. However,

for the impact of productivity shocks on misallocation to differ across sectors, we need to assume that the importance of input adjustment costs (capital and/or labor) or the degree of price rigidity vary across industries.

In order to investigate this issue further, we use equation (13) to look at the correlation between TFP shocks and the dispersion of individual wedges. This analysis allows us to tell whether the impact of TFP shocks on misallocation stems mainly from the presence of capital, of labor or of output distortions. Table (4) reports these correlations for 2008.

Table 4: Correlation between the regressors and the standard-deviations of scaled TFPR and scaled wedges

(2008)

	TFP-shocks		Sectoral structure		Share young firms	
	Manuf	Serv	Manuf	Serv	Manuf	Serv
sd(tau_k)	0.114	0.059	0.212	-0.089	0.302	0.225
sd(tau_h)	0.108	0.418	0.084	0.241	0.308	0.233
sd(tau_y)	0.342	0.393	0.299	0.163	0.398	0.160
sd($tfpr$)	0.337	0.456	0.264	0.228	0.359	0.153

Note: Entries are the correlations between the regressors and the industry-level standard-deviations of scaled TFPR and scaled wedges; sd(x) stands for the industry-level standard deviation of ln(x); TFP-shocks are proxied by the industry-level standard deviation of firms' log TFP changes between year t and year t-1; the sectoral structure is proxied by the skewness of the productivity distribution and the importance of young firms is proxied by the proportion of firms with three years of age or less.

An interesting result in Table 4 is that productivity shocks are not very correlated with the standard-deviation of the capital-wedge, especially in the service sector.[22] In contrast, productivity shocks appear more correlated with the labor-wedge and output-wedge dispersion in the service than in the manufacturing sector. This evidence is consistent with the idea of higher price rigidity and higher labor adjustment costs in the service sector. The empirical evidence in the literature has shown that price rigidity is higher in the service sector (see, for instance, Fabiani et al. (2006) for the euro area, and Dias et al. (2015) for Portugal). It is also known that price rigidity is higher in less

---

[22] Of course, this does not necessarily mean that capital adjustment costs are not present in the economy, as the adjustment costs are not the only source of dispersion of the marginal product of capital.

competitive markets (Martin (1993), Gopinath and Itskhoki (2010)). Together this evidence suggests that higher output price rigidity in the service sector might stem from lower competition in this sector, as services are typically more differentiated than manufacturing (see ECB (2006)).[23] In turn, higher informational frictions (stemming from higher spatial dispersion of firms due to local markets) might explain why labor adjustment costs appear to be higher in the service sector. Thus, together, higher output price rigidity and higher labor adjustment costs emerge as the explanations for the higher impact of productivity shocks on misallocation in the service sector.

The sectoral structure, as proxied by the skewness of the productivity distribution, emerges as the second most important factor to explaining misallocation differences between the two sectors. As explained above, we use the skewness of the productivity distribution as a way of summarizing the industry-level characteristics that may affect the impact on misallocation of size-dependent distortions.[24] The aggregate impact of size-dependent policies varies across industries according to the characteristics of the size distribution of each industry. In an economy where special lines of credit (with subsidized interest rates) or employment subsidies are available to small and medium sized firms, we would expect industries with higher skewness to exhibit higher misallocation.

From Table 3 we see that the bulk of the contribution of the sectoral structure comes from the higher impact of this regressor in the service sector (coefficients effect). According to the evidence in Table 5, which records the correlation coefficients between TFP, TFPR and its components, the higher impact of skewness on misallocation in the service sector must reflect the higher positive correlation between productivity and

---

[23]Of course, besides competition, there might be other factors that contribute to higher price rigidity in the service sector, like higher information or search costs due to higher geographical dispersion.

[24]Skewness of the productivity distribution is computed using the usual Fisher-Pearson formula: $sk_s = \sum_{i=1}^{Ms} \left(X_{si} - \overline{X_s} \right)^3 / SD(X_s)^3$, where X_{si} stands for scaled TFP_{si} and $SD(X_s)$ for the standard deviation of X_s, with $X_{si} = (A_{si} M_s^{\frac{1}{\sigma-1}} / TFP_s^*)$ and TFP_s^* is as defined above. We note that skewness is a measure of asymmetry of the distribution about its mean. Positive skewness indicates that the tail of the distribution on the right hand side is longer or fatter than the left side. In a positive skewed distribution, usually the mean is greater than the mode, which means that there are a lot of firms with low productivity levels (below the mean) and a few firms with productivity far above the mean.

TFPR in this sector (0.70 in the service sector compared to 0.43 in the manufacturing sector, in 2008). This higher positive correlation in the service sector means that there must be size-dependent distortions in this sector that are not present or are present to a less extent in the manufacturing sector.

Table 5: Correlation between scaled TFP, scaled TFPR and the scaled wedges
(average of industry-level correlations)

	2008		2010	
	Manuf	Serv	Manuf	Serv
corr(tfp, tau_k)	0.291	0.387	0.303	0.384
corr(tfp, tau_h)	0.407	0.314	0.442	0.297
corr(tfp, tau_y)	0.156	0.046	0.167	0.026
corr($tfp, tfpr$)	0.431	0.696	0.448	0.693
corr($tfp, tfpr + tau_y$)	0.437	0.424	0.462	0.419

Note: corr(tfp, x) stands for the correlation between log(scaled TFP) and log(x), where x represents scaled TFPR or one scaled wedge. Figures correspond to the (non-weighted) average of the industry-level correlations.

We may try to identify such distortions by looking at the correlations between TFP and the individual wedges. Table 5 shows that the higher correlation between TFP and TFPR in the service sector, reflects a higher correlation of productivity with the capital wedge and lower correlations with the labor and output wedges.

