THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

WP 71_12

Michele Battisti
University of Palermo, Italy
LUISS Guido Carli, Italy
The Rimini Centre for Economic Analysis (RCEA), Italy

Filippo Belloc
"G. d'Annunzio" University, Italy

Massimo Del Gatto
"G. d'Annunzio" University, Italy
University of Cagliari, Italy

UNBUNDLING TECHNOLOGY ADOPTION AND TFP AT THE FIRM LEVEL. DO INTANGIBLES MATTER?

Unbundling technology adoption and *tfp* at the firm level. Do intangibles matter?

August 5, 2012

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

Filippo Belloc
"G.d'Annunzio" University (f.belloc@unich.it)

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

Abstract

We use a panel of European firms to investigate the relationship between intangible assets and productivity. We disentangle between *tfp* and technology adoption, while available studies so far have considered only a notion of productivity conflating the two effects. To this aim, we estimate production function parameters allowing, within each sector, for the existence of multiple technologies. We find that intangible assets both push the firm towards better technologies (*technology adoption* effects) and allow for a more efficient exploitation of a given technology (*tfp* effects).

Keywords: TFP, intangible assets, firm heterogeneity, firm selection, technology adoption, mixture models.
J.E.L. Classification: C29 D24 F12 032

1 Introduction

By encompassing all those firms' assets lacking physical existence (including the quality of management, customer loyalty, information infrastructure, trade secrets, research and development (hereinafter R&D), and, more in general, the company's intellectual capital), intangible assets can be considered as the knowledge base of a firm. As such, intangibles are often found to be a relevant dimension in modern knowledge intensive production (see, among others, Bontempi and Mairesse (2008), O'Mahony and Vecchi (2009) and Marrocu *et al.* (2011)). However, once one moves beyond the general statement that intangibles are important to a firm's performance, it remains unclear whether they affect total factor productivity (hereinafter *tfp*), through improving the ability to exploit "traditional" inputs (i.e. physical capital and labour), or if they also influence firm's technology adoption, that is they help the firm at identifying and adopting more productive technologies. While it seems reasonable to expect that both channels matter, the relationship between intangibles, *tfp* and technological choices of firms has not received so far a systematic investigation.

To shed light on this issue, we use mixture models to disentangle between *technology* and *tfp* at the firm level, and then study to what extent they can be explained by the level of intangible assets, as well as by a number of other firm-, sector-, region- and country-specific variables.

In our terminology, technology and *tfp* are distinct but closely related issues. Both are measures of the ability to produce as much output as possible, given the amount of inputs, but, while the former has a firm-level nature, the latter can be either common to a number of firms or specific to a given firm, in case she owns a particular production technology.

To help the definition of terms, let us anticipate the formal description and consider the following production function:

$$Y_{i,t} = A_{i,t} \prod_{n=1}^{N} (X_{n,i,t})^{\beta_{n,s}} \tag{1}$$

where $A_{i,t}$ is firm i's *tfp* (i.e. "Solow residual") at time t, $X_{n,i,t}$ denotes the amount of input n used by firm i at time t, $\beta_{n,s}$ is the associated production coefficient, and $Y_{i,t}$ denotes produced output. Index s is introduced to refer to a specific "technology", with $s = 1, \ldots, S$ and S denoting the number of available technologies. According to (1), the world we have in mind is one in which different production functions identify different technologies by differing in their production coefficients. Several technologies are available in each sector-industry, with a number of firms using each technology.

According to (1), given the amount of inputs, how much output firm i is able to produce depends on two factors: the first one ($A_{i,t}$) is firm-specific, the other ($\beta_{n,s}$) is intrinsic to the adopted technology and is common to all the firms that use the same technology. The group of firms sharing a given technology may or may not coincide with the industry. While firms' choices concerning these two factors are sometimes referred to as technological choices and sometimes referred to as productivity choices, we will reserve the term "technology" to refer to the term $\beta_{n,s}$ and will talk about *tfp* with regard to $A_{i,t}$.

In this paper, we use data on a large sample of European manufacturing firms, observed over the 2003-2009 period and nine sectors, to estimate the production function parameters in (1) allowing, within each sector, for the existence of a higher and a lower technology. Specifically, we first employ mixture models to estimate the production function parameters in (1), then use a first order stochastic dominance criterion to identify the "higher" technology, then, for each firm, we compute the *tfp* component in (1) as the difference between actual and predicted output, given technology adoption. This allows us to identify, for each firm, her technological choice (i.e. higher versus lower technology), as well as an individual *tfp* term which measures her relative position on the *tfp* distribution, within the group of firms adopting the same technology. We then investigate, separately, the effect of intangibles on both counts.

The choice to focus on intangibles has a technical motivation. An appropriate productivity analysis at the firm level would require detailed information on firms' produced quantities or, with only firms' sales available, detailed information on firms' prices, so to be able to deflate properly. With this data, the obvious candidate for measuring the firm's effort to increase output, given input levels, would be a firm's spending in R&D. However, it is well known that, with only a few notable exceptions (e.g. Foster *et al.*, 2008), information on produced quantities is hardly available, and the data we use here, in this respect, are not exception. As in the vast majority of cases, we thus have a "revenue productivity" measure as dependent variable,[1] with the consequence that the empirical specification of the relationship of interest *de facto* mixes quantity and price effects. In this case, firm's intangibles are more consistent with the dependent variable we use and should yield more consistent results. R&D is in fact only one of the many activities available to firms to get an effect on a revenue-based dependent variable. Expenditures aimed at improving the quality of management, customer loyalty, and information infrastructure - i.e. ultimately the firm's endowment of intangible assets - are all (to a greater or lesser extent "sunk") costs that a firm can pay to try to enhance her relative market position. It is worth noting that the whole literature focusing on innovation (measured through R&D) and productivity, recently surveyed by Hall (2011), is subject to this critique.[2]

The contribution of our study is twofold.

First, we acknowledge that neglecting the presence of different (within-sector) technologies results into biased

[1] In the attempt to reduce this problem, we use country-sector specific deflators for Y and K. However, this does not fully remove the bias related to cross-firms price heterogeneity.

[2] This strand of literature mostly relies on various specification of Crepon *et al.* (1998), also known as CDM model, and usually focuses on labour productivity effects.

firm-level *tfp* estimates. To tackle this problem we suggest a novel approach to disentangle empirically between technology and *tfp* using econometric techniques, while available studies so far have considered only a notion of productivity conflating the two effects.

Second, we show that intangibles have a positive and statistically significant effect on both the probability of choosing a high technology and the *tfp* level. Moreover, we find that also other firm level variables, such as firm size and being listed, exert a significant influence, along with region and country level factors. In particular, regional levels of R&D and regional accessibility have a positive - although small - effect on high technology adoption and *tfp*. Average labour costs and corporate governance institutions further influence technology and *tfp*, while employees protection laws only affect *tfp*, given technology adoption.

The paper proceeds as follows. In Section 2, we briefly describe how our analysis contributes to economic literature. Although our theoretical motivation stems from the theoretical literature on trade with heterogenous firms, initiated by Bernard *et al.* (2003) and Melitz (2003), both our methodological and empirical contributions can easily fit into the economic literature in other fields. We thus avoid refer to the related literature in the main text, and prefer to concentrate field-specific theoretical motivations and contributions in a dedicated section. In Section 3, we apply mixture models to production function estimation and obtain firms' technology choices and *tfp* values. In Section 4 we measure the impact of intangibles on, separately, technology adoption and *tfp*. Section 5 reports several robustness checks. Section 6 concludes.

