UNIVERSITÀ
DI SIENA
1240

QUADERNI DEL DIPARTIMENTO
DI ECONOMIA POLITICA E STATISTICA

Michele Battisti
Filippo Belloc
Massimo Del Gatto

Labour Productivity and Firm-Level TFP with Technology-Specific Production Functions

n. 789 – Novembre 2018

Labour Productivity and Firm-Level TFP with Technology-Specific Production Functions*

Michele Battisti
University of Palermo, CeLEG LUISS Guido Carli and RCEA (michele.battisti@unipa.it)

Filippo Belloc
University of Siena (filippo.belloc@gmail.com)

Massimo Del Gatto
"G.d'Annunzio" University and CRENoS (massimo.delgatto@unich.it)

Abstract

We investigate the technological dimension of productivity, presenting an empirical methodology based on mixture models to disentangle the labor productivity differences associated with the firm's choice of technology (BTFP) and those related to the firm's ability to exploit the adopted technology (WTFP). The estimation endogenously determines the number of technologies (in the sector) and degree of technology sharing across firms (i.e., for each firm, the probability of using a given technology). By using comparable data for about 35,000 firms worldwide distributed across 22 (two-digit) sectors, we show BTFP to be at least as important as WTFP in explaining the labor productivity gaps across firms. Intra-sectoral and inter-sectoral heterogeneity is substantial and, even in sectors in which BTFP dominates on average, we find a considerable number of firms for which labor productivity is mostly determined by the ability to use the adopted technology. Hence, dissecting the labor productivity gaps is crucial to achieving more targeted innovation policies. The estimated number of technologies ranges from one (in only three industries) to five, being three in most cases. The suggested estimation strategy takes simultaneity into account. The BTFP measure is unaffected by omitted price bias. The presence of BTFP dispersion can be associated with the action of frictions preventing firms from switching to superior and more productive technologies. Eliminating BTFP does not eliminate misallocation.

Keywords: TFP, technology adoption, production function estimation, mixture models.

JEL Classification: D24, 033, C29.

*We thank conference participants at the 2018 ESCoE Conference on Economic Measurement, hosted by the Bank of England, the 2018 North American Productivity Workshop, hosted by the Miami Business School, and the 11th ISGEP Workshop on Firms, Trade and Productivity at "G.d'Annunzio" University.

1 Introduction

Labor productivity in OECD countries has grown by 3.7% since 1990 compared with 28.4% in emerging economies (OECD, 2015). Against the 7.5% growth in the United States, the usual productivity benchmark, the rate peaks at 63.8% in China and 38.4% in India. These trends have recently sparked widespread interest in quantifying productivity and analyzing its determinants.

Measured as the estimated residual of a sector-specific production function, total factor productivity (TFP) is usually found to be as important as capital accumulation in explaining the process of labor productivity growth (Kumar and Russell, 2002; Caselli, 2005; Hsieh and Klenow, 2009, 2010; Battisti *et al.*, 2018). However, under its conventional representation, TFP is a wide notion encompassing several hard-to-measure factors (i.e., "all the rest", respect to capital accumulation). As a result, empirical economic analysis explains no more than 50% of the productivity differentials. This fact notwithstanding, and despite the potential impact in terms of policy implications, the literature dealing with the different theoretical conceptions of TFP is unsatisfactory and it is still true that "economists should devote more effort toward modeling and quantifying TFP" (Easterly and Levine, 2001).

While these considerations apply to both aggregate and firm-level studies, the firm-level perspective has recently come to the fore, with a growing interest (see Gabler and Poschke (2013), Da Rocha *et al.* (2017), and, for an early review, Grossman and Helpman (2015)) in the link between the evolution of aggregate productivity and cross-firm technological heterogeneity, as induced by the behavior of new entrants (Sampson, 2016), technology diffusion and imitation (Luttmer, 2007; Perla and Tonetti, 2014; Perla *et al.*, 2015), innovation (Benhabib *et al.*, 2017), and international trade and competition (Bloom *et al.*, 2016).

The importance of dealing with the link between technology and productivity from a firm-level perspective has also entered the policy debate. However, the chance to translate the empirical results into effective policy prescriptions is hampered by the fact that the terms "productivity", "TFP", "labor productivity", "innovation", and "technology" are often used interchangeably. As a matter of fact, the TFP and/or labor productivity differences across firms are commonly interpreted in terms of technological differentials, but without a clear interpretation.

For instance, OECD (2015) documents the rising (labor) productivity gap between firms at the global (labor) productivity frontier and other firms. According to the report, this raises questions about the obstacles that prevent laggard firms from adopting seemingly well-known frontier technologies.[1]

The reason for this ambiguity is computational. Abstracting from the functional form (e.g., Cobb–Douglas versus translog), the technological dimension of firms' productivity should arguably reflect the estimated production function coefficients (i.e., intercept and input shares). This is a dimension that cannot be extrapolated from the data using standard econometrics, as it would require clustering firms ex-ante according to the technology they adopt. This is not possible as far as technology is the unknown dimension.

Attempts to examine this aspect in depth have been carried out using case studies. For example, by referencing the steel industry, Collard-Wexler and De Loecker (2014) study the firm-level and industry-level productivity

[1] Indeed, the report recognizes that "given the difficulties in measuring technology, the globally most productive firms are also assumed to operate with the globally most advanced technologies but it should be recognized that very technologically advanced firms might not necessarily appear as the globally most productive or the most successful in terms of profits" (OECD, 2015 page 33).

increases associated with the adoption of a particular technology (the "minimill" technology). Instead, since technologies are not observable in general, the econometrician can only estimate sector-level production function parameters. Under this standard approach, cross-firm differences in output not associated with different input choices flow entirely into the residual TFP term (i.e., the "Solow residual" measured as the difference between actual and predicted output).

In this study, we aim to shed light on the magnitude of the technological dimension of productivity in a world in which different production functions identify different technologies (by differing in terms of intercept and shape parameters). Our empirical strategy allows several technologies to be available in the same sector, with firms potentially adopting any of them. The estimation is left free to determine both the number of technologies and the degree of technology sharing across firms (i.e., for each firm, the probability of using a given technology). This approach allows us to disentangle the labor productivity differences associated with the firm's choice to adopt a certain technology, referred to herein as between-technology TFP (BTFP) and the labor productivity differences related to the firm's ability to exploit the adopted technology, referred to as within-technology TFP (WTFP). By applying our methodology to a sample of about 35,000 firms distributed across 22 (two-digit) manufacturing sectors and 76 countries, we find that the technological component is at least as important as the idiosyncratic component in explaining productivity differences. On average, the WTFP productivity of the top 5% of firms in a sector is around 31% higher than that of other firms, while the labor productivity gap between the firms using the frontier technology and those using less productive technologies amounts to around 34%. Within this general result, we document substantial intra-sectoral and inter-sectoral heterogeneity in terms of both measures, particularly when we look at their relative weight. Notably, the estimated number of technologies range from one (in a few industries) to five, being three in most cases. Hence, the hypothesis of a single technology, which is the standard implicit hypothesis in TFP estimation, is likely to hide substantial heterogeneity. In turn, this may suggest that dissecting labor productivity gaps is crucial to achieving more targeted innovation policy.

From a methodological point of view, operationalizing this multiple-technology framework entails tackling several issues.

First, as stated, we need to estimate the technology-specific production function parameters without clustering firms ex-ante. To this end, we rely on a *mixture model* (McLachlan and Peel, 2000). This approach treats the probability distribution of labor productivity as the result of the potential overlapping of several distributions, each reflecting a different technology. Adopting this approach allows us to identify the available technologies by estimating the different production functions characterized by different (cluster-specific) estimated parameters. The estimation provides us, for each sector, with the number of available technologies (i.e., clusters) and, for each firm, with the probability of belonging to each technology cluster.

Second, as a related issue, we have to square things up with the chance to potentially capture as many as possible of the worldwide available technologies. This requires using internationally comparable data with the largest possible coverage, although the country representativeness of firm-level databases tends to shrink with the number of countries included. To deal with this issue, we take advantage of the balance sheet data provided by Bureau van Dijk (2018), which show high information comparability for a large sample of worldwide distributed firms.

Third, as emphasized by the literature (Marschak and Andrews, 1944; Klette and Griliches, 1996), TFP estimation at the firm level incurs simultaneity problems as far as TFP is known to the firm at the stage of the input-level decision, which biases the OLS estimates. A well-known solution relies on the identification of an observable variable reacting to the variations in TFP perceived by the firm. To the extent that the relationship with TFP can be inverted, this proxy variable can be used to estimate the production function semi-parametrically.[2] Our multiple-technology framework amplifies this issue by bringing about additional potential simultaneity associated with the choice of technology. To deal with such simultaneity on both counts, we develop an "empirical model" of technology adoption similar in spirit to the semi-parametric approach, but that does not rely on proxy variables. This choice allows us to circumvent the problems related to inverting the relationship between the proxy variable and idiosyncratic productivity term recently highlighted by Gandhi *et al.* (2018).

Fourth, another source of bias in TFP estimates is the price dispersion across firms. Indeed, the production function should be estimated in physical terms, that is using firms' output. However, as these data are rarely available in firm-level datasets, it is common practice to base the estimation on firms' sales deflated by an industry-wide price index (individual prices are themselves usually unavailable). In this case, TFP is likely to be misstated to the extent firms apply different markups. Notably, being computed as the difference between the predicted values under different sets of technology-specific production function parameters, our BTFP measure is unaffected by firm-level differences in prices and markups.

Finally, our analysis intersects with the recent literature emphasizing the concept of "misallocation" (Alfaro *et al.*, 2008; Banerjee and Duflo, 2005; Bartelsman *et al.*, 2013; Hsieh and Klenow, 2009, 2010; Restuccia and Rogerson, 2008, 2013): when market imperfections hamper the flow of factors from less productive to more productive units, they result in lower aggregate TFP than the ideal situation of frictionless factor markets. Our empirical strategy adds to this literature by introducing an additional misallocation mechanism. Indeed, technology dispersion can be seen as one of the possible consequences of the presence of distortions preventing firms from switching to superior and more productive technologies. Two points are noteworthy. First, allowing all firms to use the frontier technology would not eliminate misallocation as long as they are not free to hire the desired amount of capital and labor Second, in the presence of technology dispersion, it is no longer possible to trace back the whole dispersion of revenue TFP to the presence of misallocation, as suggested by Hsieh and Klenow (2009) and subsequent studies aimed at quantifying misallocation.[3]

The remainder of the paper proceeds as follows. In Section 2, we introduce the basic intuitions behind our notions of the within-technology and between-technology components of TFP. In Section 3, we model the mechanism of technology adoption. In Section 4, we present our dataset and the variables we use in the empirical sections. In Section 5, we describe the mixture model used to estimate the production functions within technology clusters. In Section 6, we present our results and, in Section 7, we analyze possible country patterns and correlations. Finally, Section 8 concludes.

[2]See, among others, Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg *et al.* (2015), Wooldridge (2009), and Doraszelski and Jaumandreu (2012).

[3]See Garcia-Santana *et al.*. (2015), Gopinath *et al.* (2017), Dias *et al.* (2016), Bellone and Mallen-Pisano (2013), Crespo and Segura-Cayuela (2014), Chen and Irarrazabal (2014), and Calligaris *et al.* (2018). Asturias and Rossbach (2017) suggest that firms may face both economy-wide distortions in factor markets and distortions varying with the technology actually in use. In such a world, they find that the estimated gains from eliminating misallocation are reduced by around 60% with respect to the standard Hsieh and Klenow (2009) model.

2 Basic Intuition and Definition of the Terms

Suppose that M technologies are available in a given industry, with a number of firms using each technology. Adopting an index m to refer to technology (with $m = 1, \ldots, M$), the production function of the generic firm i can be written as

$$y_i = a_{i,m} + \alpha_m + \beta_m k_i, \tag{1}$$

where $y_i = ln(Y_i/L_i)$ and $k_i = ln(K_i/L_i)$. The parameters α_m and β_m capture the technological dimension of the production process (i.e., technology-specific intercept and slope, respectively), while $a_{i,m}$ encapsulates the firm-specific labor productivity differences across firms using the same technology.[4] To highlight the importance of disentangling these two dimensions and the role that technology can play in the evolution of output, Bernard and Jones (1996a, 1996b) use the expression "total technological productivity."

Equation (1) cannot be estimated with standard econometrics without an ex-ante assumption on the technology used by each firm. For this reason, the standard approach to production function estimation (see the surveys by Del Gatto *et al.* (2011) and Van Beveren (2012)) combines the information captured by $a_{i,m}$, α_m, and β_m into a generic TFP index computed as the difference between actual (y_i) and predicted ($\hat{y}_i$) labor productivity (i.e., as the Solow residual $\hat{a}_i = y_i - \hat{y}_i$, with $\hat{y}_i = \hat{\alpha} + \hat{\beta} k_i$ and the estimated coefficients denoted by $\hat{\alpha}$ and $\hat{\beta}$) under a one-technology assumption (i.e., $M = 1$). The cross-firm differences in this term, which are often referred to as "technology" or "productivity" interchangeably, are usually understood in terms of technological differentials but without a clear interpretation.

To understand the relationship between the conventional approach and the framework depicted by Equation (1), imagine perfectly knowing which firms, among those active in a given industry, adopt technology m, so that the production function in Equation (1) can be estimated within that specific group of firms. By yielding the technology-specific production function parameters, this allows us to express the TFP term as $\hat{a}_{i,m} = y_i - \hat{y}_{i,m}$, with $\hat{y}_{i,m} = \hat{\alpha}_m + \hat{\beta}_m k_i$. In this case, the residual is expressed in relative terms with respect to the "average" firm in the technology group rather than the industry as a whole (as conventionally estimated with $M = 1$).[5]

The relationship between the estimated TFP residuals in the two cases (i.e., with $M = 1$ and $M > 1$) can be highlighted by combining them to obtain the following decomposition:

$$\left.\begin{array}{ll} M = 1: & \hat{a}_i = y_i - \hat{\alpha} - \hat{\beta} k_i \\ \\ M > 1: & \hat{a}_{i,m} = y_i - \hat{\alpha}_m - \hat{\beta}_m k_i \end{array}\right\} \hat{a}_i = \hat{a}_{i,m} + (\hat{\alpha}_m - \hat{\alpha}) + (\hat{\beta}_m - \hat{\beta}) k_i. \tag{2}$$

Equation (2) highlights that conventional estimates of the TFP residuals are a composition of three terms: $\hat{a}_{i,m}$, which is a TFP measure estimated within the group of firms using technology m (and thus expressed in relative terms with respect to the average firm in the technology group), $\hat{\alpha}_m - \hat{\alpha}$, which can be seen as a bias in the estimated intercept of the production function, and $\beta_m - \hat{\beta}$, which can be seen as a bias in the estimated slope

[4] This term can be thought of as idiosyncratic managerial ability (Bhattacharya *et al.*, 2013).

[5] The production function coefficients estimated on the whole sector, without the technology-related clustering of firms (i.e., with $M = 1$), can be seen as a weighted (across technology groups) average of the technology-specific parameters. The relationship between our approach and conventional estimates is further discussed in the Online Appendix B.

of the production function. As a result, neglecting the presence of different (within-sector) technologies results in overstating the TFP of the firms that adopt relatively more productive technologies (i.e., $\beta_m > \hat{\beta}$ and/or $\alpha_m > \alpha$). While this bias may not be an issue in some cases, the estimation of the technology-specific production function parameters makes it possible to disentangle the labor productivity differences associated with the firm's choice to adopt a certain technology (i.e., technology adoption) and the labor productivity differences related to the firm's ability to exploit the adopted technology.

To see this, by keeping things simple, let us consider Figure 1.

[insert Figure 1 about here]

We assume that only two technologies are available in the industry and that we can perfectly identify them by estimating the technology-specific α_m and β_m. This enables us to obtain the predicted labor productivity associated with each technology at the given levels of k. Now, consider the case of firm i, producing at the K/L ratio k_i. The locally optimal (or frontier) technology (e.g., Atkinson and Stiglitz, 1969) for firm i is the technology that maximizes labor productivity at k_i. Let us refer to this technology as m^H. Assume that firm i uses instead the non-frontier technology m^L. The predicted labor productivity associated with these two technologies can be written as

$$\hat{y}_{i,m^L} = \hat{\alpha}_{m^L} + \hat{\beta}_{m^L} k_i \qquad \text{and} \qquad \hat{y}_{i,m^H} = \alpha_{m^H} + \hat{\beta}_{m^H} k_i. \tag{3}$$

The estimated Solow residual can be measured using either the frontier or the non-frontier technology as the benchmark. In the first case, it can be written as $\hat{a}_{i,m^H} = y_i - \hat{y}_{i,m^H}$. As shown in Figure 1, this measure can be decomposed into the WTFP and BTFP components, respectively given by

$$WTFP_i = \hat{a}_{i,m^L} = y_i - \hat{y}_{i,m^L} \tag{4}$$

and

$$BTFP_i = \hat{y}_{i,m^L} - \hat{y}_{i,m^H} = \hat{a}_{i,m^H} - \hat{a}_{i,m^L}. \tag{5}$$

$WTFP_i$ is the Solow residual expressed using the non-frontier technology (i.e., the technology actually chosen by firm i) as the benchmark. This can be interpreted as firm i's ability to exploit its technology compared with the other firms using the same technology. $BTFP_i$ is instead a measure of how distant firm i's predicted labor productivity under the actual technology is from the predicted labor productivity associated with the locally best technology at k_i. This quantifies the gap with respect to the labor productivity that the firm could have reached, with the given k, had it chosen technology m^H.