Figure 3, which depicts the relationship between firm-level scaled productivity and scaled wedges, allows us to characterize further the size dependent distortions prevailing in the two sectors.[25] From the figure, we conclude that less productive firms are being subsidized, on average, both in manufacturing and in the service sector (log of the scaled TFPR is negative for less productive firms in the bottom panel of Figure 3), and that these subsidies come from relatively lower capital and lower labor costs in these firms. Despite being obtained from firm-level data, the implicit correlations (steeper slope for the capital wedge in the service sector and steeper slope for the labor wedge in the manufacturing sector) appear fully consistent with the sector-level correlations in Table 5. But, what distinguishes the two sectors, in qualitative terms,

---

[25] Figures 3 and 4 are obtained by fitting a kernel-weighted local polynomial smoothing to the data, with kernel=epanechnikov and degree=1.

is the output wedge. While in the manufacturing sector, firms, on average, appear as facing output "taxes" (negative figures of the log scaled output wedge) for a large range of scaled productivity levels, which offset to some extent the capital and labor "subsidies" in terms of scaled TFPR, in the service sector firms appear to benefit from output "subsidies", on average, which add to capital and labor "subsidies" in the case of less productive firms. This evidence in Figure 3 translates to Table 5 in terms of a significantly lower correlation between productivity and the output wedge in the service sector (0.047), compared to the manufacturing sector (0.155).[26] In order to evaluate the impact of the output distortions for the correlation between TFP and TFPR, we also computed the correlation between TFP and TFPR excluding the output wedge component (see equation (13)). The numbers reported in the last row of Table 5 show that if we drop the impact of the output wedge, the correlation between TFP and TFPR changes significantly in the service sector, and becomes about the same in the two sectors.

Summing up, the capital and labor wedge have correlations with productivity that differ across sectors but that tend to offset each other, so that ultimately it is the difference between the output wedge in the two sectors that is responsible for the higher correlation between TFP and TFPR in the service sector. At first sight, it is not clear why the output wedge should be so different in the two sectors. If we look at equations (7) and (11), which define the scaled output wedge, we realize that firms that misreport sales (for tax reasons, for instance) will tend to show up in the model as less productive firms, both in terms of TFP and of TFPR and so, as benefiting from output subsidies (they appear as being producing more than what they should given their TFP levels). The anecdotal evidence suggests that informality is higher in the service sector. Part of this informality stems from characteristics of the sector that make the enforcement activity of tax collection much more difficult than in the manufacturing sector. We believe that this might be part of the story behind the