2 Theoretical motivation and related literature

Our theoretical motivation stems from the so called "New New Trade Theory" (henceforth NNTT), pioneered by Melitz (2003) and Bernard *et al.* (2003), and further developed by Bernard *et al.* (2007b), Melitz and Ottaviano (2008), Chaney (2008).[3]

This literature revitalized trade economists' interest in the Shumpeterian mechanism of firm selection, with the latter taking place on the basis of firms' *tfp*. In a typical NNTT model, firms are assumed to pay a sunk cost, usually thought of as an irreversible investment in R&D, in order to start production and obtain an initial level of *tfp* drawing it from a common *tfp* distribution. In the original NNTT formulation, firms are not allowed

[3]Several surveys of this literature have been published. See, in particular: Bernard *et al.* (2007a); Greenaway and Kneller (2007).

to re-draw their *tfp* level, thus they can afford to keep serving the market only provided that they are productive enough to price below a certain level, which is endogenously determined by the process of firm selection. In such process, which is driven by a combination of import and export market competition, winners and losers emerge, with more productive firms earning handsome profits, mediocre ones lower profits, and the worst soon vanishing, being unable to cover their production costs with revenues, due to a too low *tfp*.[4]

Thanks to the empirical evidence produced to date (see Wagner (2012) for a survey), this firm selection process is now a stylized fact. However, an issue which still requires some empirical effort is the logic behind the productivity assignment. In this respect, the original NNTT formulation and the bulk of more recent NNTT advancements provides different interpretations. As said, in the original NNTT formulation, firms are not allowed to re-draw their productivity level. A clear-cut consequence of this assumption is that firms self-select into any *status* associated with costly operational activities - e.g. exporting, doing foreign direct investments or R&D, etc. - on the basis of their productivity levels, with the most productive firms potentially involved in all the activities. Thus there is causality nexus running from productivity to e.g. export status or R&D levels. This circumstance is well documented by Bustos (2011), who relates Argentinean firms' technology upgrading to the reduction in Brazilian tariffs on entry in the export market. Differently, more recent contributions to NNTT theory have been relaxing such assumption by allowing firms to pay additional (sunk) costs over time, in order to "re-draw" their *tfp*. Such costs are modeled as irreversible investment in innovation, typically R&D. Atkeson and Burstein (2010), for instance, allow for process and/or product innovation, with firms facing a given probability of improving or worsening their productivity. Navas-Ruiz and Sala (2007) have a modified Melitz model with innovation activity resulting in lower variable costs and higher implementation costs. Costantini and Melitz (2008) model technology adoption in a dynamic framework with idiosyncratic uncertainty and sunk costs. Aw *et al.* (2008) bring to the data a model in which R&D investment, through its effect on future productivity, increases the profits from exporting, and participation in the export market raises in turn the return to R&D investments. They find that export participation can be seen as one component of a broader investment strategy by the firm. All in all, the novelty in these contributions consists of suggesting the presence of endogenous "innovation dynamics" at the firm level.

[4]In this context, the productivity of the marginal firms (i.e. those firms that generate revenues barely sufficient to cover their costs) defines the threshold below which it is impossible for a firm to survive in the market. This survival threshold is a crucial variable, since every parameter that influences the average productivity of active firms does so by shifting the threshold.

A key difference between the original NNTT formulation and the more recent specifications with innovation dynamics is that, with the latter, causality between innovation and productivity should in principle run both ways. On the contrary, under the original "once and forever draw" assumption, only the more productive firms can find it profitable to invest in innovation. As a consequence, understanding to what extent firms' innovation choices do have an impact on their productivity is key for opting in favor or against the "once and forever" assumption. The research question is thus whether firms' innovation policy do have an impact on their *tfp* or not.

We next explain how we contribute to this literature.

We start by noting that, while an economically significative impact of R&D spending on productivity is already found in the literature (see the recent survey by Hall, 2011), there are several reasons to say that further evidence is welcome in this field. First of all, most works rely on data from the Community Innovation Survey (CIS), or its imitators in other countries, which, however, provide only cross-sectional information. Second, although in principle firms always bear a some form of cost when they make choices related to technology and productivity, it is not always the case that this cost finds place in their balance sheets, as it can happen, for instance, when innovation takes the form of "in-house R&D" or when a firm simply "imitates" other firms' techniques. Thus, the explanatory variable can account only for a part of the variation in the dependent variable. Third, in NNTT models firms are assumed to differ in terms of what we named "*tfp*" (see eq. (1)), thus predictions are in "physical" terms (i.e. quantity). Conversely, as explained in the introduction, most of the literature on innovation and productivity derives results in terms of revenue productivity. We tackle this issue by focusing on "intangible fixed assets", rather than on R&D. By including R&D and other costs explicitly oriented to technology and productivity, intangible assets provide us with a wider description of the effort that a firm devotes to the final goal of improving her market position (i.e. revenues), in this being more consistent with a revenue-based dependent variable. Phrased differently, our analysis provides a better understanding of whether firms systematically use intangible assets to improve their ability to transform inputs into revenues, being this ability a composition of several ingredients: i) R&D costs borne with the objective to improve productivity, ii) "absorptive capacity", that is the Cohen and Levinthal's (1989) ability to exploit existing technology and to create new technologies,[5] iii) marketing, design and technical specification expenditures borne with the objective

[5]The literature on "absorptive capacity" usually stresses the learning channel. However, World Bank (2008) highlights that absorptive capacity also depends on country-level characteristics, such as governance, quality of regulation, legal environment,

to strengthen the market position.

Having clarified why we focus on intangibles, our main contribution to NNTT literature can be presented as the following refinement. Provided that intangible assets do have a positive impact on firms' performance, does the effect occur through pushing towards using a better (either new or already existing) technology or, instead, through the *tfp* component? We show that both channels are important. Consequently, by modeling firm heterogeneity only in terms of *tfp*, NNTT models capture only part of the story. A few papers are related to this part of our analysis. Baily *et al.* (1992) find, for example, persistence effects on the top of the productivity distribution, meaning that the more productive firms tend to remain the more productive, also in the presence of technological progress. Bartelsman and Dhrymes (1998) investigate productivity dynamics in depth, showing how the transition probabilities among different classes of productivity vary substantially by industry, plant age, and other characteristics. Differently from the latter, however, our data do not allow us to go into the details of both the transition dynamics and the aggregate technological progress. Moreover, our results are consistent with the innovation dynamics described by Klette and Kortum (2004), Luttmer (2007) and, more recently, by König *et al.* (2012). König *et al.*, in particular, reproduce real world empirical productivity distributions by allowing firms to innovate by investing in R&D and/or imitating other firms' technology. An interesting feature of their framework is that it takes into account the fact that the return of the innovation activity of the firms can differ on the basis of their actual technology, depending on its "distance" from the advanced technology, as recognized by Powell and Grodal (2006). Respect to this literature, we introduce a further dimension, that is the possibility that firms improve their *tfp*, without changing technology.

In addition, in our regressions we consider a number of country-, region- and sector-specific variables. For many of these variables NNTT literature shows theoretical predictions in terms of their final effect on aggregate productivity. For example, in NNTT models, any exogenous increase in active firms' average profits (e.g. lower labour, capital or service costs and better tangible and intangible infrastructure) has the effect to raise the productivity threshold of survival by stimulating R&D and encouraging more productive firms to enter the market.

Besides NNTT studies, our analysis closely relates to other veins of literature. First, our work fits with the body of literature on "technology adoption" (Acemoglu and Zilibotti, 2001; Aghion *et al.*, 2005; Barro and

political and macroeconomic stability, government actions that help overcome market failures, including supporting R&D and building infrastructure. In our analysis we introduce controls for many of these factors.