The following sections set out the empirical framework for the identification of WTFP and BTFP in a generalized setting, which can determine endogenously the number of available technologies, estimate the technology-specific production function parameters, and probabilistically assign the firms to the various technologies.

3 Modeling Technology Adoption

The presence of different technologies magnifies the simultaneity issues associated with productivity estimation at the firm level.

In conventional estimates, simultaneity stems from the fact that the idiosyncratic term a_i is to some extent known to the firm at the stage of the input-level decision. This distorts the OLS estimation output as long as $Cov(K_{i,t}, A_{i,t}) \neq 0$ and/or $Cov(L_{i,t}, A_{i,t}) \neq 0$ (with the index t used to denote time). This problem was first recognized by Marschak and Andrews (1944). The solution recently suggested by the literature (i.e., proxy variable or semi-parametric methods) consists of estimating the production function semi-parametrically by identifying an observed variable that is a function of a_i. Provided that such a function is invertible, the inverse relationship can be used to substitute for the idiosyncratic term in the production function (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg *et al.*, 2015; Wooldridge, 2009; Doraszelski and Jaumandreu, 2012). Our multiple-technology framework amplifies this issue by bringing about additional potential simultaneity associated with the choice of technology. In fact, using $m_{i,t}$ to denote firm i's technology at time t, $Cov(K_{i,t}, m_{i,t}) \neq 0$ and/or $Cov(L_{i,t}, m_{i,t}) \neq 0$.

To address both sources of simultaneity, we develop an empirical model of technology adoption similar in spirit to the semi-parametric models of Olley and Pakes (1996) but which, by not relying on proxy variables, can be estimated in a fully parametric setting. A key advantage of this approach is that it overcomes the problems related to inverting the relationship between the proxy variable and idiosyncratic productivity term (Gandhi *et al.*, 2018).[6]

The model proceeds under the standard (e.g., Olley and Pakes, 1996) "one period time-to-build" hypothesis, according to which the new technology is productive one period after its acquisition. Moreover, it assumes that the idiosyncratic productivity term $a_{m,i,t}$ follows the first-order Markov process $a_{m,i,t} = E[a_{m,i,t}|a_{m,i,t-1}] + \xi_{m,i,t}$, where $\xi_{m,i,t}$ is the "surprise" related to switching to a different technology (in which case, we find that $m_{i,t} \neq m_{i,t-1}$) or to idiosyncratic shocks in the firm's ability to exploit the technology in use (in which case, $m_{i,t} = m_{i,t-1}$).

Using the terminology $X[t]$ to indicate that variable X is chosen at time $[t]$, we assume the following decision timing. At the end of period $[t]$, the firm chooses $(K_{i,t+1}[t], m_{t+1}[t])$. At the beginning of period $[t+1]$, $a_{i,t+1}$ and Z_{t+1} (i.e., vectors of the exogenous market-level state variables) are observed, so that the firm freely chooses the amount of labor (i.e., $L_{i,t+1}[t+1]$). Finally, the firm chooses $K_{i,t+2}[t+1]$ and $m_{t+2}[t+1]$ at the end of period $[t+1]$ on the basis of $a_{i,t+1}$ and Z_{t+1}.

In period t, firm i maximizes the present value of its future profits, conditional on the information set $\Omega_{i,t}$:

$$\max_{(K_{i,t+1}, m_{i,j+1})} E_t \left[\sum_{j=t}^{\infty} \delta^{j-t} P_{i,j} | \Omega_{i,j} \right], \tag{6}$$

with the net profit function $P_{i,j}$ amounting to

$$P_{i,j} = \underbrace{\pi_{i,j}(K_{i,j}, a_{i,j}, m_{i,j}, Z_{i,j})}_{\text{Gross profit}} - C(K_{i,j+1}, K_{i,j}, m_{i,j+1})$$

[6] The relationship between our approach and the semi-parametric approach is discussed in the Online Appendix B.

and

$$C(K_{i,j+1}, K_{i,j}, m_{i,j+1}) = \underbrace{C^{I}_{i,j}(I_{i,j})}_{\text{Inv. cost}} + \underbrace{C^{D}_{i,j}(D_{i,j})}_{\text{Disinv. cost}} + \underbrace{C^{A}_{i,j,m}\mathbb{I}(m_{i,j+1} \neq m_{i,j})(I_{i,j})}_{\text{Tech adjustment cost}}, \tag{7}$$

where δ is the discount rate and the firm's investments $I_{i,j} = K_{i,j+1} - K_{i,j} + D_{i,j}$ encompass disinvestment costs $D_{i,j} = \epsilon_{i,j} K_{i,j}$ (with $0 \leq \epsilon_{i,j} \leq 1$). While disinvestment costs are borne independently of changing technology or not, the technology adjustment cost in the third term on the right-hand side of Equation (7) includes the costs associated with switching to a different technology in period [j+1].

Capital accumulates according to

$$K_{i,j+1} = K_{i,j} - \delta K_{i,j} + I_{i,j} - D_{i,j} \tag{8}$$

and the Bellman equation can be written as

$$V_{i,t}(\Omega_{i,t}) = \max_{(K_{i,t+1}, m_{i,t+1})} (P_{i,t} + \delta E_t \left[V_{i,t+1} | \Omega_{i,t}\right]). \tag{9}$$

The solution of (9) consists of the values of $(K_{i,t+1})$ and $(m_{i,t+1})$ that satisfy the policy function for K:

$$K_{i,t+1}(m_{i,t+1}, K_{i,t}, a_{i,t}, Z_{i,t}), \tag{10}$$

with the firm choosing, at time [t], the technology $m_{i,t+1}$ that maximizes

$$\left[\delta E_t \left(V_{i,t+1} | \Omega_{i,t}\right) - C(K_{i,t+1}, K_{i,t}, m_{i,t+1})\right] | m = m_{i,t+1}$$

among all possible $m \in \{M\}$. Such technology may coincide with the technology that would provide the maximum productivity associated with the given level of k_i (i.e., the frontier technology, as defined in Section 2).

This framework allows us to estimate the production function by taking account of the simultaneity associated with both the choice of inputs and the choice of technology. Operationally, this boils down to preliminarily estimating the system of equations consisting of the K policy function in (10) and the static condition for L:

$$\begin{cases} \ln K_{i,t}[t-1] = \rho_0 + \rho_1 \ln K_{i,t-1}[t-2] + Z_{c,t} + e^{K}_{i,t} + u^{K}_{i,t} \\ \\ \ln L_{i,t}[t] = \rho_0 + \rho_1 \ln K_{i,t}[t-1] + Z_{c,t} + e^{L}_{i,t} + u^{L}_{i,t}, \end{cases} \tag{11}$$

where $Z_{c,t}$ captures the country-year effects.

Under the assumption that $u^{K}_{i,t}$ and $u^{L}_{i,t}$ are *i.i.d.* error terms, $e^{K}_{i,t}$ includes the covariance terms $Cov(K_{i,t}[t-1], m_{i,t}[t-1])$ and $Cov(K_{i,t}[t-1], a_{i,t-1})$, while $e^{L}_{i,t}$ includes $Cov(L_{i,t}[t], a_{i,t})$. In other words, as far as $Z_{c,t}$ effectively absorbs all the country-level heterogeneity in the data, the regression residuals of the two equations in Equation (11) can be thought of as comprising the firm-level variability in input choices correlated with both the idiosyncratic productivity shock and the choice of technology. As such, they can be seen as correction factors to be included as additional regressors in the production function-estimating equation to obtain simultaneity-free production function parameters.

4 Data

An issue related to the identification of technology at the firm level is the availability of internationally comparable data with the largest possible coverage. This is crucial to potentially capture all the technologies available worldwide. To this end, we take advantage of the ORBIS database (Bureau van Dijk, 2018) in which the comparable balance items needed to carry out our estimation are provided for a sample of 35,850 firms worldwide distributed across 22 (two-digit) sectors. To be consistent with the model presented in Section 3, data for at least two consecutive years are needed. We use 2015 and 2016 (the two consecutive years with the largest coverage among the most recent ones).

From ORBIS, we draw information on the variables used in the productivity estimation: value added (VA), tangible capital (K), labor (L), and average wages within the firm (H). Value added and capital are deflated using OECD-STAN country-sector-year deflators. From ORBIS, we also draw information on the other firm characteristics used in Section 7. Table 1 describes all the variables adopted in the analysis.

[insert Table 1 about here]

Table 2 reports the descriptive statistics at the sectoral level.

[insert Table 2 about here]

The dataset covers 76 countries globally. Table 3 summarizes the geographical distribution of value added, capital, and labor as well as the share of firms. Small or peripheral economies are clustered into larger geographical aggregates. While North American countries are under-represented compared with most European countries (Germany and Great Britain, in particular) in terms of all the variables reported and even more in terms of the share of firms, the mismatches between shares of firms and shares of value added, capital, and labor partially mirror the different average firm sizes across countries. This is evident for Italy and Portugal.

[insert Table 3 about here]

Country bias is a known problem with international firm-level databases. In fact, higher country coverage always comes with lower representativeness in terms of the cross-country distribution of firms. This stems from the sometimes wide cross-country differences in the national standards of balance sheet disclosure. A trade-off, thus, may emerge. On the one hand, the larger the database, the lower are both the country representativeness and the chance to study the geographical diffusion of firm-level technologies. On the other hand, the larger the database, the higher is the chance to capture the technologies available worldwide and examine the sectoral distribution of technologies. The most important issue in our work is identifying as many technologies as possible. This entails including as many comparable data (i.e., firms) as possible in each sector, even though the country coverage is poor, since even a small number of firms from a given country may help identify the available technologies correctly.[7]

Accordingly, our discussion focuses on the sectoral articulation, with the country dimension only used in Section 7 to compare the TFP measures derived from our sample with a selection of country-level technology indicators and to describe country selection.

5 Production Function(s) Estimation

The estimation strategy, consistent with the model described in Section 3, entails two steps.

First, we use three-stage least squares to estimate the system of equations (see Equation (11)). The estimation is carried out sector by sector including the country-year effects to net the residuals of the potential effect of market frictions and other country characteristics (e.g., interest rate and wage differentials on firms' input choices). The estimation requires a minimum of two years, namely the reference year [t] and the previous year [t-1], for the lagged term $\ln K_{i,t-1}[t-2]$). To control for possible non-linearities in the choice of capital, we also include squared lagged capital in the first equation.

Second, the estimated residuals of the two equations in Equation (11), $\hat{\Phi}_i = e^K_{i,t} + u^K_{i,t}$ and $\hat{\Psi}_i = e^L_{i,t} + u^L_{i,t}$, respectively are included as additional regressors in the production function-estimating equation to obtain the simultaneity-free production function parameters. Precisely, we estimate the following equation:

$$y_i = \alpha_m + \beta_m k_i^{\delta_{i,m}} + \gamma_m h_i^{\delta_{i,m}} + \varphi \hat{\Phi}_i^{\delta_{i,m}} + \psi \hat{\Psi}_i^{\delta_{i,m}} + FE_s + \varepsilon_i, \tag{12}$$

where h_i is the average wage bill within the firm, which proxies for human capital,[8] and FE_s are the four-digit industry fixed effects meant to absorb the cross-product differences in the production processes. While the estimation in the first step is carried out over two consecutive years, the second step requires only one year; hence, the time index can be omitted from Equation (12).

In this second step, we are interested in obtaining the technology parameters α_m and β_m for each technology m avoiding any ex-ante assumption on both the degree of technology sharing across firms and the number of available technologies. To this end, we rely on mixture models (McLachlan and Peel, 2000). The basic idea is that the probability distribution of y_i can be seen as a weighted average of the M unknown segment (or cluster) distributions, each with a proper mean (μ_m) and variance (σ^2_m): $f(Y_i|\mu, \sigma^2) = \sum_{m=1}^{M} \omega_m f_m(Y_i|\mu_m, \sigma^2_m)$. The weights ω_m are given by the ex-ante probability of belonging to group m.

The fact that this probability is unknown generates a problem of missing data, solved by applying the expectation-maximization algorithm to the sector-by-sector (at the two-digit level) estimation of Equation (12) through weighted least squares (WLS), as suggested by De Sarbo and Cron (1988). The estimation starts with random values of ω_m (see below) to compute the posterior probability $p_{i,m}$ that firm i belongs to group m, and thus the observation

[7] This is even more true if the probability of being in the database is affected by, say, being listed or being a multinational and thus complying with international accounting standards.

[8] This specification assumes that the technology choice in Section 3 is independent of the level of human capital, which is taken as given by the firm. As known, the wage bill is a good proxy for human capital under the assumption of perfectly competitive labor markets and perfect signaling.

weights in Equation (12) are

$$\delta_{i,m} = \sqrt{p_{i,m}} \qquad \text{with} \quad p_{i,m} = \frac{\omega_m f_m\{y_i|\mu_m;\sigma_m^2\}}{\sum_{m=1}^{M}\omega_m f_m\{y_i|\mu_m;\sigma_m^2\}}. \tag{13}$$

This set of probabilities is then used to update the regression coefficients by changing the weights ω_m according to

$$\omega_m = \frac{\sum_i p_{i,m}}{\sum_m \sum_i p_{i,m}} \tag{14}$$

with the following constraints:

$$\omega_m \geq 0 \qquad \forall m = 1, \ldots M \qquad\qquad \text{and} \qquad\qquad \sum_{m=1}^{M} \omega_m = 1. \tag{15}$$

The algorithm iteratively alternates WLS production function estimation and computation of probabilities until a log-likelihood convergence criterion is satisfied (Grun and Leisch, 2008).[9]

To leave the routine free to set the number of available technologies M (i.e., clusters), we try different numbers of clusters and pick the optimal choice following the Bayesian information criterion (BIC) based on the following log-likelihood function:

$$M = argmax \left\{ -2 \cdot \log \left[\sum_i \omega_m f_m(y_i|\mu_m, \sigma_m^2) \right] - \zeta(m) \right\} \mid m = 1, \ldots, M \tag{16}$$

where $\zeta(m)$ is a penalty function reflecting the tension between the number of clusters and the number of parameters to be estimated.[10] The choice of the number of clusters is critical. Steele and Raftery (2010) show that the BIC provides the best results in terms of avoiding a poor description of the population, on the one hand, and overfitting, on the other. Table 4 reports the results of the BIC. In addition to using the BIC, we identify the technology clusters requiring the estimated parameters to be economically meaningful. For the group-specific estimated production function parameters to properly identify a technology, we impose the shape parameter β to be larger than 0.1 and smaller than 0.9 as well as statistically significant; the non-significant α values are set to zero.

[insert Table 4 about here]

Table 5 displays the estimated α and β parameters for each technology group and sector. The first column summarizes the number of technologies. The hypothesis of a single technology is restrictive in most cases, entailing

[9] The density functions of Equation (13) is assumed to follow a Gaussian distribution (we need to assume a distribution belonging to exponential family). Hence, the characterizing parameters are mean and variance. For the mean (a linear predictor), we use the fitted value of the WLS estimation. For the variance, we use the (weighted) variance of the segment regressions (De Sarbo and Cron, 1988). To compute the standard errors, we follow the approach suggested by Louis (1982) and Mc Lachlan and Basford (1988), based on information based matrices. Once the expectation-maximization algorithm has converged, it is common to calculate the variance-covariance matrix of the estimated parameters as the inverse of the Fisher information matrix; the standard errors are the gradients of this matrix.

[10] The penalty function is equal to the (log of the) number of observations (i.e., firms) times the number of parameters, which, for each cluster, amounts to the sum of the regression coefficients, variance, and weights, minus one.

that conventional estimates hide substantial heterogeneity. Figure 2 visualizes the resulting production functions.

[insert Table 5 about here]

[insert Figure 2 about here]

Table 6 reports the total probability of each technology group computed as $prob_m = \sum_i p_{i,m}$. This collapses to the number of firms in the sector when summed across all the technology groups (in the sector). As the table reveals, none of the groups presents a negligible probability; hence, firms are distributed across the technology groups with a certain degree of variability.

[insert Table 6 about here]

6 Measuring WTFP and BTFP

In the subsequent analysis, we focus on the cross-firm labor productivity differences associated with either being relatively less productive within a given technology group (i.e., WTFP) or not using the frontier technology (i.e., BTFP).