---

[26]Note that, according to equation (13), the lower the correlation between productivity and the output wedge the higher the correlation between productivity and TFPR.
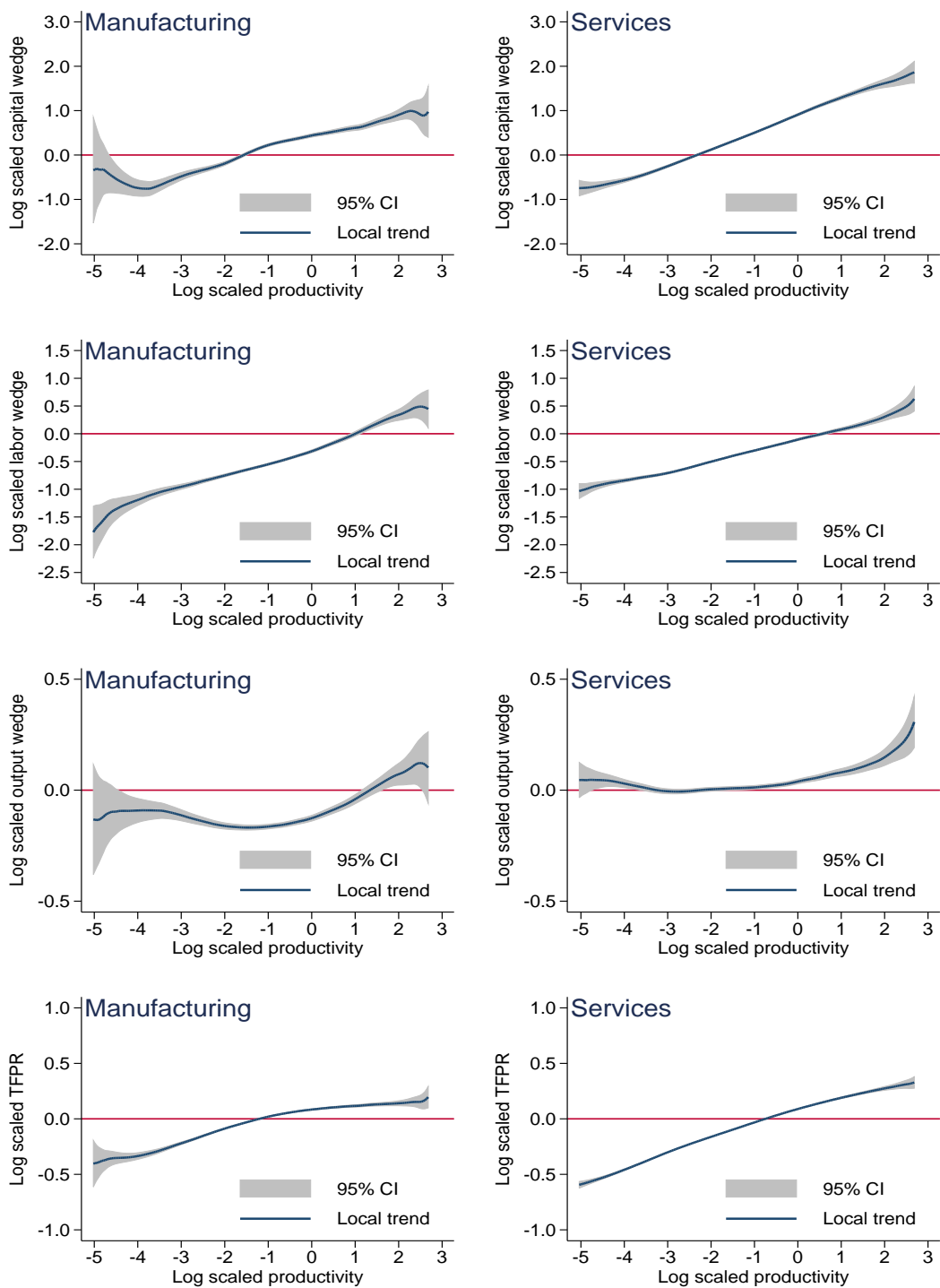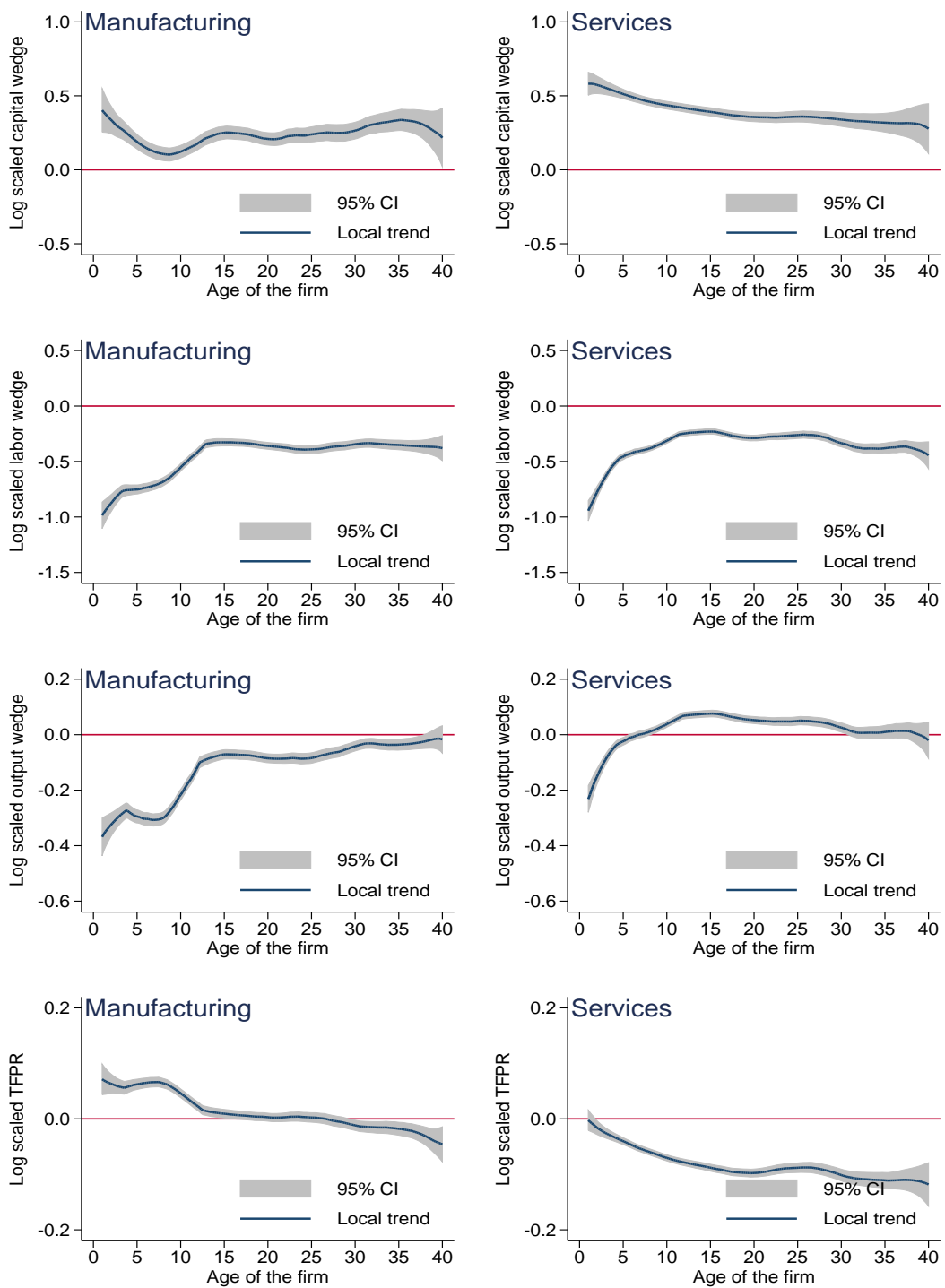
Figure 3: Wedges and productivity - 2008

Figure 4: Wedges and age of the firm - 2008

documented difference between the two sectors, but this is certainly an issue deserving further investigation.

Finally, according to the model, the proportion of young firms also has a bearing on the difference of misallocation between the two sectors.[27] From Table 3, we conclude that this regressor contributes with two opposite effects to the difference in misallocation between the manufacturing and the service sectors. On the one hand, the impact of the difference in the mean of this regressor between the manufacturing an the service sectors (characteristics effect) contributes to increasing the difference in misallocation between the two sectors (2.1 pp in 2008), but the difference in the impact of the regressor between the two sectors (coefficients effect) has the opposite effect (-5.7 pp), so that the total impact of this regressor is negative (-3.6 pp). This means that the impact on misallocation of the proportion of young firms is lower in the service sector, and that, in the absence of this effect, the difference in misallocation between the service and the manufacturing sectors would be even higher.