Sala-i Martin, 1997; Bas, 2012; Desmet and Parente, 2010; Eaton and Kortum, 2001; Hall and Jones, 1999; Howitt, 2000). Second, our empirical findings corroborate the results of several other papers that have been analyzing the relationship between intangible assets and productivity. For instance, Marrocu *et al.* (2011), in a paper in which data from the same main source as ours (i.e. Amadeus database) are employed, find a positive influence on firms' productivity exerted by knowledge capital internal to the firm, as well as by intangible assets available in the regional economy. A positive effect of intangibles on productivity is detected also by Bontempi and Mairesse (2008), Corrado *et al.* (2008), Oliner *et al.* (2008), and O'Mahony and Vecchi (2009). However, none of these articles disentangle between technology and *tfp* and recognizes the importance of allowing for multiple technologies. Third, we contribute to the literature on theory of the firm and technological change. At a firm level we provide the first empirical evidence in favor of a negative relationship between being listed and a firm's performance, as recently proposed by the theoretical model of Ferreira *et al.* (2012). Furthermore, our result of a positive link between firm's dimension and productivity relates to the well known studies by Pavitt *et al.* (1987) and Cohen and Klepper (1996a, 1996b) on the Schumpeterian hypothesis of larger firms being relatively more innovative. Finally, at a country level, we show a positive effect of stronger worker protection on *tfp*, supporting the argument that stringent labour laws protect employees from expropriation by employers and so create incentives for workers to apply their efforts to learning processes within the firm (Michie and Sheehan, 2003; Acharya *et al.*, 2010).

3 Identifying firms' technology and *tfp*

In this section we describe in details how we obtain our dependent variables for Section 4, namely firm's belonging to the high technology group, used in the technology adoption regression, and firm-level *tfp*, used in the *tfp* analysis. Our strategy consists of three steps:

- **Production function(s) estimation.** We estimate the (sector- and technology-specific) coefficients $\beta_{n,s}$ in (1) using mixture models. We do so by allowing for the existence of two technology groups (i.e. $S = 2$). For each firm, the procedure provides us with the probability to belong to one or the other group. We are thus able to assign our sample firms to the group they belong with higher probability.

- **High/low technologies.** We first compute the predicted output of each firm and then use a first order stochastic dominance criterion to identify the more productive technology.

– **Firm-level *tfp*.** We compute firm-level *tfp* as the difference between actual and predicted output (i.e. Solow residual), given firms' belonging to a given technology group.

Before to go into the details of each step, it is worth showing in what our approach allows us to depart from conventional *tfp* estimates. To this end, imagine for simplicity a one-input production function. It is common practice in productivity analysis (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; and Ackerberg *et al.*, 2006) estimating the production function in logarithms by sector and retrieve the term $A_{i,t}$ as

$$lnY_{i,t} - \hat{\beta} lnX_{i,t} = ln\hat{A}_{i,t} \tag{2}$$

where the input coefficient (β) is estimated without allowing for the presence of different (within-sector) technologies. With equation (1) in mind, this input coefficient can be seen as the average of the S technology-specific betas (β_s), each weighted by the ratio of the number of firms in the s-technology group to the total number of firms in the sector. As a consequence, common *tfp* estimates include a bias which can be easily identified by using equation (1) to substitute for output in (2) and denoting the "true" *tfp* and input coefficient as $a^*_{i,t}$ and β^*_s respectively:

$$lnA^*_{i,t} + (\beta^*_s - \hat{\beta}) lnX_{i,t} = ln\hat{A}_{i,t} \tag{3}$$

Thus, *tfp* estimates obtained without allowing for the presence of different (within-sector) production technologies conflate technology effects and "proper" *tfp*. In particular, neglecting such issue results in overstating the *tfp* of the firms that adopt relatively more productive technologies (due to underestimation of their input coefficient - i.e. $\beta^*_s > \hat{\beta}$) and understating the *tfp* of the firms that adopt relatively less productive technologies (due to overestimation of their input coefficient - i.e. $\hat{\beta} > \beta^*_s$).

Assuming that the number of technologies is correctly identified, our approach allows us to obtain *tfp* estimates which are not affected by this bias.[6]

[6]In order to keep the technology adoption analysis manageable we restrict our benchmark estimation to the presence of two technologies. However, in Section 5.2 we leave the mixture analysis free to endogenously determine the number of technologies. In that exercise we can reasonably assume that the bias vanishes (that is $\beta^*_s = \hat{\beta}^*_s$).

Production function(s) estimation. Production function estimation at the firm level raises a number of issues. There is a wide literature in this field and several stratagems have been proposed, with the choice depending on the economic focus (see Del Gatto *et al.* (2010) and Van Beveren (2012) for overviews on available methodologies).

Our focus in this paper is on the fact that firms should be allowed to use different technologies. In principle, as many production functions as the number of available technologies should be estimated. This objective suggests using *mixture models* (Mc Lachlan and Peel, 2000). Such approach enables us to cluster our sample firms across different technologies and, consistently with our framework (cfr. equation (1)), estimate the relevant coefficients allowing for technology- (or group-)specific input coefficients.

The advantage of mixture models is that they allow us to cluster firms over technologies without any *ex-ante* assumption, since the probability to belong to a given technology cluster is produced by the estimation. it is however with noting that this choice prevents us from taking into account other important issues highlighted by the literature. In particular, we are not able to deal with the simultaneity bias. In Appendix B we speculate on the effect of this omission on our estimated coefficients and *tfp*.

To describe our mixture analysis, let us start by writing the (implicit) probability distribution function of (1) as a weighted average of the S specific segment (i.e. technology) densities $f_s(Y_{i,t}|\mu_s, \sigma_s)$, each with proper mean (μ_s) and variance (σ_s^2):

$$f(Y_{i,t}|\mu, \sigma) = \sum_{s=1}^{S} \varphi_s f_s(Y_{i,t}|\mu_s, \sigma_s^2). \tag{4}$$

Weights φ_s measure the *ex-ante* probability to belong to group s.

Assuming a normal density for $f_s(Y_{i,t}|\mu_s, \sigma_s^2)$,[7] the production function coefficients can be obtained by maximizing the following loglikelihood function:

$$\ln \mathcal{L} = \sum_i \sum_{t=1,\dots T|\tau\in[0;1]} \ln \sum_{s=1,\dots,S} \varphi_s (2\pi\sigma_s^2)^{-\frac{1}{2}} \exp\left\{ \frac{-\left(Y_{i,t} - (\alpha_s + \beta_{K,s}K_{i,t} + \beta_{L,s}L_{i,t})\right)^2}{2\sigma_s^2} \right\} \tag{5}$$

where the mean of f_s has been replaced by a linear predictor - i.e. $\mu_s = \alpha_s + \beta_{K,s}K_{i,t} + \beta_{L,s}L_{i,t}$ - in which the coefficients are found by the maximization process itself and τ is an indicator variable implying that firm i can

[7] Or another density belonging to the exponential family (see Wedel and De Sarbo, 1995).

exists or not at year t.

As the individual probabilities that compose the weights φ_s are unknown, there is basically a problem of missing data in the maximization of (4). Operatively, the problem is solved through the EM (expectation-maximization) algorithm of Dempster *et al.* (1977): the process is run choosing random starting points for φ_s, then posterior probabilities are computed at each step, which in turn update the regression coefficients (since weights change).

Posterior probabilities of firm i belonging to group s at time t are thus computed as:

$$p_{i,s,t} \equiv pr(i_t \in s) = \frac{\varphi_s f_s(Y_{i,t}|K_{i,t}, L_{i,t}; \hat{\sigma}_s^2, \hat{\beta}_{K,s}, \hat{\beta}_{L,s}, \hat{\alpha}_s)}{\sum_{s=1}^{S} \varphi_s f_s(Y_{i,t}|K_{i,t}, L_{i,t}; \hat{\sigma}_s^2, \hat{\beta}_{K,s}, \hat{\beta}_{L,s}, \hat{\alpha}_s)} \tag{6}$$

De Sarbo and Cron (1988) show that maximizing (4) is equivalent to performing a weighted least squares (WLS) regression with weights given by the segments' probabilities. We thus solve the maximization problem by iteratively alternating computation of WLS and probabilities until a loglikelihood convergence criterion is reached. We repeat the process many times to avoid local optima.