To this end, let us generalize the definitions provided in Section 2 by referring to the frontier technology as the technology m^H that maximizes labor productivity at the capital-to-labor ratio actually chosen by the firm, among all possible alternative technologies: $y_{i,m^H}|k = k_i > y_{i,m}|k = k_i \quad \forall m \neq m^H$.

The BTFP measure can now be expressed as a weighted average of the difference between the predicted labor productivity associated with each possible technology m and the frontier labor productivity at the given level of k (or, alternatively, as the inverse difference between the two Solow residuals):

$$BTFP_i = \sum_{m=1}^{m^H} p_{i,m} \cdot (\hat{y}_{i,m} - \hat{y}_{i,m^H}) = \sum_{m=1}^{m^H} p_{i,m} \cdot (\hat{a}_{i,m^H} - \hat{a}_{i,m}), \tag{17}$$

where the weights $p_{i,m}$ correspond to the estimated probability of firm i belonging to the technology group m defined in Equation (13).

The BTFP in Equation (17) is a probabilistic quantification of the labor productivity gap with respect to the labor productivity associated with the frontier technology at the given k (i.e., the productivity that firm i could have reached had it used k_i with the best technology for that level of k). $BTFP_i$ is zero when only one technology is available in the sector (i.e., $M = 1$). By contrast, when $M > 1$ technologies are available, $BTFP_i$ is zero only if the firm uses the frontier technology with probability one (i.e., $p_{i,m^H} = 1$). In all the other cases, BTFP is negative.

Similarly, the WTFP measure can be expressed as a weighted average of the Solow residuals across all the available technologies in the sector:

$$\text{WTFP}_i = \sum_{m=1}^{m^H} p_{i,m} \cdot \hat{a}_{i,m}. \tag{18}$$

This term is the empirical equivalent of the idiosyncratic productivity term $a_{i,m}$ in Equation (1).

We next assess the relative importance of the within-technology and between-technology dimensions at the sectoral level. If the technological dimension turns out to be relatively large with respect to the idiosyncratic dimension in most sectors, then this may suggest that the "distance" across technologies in terms of labor productivity is economically significant. To this end, it is convenient to rescale the WTFP and BTFP measures in terms of the labor productivity "gaps" with respect to the frontier to have positive values only.

For BTFP, we just multiply by minus one. For the WTFP term, since the benchmark is the average firm within the technology group, we observe a zero-mean distribution. Thus, we rescale the firm values using the average WTFP of the best 5% in the WTFP distribution: $WTFP_{\text{best5\%}} - WTFP_i$.[11]

Figure 3 shows the sectoral distribution of the relative weights of WTFP and BTFP by plotting the ratio $BTFP_i/WTFP_i$ for all the firms in each sector. The concentration in the ratio seems to be uncorrelated with the estimated number of technologies, as one might suspect.

[insert Figure 3 about here]

We can then obtain aggregate figures by computing the sectoral averages and standard deviation of $WTFP_i$, $BTFP_i$, and $BTFP_i/WTFP_i$ (see Table 7). A first result is that the contribution of WTFP and BTFP to the cross-firm labor productivity differences is balanced, with the technological dimension slightly dominating: the WTFP of the top 5% of firms in the WTFP distribution is on average 31% higher than the other firms, while the average BTFP is around 34%.

[insert Table 7 about here]

Substantial intra-sectoral and inter-sectoral heterogeneity emerges in terms of both measures. The BTFP sectoral average ranges from less than 7% for textiles to around 109% for rubber and plastic products, while for beverages, fabricated metals, and transport equipment the BTFP average is zero since a single technology is obtained. The sectoral averages of WTFP appear less dispersed across sectors, with values ranging from 17% in the textiles industry to around 79% in the basic metals industry.

The intra-sectoral and inter-sectoral heterogeneity of the $BTFP_i$-to-$WTFP_i$ ratio is informative. Overall, the between-technology dimension is more than twice as important as the within-technology dimension. Apart from the three sectors in which BTFP is zero, the ratio is on average lower than one in eight sectors of the 19 under consideration. Notably, average BTFP is over 10 times average WTFP in the rubber and plastics sector, eight times in the case of electrical products, and over six times in the pharmaceutical industry. By contrast, the within-technology dimension takes over in industries such as paper and paper products, machinery and equipment, and basic metals.

The average values in Table 7 and firm-level distributions in Figure 3 allow us to establish that taking into account the roots of the firm-level labor productivity differences to fine-tune innovation policy might result in

[11] This represents a firm-level measure of the probabilistic productivity gain that would be obtained by eliminating the WTFP dispersion within the technology cluster.

more effective and targeted choices. In fact, even in sectors in which BTFP dominates on average, there are firms for which labor productivity is mostly determined by the ability to use the adopted technology. In these cases, increasing workers' skills through, say, lifelong learning, managerial improvements, and investment in organizational innovation might be more effective than trying to stimulate the adoption of new production technologies.

Three considerations are in order. First, an important issue in productivity estimation at the firm level is the dispersion of prices across firms.[12] This (see Klette and Griliches, 1996) yields biased TFP estimates as long as firm sales are deflated using industry-wide price indexes rather than firm-level prices. In our case, as the difference between predicted values, the BTFP term should not be affected by firms' markups. If any, only the (average) systematic differences across prices applied by the firms in different technology clusters would eventually flow into the BTFP term. Conversely, cross-firm differences in markups are entirely reflected in WTFP differences.

Second, as far as we are effectively controlling for human capital by including the wage bill in Equation (12), we can attribute BTFP differences across firms to the differences in technology in a strict sense (e.g., different machinery and software). To check the validity of this strategy, we construct a country-sector human capital endowment variable by interacting Cohen and Soto's (2007) measure of education in each country in 1970 and 2005 with the industry schooling intensity of the United States in 1980, drawn from IPUMS as in Ciccone and Papaioannou (2009). The regressions of our WTFP and BTFP terms on this country-sector-specific human capital variable (plus country and sector controls) yield an insignificant effect for BTFP and a negative and significative effect for WTFP (the results of this check are available upon request). Under the assumption that sectoral schooling intensity is the same in all countries and keeping in mind the limits involved in such a measure of human capital, this might suggest that the largest part of the technological differences across firms detected in the analysis does not stem from the omission of human capital.

Third, the literature has largely shown that when market imperfections interfere with the resource reshuffling across firms (from less productive to more productive firms), the allocation of inputs is inefficient and aggregate productivity can be substantially lower than in an ideal situation of frictionless factor markets (Alfaro *et al.*, 2008; Banerjee and Duflo, 2005; Bartelsman *et al.*, 2013; Hsieh and Klenow, 2009, 2010; Restuccia and Rogerson, 2008). The BTFP heterogeneity documented in our analysis introduces an additional misallocation mechanism: market imperfections can also prevent firms from switching to superior technologies, thereby remaining with the misaligned marginal revenue product (MRP) of inputs (i.e., inefficient allocation of inputs). Hence, BTFP dispersion is a possible consequence of the presence of frictions at the level of factor markets. However, as discussed in Online Appendix C, bringing all firms to the technology frontier would eliminate BTFP dispersion but not misallocation as long as frictions prevent firms from freely setting the amount of inputs. Moreover, once cross-firm technological heterogeneity is explicitly taken into account, Hsieh and Klenow's (2009) suggestion of using the dispersion of revenue TFP to infer the presence of misallocation[13] is no longer valid, as revenue TFP can differ across firms not

[12]Firms' output, which is needed to estimate the production function parameters, is commonly unavailable in physical terms. This forces the econometrician to use a proxy that, in the vast majority of cases, consists of sales deflated by an industry-wide price index, given that individual prices are themselves commonly unavailable. Such a circumstance has no relevance under perfect competition as all firms quote the same price. On the contrary, when markets are imperfectly competitive, estimated firm-level productivity is likely to be misstated. Since the problem is caused by omitting the individual price from the estimation, this problem is usually referred to as omitted price bias.

[13]This approach has been widely used to quantify the extent of misallocation (Garcia-Santana *et al.*, 2015; Gopinath *et al.*, 2017; Dias *et al.*, 2016; Bellone and Mallen-Pisano, 2013; Crespo and Segura-Cayuela, 2014; Chen and Irarrazabal, 2014; Calligaris *et al.*,

only because of the different MRP of inputs but also because of the different technology adopted.

7 Broad Validation

Country-sector correlation with international exchanges of technology. To help interpret the documented WTFP and BTFP dimensions, we consider their relationship with the country-sector international flows of knowledge and technology in a simple cross-sectional regression analysis. While not establishing causality, a statistically significant correlation between our technological dimension of TFP and the technological balance of payments might support our emphasis on the opportunity to isolate the technological component of the labor productivity differentials by adopting an international perspective.

In this analysis, we rely on OECD Stat (2015) data from the technology balance of payments, measured as the difference between international technology receipts, namely, technology outflows (variable *Tech Receipts*), and payments, namely, technology inflows (variable *Tech Payments*). The data cover license fees, patents, purchases and royalties paid, know-how, and research and technical assistance.[14] In the econometric analysis, we also consider a vector of firm-level characteristics (intangible asset intensity, firm age and size, liquidity, and whether the firm is listed on a stock market and is part of a multinational group; drawn from ORBIS).Table 8 presents the results.

[insert Table 8 about here]

As expected, we find our BTFP measure to be strongly correlated with the technological balance of payments. This finding supports the validity of our estimation strategy. Moreover, the relationship is positive, suggesting that the country sectors with relatively high technology outflows are those in which firms' labor productivity is closer to the local frontier. On the contrary, the correlation with WTFP is only slightly significant and negative. This finding might suggest that a relatively high ability to use the actual technology is more likely to be associated with higher technology inflows.

Although the policy dimension is outside the scope of this analysis, an intuitive implication of the above results is that net exporting country sectors, which are more likely to be associated with the adoption of advanced technologies, may benefit more from initiatives aimed at improving the WTFP component, while net importing country sectors, which are more likely to be associated with greater abilities to use available technologies, may more largely benefit from technology upgrades.

Interestingly, the firm-level dimension also seems to suggest some important differences with respect to the relative positioning of the firm in the WTFP and BTFP distributions, with younger firms and multinational partners showing smaller WTFP gaps. This fact might reflect that multinational chains tend to be vehicles of know-how rather than a channel for transferring physical technologies. The other firm-level variables, namely

2018)

[14]Technology receipts depend on firms' R&D effort and correspond to foreign sales of the marketable results of that effort. Technology payments correspond to the acquisition of technology inputs that are immediately useable by firms. Thus, technology receipts and payments may reflect different dimensions, including the ability of firms in a country sector to sell their intangible technology abroad and the extent to which they use foreign technologies, the degree of technological autonomy (i.e., ability to assimilate foreign technologies), and, more in general, the choice between the domestic production of technology and foreign absorption, which is a crucial dimension of globalization and growth (Perla *et al.*, 2015).

intangible asset intensity, firm size, and liquidity, affect the WTFP and BTFP gaps with the same sign.

Cross-country correlation with standard indicators of technology. Here, we relate our estimation results to the available data on the cross-country characterization of technology adoption. In particular, we compare the country average of firm-level BTFP gaps with standard measures that reflect the level and structure of the efforts undertaken by countries in the field of science and technology. We consider the following: gross domestic expenditure on research and development as a percentage of GDP, total researchers per thousand employees, the number of patents in the ICT sector per thousand employees, and the number of triadic patents per thousand employees (i.e., patents filed at the European Patent Office, the United States Patent and Trademark Office, and the Japan Patent Office by the same applicant for the same invention). These variables are obtained from the STAN database (OECD, 2018) and cover OECD member countries and a selection of non-member economies including China, the Russian Federation, Argentina, South Africa, Romania, and South Korea.

Figure 4 plots the country-level BTFP gaps against these four measures of technology effort averaged over 2014–2016. Countries not covered by the STAN database are excluded. In the figure, relatively small economies are averaged in restricted geographical categories.

[insert Figure 4 about here]

We observe that countries with lower average technology gaps also show higher gross domestic expenditure on research and development and higher numbers of ICT and triadic patents and researchers per thousand employees. The United States, Israel, Germany, and northern European countries are at the top of the cross-country distribution for these dimensions followed by other European OECD countries (France and the United Kingdom, in particular), while Mediterranean countries, countries in eastern Europe, and other non-OECD member economies lag behind. These cross-country comparisons show substantial coherence between the results of our estimates and more traditional measures of a country's technology patterns.

Country selection. Throughout the analysis, we have highlighted that using international (comparable) firm-level data, with a cross-country coverage as large as possible, is crucial to allow our mixture analysis to identify all the available technologies.

An intuitive way to check whether a reduced cross-country coverage affects our technology clustering by reducing the number of technologies possibly identified is to rerun our mixture analysis on a subsample of countries. As a result, we would expect lower technological heterogeneity. To allow a comparison with the previous literature on cross-country production function estimations, we employ the subsample of countries used by Bos *et al.* (2010). They use data for six countries (Finland, Italy, Germany, France, the Netherlands, and Spain) and 21 industries, finding strong evidence in favor of two classes based on a stochastic frontier production model.[15] Table 9 reports our results obtained on a subsample of firms from the same six countries. We find one or two clusters in most of the sectors (only two sectors show a larger number of clusters), with an average number of clusters of 1.6, which is

[15]Similarly, Asturias and Rossback (2017) find two technological groups in most of the industries in a sample of Chilean firms, using a cluster analysis based on firm-level data on revenues and factor input expenditures. This is also coherent with Battisti *et al.*. (2015), who find two to three clusters in a sample of European firms.

close to the number of technology clubs identified by Bos *et al.* (2010).

[insert Table 9 about here]

More importantly, we verify that our mixture model can identify only a reduced number of technologies when the cross-country coverage of the data is abridged.

8 Conclusions

We investigated the technological dimension of productivity in a framework in which the various technologies available in a given sector are represented in terms of different production function parameters (i.e., intercept and input coefficients) and presented an empirical methodology to disentangle the labor productivity differences associated with the firm's choice of technology (BTFP) and those related to the firm's ability to exploit the adopted technology (WTFP).

While the policy debate has recently recognized the opportunity to tackle issues such as technology and productivity from a firm-level perspective, our analysis emphasizes the importance of contributing to such a debate by providing the policymaker with theoretical conceptions and empirical results aimed at clarifying the relationship between notions such as "productivity," "TFP," "innovation," and "technology", which are often used interchangeably.

We argued that the reason behind this terminological ambiguity is computational, as extrapolating the technological dimension from the production function is problematic with standard econometrics, and suggested an alternative strategy relying on mixture models. The estimation algorithm is consistent with a world in which several technologies are available in the same sector, with firms potentially adopting any of them. The estimation is left free to determine both the number of technologies in the given sector and the degree of technology sharing across firms (i.e., for each firm, the probability of using a given technology).

We stress in the analysis that the chance to potentially capture all the technologies available grows with the size of the database and extent of country coverage, even though this is detrimental to sample representativeness. Thus, in bringing the methodology to the data, we take advantage of the comparable balance information provided by the ORBIS database for a sample of 35,850 firms distributed across 22 (two-digit) sectors and 76 countries.

We find BTFP to be at least as important as WTFP in explaining the labor productivity differences across firms. On average, the WTFP productivity of the top 5% of firms in a sector is around 31% higher than that for other firms, while the labor productivity gap between the firms using the frontier technology and those using less productive technologies amounts to around 34%. Within this general result, we document substantial intra-sectoral and inter-sectoral heterogeneity in terms of both measures, particularly when we assess the relative weight of the two. In particular, average BTFP is over 10 times average WTFP in the rubber and plastics sector, eight times in the case of electrical products, and over six times in the pharmaceutical industry. By contrast, the within dimension takes over in industries such as paper and paper products, machinery and equipment, and basic metals. Moreover, we show that even in sectors in which BTFP dominates on average, there are firms for which labor productivity is mostly determined by the ability to use the adopted technology. In these cases, increasing workers' skills through, say, lifelong learning, managerial improvements, and investment in organizational innovation might be more effective than trying to stimulate the adoption of new production technologies. Hence, dissecting the labor productivity gaps is crucial to achieving more targeted innovation policies.

The suggested estimation strategy tackles the two main issues highlighted by the literature on TFP estimation at the firm level: (i) simultaneity, by dealing with the additional problem of the potential correlation between the technological and input choices, and (ii) price dispersion, with our BTFP being unaffected by omitted price bias.

Finally, while the presence of technology dispersion can be a possible consequence of the presence of distortions preventing firms from switching to superior and more productive technologies, we show that (i) eliminating BTFP by bringing all firms to the technology frontier would not eliminate misallocation as long as firms are not free to hire the desired amount of capital and labor and (ii) the presence of technology dispersion introduces an additional source of dispersion in revenue TFP, which prevents us from using the latter to infer the presence of misallocation (as in Hsieh and Klenow (2009)).

References

[1] Ackerberg, D. & Caves, K. & Frazer, G. (2015). Identification Properties of Recent Production Function Estimators, *Econometrica*, 83 (6), 2411-2451.