Figure 4, which depicts the relationship between firms' age and the scaled wedges, shows that despite benefiting from lower labor costs (lower scaled labor wedge), young firms, on average, face higher distortions (higher TFPR) than older firms, stemming from higher capital costs and higher output distortions. We link the higher capital costs to the presence of credit constraints imposed by financial institutions on young firms, because they have no credit history or because they have insufficient guaranties.[28] From Table 4, we see that the correlation between the proportion of young firms and the standard deviation of the individual wedges is lower in the service sector for the three wedges, suggesting that all the three distortions contribute to the lower impact of this regressor in the service sector (negative coefficients effect). Nevertheless, as in the case of the skewness regressor, the output wedge appears to be the main responsible for the differences in the impact of this regressor between the two sectors: output distortions

---

[27]We proxy the importance of young firms at the industry level, by the ratio of the number of firms with 3 years of age or less to the total number of firms.

[28]Note the qualitative differences vis-à-vis the evidence in Figure 3, where small firms appear as benefitting from capital subsidies, on average.

in young firms are much less important in the service than in the manufacturing sector, contributing significantly for the smaller contrast between younger and older firms in the former.

6   Conclusions

The empirical literature on misallocation has recently documented that the level of allocative efficiency in the service sector is significantly lower than that of the manufacturing sector. Because services are, by far, the most important sector of activity in developed economies, significantly higher levels of misallocation in this sector have important implications for aggregate TFP and aggregate GDP.

Using firm-level data for the Portuguese economy, we document that the significantly higher levels of allocative inefficiency in the service sector are not the result of a small number of industries with abnormal levels of inefficiency, but rather the outcome of a strong regularity. The great majority of the industries belonging to the manufacturing sector rank among the industries with the lowest misallocation. Conservative estimates for Portugal suggest that resource misallocation in 2008 was around 24 percentage points higher in the service than in the manufacturing sector. Closing this misallocation gap, that is, reducing misallocation in the service sector to the manufacturing levels would lead to a 12 percent boost in aggregate gross output (or aggregate TFP) and a 31 percent boost in aggregate value added (GDP).

Using regression analysis, we are able to fully explain the difference between efficiency gains in the two sectors. Productivity shocks, which capture the impact of (capital/labor) adjustment costs and/or output-price rigidity on misallocation, emerge as the most important factor contributing to the differences of misallocation between the two sectors. Such contribution stems more from differences of the impacts, than from differences of the shocks between the two sectors. The bulk of the difference in misallocation due to productivity shocks is likely to originate from the presence of

higher output-price rigidity and higher labor adjustment costs in the service sector.

The sectoral structure, which captures the impact on misallocation of size-dependent distortions, and is proxied by the skewness of the productivity distribution, emerges as the second most important factor to explaining the difference of misallocation between the two sectors. Also in this case, the bulk of the contribution comes from differences of the impact and not from differences of the mean of the regressor between the two sectors. The higher impact of this regressor in the service sector stems mainly from a higher level of informality that makes the enforcement activity of tax collection more difficult than in the manufacturing sector.

Finally, the empirical model suggests that the proportion of young firms also has a bearing on the difference of misallocation between the two sectors, but its impact in the service sector is lower. We link this regressor to the presence of barriers to growth, stemming, for instance, from credit constraints imposed by financial institutions on young firms, because they have no credit history or because they have insufficient guaranties.

The fact that we are able to fully explain the differences of allocative efficiency between the manufacturing and the service sector suggests that such differences originate from identifiable theoretical sources of misallocation and not from higher unexplainable heterogeneity in the service sector that would make the Hsieh and Klenow (2009, 2011) methodology not applicable to the service sector.

Our findings have important implications for economic policy. A significant part of the difference of allocative efficiency between the two sectors may be attributed to higher output price rigidity in the service sector. To the extent that such higher price rigidity stems from lower competition, measures aimed at increasing product market competition in the service sector will contribute to increase allocative efficiency in this sector and thus, will boost aggregate productivity.

Less productive firms appear as benefitting from capital and labor subsidies, which suggests that there might be a trade-off between employment creation and aggregate

productivity. Thus, size-contingent laws passed by the economic authority and aimed at boosting employment creation in small or medium-sized firms (special lines of credit with subsidized interest rates and/or labor subsidies), to the extent that they contribute to the survival of unproductive firms, especially in the service sector, where competition is weaker, will increase misallocation and have a strong negative impact on aggregate productivity. Finally, eliminating or minorating the impact of financial constraints on young firms would also contribute to increase aggregate productivity.

Appendix : Misallocation under alternative assumptions

In this Appendix we investigate how changes in assumptions about the model impact the estimates for misallocation in the manufacturing and service sectors. In order to take the model to the data it is necessary to choose a value for the elasticity of substitution parameter (σ), decide how to treat outliers, and choose the group of firms included in the analysis. Although it is already known that these assumptions will impact the estimated level of misallocation (see Hsieh and Klenow (2009) or Dias et al. (2016)), our interest is knowing whether the effect is homogeneous across industries and thus whether they significantly affect the difference of misallocation between the manufacturing and service sectors.