Concerning the number of segments S, which the algorithm is not able to identify, we have to set it *ex-ante* and then, after the maximization procedure, to test for the value of S which is most informative. The trade-off is between a parsimonious solution, without unuseful components, and a good fit of the data, so that there is a penalization for the additional components that imply more estimated parameters. Here we adopt a two segments solution (i.e. $S = 2$). This simplifies the analysis without loss of generality in terms of results. In Section 5.2 we relax this assumption following Mc Lachlan (1987) and Fraley and Raftery (2002).

The estimation takes advantage of detailed information on value added, tangible fixed assets and number of employees in the Amadeus database, provided by "Bureau van Dijk". Descriptive statistics are reported in Table 1.[8]

The values of $\beta_{K,s}$ and $\beta_{L,s}$ obtained for two groups are reported in Table 2. There is a large difference between the two clusters, with the first one having higher physical capital coefficient and smaller labour and intercept coefficients.[9] The coefficients are highly significant.

[8]The data cover the following countries: Austria, Belgium, Czech Republic, Germany, Spain, Finland, France, UK, Hungary, Italy, the Netherlands, Norway, Portugal, Sweden, Slovenia, Slovak Republic.

[9]At a country level, Battisti and Parmeter (2011) found evidence of clusters with high and low physical and human capital returns.

As well as with the production function coefficients for each technology s, the procedure provides us, for each observation (i.e. for each firm in each year), with an attached probability to stay in cluster 1 or 2. Since these probabilities sum to one, we are able to assign a firm to cluster 1 at time t if $p_{i,1,t} > p_{i,2,t}$ and viceversa for cluster 2 (hard assignment). It worth noting that the percentage of firms that change group in the period under consideration is around 10%. None of these firms changes group more than once. We identify which group of firms uses the best technology in the next step.

High/low technologies. With the estimated coefficients of the two groups in our hands, we can compute each firm's predicted output by applying the latter to her actual amount of capital and labour (i.e. $ln\hat{Y}_{i,s,t} = \alpha_s + \hat{\beta}_{K,s} lnK_{i,t} + \hat{\beta}_{L,s} lnL_{i,t}$). This allows us to derive the cumulative distribution function of the two groups and use a first order stochastic dominance criterion to identify the more productive technology as the technology with higher mean and variance. For each sector, the cumulative distribution function of the two identified groups is reported in Figure 1.

It is worth noting how the choice to allow for only two technologies makes the identification of the dominant technology straightforward, as group 2 first order stochastically dominates group 1 in all sectors. We thus do not need more sophisticated stochastic dominance tests to start referring to the technology used by cluster 2 as the higher technology (hereinafter technology H), in contrast to the lower technology (hereinafter technology L), used by firms in group 1.[10]

This given, we define the two sets of firms adopting technology H and L at time t respectively as:

$$\Theta_H \equiv \{i_t : p_{i,2,t} > p_{i,1,t}\} \qquad \text{and} \qquad \Theta_L \equiv \{i_t : p_{i,1,t} > p_{i,2,t}\} \tag{7}$$

The regional distribution of the ratio Θ_H/Θ_L is reported in Figure 2.

Firm-level *tfp*. Having computed the predicted output for all the firms on the basis of their group-specific production coefficients, we are able to calculate each firm's *tfp* at time t (i.e. the Solow residual) as $A_{i,t} = exp(lnY_{i,t} - ln\hat{Y}_{i,s,t}) = exp(lnY_{i,t} - \alpha_s - \hat{\beta}_{K,s} lnK_{i,t} - \hat{\beta}_{L,s} lnL_{i,t})$.

The regional *tfp* distribution is reported in Figure 3.

[10] It is worth noting how, in two out of our nine sectors, the technology which our criterium identifies as L presents higher estimated returns to scale (RSC), compared to technology H (see Table 11), meaning that comparing the RSC of the two clusters could not be an alternative criterium for the identification of the higher technology.

Note that, although we assumed the *tfp* term $A_{i,t}$ to be only firm-specific in (1), its empirical equivalent is *de facto* influenced by the type of technology adopted, as the *tfp* of the average firm (i.e. the firm whose observed output exactly matches the output predicted on the basis of her group-specific coefficients $\beta_{n,s}$) amounts to one (i.e. the exponential of zero), with the *tfp* of all the other firms in the same technology-group expressed with respect to that benchmark. This is a key point for our contribution. As equation (2) points out, controlling for the fact that sample firms might use different technologies is in fact a central issue in productivity estimation. By neglecting this, conventional methods of *tfp* estimation at the firm-level (e.g. the semi-parametric methods suggested by Olley and Pakes, 1996; Levinsohn and Petrin, 2003; and Ackerberg *et al.*, 2006) implicitly proceed by assuming that all the firms in the sample use the same technology and express the firms' *tfp* taking the average firm in the sector, rather than in the technology group, as benchmark. This is at the basis in (3).[11]

4 Technology adoption and *tfp* effects

In this Section we study whether intangible fixed assets, and a number of other firm-, sector-, region- and country-specific variables, do have an impact on i) firms' technology adoption, by pushing firms towards better (either new or already existing) technologies, and ii) *tfp*, by allowing for a more efficient exploitation of a given technology.

Information on intangible fixed assets is drawn from Amadeus (see Appendix A). From an accounting point of view, intangible fixed assets cover patents, copyrights, franchises or licenses, trademarks or trade names, and goodwill.

We first study firms' technological choices. With our sample firms allocated to either the high or the low technology group, firm's belonging to the high technology group can be used as the dependent variable in a technology adoption regression. Intangible assets should positively affect a firm's choice for a high technology, because they enhance the firm's knowledge base and consequently improve the firm's ability to identify and adopt technology H, i.e. - with reference to equation (1) - the firm's ability to produce a higher $Y_{i,t}$ keeping

[11] As known, *tfp* is a relative notion, which, as such, only makes sense if expressed with respect to a benchmark. Which benchmark to refer to depends, however, on the reference approach. In this respect, it is key the difference between "frontier" and "non-frontier" approach. While the former relies on the identification of "best-practice" firms, non-frontier methods provide us with productivity measures expressed in relative terms with respect to the "average" firm. Although recognizing the importance of the best-practice analysis, here we mainly focus on the non-frontier approach, as it is the basis of the "semi-parametric" methods suggested by Olley-Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg *et al.* (2006), which are the most widely used measures in our reference literature.

$A_{i,t}$ equal.

We then study the effect of intangibles on firms' *tfp*. As the *tfp* measure obtained in Section 3 is expressed in relative terms with respect to the average firm in the technology group, it can be used as such to study the impact of intangible assets. Also in this case, we expect to estimate a positive link. With reference to equation (1), the idea behind this relationship is that, all else being equal, intangible assets should raise the inputs' marginal contribution to revenues through the *tfp* term $A_{i,t}$. In other words, intangibles are expected to increase the firm's ability to produce a higher $Y_{i,t}$ keeping $\beta_{n,s}$ equal.

Needless to say, other factors, besides intangibles, can have an effect on the two dependent variables we use in this analysis. We thus consider a set of controls at a firm, region and country level.