[2] Alfaro, L. & Kalemli-Ozcan, S. & Volosovych, V. (2008). Why doesn?t Capital Flow from Rich to Poor Countries? An Empirical Investigation, *Review of Economics and Statistics*, 90(2), 347-368.

[3] Asturias, J. & Rossback, J. (2017). Misallocation in the Presence of Multiple Production Technologies, Georgetown University Qatar working paper.

[4] Atkinson, A.B. & Stiglitz, J.E. (1969). A New View of Technological Change. *Economic Journal*, 79(315), 573-578.

[5] Banerjee, A. & Duflo, E. (2005). Growth Theory through the Lens of Development Economics, *Handbook of Development Economics*, Vol. 1a. Amsterdam: Elsevier, 473-552.

[6] Bartelsman, E. & Haltiwanger, J. & Scarpetta, S. (2013), Cross-Country Differences in Productivity: The Role of Allocation and Selection, *American Economic Review*, 103(1), 305-334.

[7] Battisti, M. & Belloc, F. & Del Gatto, M. (2015). Unbundling technology adoption and TFP effects: do firms intangibles matter? *Journal of Economics and Management Strategy*, 24(2), 386-410.

[8] Battisti, M. & Del Gatto, M. & Parmeter, C.F. (2018). Labor productivity growth: disentangling technology and capital accumulation, *Journal of Economic Growth*, 23(1), 111-143.

[9] Bellone, F. & Mallen-Pisano, J. (2013). Is misallocation higher in France than in the United States?, GREDEG Working Paper, N. 38.

[10] Benhabib, J. & Perla, J. & Tonetti, C. (2017). Reconciling models of diffusion and innovation: A theory of the productivity distribution and technology frontier. *NBER Working Papers*, 23095.

[11] Bernard, A. B. & Jones, C. I. (1996a). Comparing apples to oranges: productivity convergence and measurement across industries and countries. *American Economic Review*, 86(5), 1216-1238.

[12] Bernard, A. B. & Jones, C. I. (1996b). Technology and convergence. *Economic Journal*, 106(437), 1037-1044.

[13] Bhattacharya, D. & Guner, N. & Ventura, G. (2013). Distortions, endogenous managerial skills and productivity differences, *Review of Economic Dynamics*, 16, 11-25.

[14] Bloom, N. & Draca, M. & Van Reenen, J. (2016). Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity. *Review of Economic Studies*, 83(1), 87-117.

[15] Bos, J.W.B & Economidou, C. & Koetter, M. (2010). Technology clubs, RD and growth patterns: Evidence from EU manufacturing, *European Economic Review*, 54, 60-79.

[16] Bureau van Dijk (2018) Orbis Database. Amsterdam.

[17] Calligaris, S. & Del Gatto M. & Hassan, F. & Ottaviano, G. & Schivardi, F. (2018). Global Productivity Puzzles and Misallocation, an Italian Tale), *Economic Policy*, forthcoming.

[18] Caselli, F. (2005). Accounting for income differences across countries. In P. Aghion & S. Durlauf (Eds.), *Handbook of economic growth*. North Holland, Elsevier.

[19] Chen, K. & Irarrazabal A. (2014). The role of allocative efficiency in a decade of recovery, *Review of Economic Dynamics*, 18, 523-550.

[20] Ciccone, A. & Papaioannou, E. (2009). Human capital, the structure of production and growth. *Review of Economics and Statistics*, 91(1), 66?82.

[21] Cohen, D. & Soto, M. (2007). Growth and Human Capital: Good Data, Good Results. *Journal of Economic Growth*, 12(1), 51-76.

[22] Collard-Wexler, A. & De Loecker, J. (2014). Reallocation and technology: evidence from the US steel industry. *American Economic Review*, 105(1), 131-171.

[23] Crespo, A. & Segura-Cayela, R. (2014). Understanding competitiveness, EUI Working Papers, 2.

[24] Da Rocha, J.M. & Mendes Tavares, M. & Restuccia, D. (2017). Policy Distortions and Aggregate Productivity with Endogenous Establishment-Level Productivity, *NBER Working Paper*, 23339.

[25] De Sarbo, W.S. & Cron, W.L. (1988). A Maximum Likelihood Methodology for Clusterwise Linear Regression, *Journal of Classification*, 5, 249-282.

[26] Del Gatto, M. & Di Liberto, A. & Petraglia, C. (2011). Measuring productivity, *Journal of Economic Surveys*, 25(5), 952-1008.

[27] Dias, D. & Marques, C.R. & Richmond, C. (2016). Misallocation and productivity in the lead up to the Eurozone crisis, *Journal of Macroeconomics*, 49, 46-70.

[28] Doraszelski, U. & Jaumandreu, J. (2013) R&D and Productivity: Estimating Endogenous Productivity, *Review of Economic Studies*, 80(4), 1338-1383.

[29] Duclos, J.-Y. & Esteban, J. & Ray, D. (2004). Polarization: Concepts, Measurement, Estimation, *Econometrica*, 72(6), 1737-1772.

[30] Easterly, W. & Levine, R. (2001). What have we learned from a decade of empirical research on growth? It's not factor accumulation: Stylized facts and growth models, *World Bank Economic Review*, 15(2), 177-219.

[31] Gabler, A. & Poschke, M. (2013). Experimentation by firms, distortions, and aggregate productivity. *Review of Economic Dynamics*, 16(1), 26-38.

[32] Gandhi, A. & Navarro, S. & Rivers, D. (2018). On the Identification of Gross Output Production Functions, *Journal of Political Economy*, forthcoming.

[33] Garcia-Santana, M. & Moral-Benito, E. & Pijoan-Mas, J. & Ramos, R. (2016). Growing like Spain: 1995-2000, Banco de Espana Working Paper No. 1609.

[34] Gopinath, G. & Kalemli-Ozcan, S. & Karabarbounis, L. & Villegas-Sanchez, C. (2017). Capital Allocation and Productivity in South Europe, *Quarterly Journal of Economics*, 132(4), 1915-1967.

[35] Grossman, G.M. & Helpman, H. (2015). Globalization and Growth. *American Economic Review: Papers & Proceedings*, 105(5), 100-104

[36] Grun, B. & Leisch, F. (2008). FlexMix version 2: finite mixtures with concomitant variables and varying and constant parameters, *Journal of Statistical Software*, 28(4), 1-35.

[37] Hsieh, C.T. & Klenow, P.J. (2009). Misallocation and Manufacturing TFP in China and India, *Quarterly Journal of Economics*, 124(4), 1403-1448.

[38] Hsieh, C.T. & Klenow, P.J. (2010). Development Accounting, *American Economic Journal: Macroeconomics* 2(1), 207-223.

[39] Klette, T. J., Griliches, Z. (1996). The inconsistency of common scale estimators when output prices are unobserved and endogenous, *Journal of Applied Econometrics*, 11(4), 343-361.

[40] Kumar, S. & Russell, R.R. (2002). Technological Change, Technological Catch-up, and Capital Deepening: Relative Contributions to Growth and Convergence, *American Economic Review*, 92(3), 527-548.

[41] Levinsohn, J. & Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables, *Review of Economic Studies*, 70, 317-342.

[42] Louis, T.A. (1982). Finding the observed information matrix when using the EM algorithm, *Journal of Royal Statistical Society B*, 44, pp. 226?33.

[43] Luttmer, E. G. (2007). Selection, Growth, and the Size Distribution of Firms. *Quarterly Journal of Economics*, 122(3), 1103-1144.

[44] Marschak, J., Andrews, W. H. (1944). Random simultaneous equations and the theory of production, *Econometrica*, 12(3), 143-205.

[45] Mc Lachlan, G.J. & Basford, K.E. (1988). Mixture Models: Inference and Applications to Clustering, *Applied Statistics*, 38(2).

[46] Mc Lachlan, G.J. & Peel, D. (2000) *Finite Mixture Models*, John Wiley and Sons, NY.

[47] OECD (2015). *The Future of Productivity*, OECD Publishing, Paris.

[48] Olley, S. & Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64, 1263-1297.

[49] Perla, J. & Tonetti, C. (2014). Equilibrium imitation and growth. *Journal of Political Economy*, 122(1), 52-76.

[50] Perla, J. & Tonetti, C. & Waugh, M. E. (2015). Equilibrium Technology Diffusion, Trade, and Growth. NBER Working Paper, National Bureau of Economic Research, 20881.

[51] Restuccia, D. & R. Rogerson (2008). Policy distortions and aggregate productivity with heterogeneous establishments, *Review of Economic Dynamics*, 11(4), 707-720.

[52] Restuccia, D. & Rogerson, R. (2013) Misallocation and Productivity. *Review of Economic Dynamics*, 16, 1-10.

[53] Sampson, T. (2016). Dynamic Selection: An Idea Flows Theory of Entry, Trade and Growth. Quarterly *Quarterly Journal of Economics*, 131(1), 315-380.

[54] Steele, R.J. & Raftery, A.E. (2010). Performance of Bayesian Model Selection Criteria for Gaussian Mixture Models. In *Frontiers of Statistical Decision Making and Bayesian Analysis* (edited by M.-H. Chen *et al.*), 113-130, New York: Springer.

[55] Van Beveren, I. (2012). Total Factor Productivity Estimation: A Practical Review, *Journal of Economic Surveys*, 26(1), 98-128.

[56] Wooldridge, J. (2009). On Estimating Firm-level Production Functions Using Proxy Variables to Control for Unobservables. *Economics Letters*, 104(3), 112-114.

Table 1: Description of the variables.

Variable	Description	Source
Added Value	Log of added value. Added value is defined as profit for the period + depreciation + taxation + interest paid + cost of employees. Firm-level variable deflated using the OECD-Stan sector-country-specific deflators (firm-level variable)	Orbis (2015)
Capital Input	Log of tangible assets. Tangible assets include buildings, machinery, and all other tangible assets. Firm-level variable deflated using the OECD-Stan sector-country-specific deflators (firm-level variable)	Orbis (2015)
Labor Input	Log of the total number of employees included in the company's payroll. Firm-level variable deflated using the OECD-Stan sector-country-specific deflators (firm-level variable)	Orbis (2015)
Average Wage	Log of the average wage bill within the firm. Firm-level variable	Orbis (2015)
Firm Intangibles	Log of intangible to tangible asset ratio. Intangible assets include formation expenses, research expenses, goodwill, and development expenses. Tangible assets include buildings, machinery, and all tangible assets (firm-level variable)	Orbis (2015)
Firm Age	Age of the firm (years) (firm-level variable)	Orbis (2015)
New Entrant	Dummy variable (1 = firm age is below or equal to five years, 0 = otherwise) (firm-level variable)	Orbis (2015)
Listed Firm	Dummy variable (1 = the firm is listed on the stock market, 0 = otherwise) (firm-level variable)	Orbis (2015)
Multinational	Dummy variable (1 = the firm is part (as a controller or controlled enterprise) of a multinational group (firm-level variable)	Orbis (2015)
Liquidity Ratio	Cash and cash equivalent as a percentage of total assets (firm-level variable)	Orbis (2015)
Tech Balance	Technological balance of payments, as the difference between country-sector receipts and payments: receipts are measured as international technology receipts for license fees, patents, purchases and royalties paid, know-how, and research and technical assistance; payments are measured as international technology payments for license fees, patents, purchases and royalties paid, know-how, and research and technical assistance (both payments and receipts are weighted by sectoral value added) (country-sector variable)	Technological payments and receipts: OECD Stat, Technology Balance of Payments database; sectoral value added: WDI database

Table 2: Descriptive statistics: Sectoral distribution.

Sector description	Sector code	VA (% of tot)	K (% of total)	# firms	VA/L (avg)	K/L (avg)
Food products	Fd	4.47	3.95	4332	25.8	37.3
Beverages	Bv	3.19	6.85	583	45.6	131.8
Tobacco products	Tb	2.48	2.47	16	58.3	98.8
Textiles	TX	0.79	0.83	1435	26.0	29.3
Apparel	WA	0.92	0.54	2018	15.4	9.3
Leather and related products	LP	0.35	0.13	1151	20.8	11.4
Wood and products of wood and cork	Wo	0.35	0.26	2051	23.1	26.8
Paper and paper products	Pa	3.38	3.83	769	41.0	60.4
Printing and reproduction of recorded media	Pr	1.37	1.74	1541	30.1	29.3
Coke and refined petroleum products	PC	1.80	2.08	84	105.4	251.3
Chemicals and chemical products	Ch	7.68	9.95	1191	55.5	85.4
Basic pharmaceutical products and preparations	Ph	13.26	15.28	294	61.6	85.2
Rubber and plastic products	RP	4.37	3.04	2026	36.0	41.4
Other non-metallic mineral products	NM	6.83	10.26	2137	30.4	53.2
Basic metals	BM	6.21	7.15	710	48.1	74.6
Fabricated metal products, except machinery	MP	3.82	2.07	7060	33.3	29.6
Computer, electronic, and optical products	EP	12.51	9.86	1233	42.9	38.2
Electrical equipment	El	4.54	3.84	1309	38.3	30.9
Machinery and equipment nec	Ma	12.42	8.47	3136	45.8	36.9
Motor vehicles, trailers, and semi-trailers	MV	7.72	6.36	881	33.9	43.2
Other transport equipment	Tr	1.15	0.86	330	38.6	46.0
Furniture	Fu	0.36	0.21	1563	19.8	20.2
Total		-	-	35850	-	-

Table 3: Descriptive statistics: Country coverage (% of total).

Country/Country Aggregate	VA	K	L	# firms
DE	27.81	23.98	17.82	4.35
ES	0.39	0.63	0.35	1.48
FR	9.31	7.47	7.69	7.20
GB	11.25	13.02	5.43	2.33
IT	5.11	4.17	4.68	25.06
PT	1.12	0.80	2.34	26.87
US	1.96	2.28	1.54	0.10
IL	0.47	1.24	0.49	0.09
OECD northern Europe	9.76	8.43	6.66	5.67
Other European OECD	1.80	1.52	1.22	0.44
Other non-European OECD	7.92	7.66	3.41	0.26
Eastern Europe	2.68	2.23	6.77	19.36
Other non-OECD	20.43	26.57	41.60	6.80

Note. OECD northern Europe includes NO, FI, SE, DK, and NL. Eastern Europe includes EE, GR, HU, LV, PL, SK, SI, CZ, RO, UA, BA, MK, HR, BG, RS, and CY. Other European OECD includes IS, IE, LU, AT, BE, and LT. Other non-European OECD includes NZ, CA, CH, AU, JP, MX, and TR. Other non-OECD includes RU, AR, SG, CL, ZA, BM, IQ, MY, KY, IN, PH, OM, BD, JO, RS, SA, UY, TZ, TN, PK, NA, LK, KE, EG, CV, BR, BH, AE, KR, CN, BW, CI, FJ, IR, KW, MA, NG, QA, TH, TT, HK, and TW.

Table 4: BIC values from the mixture analysis.

Sector	BIC_1	BIC_2	BIC_3	BIC_4	BIC_5	BIC_6	BIC_7	BIC_8	BIC_9	BIC_{10}	BIC_{min}
Fd	5010.87	3898.95	3741.37	3577.64	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	3577.64
Bv	1132.48	974.24	993.91	1010.26	1041.50	n.c.	n.c.	n.c.	n.c.	n.c.	974.24
Tb	38.49	25.05	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	25.05
TX	2170.23	1902.11	1907.00	1820.22	1781.37	n.c.	n.c.	n.c.	n.c.	n.c.	1781.37
WA	2059.63	1343.36	1272.99	1212.16	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	1212.16
LP	1189.45	814.69	755.69	711.76	701.95	656.92	n.c.	n.c.	n.c.	n.c.	656.92
Wo	2815.85	2280.44	2215.38	2143.86	2112.48	2124.15	2073.52	n.c.	n.c.	n.c.	2073.52
PC	168.49	148.46	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	148.46
Pr	2111.21	1904.85	1891.29	1907.51	1840.34	n.c.	n.c.	n.c.	n.c.	n.c.	1840.34
Pa	1060.93	971.59	983.96	987.68	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	971.59
Ch	2058.61	1905.79	1936.74	1867.89	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	1867.89
Ph	528.28	485.00	484.01	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	484.01
RP	2726.69	2357.59	2329.68	2287.84	2248.99	2242.98	n.c.	n.c.	n.c.	n.c.	2242.98
NM	3438.23	2904.53	2821.25	2767.11	2708.72	2613.42	n.c.	n.c.	n.c.	n.c.	2613.42
BM	1255.25	1171.33	1186.54	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	1171.33
MP	9011.74	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	9011.74
El	1914.06	1695.30	1691.26	1655.78	1577.49	n.c.	n.c.	n.c.	n.c.	n.c.	1577.49
EP	1984.22	1758.36	1712.58	1651.17	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	1651.17
Ma	4799.58	4137.46	4080.23	4046.61	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	4046.61
MV	1431.00	1303.53	1278.07	1231.85	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	1231.85
Tr	602.40	584.83	555.29	530.60	n.c.	n.c.	n.c.	n.c.	n.c.	n.c.	530.60
Fu	2069.80	1523.02	1472.50	1482.03	1448.46	1420.63	1396.22	n.c.	n.c.	n.c.	1396.22

Note. n.c. = not converged.
Legend. Fd: Food products; Bv: Beverages; Tb: Tobacco products; TX: Textiles; WA: Apparel; LP: Leather and related products; Wo: Wood and products of wood and cork; Pa: Paper and paper products; Pr: Printing and reproduction of recorded media; PC: Coke and refined petroleum products; Ch: Chemicals and chemical products; Ph: Basic pharmaceutical products and pharmaceutical preparations; RP: Rubber and plastic products; NM: Other non-metallic mineral products; BA: Basic metals; MP: Fabricated metal products, except machinery and equipment; EP: Computer, electronic, and optical products; El: Electrical equipment; Ma: Machinery and equipment nec; MV: Motor vehicles, trailers, and semi-trailers; Tr: Other transport equipment; Fu: Furniture.