As is common in the literature (see, for instance, Hsieh and Klenow (2009), Dias et al. (2016), Gopinath et al. (2015)), we use the wage bill paid by the firm (total labor costs) to measure labor input. Implicitly, we are assuming that wages per worker adjust for firm differences in hours worked per worker and worker skills. For the rental price of capital, we define an industry specific price equal to the industry depreciation rate plus a 5 percent real interest rate, so that $R_s = \delta_s + 0.05$. For the intermediate inputs, we make a similar assumption as in the case of the labor input, and assume that the price of intermediate products, Z_s, is equal to 1, so that the expenditure on

intermediate inputs reflects not only the amount of inputs but also their quality.[29]

In line with other studies (see, for instance, Hsieh and Klenow (2009), Ziebarth (2013) and Dias et al. (2016)), we define a baseline by making the following set of assumptions: i) the elasticity of substitution, σ, equal to 3; ii) trimming the top and bottom 1.0 percent tails of scaled TFPR and TFP distributions across industries;[30] iii) inclusion of all firms in the retained industries.[31]

Next, we change assumptions i), ii) and iii) to gauge the impact each one and altogether have on the estimated efficiency gains and, more importantly, on the difference of misallocation between the service and manufacturing sectors. Specifically, we will investigate the implications of changing σ and the level of trimming, as well as of leaving aside smaller firms.

The efficiency gains for 2008 and 2010, obtained under the baseline assumptions, are recorded in Table A1. We can see from the first row that, if distortions in the economy were eliminated (by equalizing TFPR across firms in each industry and keeping industry level factor demand constant), the gross-output efficiency gains (or TFP gains) for the whole economy would be around 43 percent in 2008 and 49 percent in 2010 (this figure also includes firms from Agriculture). Efficiency gains are also clearly higher in the service sector (around 59 percent in 2008 and 66 percent in 2010) than in the manufacturing sector (around 16 and 17 percent in 2008 and 2010, respectively). Thus,

---

[29]Note that the values of W_s, R_s and Z_s affect the corresponding average wedges, but not the relative comparison between firms in a given industry. In other words, the choices of W_s, R_s and Z_s affect the estimates of the wedges (capital, labor, and output wedges), but not the efficiency gains calculated in this paper. Note also that to compute efficiency gains we only need to compute $TFPR_{si}$, $TFPQ_{si}$ (or A_{si}) and $TFPR_{si}^*$, and that factor prices that are common to all firms in an industry cancel out.

[30]That is, the distributions of $\ln(TFPR_{si}/TFPR_s^*)$ and $\ln(A_{si}M_s^{\frac{1}{\sigma-1}}/TFP_s^*)$ where TFP_s^* is the industry efficient TFP defined in Section 1. More specifically, we pool all industries and omit the top and bottom 1 percent of firms in each of the pools to eliminate outliers and to control for possible measurement error. We then recalculate the relevant industry level aggregates such as L_s, K_s, Q_s, P_sY_s, $TFPR_s^*$ and TFP_s^*.

[31]In order to avoid computing misallocation with a very small number of firms, we drop industries that are left with less than 10 firms after the trimming. This condition is imposed in all the variants considered in Table A1 below, to ensure comparability. After excluding industries with less than 10 firms, we are left with 162 different industries for 2008 (7 for agriculture, 80 for manufacturing and 75 for services) and 163 industries for 2010 (8 for agriculture, 79 for manufacturing and 76 for services).

Table A1: Efficiency gains under alternative assumptions

	2008				2010			
Assumptions	Total	M	S	S-M	Total	M	S	S-M
1) Baseline model	43.36	16.02	59.19	43.18	49.33	16.81	66.46	49.65
2) $\sigma = 4.5$	63.48	19.93	90.65	70.72	71.55	20.33	100.84	80.51
3) Trimming=2.5%	36.35	16.36	47.43	31.07	40.00	17.12	51.42	34.30
4) Employment>10	28.31	12.92	38.33	25.41	31.37	13.43	41.68	28.25
5)= 2)+3)	48.98	19.73	66.01	46.29	53.33	19.61	71.09	51.49
6)= 2)+4)	34.97	15.13	48.18	33.05	38.85	15.71	52.52	36.82
7)= 3)+4)	23.56	12.08	30.88	18.80	26.05	12.50	33.65	21.15
8) Final model	28.46	14.15	37.66	23.51	31.28	14.43	40.82	26.39

Note: Efficiency gains in the baseline model are computed taking all the firms in the dataset, assuming σ=3.0 and trimming 1 percent tails of $log(TFPR_{si}/TFPR_s^*)$ and $log(A_{si}M^{\frac{1}{\sigma-1}}/TFP_s^*)$. M stands for manufacturing and S for services. S-M is the difference between the service and the manufacturing sectors. The total also includes firms from agriculture.

the service sector emerges as far more inefficient than the manufacturing sector in line with the results in Dias et al. (2016). However, one question that may arise here is whether the documented difference in misallocation between the two sectors can be explained by one or more of the assumptions that underlie the baseline results. We thus now consider the implications of alternative assumptions to the baseline.

Elasticity of substitution

It is well known that the assumed value of σ has important implications for the estimated efficiency gains in the context of the Hsieh and Klenow (2009) model (see, for instance, Hsieh and Klenow (2009), Dias et al. (2016)), and that higher values of σ are associated with higher levels of efficiency gains. In the absence of country specific estimates, the empirical literature usually assumes σ=3 as a way of computing a conservative estimate for the importance of misallocation. However, the empirical estimates for industry markups usually suggest higher values for σ, both in the Eurozone as in the U.S. (see, for instance, Christopoulou and Vermeulen (2012)). The estimates available for Portugal for the 2004-2009 period (see Amador and Soares (2013)) are also on average above 3. In particular, the 3-digit industry level estimates for the markups, under the assumption of imperfect labor markets, imply an (non-weighted) average

of σ=4.5 for the Portuguese economy. Moreover, there is not a significant difference between the average estimates for manufacturing and the service sector. Thus, in what follows, we use σ=4.5, which is a more realistic number for the Portuguese economy. From Table A1 (second row), we see that the estimated efficiency gains increase vis-à-vis the baseline in the two sectors. Moreover, the increase in the service sector is significantly higher implying that the difference between the two sectors increases from around 43 percentage points (pp) to around 71 pp in 2008, and from around 50 pp to around 81 pp in 2010.