At a firm level, as well as the intangible to tangible assets ratio ($Firm\ Intangibles$), which is our main variable of interest, we consider firms' age ($Firm\ Age$), firms' dimension, proxied by the level of sales ($Firm\ Size$), and a dummy variable indicating whether the firm is listed in a stock market ($Listed\ Firm$). Age and size should both play a positive influence on high technology adoption and *tfp*. On the one hand, a higher firm's age should imply a greater cumulate knowledge of the technological alternatives available in the industry, on the other, a larger size should increase the company's capability to exploit sizeable development laboratories and equipments (Pavitt *et al.*, 1987; Cohen and Klepper, 1996a, 1996b). Moreover, if firm size is positively related to market power, it may also increase the firm's incentives to employ more productive technologies because of pre-emption motives (Gilbert and Newbery, 1982). Differently, the effect of being listed is expected to be negative. Under private ownership (i.e. the firm is not listed), insider shareholders, such as a manager-entrepreuner, can time the market by choosing an early exit after receiving bad signals about production; therefore, managers are more tolerant of early failures and more inclined to invest in new and more profitable - even if riskier - projects (Ferreira *et al.*, 2012). In addition, listed companies are relatively more vulnerable to the adoption of sub-optimal business strategies by activist short-termist shareholders (Kochnar and David, 1996; Sherman *et al.*, 1998; Hoskisson *et al.*, 2002).

At a regional level, we first of all include the R&D levels of the given region (*Regional R&D*) and the neighbouring regions (*Neighbouring Regions R&D*). With these variables, we aim at measuring R&D spillovers within and between regions. Previous studies on this issue are numerous and suggest a positive relationship between R&D spillovers and both technology dynamics and firms' *tfp* (e.g., Cohen and Levinthal, 1989; Griliches, 1992; Ciccone, 2002; Autant-Bernard and Mairesse, 2007). We also include a control for the region's accessibility (*Region Accessibility*). This variable is a measure of the total population reachable from the region, weighted by the ease of getting the other regions. The rationale for its inclusion is that accessibility might be seen as a measure of market potential and degree of competition at the same time, with the expected sign depending on which dimension takes over. On the one hand, in fact, higher accessibility means higher exposure to both export and import market competition. As NNTT points out (see Section 2), when competition gets fiercer, the less productive firms are forced to leave the market. This suggests an expected positive sign in the regression. On the other hand, a higher accessibility entails the chance to target a larger market for the firms located in the region, with more space for the less productive firms. This induces an expected negative relationship in the regression.

At a country level, we include a measure of labour cost (*Labour Cost*) and a set of three institutional variables. Labour cost can affect the observed technology adoption and *tfp* through inducing firms selection. The rationale for including labour costs lies again in the expected more intense selection effects associated with higher production costs. By cutting firms' profits, higher labour costs should be associated with a stronger firm selection and a higher probability to observe a larger share of firms in the H-technology group and/or in the right tail of the *tfp* distribution. This would suggest a positive sign in the regressions. As institutional variables, we consider a measure of employment protection (*Country EPL*), an index of minority shareholder rights (*Shareholder Rights*) and the percentage of independent directors that must be in the board by law (*Independent Directors*). The effect of employment protection legislation is difficult to predict *ex-ante*. Stringent labour laws may stimulate the adoption of more efficient knowledge intensive production methods, because such laws protect employees from expropriation by employers and so create incentives for workers to apply their efforts to learning processes (Michie and Sheehan, 2003; Acharya *et al.*, 2010). However, strong employment protection may favor employees' resistance to the use of innovative technologies, if such new production technologies imply job losses or an increase in the labour burden (Zwick, 2002). Stronger minority shareholder rights should have a negative effect on our dependent variables, to the extent that minority shareholders use increased voice opportunity for private gain-seeking causing sub-optimal business strategies (Belloc, 2012). Finally, the presence of independent directors is expected to exert a positive influence on both technology H adoption and *tfp*. The involvement of outside directors in the board should reduce agency costs and consequently improve strategic business decisions

(Kaplan and Minton, 1994).

While a detailed description of the variables used is reported in Appendix A, we now illustrate our results.

Technology adoption effects. To study whether intangible assets affect firms' technology adoption, by pushing firms towards better technologies, we define our dependent variable *Group H* as follows

$$Group\ H_{i,t} = \begin{cases} 1 & \text{if} \quad i_t \in \Theta_H \\ 0 & \text{if} \quad i_t \in \Theta_L \end{cases} \tag{8}$$

and, making covariates explicit, we estimate the following logit model:

$$\begin{aligned} Group\ H_{i,t} = {} & \delta_0 + \delta_1 Firm\ Intangibles_{i,t} + \delta_2 Firm\ Age_{i,t} + \delta_3 Listed\ Firm_{i,t} + \\ & + \delta_4 Firm\ Size_{i,t} + \delta_5 Regional\ R\&D_{r,t} + \delta_6 Region\ Accessibility_{r,t} + \\ & + \delta_7 Neighbouring\ Regions\ R\&D_{r,t} + \delta_8 Labour\ Cost_{c,t} + \\ & + \delta_9 Country\ EPL_{c,t} + \delta_{10} Shareholder\ Rights_{c,t} + \\ & + \delta_{11} Independent\ Directors_{c,t} + \delta_{12,\ldots,k} Sector_s + \\ & + \delta_{k+1,\ldots,q} Country_c + \delta_{q+1,\ldots,w} Year_t + u_i + \varepsilon_{i,t} \end{aligned} \tag{9}$$

where u_i are unobservable firm specific effects and $\varepsilon_{i,t}$ indicates the residuals. Subscript i refers to firms, r to regions, c to countries, and t to years.

Results of the logit regression analysis are collected in Table 3. Control variables are added progressively. Only firm level variables are included in model version 1; both firm level and region level variables are included in model 2; firm, region and country level variables are included in model 3. In models 4 and 5, we tackle possible endogeneity of intangibles with respect to technology choice, i.e. reverse causality running from a high technology adoption in t to the level of intangibles in t. In model version 4, the full set of controls is included, and the intangibles variable is one-year lagged, so as to circumvent possible simultaneous determination of technology adoption and intangible assets. Similarly, in model version 5, again, the full set of controls is included, and the intangibles variable is instrumented by its one-year lagged values.

Estimates are stable across different model specifications. The level of a firm's intangibles is shown to push the firm towards the adoption of the more productive technology, regression coefficients being positive and strongly significative.

Also all other firm level variables have a statistically significant coefficient. Firm's age and size have a positive effect on the firm's probability of choosing technology H (at 1 percent level of statistical significance), while being a listed firm has a negative effect (at a level of statistical significance between 1 and 5 percent). The detected negative effect of being listed on high technology adoption is, to the best our knowledge, the first empirical evidence in favor of the theoretical predictions recently proposed by Ferreira *et al.* (2012).

At a regional level, we find that regional R&D has a positive and statistically significant effect (at 1 percent level), this confirming the relevance of within region R&D activities to firms' technology choices. At the opposite, R&D levels of neighbouring regions do not exert a statistically significant influence, while a region's accessibility shows only a negligible impact.

At a country level, the average labour cost has an impact, positively affecting the probability of the firm's adopting technology H at 1 percent level of statistical significance. This result seems to confirm the presence of a self-selection process, in which lower technology firms withdraw from the market due to insufficient revenues when - all else being equal - labour costs are higher. As for institutional country variables, finally, we find that the mandatory presence of independent directors in the board has a positive effect on the firm's choice of high production technology and that minority shareholder rights have a negative effect (again, at 1 percent level of statistical significance), while employment protection legislations result uninfluential.[12]

***Tfp* effects.** We now study the effect of intangibles using firm-level *tfp* as the dependent variable and employing the same vector of covariates and model structure used in the technology adoption analysis. A standard Generalized Least Square (GLS) regression is used. Being the dependent variable a relative measure, estimated coefficients provide information on the direction that an explanatory variable's change gives to the firm's move

[12] In unreported estimations, we have verified that our main findings remain statistically significant also in sector-specific regressions. From this, we can conclude that sector size effects, if present, do not drive our results.

in the (sector- and technology-specific) *tfp* distribution.

$$
\begin{aligned}
tfp_{i,t} = {} & \gamma_0 + \gamma_1 Firm\ Intangibles_{i,t} + \gamma_2 Firm\ Age_{i,t} + \gamma_3 Listed\ Firm_{i,t} + \\
& + \gamma_4 Firm\ Size_{i,t} + \gamma_5 Regional\ R\&D_{r,t} + \gamma_6 Region\ Accessibility_{r,t} + \\
& + \gamma_7 Neighbouring\ Regions\ R\&D_{r,t} + \gamma_8 Labour\ Cost_{c,t} + \\
& + \gamma_9 Country\ EPL_{c,t} + \gamma_{10} Shareholder\ Rights_{c,t} + \\
& + \gamma_{11} Independent\ Directors_{c,t} + \gamma_{12,\dots,k} Sector_s + \\
& + \gamma_{k+1,\dots,q} Country_c + \gamma_{q+1,\dots,w} Year_t + u_i + \eta_{i,t} \qquad (10)
\end{aligned}
$$

where u_i are unobservable firm specific effects and $\eta_{i,t}$ indicates the residuals. Subscript i refers to firms, r to regions, c to countries, and t to years.