Table 5: Mixture regressions: Estimated production function parameters.

Sector	# Tech	α_1	β_1	α_2	β_2	α_3	β_3	α_4	β_4	α_5	β_5
Fd	4	0.190	0.121	0.767	0.379	0.000	0.147	0.000	0.180		
Bv	1	0.055	0.199								
TX	2	-0.093	0.131	0.000	0.238						
WA	3	0.326	0.120	0.000	0.267	-0.018	0.150				
LP	5	-0.251	0.101	-1.232	0.334	-0.687	0.153	0.458	0.313	0.000	0.119
Wo	3	-0.321	0.246	-0.525	0.499	0.000	0.149				
Pa	2	-0.055	0.175	0.000	0.207						
Pr	3	-0.130	0.134	-0.433	0.343	0.000	0.202				
Ch	3	-0.087	0.165	0.000	0.348	0.352	0.211				
Ph	2	1.879	0.126	-0.306	0.187						
RP	3	0.142	0.126	0.584	0.134	0.000	0.238				
NM	3	-0.076	0.155	-0.109	0.289	0.000	0.319				
BM	2	0.144	0.214	0.000	0.325						
MP	1	0.285	0.165								
EP	2	0.194	0.170	0.000	0.442						
El	3	-0.461	0.121	-0.291	0.185	0.000	0.186				
Ma	2	-0.317	0.117	0.000	0.174						
MV	3	-0.345	0.127	0.000	0.250	0.000	0.439				
Tr	1	0.350	0.217								
Fu	3	-0.100	0.189	-1.495	0.453	0.000	0.145				

All reported β parameters are statistically significant at the 1% level. Clusters are dropped (and the corresponding coefficients are not reported) when $\beta < 0.1$ or $\beta > 0.9$. Both α and β are considered to be equal to zero when they are not statistically different from zero at the 1% level of statistical significance.
Legend. Fd: Food products; Bv: Beverages; Tb: Tobacco products; TX: Textiles; WA: Apparel; LP: Leather and related products; Wo: Wood and products of wood and cork; Pa: Paper and paper products; Pr: Printing and reproduction of recorded media; PC: Coke and refined petroleum products; Ch: Chemicals and chemical products; Ph: Basic pharmaceutical products and pharmaceutical preparations; RP: Rubber and plastic products; NM: Other non-metallic mineral products; BA: Basic metals; MP: Fabricated metal products, except machinery and equipment; EP: Computer, electronic, and optical products; El: Electrical equipment; Ma: Machinery and equipment nec; MV: Motor vehicles, trailers, and semi-trailers; Tr: Other transport equipment; Fu: Furniture.

Table 6: Mixture regressions: Total probability by technology group.

SECTOR	$prob_1$	$prob_2$	$prob_3$	$prob_4$	$prob_5$
Fd	1331	487	1525	988	-
Bv	422	-	-	-	-
TX	453	498	-	-	-
WA	70	559	905	-	-
LP	464	11	15	207	438
Wo	290	8	582	-	-
Pa	440	329	-	-	-
Pr	503	277	577	-	-
Ch	514	449	80	-	-
Ph	14	250	-	-	-
RP	831	14	789	-	-
NM	1005	26	573	-	-
BM	368	342	-	-	-
MP	7060	-	-	-	-
EP	552	139	-	-	-
El	744	26	443	-	-
Ma	1438	1371	-	-	-
MV	251	127	459	-	-
Tr	244	-	-	-	-
Fu	430	52	448	-	-

For each technology m (with $m = 1, \ldots, 5$), the reported values represent the sum, over all the firms in the sector, of the probability of using technology m.
Legend. Fd: Food products; Bv: Beverages; Tb: Tobacco products; TX: Textiles; WA: Apparel; LP: Leather and related products; Wo: Wood and products of wood and cork; Pa: Paper and paper products; Pr: Printing and reproduction of recorded media; PC: Coke and refined petroleum products; Ch: Chemicals and chemical products; Ph: Basic pharmaceutical products and pharmaceutical preparations; RP: Rubber and plastic products; NM: Other non-metallic mineral products; BA: Basic metals; MP: Fabricated metal products, except machinery and equipment; EP: Computer, electronic, and optical products; El: Electrical equipment; Ma: Machinery and equipment nec; MV: Motor vehicles, trailers, and semi-trailers; Tr: Other transport equipment; Fu: Furniture.

Table 7: WTFP and BTFP estimates: Sectoral averages.

	Average			Standard Deviation		
Sector	$WTFP_i$*	$BTFP_i$*	$BTFP_i/WTFP_i$	$WTFP_i$*	$BTFP_i$*	$BTFP_i/WTFP_i$
Fd	31.906	30.451	0.88	23.1254	29.84432	2.33
Bv	16.591	0.000	0.00	20.694	0.000	0.00
TX	17.427	6.233	0.31	10.604	10.203	0.46
WA	34.990	26.717	0.52	30.113	16.392	0.85
LP	32.180	59.749	2.67	22.695	31.852	4.78
Wo	32.757	11.851	1.21	22.008	25.793	4.39
Pa	52.357	9.740	0.20	37.872	15.933	0.36
Pr	27.059	23.949	0.99	23.592	19.273	5.59
Ch	29.539	43.144	1.32	18.365	61.048	5.88
Ph	39.361	65.329	6.47	28.887	98.280	109.29
RP	18.103	109.059	10.61	15.526	51.932	29.35
NM	20.216	44.674	2.24	12.750	22.421	8.95
BM	78.822	24.968	0.29	46.634	16.643	0.30
MP	63.469	0.000	0.00	40.706	0.000	0.00
EP	19.246	45.203	8.13	21.112	28.217	11.37
El	33.731	55.588	2.35	19.584	27.540	4.17
Ma	34.916	9.289	0.24	22.286	17.506	1.09
MV	29.334	46.874	1.73	19.409	49.955	74.14
Tr	51.144	0.000	0.00	30.365	0.000	0.00
Fu	40.039	8.845	0.38	23.968	21.641	2.31
Total	31.265	34.191	2.11	24.59	27.09	13.86

* % of frontier values.

Total average BTFP is calculated over positive sectoral averages only (1-technology sectors are omitted from the total average). Both WTFP and BTFP are weighted by the sectoral share of employees over total employees.

Legend. Fd: Food products; Bv: Beverages; Tb: Tobacco products; TX: Textiles; WA: Apparel; LP: Leather and related products; Wo: Wood and products of wood and cork; Pa: Paper and paper products; Pr: Printing and reproduction of recorded media; PC: Coke and refined petroleum products; Ch: Chemicals and chemical products; Ph: Basic pharmaceutical products and pharmaceutical preparations; RP: Rubber and plastic products; NM: Other non-metallic mineral products; BA: Basic metals; MP: Fabricated metal products, except machinery and equipment; EP: Computer, electronic, and optical products; El: Electrical equipment; Ma: Machinery and equipment nec; MV: Motor vehicles, trailers, and semi-trailers; Tr: Other transport equipment; Fu: Furniture.

Table 8: Markers of WTFP and BTFP (OLS regressions).

	WTFP_i	BTFP_i
COUNTRY-SECTOR VARIABLES		
Tech Balance	1.935*	-2.674***
	(0.566)	(0.355)
FIRM-LEVEL VARIABLES		
Firm Age	3.696***	0.070
	(0.483)	(0.350)
Listed	-2.932	15.502
	(8.391)	(11.644)
Firm Intangibles	-0.408**	-0.367***
	(0.153)	(0.125)
Liquidity Ratio	-6.836***	-2.752***
	(0.504)	(0.365)
Multinational	-2.936***	-0.475
	(0.982)	(0.751)
Labor Input	-0.709**	-0.746**
	(0.331)	(0.322)
Constant	30.222***	13.683***
	(4.180)	(4.067)
# obs.	4714	4714
Country FE	yes	yes
Sector FE	yes	yes
R^2	0.414	0.366

Robust standard errors are in parentheses. All variables are in logs. Statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: WTFP and BTFP from the subsample: Sectoral averages.

Sector	WTFP_i* gap	BTFP_i*	$\text{BTFP}_i/\text{WTFP}_i$	# technology clusters	# firms
Fd	40.367	44.648	1.506	4	1214
Bv	14.301	0	0	1	182
TX	27.684	11.887	0.421	2	391
WA	34.454	0	0	1	278
LP	31.416	67.214	2.872	3	267
Pa	13.815	18.803	2.081	2	355
Pr	40.784	1.447	0.039	2	455
Ch	29.127	5.808	0.228	2	560
Ph	24.413	21.393	0.942	2	123
RP	39.462	0	0	1	785
NM	14.891	0	0	1	664
BM	49.400	0	0	1	317
MP	18.812	0	0	1	2609
El	6.461	0	0	1	564
Ma	14.819	0	0	1	1650
MV	17.412	3.905	0.249	2	272
Tr	32.468	0	0	1	171
Fu	8.020	0	0	1	363
Total	31.820	25.042	1.052	–	11220 (sum)

* % of frontier values.

Note. Subsample mixture estimations are run on firm-level data from Finland, Italy, Germany, France, the Netherlands, and Spain.

Legend. Fd: Food products; Bv: Beverages; Tb: Tobacco products; TX: Textiles; WA: Apparel; LP: Leather and related products; Wo: Wood and products of wood and cork; Pa: Paper and paper products; Pr: Printing and reproduction of recorded media; PC: Coke and refined petroleum products; Ch: Chemicals and chemical products; Ph: Basic pharmaceutical products and pharmaceutical preparations; RP: Rubber and plastic products; NM: Other non-metallic mineral products; BA: Basic metals; MP: Fabricated metal products, except machinery and equipment; EP: Computer, electronic, and optical products; El: Electrical equipment; Ma: Machinery and equipment nec; MV: Motor vehicles, trailers, and semi-trailers; Tr: Other transport equipment; Fu: Furniture.

Figure 1: WTFP and BTFP: Graphical representation with two technologies.

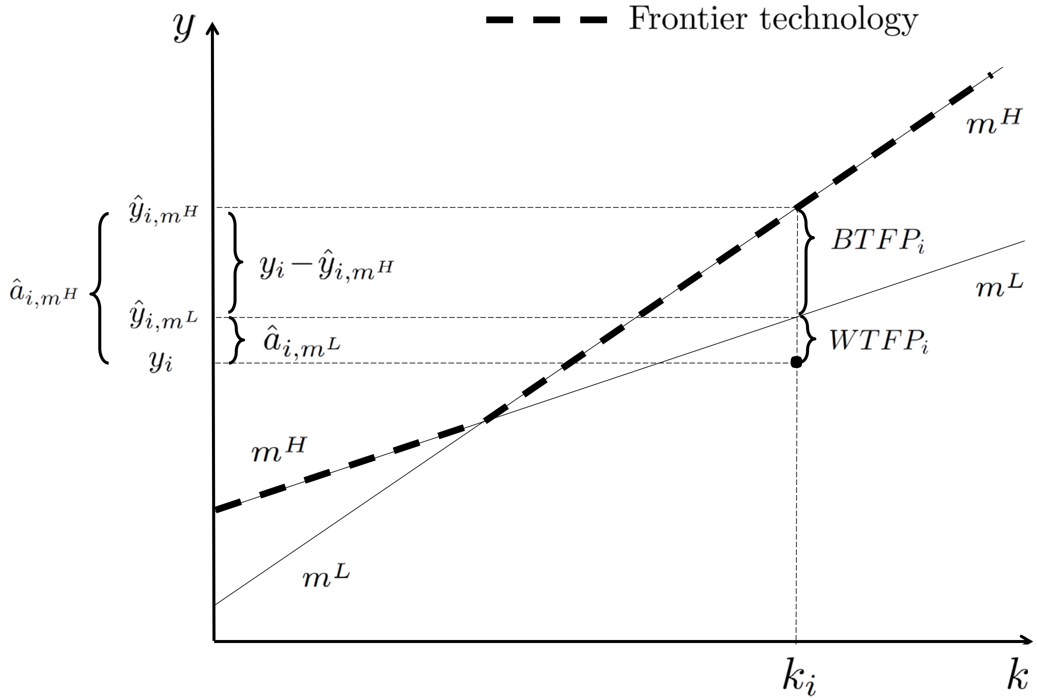
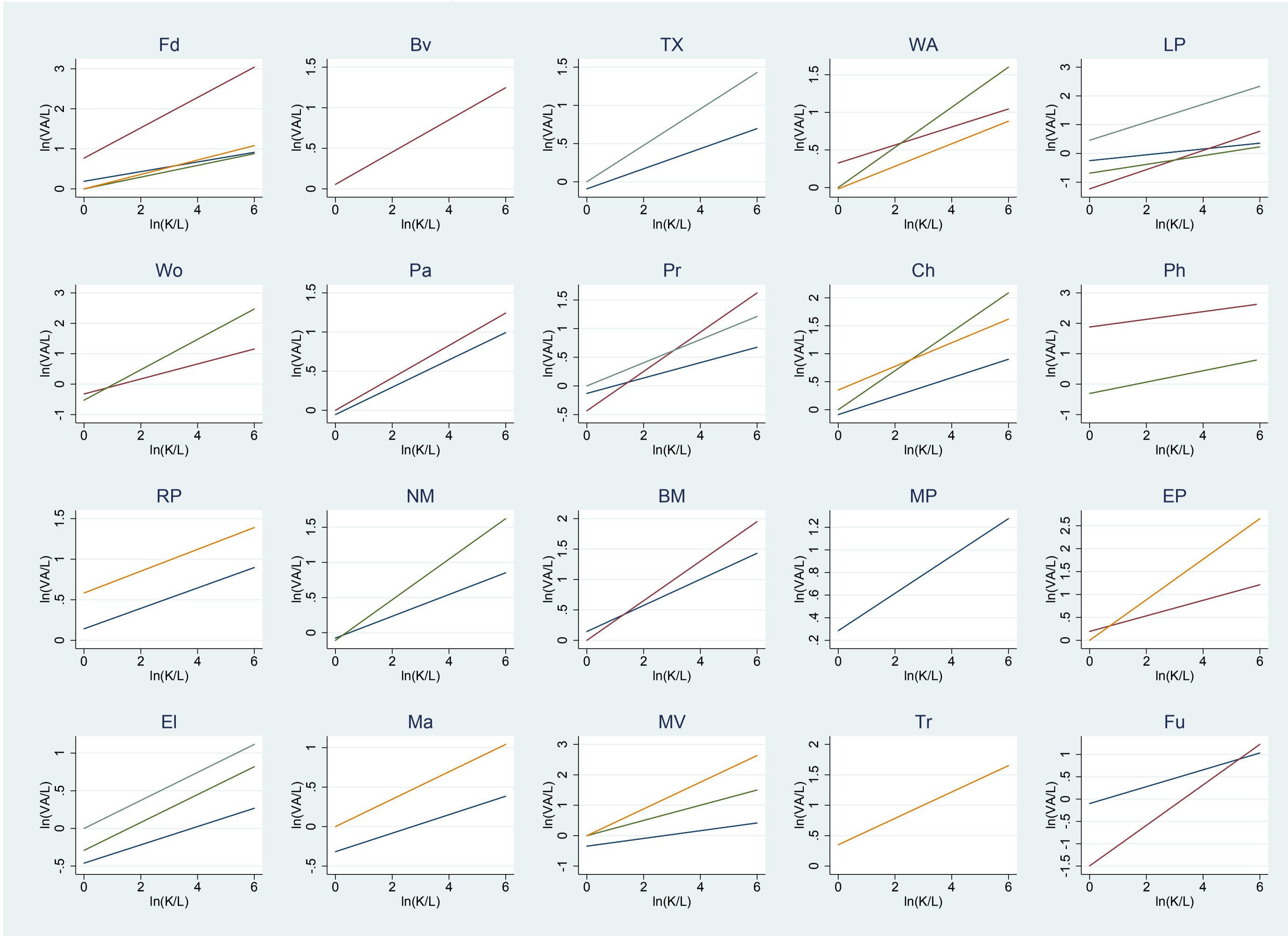
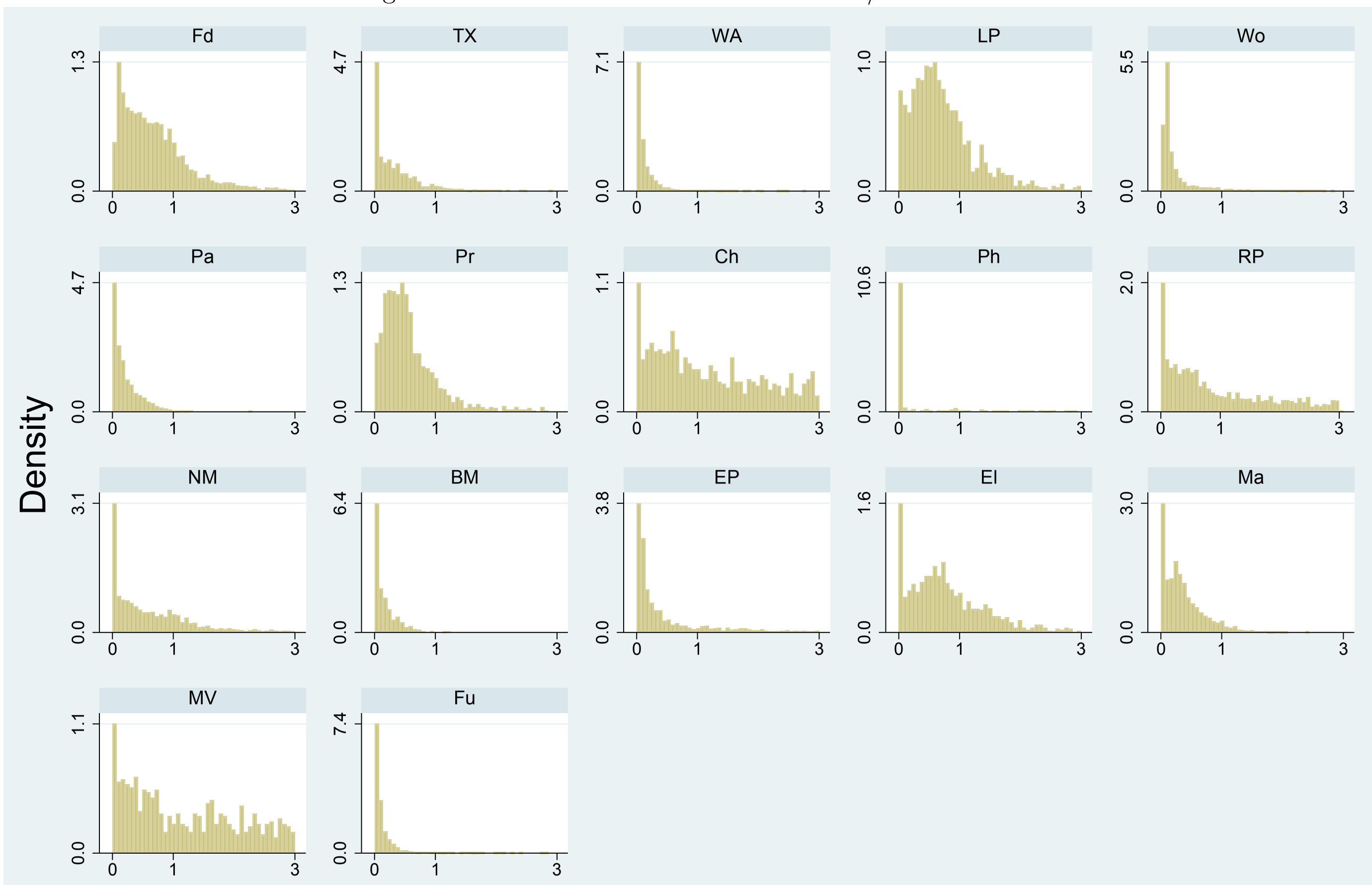