Treatment of outliers

The presence of outliers also has strong implications for the empirical estimates of efficiency gains. For example, if a firm mistakenly reports very low input factors it will generate very large numbers for TFP and TFPR, giving rise to spurious misallocation.[32] One way to deal with this problem is to trim the distributions of scaled TFP and scaled TFPR.[33] Of course, the choice of the trimming is largely ad-hoc and it has implications for the results, especially in cases of large measurement errors. The way changes in the trimming affects the difference of efficiency gains between the service and the manufacturing sectors will depend on how outliers are distributed across the two sectors.

Table A1 (third row) also reports the estimates for TFP gains when we trim 2.5 percent on each tail of the scaled TFP and TFPR distributions. As could be expected the estimated TFP gains for the whole economy are lower. For instance, for 2008, efficiency gains drop from 63 to 49 percent if we use σ=4.5 (rows 2 and 5), or from 43 to 36 percent if we use σ=3 (rows 1 and 3). More interestingly, however, we see that estimated efficiency gains decrease in the service sector, but remain basically unchanged in the manufacturing sector, so that the difference between the two sectors is reduced

---

[32]Note that, in the model, both TFP and TFPR are computed using the production function, i.e., $TFP_{si} = A_{si} = Y_{si}/\left(K_{si}^{\alpha_s} H_{si}^{\beta_s} Q_{si}^{1-\alpha_s-\beta_s}\right)$ and $TFPR_{si} = P_{si}A_{si} = P_{si}Y_{si}/\left(K_{si}^{\alpha_s} H_{si}^{\beta_s} Q_{si}^{1-\alpha_s-\beta_s}\right)$. Thus, reporting mistakes involving the inputs or the gross output (sales) translate into abnormal TFP and TFPR figures, which directly affect estimates of efficiency gains.

[33]Another one could be to compute winsorized efficiency gains at a given percent level.

from 71 to 46 percent in the model with $\sigma=4.5$ (or from 43 to 31 percent in the model with $\sigma=3$), thanks only to changes in the service sector. This result shows that a significant part of the baseline difference between the two sectors is due to outliers that are concentrated in the service sector.

Minimum firm size

Another important issue is that of the minimum firm size to consider in the exercise. The results in Dias et al. (2016) suggest that an important part of the misallocation problems due to distortions in the Portuguese economy is concentrated in the micro and small service sector firms. Technically, it is possible to compute the efficiency gains stemming from labor distortions for firms with one or more employees. However, it is unclear whether micro or very small firms should be included in the computation of efficiency gains. In fact, the reallocation gains in some of these firms may be somewhat overestimated, either because the model does not allow for indivisibilities in the input factors (labor force or capital stock), which mainly affect micro and small firms, or because some of these units (young firms) might be growing at a faster pace, as they are in the process of converging to their optimal size. Furthermore, because enforcement activity of tax collection focuses more on large firms, micro and small firms are more likely to exhibit noncompliance with taxes by misreporting some relevant items like sales or gross output, also implying higher distortions.

Table A2 reports the estimated efficiency gains corresponding to different restrictions on firm's size, as measured by the number of employees, under the baseline assumptions. In particular, we compute the efficiency gains within micro and small firms vis-à-vis efficiency gains within larger firms, for different cut-offs of employment. A very important conclusion is that efficiency gains are much higher within micro and small firms, than within larger firms. For instance, efficiency gains for the whole economy in 2008, within firms with 5 or less employees, are around 87 percent, compared to 32 percent within firms with more than 5 employees (43 percent if we take the firms altogether). If we restrict the sample to firms with between 6 and 10 employees, ef-

ficiency gains are about 45 percent, i.e., significantly lower than within firms with a number of employees between 1 and 5, but, nevertheless, significantly higher than the efficiency gains within firms with more than 10 employees (28 percent). As the size of firms (employment cut-offs) increases, the difference in the efficiency gains between small and big firms shrinks. In particular, efficiency gains for firms with employment between 10 and 15 employees are close to the efficiency gains for firms with more than 15 employees, especially in the service sector.

Table A2: Efficiency gains for different employment cut-offs
(Baseline assumptions)

	2008			2010		
	Total	Manuf	Serv	Total	Manuf	Serv
All firms	43.36	16.02	59.19	49.33	16.81	66.46
Empl<=5	87.04	34.95	95.10	89.74	36.91	97.89
Empl>5	31.54	13.74	42.39	34.21	14.59	44.83
5 <Empl<=10	44.54	22.03	49.76	48.50	23.46	53.88
Empl<=10	71.15	28.03	79.34	77.00	29.77	86.14
Empl>10	28.31	12.92	38.33	31.37	13.43	41.68
10 <Empl<=15	34.91	19.57	39.91	42.96	20.71	50.17
Empl<=15	64.74	25.63	73.43	71.55	27.32	81.44
Empl>15	26.79	12.47	36.44	29.66	12.73	39.72
15 <Empl<=20	31.78	14.83	39.77	35.22	17.76	42.03
Empl<=20	62.05	24.98	71.64	67.14	25.64	77.45
Empl>20	25.89	12.01	35.42	29.55	13.96	38.92

Note:Baseline assumptions: σ=3.0 and trimming 1 percent tails of $\log(TFPR_{si}/TFPR_s^*)$ and $log(A_{si}M^{\frac{1}{\sigma-1}}/TFP_s^*)$.