Results are presented in Table 4. Estimates are broadly similar to those obtained in the technology adoption analysis. In particular, we find that an increase in the intangible to tangible assets ratio has a positive and statistically significant effect on *tfp*, as suggested by previous studies (Bontempi and Mairesse, 2008; Corrado *et al.*, 2008; Oliner *et al.*, 2008; O'Mahony and Vecchi, 2009). Hence, intangibles are shown to be not only relevant to firms' technological choices but also a crucial channel through which firms can increase their ability to exploit their given technology.

Firm's age exerts a statistically insignificant influence on *tfp*, while, similarly to the technology adoption regressions, being listed results associated to a negative and firm's size to a positive parameter. The two latter results are consistent, respectively, with the Ferreira *et al.*'s (2012) theory of public ownership increasing incentives to choose conventional projects and with the Schumpeterian argument of larger firms being more productive. Both a region's R&D and accessibility have a positive and statistically significant - albeit small - impact on *tfp*, i.e. within region R&D spillovers and a region's infrastructures to access other regions' markets boost firms' production performance given the technology. At a country level, we observe a positive effect of stronger employment protection laws and of the presence of mandatory independent directors in the company's board, and a negative effect of minority shareholder rights, thus confirming the similar results of, respectively, Acharya *et al.* (2010), Kaplan and Minton (1994) and Belloc (2012). Specifically, Acharya *et al.* (2010) present cross-country evidence showing that stringent dismissal laws provide firms a commitment device to not punish short-run failures and thereby spur employees' effort to pursue value-enhancing innovative practices in the production process; Kaplan and Minton (1994) show that outside directors significantly decrease agency costs in the firm, by improving monitoring and discipline of managers; Belloc (2012) finds that stronger minority shareholder intervention power in the general meetings causes coordination failures in strategic business decisions, consequently reducing firms' ability to develop innovative products. Finally, analogously to the technology effects analysis we find that higher labour costs are associated to a positive and significant parameter. Again, this suggests that firms self-select, withdrawing from the market if their *tfp* is not high enough to compensate for relatively high labour costs.

5 Robustness checks

5.1 Clustering over sub-periods

We test here an alternative specification of the mixture model we used in the production function analysis. We analyze whether the panel nature of the data affects our estimated capital and labour input coefficients, and in turn the technology adoption and *tfp* regressions.

A problem may arise in our benchmark estimation due to the fact that we pooled observations together - as they were independent - in order to allow free transitions over time. We now run a robustness check estimation by introducing firm-specific effects accounting for time dependency at the firm level.[13] At the same time, however, we again want to allow for transitions across groups. We therefore split the sample in two sub-periods (2003-2005 and 2006-2009), so that firms are allowed to change technology (once) in the whole period. This two sub-periods structure holds under the assumption that production function's changes, if any, do not occur in the short run. This strategy is consistent with the low number of transitions we found in the benchmark full i.i.d. estimation.

Results obtained from this robustness check are presented in Table 5 and Table 6. Table 5 presents the estimated inputs' coefficients by sector, while Table 6 shows the technology adoption and *tfp* regressions re-run

[13]This follows the same logic of the longitudinal clustering of McNicholas and Murphy (2010), although their approach doesn't work for multivariate density or regressions.

by using the new clustering and inputs' coefficients. First, we find that coefficients' estimates presented in Table 5 are virtually unchanged with respect to the basic results showed in Table 2. We observe a slight reduction of the capital coefficients from the first to the second sub-period. Nevertheless, this reduction is rather small, with the only exception of sector OTR, where the drop in the capital coefficient's size is due to the small number of observations (85) we have in the first sub-period. Second, we observe from Table 6 that the estimated effects of the considered explanatory variables on both technology adoption and *tfp* are substantially similar to those estimated in the basic model specifications presented in Table 3 and Table 4. Notice that, in the robustness check presented in Table 6, we have employed a model specification in which the full set of controls is included and the intangibles variable is instrumented by its one-year lagged values.[14]

5.2 Flexible number of clusters

We used two clusters in our main specification of the mixture analysis, for sake of simplicity and because they are much more interpretable in the second step (i.e. technology adoption regression). Here we show how, by following a sound statistical criterion and choosing a different number of clusters, things do not change dramatically, so that by a qualitative point of view the implications are still the same and by a quantitative point of view there are only minor changes. We follow here a log-likelihood ratio (LR) boostrapped test (Mc Lachlan, 1987)[15]. The LR test proceeds to boostrap replications of the model and then tests as a chi-square the difference about the log-likelihoods of two solutions (e.g. 2 and 3 segments).

The findings are as follows. Table 7 shows LR test results with respect to the alternative hypothesis that the number of segments is greater than a given S, with decisive highlighted. We observe that the LR test suggests to adopt two components in one out of nine sectors, three components in five sectors, and four components in three sectors. Table 8, then, shows the inputs' coefficients for each cluster and sector, which result from the estimation of as much production functions as the number of segments suggested by the LR test. Inputs' coefficients show some variability across clusters, but they generally remain on economically meaningful values.

[14] In an additional robustness check we have verified also if *ex-ante* probabilities are explained by omitted drivers, so-called concomitant variables (Dayton and Macready, 1988). If omitted variables drive the clustering process, then the groups composition may be altered. In particular, we consider a vector of concomitant variables containing intangible fixed assets and firm's age. We thus can test whether underlying sources of clustering, omitted in our basic mixture analysis, cause probability values substantially different from the random maximised ones we used before. We observe that concomitant variables act on clustering probabilities differently across sectors. Nonetheless, we also find that the relative size of the clusters is very similar to the one showed in Table 2. This suggests that initial random assignment does not affect the final clustering. Tables of results of this robustness check are available upon request.

[15] Other alternatives linked to log-likelihood penalization as the MAIC - Modified Akaike Information Criterion (Hawkins *et al.*, 2001) and the classical Bayesian Information Criterion (BIC) (Fraley and Raftery, 2002) still give equivalent results.

Finally, we re-obtained for each firm her *tfp*, by using the new inputs' coefficients, and re-run *tfp* regressions. Results are presented in Table 9. We find that the estimated coefficients of our main explanatory variables are broadly similar, in their same sign and statistical significance, to those obtained in our basic estimation (see Table 4). Hence, adding components does not seem to introduce significant gains in the analysis of the relation between intangibles and *tfp*.

Since a number of segments higher than two makes it difficult to order the technology groups (see Section 3), we can conclude from this robustness check that assuming $S = 2$ is the best option among a slightly better description of the relationships of interest - i.e. more components - and a better interpretation - i.e. less components -.