Figure 2: Estimated production functions.



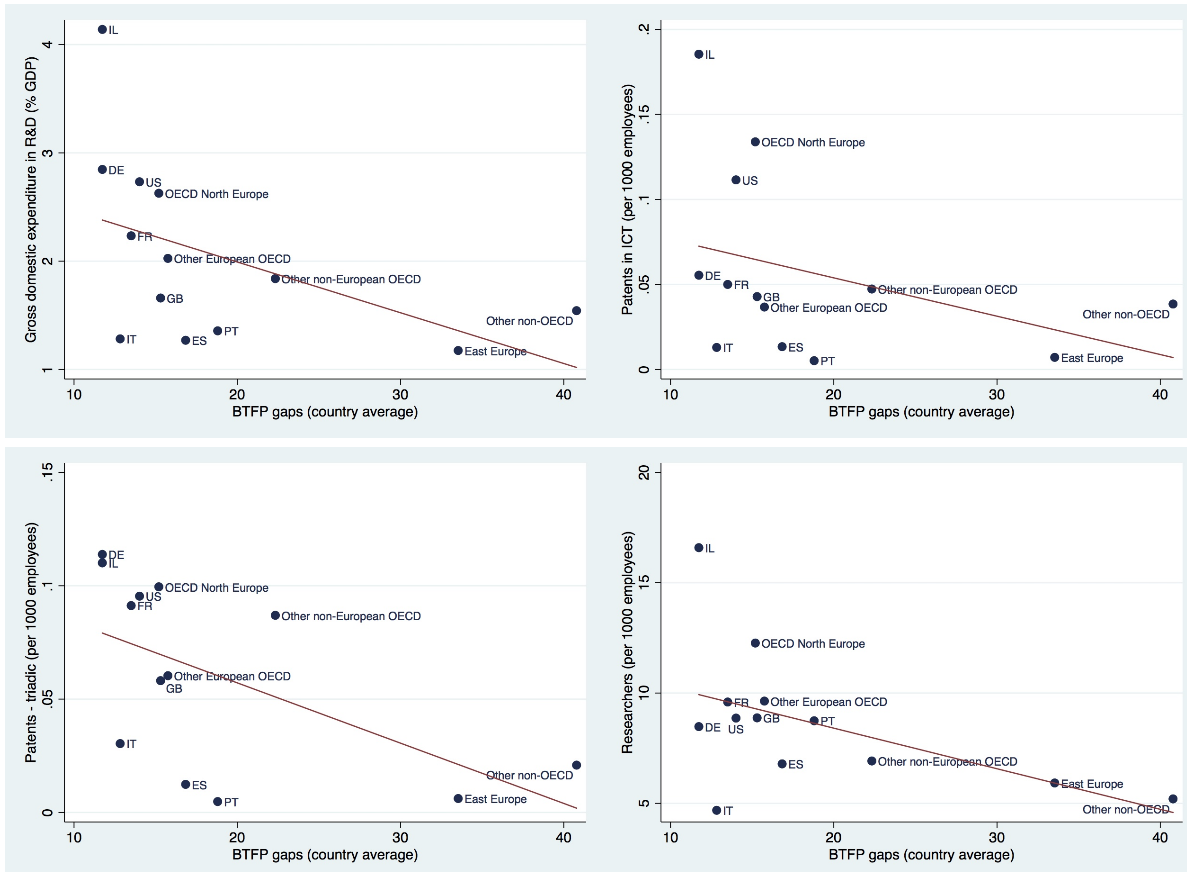
Legend. Fd: Food products; Bv: Beverages; Tb: Tobacco products; TX: Textiles; WA: Apparel; LP: Leather and related products; Wo: Wood and products of wood and cork; Pa: Paper and paper products; Pr: Printing and reproduction of recorded media; PC: Coke and refined petroleum products; Ch: Chemicals and chemical products; Ph: Basic pharmaceutical products and pharmaceutical preparations; RP: Rubber and plastic products; NM: Other non-metallic mineral products; BA: Basic metals; MP: Fabricated metal products, except machinery and equipment; EP: Computer, electronic, and optical products; El: Electrical equipment; Ma: Machinery and equipment nec; MV: Motor vehicles, trailers, and semi-trailers; Tr: Other transport equipment; Fu: Furniture.

Figure 3: Sectoral distribution of the BTFP/WTFP ratio.



Legend. Fd: Food products; Bv: Beverages; Tb: Tobacco products; TX: Textiles; WA: Apparel; LP: Leather and related products; Wo: Wood and products of wood and cork; Pa: Paper and paper products; Pr: Printing and reproduction of recorded media; PC: Coke and refined petroleum products; Ch: Chemicals and chemical products; Ph: Basic pharmaceutical products and pharmaceutical preparations; RP: Rubber and plastic products; NM: Other non-metallic mineral products; BA: Basic metals; MP: Fabricated metal products, except machinery and equipment; EP: Computer, electronic, and optical products; El: Electrical equipment; Ma: Machinery and equipment nec; MV: Motor vehicles, trailers, and semi-trailers; Tr: Other transport equipment; Fu: Furniture.

Figure 4: Country-level BTFP gaps and technological patterns.



Note. Proxies of technological patterns (vertical axis) are averaged over 2014–2016 (source: STAN database; OECD, 2018).

Labour productivity and firm-level TFP with technology-specific production functions

(Online Appendix)

A Mixture Regressions: Additional Statistical Output

In this appendix, we provide additional graphical output of the mixture analysis of the technology-specific production functions.

Figure 5 and Figure 6 show the within-sector distribution of the WTFP and BTFP gaps, respectively. Recall that the WTFP gaps are obtained at the firm level as $WTFP_{\text{best5\%}} - WTFP_i$, where $WTFP_{\text{best5\%}}$ is the average $WTFP_i$ of the best 5% of firms in the WTFP distribution in a sector, while the BTFP gaps are obtained as $BTFP_i$ multiplied by minus one. The sectoral panels in both figures are censured on the right-hand side where convenient to ensure graphical clearness. For both the WTFP and the BTFP gaps, we observe substantial heterogeneity within and between sectors. For the BTFP gaps, in particular, sectors characterized by a relatively large group of firms using the locally optimal technology also show distributions with higher skewness and a thinner right-hand tail.

[insert Figure 5 about here]

[insert Figure 6 about here]

Figure 7 provides a sectoral example of the probabilistic clustering obtained by means of a mixture model estimation of the technology-specific production functions. For clarity, we choose the basic metals (BM) sector, where only two technological clusters emerge from the analysis, and plot the firm-level observations with a color scale reflecting the probability of belonging to cluster 1 (i.e., $Prob_1$). Given that only two clusters are obtained for this sector, the firm probability of belonging to cluster 2 can be obtained as $1 - Prob_1$. The figure shows that firms with a relatively high probability of belonging to cluster 1 are in fact clustered in a defined region of the distribution, which we interpret as the pattern of a technology-specific production function.

[insert Figure 7 about here]

Figure 8 plots the correlation between the benchmark and re-estimated BTFP and WTFP values discussed in Section 7. The benchmark values are those obtained by running the mixture model estimation of the technology-specific production functions on worldwide international data, while the re-estimated values are obtained on a restricted sample of data for six countries (Finland, Italy, Germany, France, the Netherlands, and Spain), as in Bos *et al.* (2010). Reduced cross-country coverage affects our technology clustering by reducing the number of technologies possibly identified (when run on the subsample, our mixture model can identify one or two clusters in most of the sectors, with only two sectors showing a larger number of clusters). Figure 8 shows that a positive correlation emerges between the benchmark and re-estimated BTFP and WTFP values, suggesting that the rank correlation between the BTFP and WTFP values obtained in the two cases is generally preserved. As expected, however, the measures obtained in the two estimation exercises do not precisely overlap. This confirms that only large data coverage allows our mixture model to identify all the technologies possibly available worldwide.

[insert Figure 8 about here]

B Comparison with Conventional TFP Estimates

In this section, we re-estimate the production functions at the sectoral level without taking the presence of different technologies into account. We present the results with and without controlling for simultaneity.

As a first comparison, Figure 9 reports the production functions estimated by simple OLS (dashed line) together with our technology-specific production functions. Essentially, the former can be seen as one-technology mixture regression ($M = 1$). More precisely, being the mixture regression carried out through weighted least squares, we might see the OLS-estimated coefficients as a weighted average of the M technology-specific coefficients, with weights (i.e., Θ_m) given by the ratio of the number of firms in the m-technology group to the total number of firms in the sector, namely, $\hat{\beta} = \sum_{m=1}^{M} \hat{\beta}_m \frac{\Theta_m}{\sum_{m=1}^{M} \Theta_m}$. The same reasoning applies to α.

[insert Figure 9 about here]

Consistently, Figure 9 shows that the dashed line lies essentially in between the technology-specific production functions. Columns (2) to (5) of Table 10 report the OLS-estimated coefficients.

[insert Table 10 about here]

As discussed in Section 3, among the issues highlighted by the literature on estimating production functions, recent works have focused on simultaneity bias. The source of simultaneity bias is the fact that information on actual productivity, although unknown to the econometrician, is known to the firm when it decides on the amount of inputs. This biases the production function parameters obtained through the least squares estimates because of the potential correlation between the regressors and error term.

A successful stratagem suggested by the literature for addressing this issue consists of recovering the a_i component from the traces it leaves in the observed behavior of the firm. Key studies examining this approach, which is commonly referred to as the proxy variable method, include the semi-parametric approach put forward by Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg *et al.* (2006) as well as the GMM estimation adopted by Wooldridge (2009). The basic idea of this methodology consists of identifying a (proxy) variable that reacts to the changes in TFP observed by a firm and is thus a function of these changes. Insofar as this function is invertible, its inverse may be calculated and plugged into the production function-estimating equation. Olley and Pakes (1996) suggest resorting to investment as a proxy, whereas Levinsohn and Petrin (2003) use intermediates. Doraszelski and Jaumandreu (2012), who extend Olley and Pakes (1996), show that a firm's TFP is stochastically affected by its investment in knowledge (considered in terms of R&D).[16]

The implementation of such a semi-parametric approach within our mixture model framework raises identification problems. In particular, the way in which the proxy variable (either investment or materials) reacts to changes in technology and TFP should be specified separately because firms' input choices may be differently correlated with the technology parameters and firms' TFP. Our model of technology adoption developed in Section 3 is meant to deal with these two sources of simultaneity separately, without the need to rely on a specific proxy variable.

Similarly to the semi-parametric strategy, our approach to simultaneity entails a two-step estimation. However, instead of trying to include the whole productivity term (as proxied by the inverted investment or intermediates demand function) in the production function, our strategy seeks to take its effect (on the capital and labor choice) into account by controlling for its covariance with labor and capital. We exploit the fact that the correlations between capital and technology as well as between either capital or labor and TFP flow into the residuals of the system of equations in Equation (11), consisting of the K policy function and static condition for L. Once estimated, these residuals allow us to control for simultaneity in the mixture analysis. As well as controlling for the simultaneity caused by idiosyncratic TFP, this strategy can control for any other factor affecting the input choice through the relationships in Equation (11). This aspect is particularly relevant in light of the results provided by Gandhi *et al.* (2018) on the identification problems characterizing the semi-parametric approach. Such problems stem from the circumstance that even when the invertibility conditions (for the proxy variable demand function) identified by the literature are satisfied, it is still possible to construct a continuum of observationally equivalent production functions that cannot be distinguished from the true production function in the data.[17] By not resorting to any proxy variable, our approach is in principle immune to this order of problems.

Second, our timing of input choice is strictly in line with Olley and Pakes (1996), with labor dealt with as a "perfectly variable" input decided at the time of production. While this is also true in Levinsohn and Petrin (2003), Ackerberg *et al.* (2015) highlight the collinearity problems associated with such decision timing and suggest a slightly different specification in which labor is a "less variable" input than intermediates chosen one "subperiod" before productivity is observed. This implies that differently from Olley and Pakes (1996) and Levinsohn and Petrin (2003), in which the labor and capital coefficients are identified in the first and second steps respectively, the first step is only used to net the output function of the idiosyncratic productivity (and in our case also technology) shock, with both the capital and the labor coefficients retrieved in the second stage.[18] Although we use the same timing of input choice as Olley and Pakes (1996) and Levinsohn and Petrin (2003), our approach is not subject to the Ackerberg *et al.* (2015) critique, as both coefficients are estimated in the second stage in our case.

To understand how the different estimation strategies are reflected in the firm TFP distribution, Figure 10 compares our WTFP with the OLS-estimated TFP and TFP estimated through the Olley and Pakes (1996) procedure (the coefficients are reported in the four last columns of Table 10), used as a benchmark estimation

[16] Firms' productivity is assumed to evolve according to a Markov process, which is "shifted" (either positively or negatively) by R&D expenditure. The R&D choice gives rise to an additional policy function (besides the policy function for investment in physical capital) that, under the crucial assumption that the error in t is uncorrelated with the innovation choice in $t-1$, may be exploited in the production function estimation to purge the estimates from the part of the error correlated with the input choice. Loosely speaking, this approach allows us to estimate firms' TFP while controlling for simultaneity and the effect of innovation choices at the same time.