Overall, the results in Table A2 show that heterogeneity (efficiency gains) within small firms is higher than heterogeneity between small and large firms. This is an interesting result on its own. Indivisibilities, faster growth or financial constraints faced by small firms could explain heterogeneity between small firms and large firms, and thus explain higher efficiency gains in industries with a higher proportion of small firms. However, these factors cannot explain the high heterogeneity uncovered within small firms. We believe that misreporting of some relevant items, like sales or gross output might be an explanation for some results in Table A2.

Thus, for the purpose of the present exercise, we restrict the analysis to firms with more than 10 employees. The chosen cut-off is somewhat ad-hoc but we believe that given the importance of small and medium-sized firms in the Portuguese economy this solution strikes a balance between the need to reduce the importance of spurious misallocation and the representativeness of the final sample. By dropping firms with 10 or less employees, the number of firms in the dataset is reduced from 236,022 to 41,123 in 2008 and from 230,157 to 38,675 in 2010. Despite representing around 83 percent of the total number of firms, firms with 10 or less employees account only for 16.8 percent of total gross output and 25.4 percent of total employment in 2008 (the figures for 2010 are similar). From Table A1, we see that if we drop firms with 10 or less employees from the dataset, the efficiency gains for the whole economy are reduced from around 43 to about 28 percent (row 4), and the difference between the two sectors is reduced from 43 pp to 25 pp in 2008 and from 50 pp to 28 pp in 2010.

Finally, if we consider the three changes to the baseline altogether (σ=4.5, trimming=2.5 and employment>10), we get what we denote the "final model" assumptions. Under this set of assumptions, the efficiency gains for the whole economy drop to about 28 percent in 2008, and to about 31 percent in 2010 (see last row in Table A1). In turn, the differences between the service and the manufacturing sectors drop to about 24 pp in 2008 and to about 26 pp in 2010.

Summing up, the evidence in this Appendix shows that the existence of a significant difference in allocative efficiency between the manufacturing and service sectors does not rest on possible the assumptions about the elasticity of substitution parameter (σ), the way outliers are dealt with or the type of firms included in the analysis.[34]

---

[34]It has been claimed (see, for instance, Vollrath (2015)) that, for historical reasons, data collection and classification is skewed towards manufacturing in the sense that it provides a much more detailed industry classification in manufacturing than in the service sector. In order to understand how the estimates of the difference of efficiency gains between the two sectors depend on the level of industry classification, we computed the efficiency gains using a 5-digit industry definition. Compared to the 3-digit results in Table A1, we conclude that global efficiency gains decrease, but not by much, which means that misallocation is not a spurious outcome of aggregation. Efficiency gains stay virtually unchanged in the manufacturing sector, but decrease somewhat in the service sector, reflecting the impact of a higher increase in disaggregation in this sector. Nevertheless, the difference of efficiency gains between the two sectors remains very high. For instance, in the "final model" case of Table

References

Adamopoulos, T. and Restuccia, D. (2014), 'The size distribution of farms and international productivity differences', *American Economic Review* **104**(6), 1667–97.

Amador, J. and Soares, A. C. (2013), Competition in the portuguese economy: Estimated price-cost margins under imperfect labour markets, Working Paper 8, Banco de Portugal.

Asker, J., Collard-Wexler, A. and De Loecker, J. (2014), 'Dynamic inputs and resource (mis)allocation', *Journal of Political Economy* **122**(5), 1013 – 1063.

Banerjee, A. V. and Moll, B. (2010), 'Why does misallocation persist?', *American Economic Journal: Macroeconomics* **2**(1), 189–206.

Bartelsman, E., Haltiwanger, J. and Scarpetta, S. (2013), 'Cross-country differences in productivity: The role of allocation and selection', *American Economic Review* **103**(1), 305–334.

Bellone, F. and Mallen-Pisano, J. (2013), Is misallocation higher in France than in the United States?, GREDEG Working Papers 2013-38, Groupe de REcherche en Droit, Économie, Gestion (GREDEG CNRS), University of Nice-Sophia Antipolis.

Benkovskis, K. (2015), Misallocation of resources in Latvia: did anything change after the crisis?, Working Paper 5, Bank of Latvia.

Blinder, A. S. (1973), 'Wage discrimination: Reduced form and structural estimates', *Journal of Human Resources* **8**(4), 436–455.

Buera, F. J., Kaboski, J. and Shin, Y. (2011), 'Finance and development: A tale of two sectors', *American Economic Review* **101**(5), 1964–2002.

A1, the difference in the 5-digit case is 17.64 pp, in 2008 (compared to 23.51 pp in the 3-digit case), and 21.67 pp in 2010 (compared to 26.39 pp in the 3-digit case). Thus, in the paper we focus on the 3-digit case, not only because it does not make a significant difference in quantitative terms, but also because in the 5-digit case, industries have, on average, a smaller number of firms, implying more volatile estimates of industry-level efficiency gains and of the regressors used in the analysis.

Calligaris, S. (2015), Misallocation and total factor productivity in Italy: Evidence from firm-level data, Research Paper Series 357, Centre for Economic and International Studies.

Camacho, A. and Conover, E. (2010), Misallocation and productivity in Colombiaï¿½s manufacturing industries, Working Paper IDB-WP-123, Inter-American Development Bank.

Chen, K. and Irarrazabal, A. (2015), 'The role of allocative efficiency in a decade of recovery', *Review of Economic Dynamics* **18**(3), 523–550.