This notwithstanding, it is worth noting that, if one were only interested in the best identification of the *tfp* term, the procedure suggested in this Section provides with the best solution to the bias problem described by equation (3), as we can reasonably assume that the estimated production function coefficients are such that $\beta_s^* \simeq \hat{\beta}_s$.

6 Conclusions

Intangible assets play an increasingly important role in modern knowledge intensive production (Bontempi and Mairesse, 2008; Corrado *et al.*, 2008; Oliner *et al.*, 2008; O'Mahony and Vecchi, 2009). For example, Marrocu *et al.* (2011) have recently showed that intangibles could act as an input - like tangible fixed assets - in a production function, contributing with a positive and statistically significant coefficient. However, despite this growing interest by economists, it has been unclear so far whether intangibles can allow for a more efficient exploitation of a firm's "traditional" inputs (i.e. tangible fixed capital and labour) or they help the firm at identifying and adopting more productive technologies, or both. Such gap in previous studies stems from the fact that traditional estimates of *tfp* are obtained under the assumption of firms using a single given technology. This, besides imposing a strong limitation to the *tfp* estimation (i.e. a unique production function for all firms), also precludes the possibility to analyze firms' technological choices. In this paper, we have proposed an empirical strategy that allows for multiple technologies. We have found that intangibles have a positive and statistically significant effect on both the firm's probability of choosing relatively more productive technologies and *tfp*. This result adds at least to two veins of literature. On the one hand, we contribute to the discussion currently taking place within the NNTT literature on if and how firms can influence their *tfp* in a dynamic innovation framework by paying additional costs. We show that a firm's effort, in terms of increasing her endowment of intangible assets, is worth it. On the other hand, we refine the literature on intangibles and firms' productivity by disentangling the positive influences that intangible assets have independently on technology adoption and *tfp*.

Apart from unveiling the effects of intangibles on firm's production abilities, our empirical strategy provides a methodological contribution to *tfp* estimation. We have showed that, in our sample of European manufacturing firms, the traditional assumption of all firms using a unique production function is rejected by the mixture analysis of the data. Even within the same sector, firms choose different technologies (i.e. different production functions).

Admittedly, the advantage of allowing for multiple production functions through mixture models comes at a price. The fact that mixture analysis operates by means of WLS hampers the possibility to control for other sources of bias highlighted by the recent literature on productivity estimation. It is thus in our agenda to enrich our approach by controlling for the simultaneity bias.

References

[1] Acemoglu, D., & Zilibotti, F. (2001). Productivity Differences. *Quarterly Journal of Economics*, **116**(2), 563-606.

[2] Acharya, V.V., Baghai, R.P. & Subramanian, K.V. (2010). Labor Laws and Innovation. *NBER* Working Paper 16484, Cambridge (MA).

[3] Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and Innovation: An Inverted-U Relationship. *Quarterly Journal of Economics*, **120** (2), 701-728.

[4] Anselin, L. (1995). Local Indicators of Spatial Association LISA, *Geographical Analysis*, **27**, 93-115.

[5] Atkeson A. & Burstein A. (2010). Innovation, Firm Dynamics, and International Trade, *Journal of Political Economy*, **118**(3), 433-484.

[6] Autant-Bernard, C., Mairesse, J. & Massard, N. (2007). Spatial Knowledge Diffusion through Collaborative Networks: An Introduction. *Papers in Regional Science*, **86**(3), 341-350.

[7] Aw, B.Y., Roberts, M.J. & Yi Zu, D. (2008). R&D Investment, Exporting, and Productivity Dynamics. *American Economic Review Papers and Proceedings*, **98**(2), 451-456.

[8] Bas M. (2012). Technology Adoption, Export Status, and Skill Upgrading: Theory and Evidence. *Review of International Economics*, **20**(2), 315-331.

[9] Baily, M., Hulten, C., & Campbell, D. (1992). Productivity Dynamics in Manufacturing Plants. *Brookings Papers on Economic Activity: Microeconomics*, **4**, 187-267. Brookings Institute.

[10] Battisti, M. & Parmeter, C.F. (2011). Income Polarization, Convergence Tools and Mixture Analysis. *Working paper Miami Business School*, WP2011/17.

[11] Belloc, F. (2012). Law, Finance and Innovation: The Dark Side of Shareholder Protection. University of Chieti-Pescara, mimeo.

[12] Bernard, A., Eaton, J., Jensen, J. & Kortum, S. (2003). Plants and Productivity in International Trade. *American Economic Review*, **93**, 1268-1290.

[13] Bernard, A., Jensen J.B, Redding S.J. & Schott P.K. (2007a). Firms in International Trade. *Journal of Economic Perspectives*, **21**(3), 105130.

[14] Bernard, A., Redding S.J. & Schott P.K. (2007b). Comparative Advantage and Heterogeneous Firms. *Review of Economic Studies*, **74**, 31-66.

[15] Bontempi, M.E. & Mairesse, J. (2008). Exploring the Relationship between Intangible Capital and Productivity in Italian Manufacturing Firms. *NBER* Working Paper 14108, Cambridge (MA).

[16] Bustos P. (2011). Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms. *American Economic Review*, **101**, 304-340.

[17] Ciccone, A. (2002). Agglomeration Effects in Europe. *European Economic Review*, **46**, 213-227.

[18] Cohen, W.M. & Klepper, S. (1996a). A Reprise of Size and R&D. *Economic Journal*, **106**(437), 925-951.

[19] Cohen, W.M. & Klepper, S. (1996b). Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D. *Review of Economics and Statistics*, **78**(2), 232-243.

[20] Cohen, W.M. and Levinthal, D.A. (1989). Innovation and Learning: The Two Faces of R&D. *The Economic Journal*, **99**, 569-596.

[21] Corrado, C., Hulten, C. & Sichel, D. (2008). Intangible Capital and Economic Growth. *Review of Income and Wealth*, **55**(3), 661-685.

[22] Costantini & Melitz, M. (2008). The Dynamics of Firm-Level Adjustment to Trade Liberalization. In Helpman, E., Marin, D. and Verdier, T. (eds.), *The Organization of Firms in a Global Economy*, Harvard University Press, 2008

[23] Crepon, B., Duguet, E. & Mairesse, J. (1998). Research, Innovation, and Productivity: An Econometric Analysis at the Firm Level. *Economics of Innovation and New Technology* **7**, 115-156.

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

[25] Dayton, C. M. & Macready, G.B. (1988). Concomitant-Variable Latent-Class Models. *Journal of the American Statistical Association*, **78**, 173-178.

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

[27] Dempster, A.P, Laird, N.M. & Rubin, D.B. (1977). Maximum Likelihood for Incomplete Data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, **39**(1), 1-38.

[28] Desmet, J. & Parente, S. (2010). Bigger is Better: Market Size, Demand Elasticity and Innovation. *International Economic Review*, *51*(2), 319-333.

[29] Eaton, J. & Kortum, S. (2001). Trade in Capital Goods. *European Economic Review* **45**(7), 1195-1235.

[30] Ferreira, D., Manso, G. & Silva, A.C. (2012). Incentives to Innovate and the Decision to Go Public or Private. *Review of Financial Studies*, forthcoming.

[31] Foster, L., Haltiwanger, J. & Syverson, C. (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?. *American Economic Review*, **98**, 394-425.

[32] Fraley, C. & Raftery, A. (2002). Model-based Clustering, Discriminant Analysis, and Density Estimation. *Journal of the American Statistical Association*, **97**, 611-631.

[33] Gilbert, R.J. & Newbery, D.M.G. (1982). Preemptive Patenting and the Persistence of Monopoly, *American Economic Review*, **72**(3), 514-526.

[34] Greenaway D. & Kneller R. (2007). Firm Heterogeneity, Exporting and Foreign Direct Investment. *The Economic Journal*, **117**, F134F161.

[35] Griliches, Z. (1992). The Search for R&D Spillovers. *Scandinavian Journal of Economics*, **94**, 24-47.