[17] The reason why is shown to be the assumption of flexible elasticity for the input used as a proxy.

[18] The most obvious hypothesis of the data-generating process for labor demand is for it to be a function of capital and productivity. Thus, labor demand is only a function of capital plus the variable chosen as the proxy (i.e., investment in Olley and Pakes (1996) and intermediates in Levinsohn and Petrin (2003)), showing that the labor coefficient cannot be identified in the first stage, as one cannot simultaneously estimate a fully non-parametric function of two variables along with a coefficient of a variable (labor) that is only a function of those same variables.

within the semi-parametric approach. The OLS approach tends to fatten the tails of the distribution, particularly by overstating the TFP of the most productive firms. Our methodology results in a distribution that lies in the middle of the OLS and Olley–Pakes ones because part of the correction imposed by the Olley–Pakes method is recognized to be related to firms' choice of technology rather than to TFP, and this is thus captured by our BTFP term.

[insert Figure 10 about here]

The OLS estimates in Figure 9 and Figure 10 include a correction for the simultaneity suggested by our model. To illustrate the effect of this correction, Table 10 reports the estimated coefficients and Figure 11 illustrates the distribution of the difference between the OLS estimates with and without the correction. This distribution looks reasonable and suggests the absence of specific patterns (e.g., a stronger effect on a more productive firm). Table 10 reports the estimated coefficients obtained without the correction.

[insert Figure 10 about here]

[insert Figure 11 about here]

C Relationship with the Misallocation Literature

A growing body of the literature studies how firm-level differences in TFP are reflected in aggregate productivity, including Alfaro *et al.* (2008), Banerjee and Duflo (2005), Bartelsman *et al.* (2013), Hsieh and Klenow (2009, 2010), Jones (2011), Restuccia and Rogerson (2008), and Olley and Pakes (1996). A key word in this stream of the literature is misallocation, namely lower aggregate TFP due to distortions in the allocation of inputs across units (Restuccia and Rogerson, 2013). The extent of misallocation is easily understood in terms of the deviation from the efficiency condition (i.e., in efficient markets, the MRP of an input equals its price) as well as, according to Hsieh and Klenow (2009, 2010), of the dispersion of revenue TFP.

Several authors have focused on the misallocation patterns in single countries: misallocation has increased in Spain (Garcia-Santana *et al.*, 2015; Gopinath *et al.*, 2015), Portugal (Dias *et al.*, 2016), and Italy (Calligaris *et al.*, 2018); remained constant in France (Bellone and Mallen-Pisano, 2013); and declined in Germany (Crespo and Segura-Cayuela, 2014) and Chile (Chen and Irarrazabal, 2014). The general conclusion reached by these studies is that the aggregate productivity losses associated with misallocation are substantial and that, as a consequence, eliminating or reducing the within-industry distortions is crucial to improve the performance of countries.

One may wonder how our approach with technological heterogeneity is reflected in the quantification of misallocation. As stated, technological heterogeneity is not a source but a manifestation of misallocation.

To understand why, it is useful to briefly resume the Hsieh and Klenow (2009) model. The rational for misallocation is that from standard profit maximization, firms choose the amount of capital K and labor L by equalizing the MRP of each input to its marginal cost. While this process equalizes the MRP of capital (MRPK) and MRP of labor (MRPL) across firms, when all firms face the same input cost, the presence of market distortions (e.g., access to credit) can drive wedges between MRPK and MRPL across firms. In this case, we say that capital and labor are "misallocated" across firms.

Let us start with a standard Cobb–Douglas technology with sector-specific production coefficients:

$$Y_{si} = A_{si} K_{si}^{\alpha_s} L_{si}^{1-\alpha_s} \tag{19}$$

and follow Hsieh and Klenow (2009) (hereafter HK) in denoting distortions that increase the marginal products of capital and labor by the same proportion as an output distortion (e.g., government restrictions on size, transportation costs, public output subsidies) by τ_{si}^{Y} and distortions that raise the marginal product of capital relative to labor (i.e., capital distortions) by τ_{si}^{K} (e.g., different access to credit). From the first-order condition of firm i, active in sector s, we find that

$$MRPK_{si} = P_{si} \frac{\partial \tilde{Y}_{si}}{\partial K_{si}} = \beta_s^K P_{si} \frac{Y_{si}}{K_{si}} = \tilde{W}^K \tag{20}$$

and

$$MRPL_{si} = P_{si} \frac{\partial \tilde{Y}_{si}}{\partial K_{si}} = (\beta_s^L) P_{si} \frac{Y_{si}}{L_{si}} = W^L, \tag{21}$$

where $\tilde{Y}_{si} = (1 - \tau^Y_{si})Y_{si}$ and $\tilde{W}^K = (1 + \tau^K_{si})W^K$, with W^K and W^L referring to the costs of capital and labor, respectively.

If $\tau^Y_{si} = \tau^K_{si} = 0 \ \forall i \in s$, firms face the same input cost, and the MRP of the two inputs is equalized across firms. In this case, capital and labor are perfectly (i.e., efficiently) allocated. When this happens, the within-sector distributions of MRPK and MRPL exhibit zero dispersion around the mean, as average MRPK in sector s (i.e., $\overline{MRPK}_s$) equals $MRPK_{si} \ \forall i \in s$ (and the same for MRPL). No misallocation emerges in this case.

The MRP equalization condition holds true independently of the way in which firms set P_{si} (i.e., independently of market structure), the only condition being the absence of distortions in the capital and labor markets.

Since the higher the dispersion, the more effective are the distortions, it would be relatively easy to investigate the presence, and magnitude, of resource misallocation by looking at the within-industry dispersion of MRPK and MRPL. However, if one is interested in the aggregate effects of the documented distortions, more structure is needed.

To this end, a useful strategy is suggested by HK, whose approach enables us to study the effect of misallocation on aggregate TFP. The intuition is simple and rests on the proportionality between firm TFP and the MRP of inputs.

Using (19), it is possible to write firm i's TFP as

$$TFP_{si} = A_{si} = \frac{Y_{si}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}}. \tag{22}$$

As statistical information on the physical output Y_{si} or firm price P_{si} is hardly available (see Foster *et al.*, 2008), TFP is usually calculated on the basis of firms' revenues (TFPR hereafter). Using (19), this yields

$$TFPR_{si} = P_{si}A_{si} = \frac{P_{si}Y_{si}}{K_{si}^{\alpha_s} L_{si}^{1-\alpha_s}}. \tag{23}$$

While using $TFPR_{si}$ instead of TFP_{is} usually represents a shortcoming in the analysis, this is not the case in the HK framework. Under specific assumptions on market structure, $TFPR_{si}$ can be shown to be not influenced by other firm-specific factors than the distortions τ^Y_{si} and τ^K_{si}. In particular, if each sector s is monopolistically competitive, firms' set prices according to

$$P_{si} = (\frac{\sigma}{\sigma-1})\beta_s (W^K)^{\beta^K_s}(W^L)^{\beta^L_s}\frac{(1+\tau^K_{si})^{\beta^K_s}}{(1-\tau^Y_{si})}\frac{1}{A_{si}}, \tag{24}$$

where $\frac{\sigma}{\sigma-1}$ is the markup and $\beta_s = \beta_s^{K-\beta^K_s}(\beta^L_s)^{-\beta^L_s}$ is the bundle of parameters associated with the Cobb–Douglas production function (19). Apart from A_{si}, the only firm-specific terms in Equation (24) are the distortions. When substituted into Equation (23), the pricing rule in Equation (24) yields

$$TFPR_{si} = \frac{\sigma}{\sigma-1}\beta_s (W^K)^{\beta^K_s}(W^L)^{\beta^L_s}\frac{(1+\tau^K_{si})^{\beta^K_s}}{(1-\tau^Y_{si})}. \tag{25}$$

According to Equation (25), the cross-firm variability of $TFPR_{si}$ is not influenced by other firm-specific factors than τ^K_{si} and τ^Y_{si} (in fact, term A_{si} cancels out).[19]

As a result, *the extent of misallocation can be studied by looking at the dispersion of the $TFPR_{si}$ distribution* instead of considering the distributions of $MRPK_{si}$ and $MRPL_{si}$.

Our framework introduces an additional misallocation mechanism. Indeed, in the presence of frictions in the capital and/or labor market(s), firms can be unable to switch to superior technologies and thus retain misaligned MRPs. Hence, the presence of different technologies can be a possible consequence of the firm-level distortions preventing firms from being located on the technological frontier.

To gain an intuition of the relationship with the HK approach, consider the simplified version of our setting presented in Section 2 and imagine two firms identical in everything to firm i except for the technology they adopted, with one of them (say firm h) using the frontier technology H and the other (firm l) using the non-frontier

[19]Moreover, HK show that $TFPR_{si}$ is proportional to the geometric average of $MRPK_{si}$ and $MRPL_{si}$:

$$TFPR_{si} \propto (MRPK_{si})^{\beta^K_s} \propto (MRPL_{si})^{\beta^L_s} \propto \frac{(1+\tau^K_{si})^{\beta^K_s}}{(1-\tau^Y_{si})}. \tag{26}$$

technology l. Now (abstracting, for ease of comparison, from the fact that description in Section 2 is carried out in intensive form) focus on the efficiency condition of capital and imagine that

$$MRPK_{s,h} = \beta_H^K P_{si} \frac{Y_{si}}{K_{si}} = \tilde{W}^K > W^K \tag{27}$$

and

$$MRPK_{s,l} = \beta_L^K P_{si} \frac{Y_{si}}{K_{si}} = \tilde{W}^K < W^K. \tag{28}$$

While eliminating misallocation would require shifting capital from firm l to firm h, *allowing the former to switch to the superior technology would increase $MRPK_{s,l}$ but it would not solve the misallocation issue*, as the inefficiency condition would persist until the two firms are not allowed to employ a higher amount of capital. Thus, *eliminating technological heterogeneity by bringing all firms to the technology frontier would not eliminate misallocation.*

On the contrary, the HK strategy of measuring misallocation through the lens of TFPR dispersion is crucially affected. In fact, in terms of Equation (25), we would have

$$TFPR_{sh} = \frac{\sigma}{\sigma-1}\beta_h (W^K)^{\beta_s^K}(W^L)^{\beta_s^L}\frac{(1+\tau_{si}^K)^{\beta_s^K}}{(1-\tau_{si}^Y)} > \frac{\sigma}{\sigma-1}\beta_l (W^K)^{\beta_s^K}(W^L)^{\beta_s^L}\frac{(1+\tau_{si}^K)^{\beta_s^K}}{(1-\tau_{si}^Y)} = TFPR_{sl}. \tag{29}$$

With firm l switching to the superior technology, TFPR dispersion would drop to zero even if, as seen, misallocation is still present. Thus, taking the presence of multiple technologies into account in the HK framework entails that *TFPR dispersion is no longer a sufficient statistic for the dispersion of the MRP of factors and thus cannot be used to quantify the extent of misallocation.*[20] The reason why is that TFPR can also differ across firms due to technological differences as well as because of differences in the MRPs of inputs.

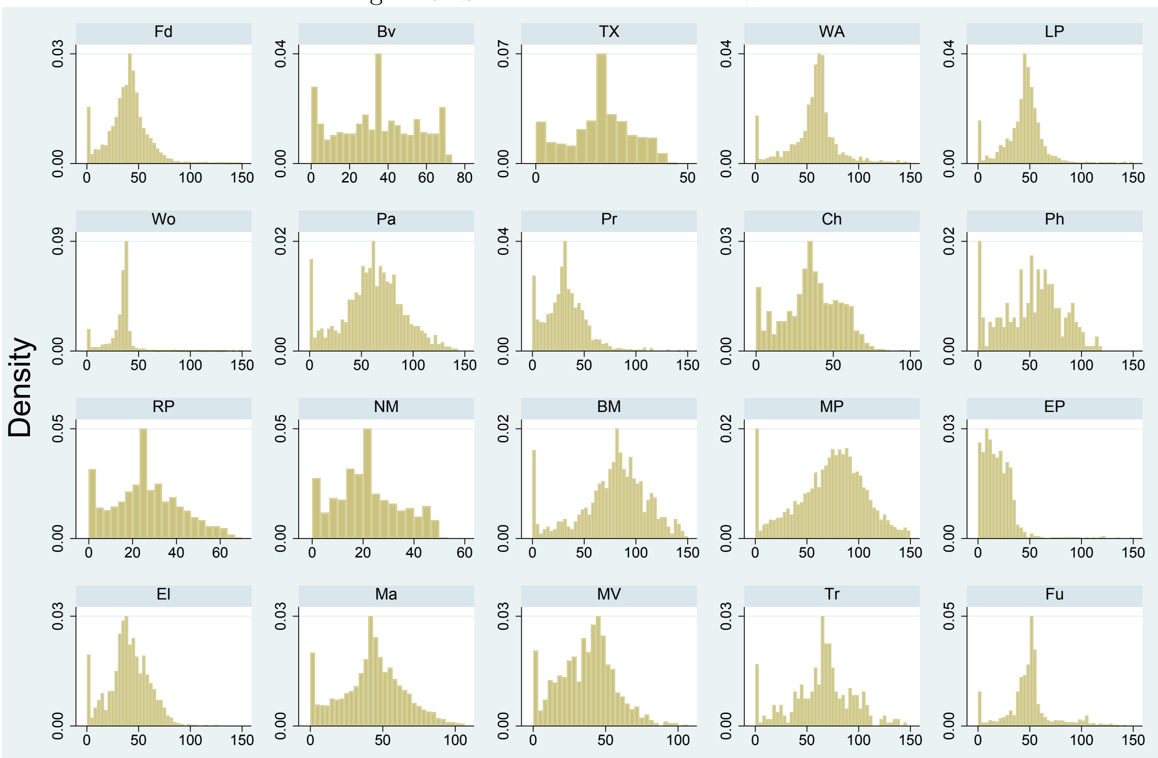
[20]The statistical counterpart of this argument may be drawn by Duclos *et al.* (2004). In a unimodal distribution, a reduced dispersion (a single squeeze) is indicative of less inequality, which, in the HK setting, would mean less misallocation. In a bimodal (or higher-order) distribution, reduced dispersions around the modes (a double squeeze) would imply less dispersion but more polarization, so only firms within a technological cluster would become more similar.

Table 10: Sectoral production functions estimated using different methodologies.

	OLS with correction				**OLS without correction**				**Olley–Pakes**			
SECTOR	α	Std. Err.	β	Std. Err.	α	Std. Err.	β	Std. Err.	α	Std. Err.	β	Std. Err.
Fd	-0.031	0.304	0.429	0.012	1.532	0.331	0.309	0.009	-0.198	0.019	0.370	0.106
Bv	0.145	0.069	0.525	0.029	1.734	0.137	0.461	0.026	-0.046	0.066	0.337	0.291
TX	1.420	0.005	0.314	0.020	3.896	0.029	0.179	0.011	-0.102	0.030	0.562	0.252
WA	-0.015	0.209	0.329	0.014	2.131	0.196	0.182	0.008	-0.100	0.019	0.285	0.224
LP	0.257	0.058	0.448	0.020	2.771	0.055	0.253	0.011	-0.104	0.026	0.492	0.138
Wo	-0.272	0.021	0.375	0.015	1.932	0.045	0.270	0.011	-0.134	0.026	0.575	0.153
Pa	-0.387	0.127	0.451	0.029	2.202	0.118	0.279	0.018	-0.072	0.054	0.290	0.146
Pr	-0.026	0.189	0.253	0.019	2.958	0.159	0.148	0.011	-0.175	0.028	0.312	0.139
Ch	0.512	0.169	0.414	0.023	3.227	0.175	0.284	0.016	0.049	0.043	0.330	0.096
Ph	0.994	0.420	0.312	0.051	4.096	0.457	0.238	0.039	0.052	0.191	-0.092	0.522
RP	0.002	0.246	0.326	0.018	2.425	0.227	0.221	0.012	0.011	0.033	0.378	0.147
NM	-0.006	0.355	0.428	0.016	2.230	0.362	0.284	0.011	-0.038	0.025	0.673	0.196
BM	0.820	0.127	0.384	0.029	3.456	0.126	0.255	0.019	0.082	0.076	0.191	0.179
MP	-0.022	0.302	0.307	0.009	3.126	0.253	0.168	0.005	-0.114	0.013	0.283	0.128
EP	-0.895	0.308	0.098	0.021	2.468	0.297	0.063	0.013	0.057	0.039	-0.063	0.164
El	-0.310	0.011	0.217	0.020	2.820	0.034	0.130	0.012	-0.094	0.031	0.075	0.077
Ma	0.411	0.081	0.178	0.014	3.760	0.065	0.093	0.007	-0.049	0.020	0.110	0.080
MV	0.310	0.079	0.221	0.028	3.211	0.080	0.128	0.016	0.063	0.059	0.152	0.151
Tr	0.714	0.027	0.284	0.041	3.661	0.075	0.161	0.023	-0.023	0.065	-0.047	0.177