Christopoulou, R. and Vermeulen, P. (2012), 'Markups in the euro area and the us over the period 1981-2004: a comparison of 50 sectors', *Empirical Economics* **42**(1), 53–77.

David, J. M., Hopenhayn, H. A. and Venkateswaran, V. (2014), Information, misallocation and aggregate productivity, NBER Working Paper 20340, National Bureau of Economic Research, Inc.

Dias, D. A., Marques, C. R., Martins, F. and Santos Silva, J. M. (2015), 'Understanding price stickiness: Firm level evidence on price adjustment lags and their asymmetries', *Oxford Bulletin of Economics and Statistics* **77**(5), 701–718.

Dias, D. A., Marques, C. R. and Richmond, C. (2016), 'Misallocation and productivity in the lead up to the Eurozone crisis', *Journal of Macroeconomics* **49**, 46–70.

Duarte, M. and Restuccia, D. (2010), 'The role of the structural transformation in aggregate productivity', *The Quarterly Journal of Economics* **125**(1), 129–173.

ECB (2006), Competition, productivity and prices in the euro area services sector, Working Paper Series 44, European Central Bank.

Ellison, G. and Glaeser, E. L. (1997), 'Geographic concentration in u.s. mnufacturing industries: A dartboard approach', *Journal of Political Economy* **105**(5), 889–927.

Fabiani, S., Druant, M., Hernando, I., Kwapil, C., Landau, B., Loupias, C., Martins, F., Mathä, T., Sabbatini, R., Stahl, H. and Stokman, A. (2006), 'What firms' surveys tell us about price-setting behavior in the Euro Area', *International Journal of Central Banking* **2**(3).

Garcia-Santana, M., Moral-Benito, E., Pijoan-Mas, J. and Ramos, R. (2015), Growing like Spain: 1995-2007, mimeo.

Garicano, L., LeLarge, C. and Reenen, J. V. (2013), Firm size distortions and the productivity distribution: Evidence from france, Working Paper 18841, National Bureau of Economic Research.

Gelbach, J. B. (2016), 'When do covariates matter? And which ones, and how much?', *Journal of Labor Economics* **34**(2), 509–543.

Gilchrist, S., Sim, J. W. and Zakrajsek, E. (2013), 'Misallocation and financial market frictions: Some direct evidence from the dispersion in borrowing costs', *Review of Economic Dynamics* **16**, 159–176.

Gopinath, G. and Itskhoki, O. (2010), 'Frequency of price adjustment and pass-through', *The Quarterly Journal of Economics* **125**(2), 675–727.

Gopinath, G., Kalemli-Ozcan, S., Karabarbounis, L. and Villegas-Sanchez, C. (2015), Capital allocation and productivity in south europe, Working Paper 21453, National Bureau of Economic Research.

Gourio, F. and Roys, N. (2014), 'Size-dependent regulations, firm size distribution, and reallocation', *Quantitative Economics* **5**, 377–416.

Guner, N., Ventura, G. and Yi, X. (2008), 'Macroeconomic Implications of Size-Dependent Policies', *Review of Economic Dynamics* **11**(4), 721–744.

Holmes, T. J. and Schmitz, J. A. (2010), Competition and productivity: A review

of evidence, Research Department Staff Report 439, Federal Reserve Bank of Minneapolis.

Hosono, K. and Takizawa, M. (2012), Do financial frictions matter as a source of misallocation? Evidence from Japan, PRI Discussion Paper Series 12A-17, Policy Research Institute.

Hsieh, C.-T. and Klenow, P. J. (2009), 'Misallocation and manufacturing TFP in China and India', *The Quarterly Journal of Economics* **124**(4), 1403–1448.

Hsieh, C.-T. and Klenow, P. J. (2011), 'Misallocation and manufacturing TFP in China and India: Correction appendix', *Stanford University. Unpublished* .

Jann, B. (2008), 'The Blinder-Oaxaca decomposition for linear regression models', *The Stata Journal* **8**(4), 453–479.

Machicado, C. G. and Birbuet, J. C. (2012), 'Misallocation and manufacturing TFP in Bolivia during the market liberalization period', *The B.E. Journal of Macroeconomics* **12**(1), 1–45.

Martin, C. (1993), 'Price adjustment and market structure', *Economics Letters* **41**(2), 139–143.

Midrigan, V. and Xu, D. (2014), 'Finance and misallocation: Evidence from plant-level data', *American Economic Review* **104**(2), 422–458.

Moll, B. (2014), 'Productivity losses from financial frictions: Can self-financing undo capital misallocation?', *American Economic Review* **104**(10), 3186–3221.

Oaxaca, R. (1973), 'Male-female wage differentials in urban labor markets', *International Economic Review* **14**(3), 693–709.

Peters, M. (2013), Heterogeneous mark-ups, growth and endogenous misallocation, mimeo, London School of Economics.

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

Song, Z. and Wu, G. (2013), A structural estimation on capital market distortions in Chinese manufacturing, Economic Growth Centre Working Paper Series 1306, Nanyang Technological University, School of Humanities and Social Sciences, Economic Growth Centre.

Syverson, C. (2004*a*), 'Market structure and productivity: A concrete example', *Journal of Political Economy* **112**(6), 1181–1222.

Syverson, C. (2004*b*), 'Product substitutability and productivity dispersin', *The Review of Economics and Statistics* **86**(2), 534–550.

Vollrath, D. (2015), There is more to life than manufacturing, The growth economics blog, https://growthecon.wordpress.com/.

Ziebarth, N. (2013), 'Are China and India backwards? Evidence from the 19th century U.S. Census of Manufactures', *Review of Economic Dynamics* **16**(1), 86–99.