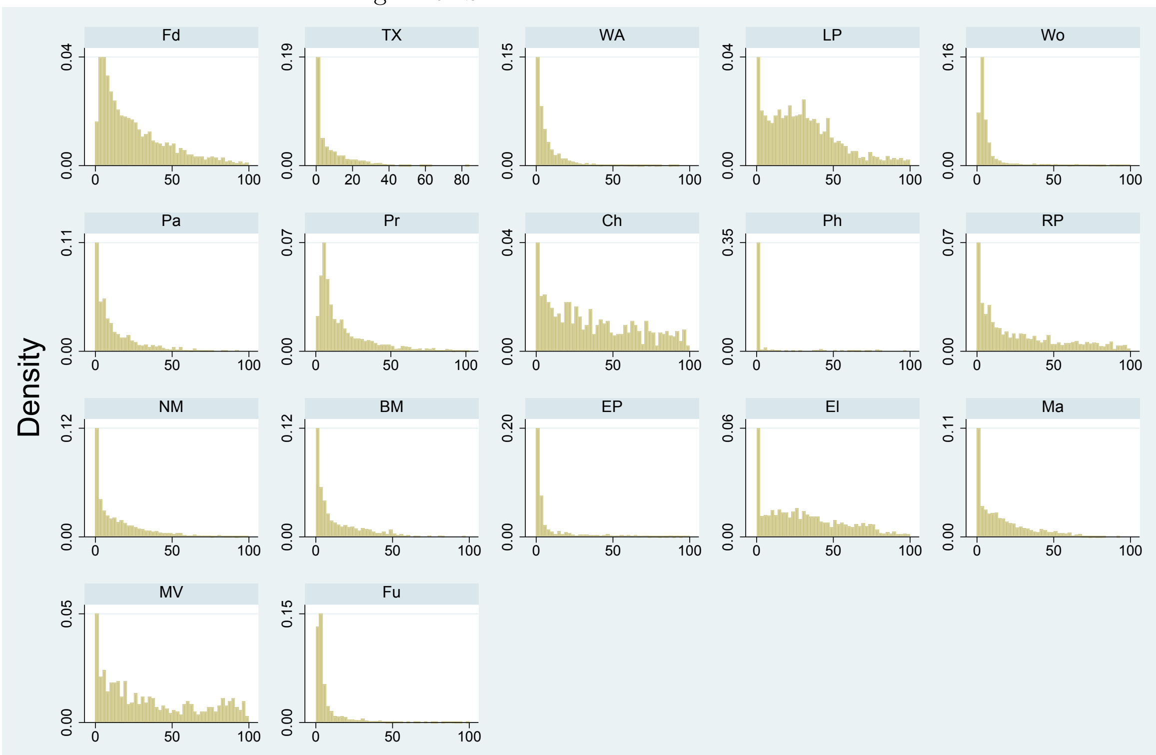
Legend. Fd: Food products; Bv: Beverages; Tb: Tobacco products; TX: Textiles; WA: Apparel; LP: Leather and related products; Wo: Wood and products of wood and cork; Pa: Paper and paper products; Pr: Printing and reproduction of recorded media; PC: Coke and refined petroleum products; Ch: Chemicals and chemical products; Ph: Basic pharmaceutical products and pharmaceutical preparations; RP: Rubber and plastic products; NM: Other non-metallic mineral products; BA: Basic metals; MP: Fabricated metal products, except machinery and equipment; EP: Computer, electronic, and optical products; El: Electrical equipment; Ma: Machinery and equipment nec; MV: Motor vehicles, trailers, and semi-trailers; Tr: Other transport equipment; Fu: Furniture.

Figure 5: Sectoral distribution of WTFP.



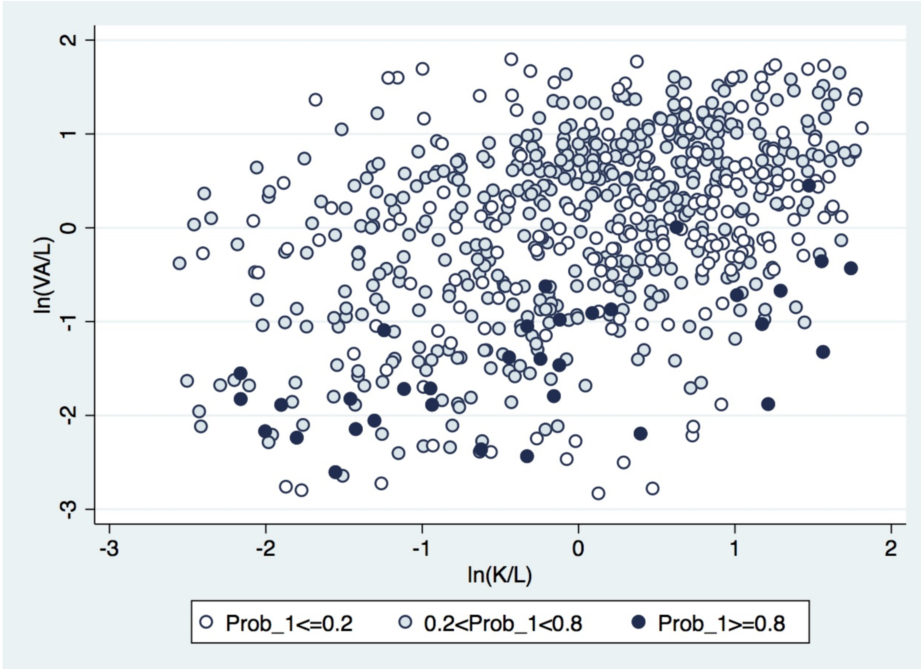
Legend. Fd: Food products; Bv: Beverages; Tb: Tobacco products; TX: Textiles; WA: Apparel; LP: Leather and related products; Wo: Wood and products of wood and cork; Pa: Paper and paper products; Pr: Printing and reproduction of recorded media; PC: Coke and refined petroleum products; Ch: Chemicals and chemical products; Ph: Basic pharmaceutical products and pharmaceutical preparations; RP: Rubber and plastic products; NM: Other non-metallic mineral products; BA: Basic metals; MP: Fabricated metal products, except machinery and equipment; EP: Computer, electronic, and optical products; El: Electrical equipment; Ma: Machinery and equipment nec; MV: Motor vehicles, trailers, and semi-trailers; Tr: Other transport equipment; Fu: Furniture.

Figure 6: Sectoral distribution of BTFP.



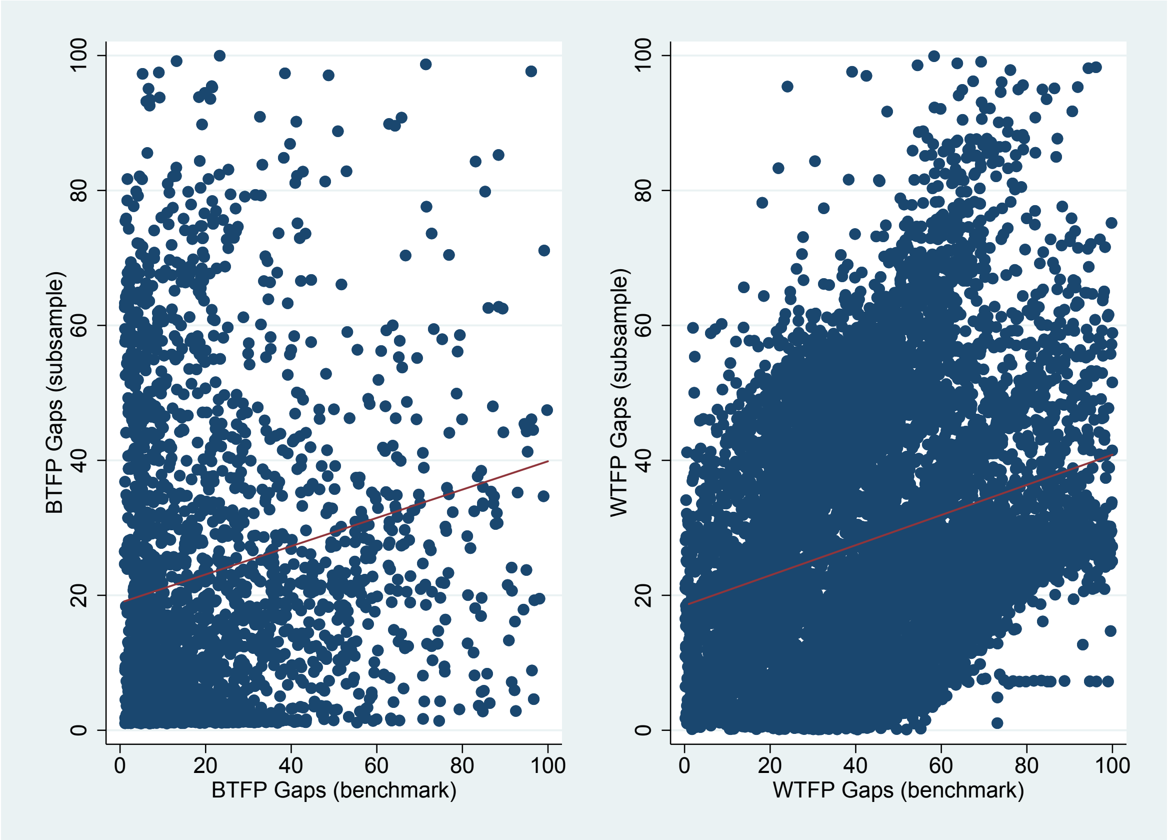
Legend. Fd: Food products; Bv: Beverages; Tb: Tobacco products; TX: Textiles; WA: Apparel; LP: Leather and related products; Wo: Wood and products of wood and cork; Pa: Paper and paper products; Pr: Printing and reproduction of recorded media; PC: Coke and refined petroleum products; Ch: Chemicals and chemical products; Ph: Basic pharmaceutical products and pharmaceutical preparations; RP: Rubber and plastic products; NM: Other non-metallic mineral products; BA: Basic metals; MP: Fabricated metal products, except machinery and equipment; EP: Computer, electronic, and optical products; El: Electrical equipment; Ma: Machinery and equipment nec; MV: Motor vehicles, trailers, and semi-trailers; Tr: Other transport equipment; Fu: Furniture.

Figure 7: Estimated probability of belonging to a given technology cluster: The basic metals (BM) sector.



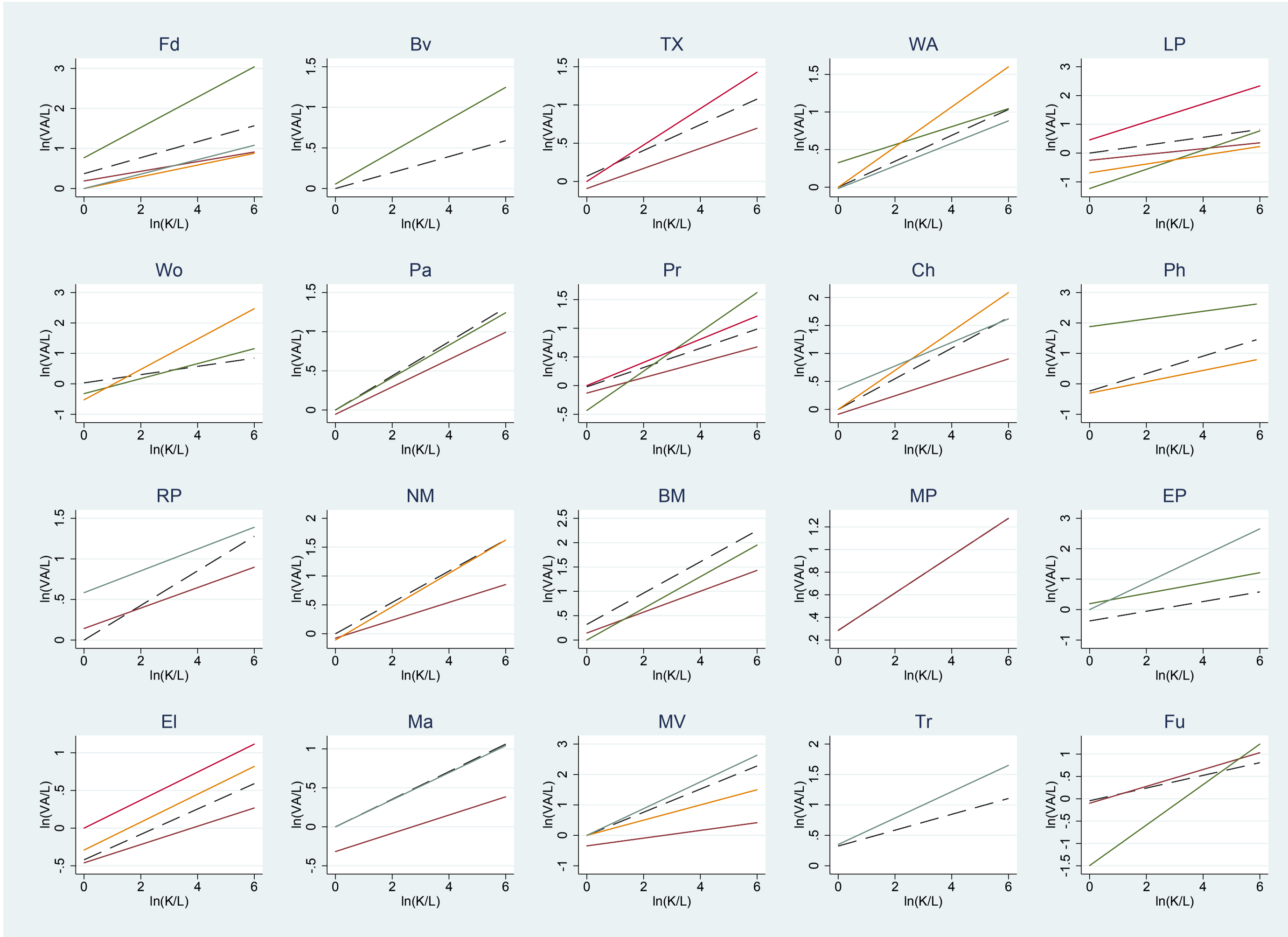
Note. Firm-level observations are plotted. Color scale reproduces the probability classes of belonging to technology cluster 1 (for BM, two clusters are obtained from the mixture analysis).

Figure 8: Correlation between the benchmark and re-estimated (subsample) BTFP and WTFP.



Note. Subsample mixture estimations are run on firm-level data from Finland, Italy, Germany, France, the Netherlands, and Spain.

Figure 9: Estimated production functions: Comparison with standard OLS (dashed line)



Legend. Fd: Food products; Bv: Beverages; Tb: Tobacco products; TX: Textiles; WA: Apparel; LP: Leather and related products; Wo: Wood and products of wood and cork; Pa: Paper and paper products; Pr: Printing and reproduction of recorded media; PC: Coke and refined petroleum products; Ch: Chemicals and chemical products; Ph: Basic pharmaceutical products and pharmaceutical preparations; RP: Rubber and plastic products; NM: Other non-metallic mineral products; BA: Basic metals; MP: Fabricated metal products, except machinery and equipment; EP: Computer, electronic, and optical products; El: Electrical equipment; Ma: Machinery and equipment nec; MV: Motor vehicles, trailers, and semi-trailers; Tr: Other transport equipment; Fu: Furniture.

Figure 10: Comparison across TFP densities estimated with different methods

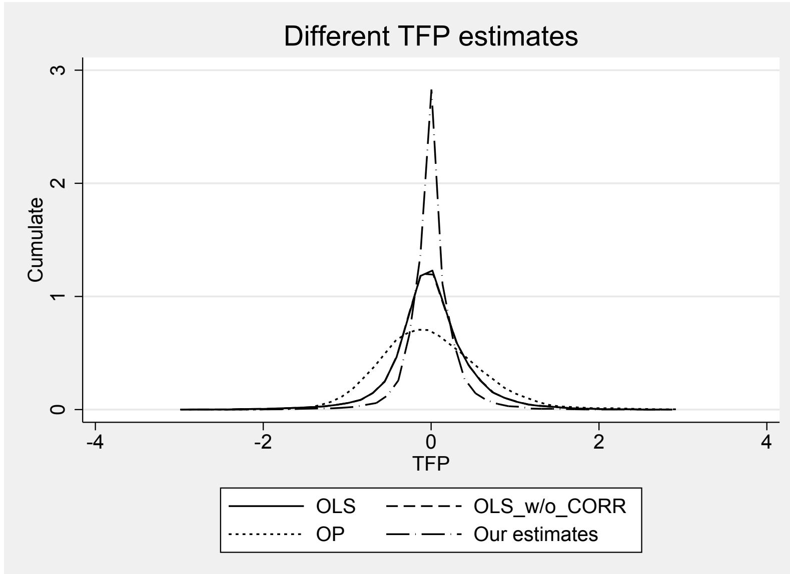


Figure 11: Difference between corrected and non-corrected OLS-estimated TFP