[36] Hall, B.H. (2011). Innovation and Productivity. *NBER Working Paper*, No. 17178.

[37] Hall, R. & Jones, C. (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others? *Quarterly Journal of Economics* **114**(1), 83-116.

[38] Hawkins, D. S., Allen, D. M. & Stromberg, A. J. (2001). Determining the Number of Components in Mixtures of Linear Models. *Computational Statistics and Data Analysis*, **38**, 15-48.

[39] Hoskisson, R.E., Hitt, M.A., Johnson, R.A. & Grossman, W. (2002). Conflicting Voices: The Effects of Institutional Ownership Heterogeneity and Internal Governance on Corporate Innovation Strategies. *Academy of Management Journal*, **45**(4), 697-716.

[40] Howitt, P. (2000). Endogenous Growth and Cross-country Income Differences. *American Economic Review* **90**(4), 829-846.

[41] Kaplan, S. & Minton, B. (1994). Appointments of Outsiders to Japanese Boards: Determinants and Implications for Managers. *Journal of Financial Economics*, **36**, 225-257.

[42] Kochnar, R. & David, P. (1996). Institutional Investors and Firm Innovation: A Test of Competing Hypothesis. *Strategic Management Journal*, **17**(1), 73-84.

[43] Mc Lachlan, G. (1987). On Bootstrapping the Likelihood Ratio Test Statistic for the Number of Components in a Normal Mixture, *Journal of the Royal Statistical Society C*, **36**, 31824.

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

[45] McNicholas, P. D. & Murphy, T.B.(2010). Model-based Clustering of Longitudinal Data. *The Canadian Journal of Statistics*, **38**(1), 153-168.

[46] Melitz, M. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, **71**, 1695-1725.

[47] Melitz, M. & Ottaviano, G. (2008). Market Size, Trade, and Productivity. *Review of Economic Studies*, **75**, 295-316.

[48] Michie, J. & Sheehan, M. (2003). Labour Market Deregulation, 'Flexibility' and Innovation. *Cambridge Journal of Economics*, **27**(1), 123-143.

[49] Navas-Ruiz & Sala (2007). Technology Adoption and the Selection Effect of Trade. Working Paper of the European University Institute, 2007/58.

[50] O'Mahony, M. & Vecchi, M. (2009). R&D, Knowledge Spillovers and Company Productivity Performance. *Research Policy*, **38**, 25-44.

[51] Oliner, S.D., Sichel, D. & Stiroh, K.J. (2008). Explaining a Productive Decade. *Journal of Policy Modeling*, **30**(4), 633-673.

[52] Pavitt, K., Robson, M. & Townsend, J. (1987). The Size Distribution of Innovating Firms in the UK: 1945-1983. *Journal of Industrial Economics*, **35**(3), 291-316.

[53] Sherman, H., Beldona, S. & Joshi, M.P. (1998). Institutional Investor Heterogeneity: Implications for Strategic Decisions. *Corporate Governance*, **6**(3), 166-173.

[54] Siems, M., Lele, P., Iglesias-Rodriguez, P. et al. (2009). CBR Extended Shareholder Protection Index. Centre for Business Research, University of Cambridge, Cambridge (UK). Retrived from http://www.cbr.cam.ac.uk/research/programme2/project2-20output.htm.

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

[56] Wedel, M. & De Sarbo, W.S. (1995). A Mixture Likelihood Approach for Generalized Linear Models. *Journal Classification*, **12**, 21-55.

[57] World Bank (2008). Determinants of Technological Diffusion: Recent Trends and Prospects. In *Global Economic Prospects 2008: Technology Diffusion in the Developing World*, The World Bank, Washington, D.C.

[58] Zwick, T. (2002). Employee Resistance against Innovations. *International Journal of Manpower*, **23**(6), 542-552.

A Variables description

Added Value. Log of added value. Added value is defined as profit for period + depreciation + taxation + interests paid + cost of employees. This is a firm-level variable, covering years from 2003 to 2009, which we deflated using the OECD-Stan sector-country specific deflators (source: Amadeus (2012))

Labour Input. Log of total number of employees included in the company's payroll. This is a firm-level variable, covering years from 2003 to 2009, which we deflated using the OECD-Stan sector-country specific deflators (source: Amadeus (2012))

Capital Input. Log of tangible assets. Tangible assets include buildings, machinery and all other tangible assets. This is a firm-level variable, covering years from 2003 to 2009, which we deflated using the OECD-Stan sector-country specific deflators. (source: Amadeus (2012))

Firm Intangibles. Log of intangible to tangible assets ratio. Intangible assets include formation expenses, research expenses, goodwill, development expenses. Tangible assets include buildings, machinery and all tangible assets. This is a firm-level variable, covering years from 2003 to 2009. (source: Amadeus (2012))

Firm Age. Age of the firm (years). This is a firm-level variable, covering years from 2003 to 2009. (source: Amadeus (2012))

Listed Firm. Dummy variable (1 = the firm is listed in the stock market, 0 = otherwise). This is a firm-level variable, covering years from 2003 to 2009. (source: Amadeus (2012))

Sales. Log of net sales. This is a firm-level variable, covering years from 2003 to 2009, which we deflated using the OECD-Stan sector-country specific deflators. (source: Amadeus (2012))

Regional R&D. Total intramural R&D expenditure. This is a region-level variable, covering years from 2003 to 2009, expressed in Purchasing Power Standard per inhabitant at constant 2000 prices. "Total" refer to the fact that the variable cover: i) business enterprise sector, ii) government sector, iii) higher education sector, iv) private non-profit sector. (source: Eurostat (2012))

Region Accessibility. Multi-modal potential accessibility, std. This is a region-level variable, covering 2001 and 2006 (source: Espon database). The variable is defined as $Acc_j = \sum_r Pop_r exp(-\beta \bar{c}_{jr})$, where the term $\bar{c}_{jr}$

refers to the aggregation over transport modes (air, rail, road) of the cost of reaching region r from region j - i.e. $\bar{c}_{jr} = -(1/\lambda) ln \sum_m exp(-\lambda c_{jrm})$, where index m refers to the transportation mode, and λ is a parameter indicating the sensitivity to travel cost. The interpretation is that the greater the number of "attractive destinations" in areas r is, and the more accessible areas r are from area j, the greater is the accessibility of area j.

Neighbouring Regions R&D. This is a region-level variable, covering years from 2003 to 2009. With R&D defined as above, this variable is derived as a local clusters indicator, of the class known as "$Gi*$ statistics" (Anselin, 1995). These are defined as indexes of spatial clustering of a set of observations. They are the ratio of the sum of a variable over the neighbours to the total sum of the variable within the sample. The variable we use is the R&D expenses and the spatial framework are the EU27 regions plus Norway and Switzerland. The numerator is computed over a defined neighborhood that depends by the weight distance matrix. In our case, the number of allowed neighbours is set as $k = 4$ (we obtain similar results with k=3,5,...,10). (source: authors' own calculation on Eurostat (2012) data)

Labour Cost. Hourly labour costs, manufacturing. This is a country-level variable, covering years from 2003 to 2008 (source: Eurostat (2012)).

Country EPL. Unweighted average of sub-indicators for regular contracts (EPR) and temporary contracts (EPT). EPR covers notification procedures, delay involved before notice can start, length of notice period at 9 months / 4 years / 20 years of tenure, severance pay at 9 months / 4 years / 20 years of tenure, definition of justified or unfair dismissal, length of trial period, compensation following unfair dismissal, possibility of reinstatement following unfair dismissal. EPT covers valid cases for use of fixed-term contracts, maximum number of successive fixed-term contracts, maximum cumulated duration of successive fixed-term contracts, types of work for which temporary work agency employment is legal, restrictions on number of renewals of temporary work agency contracts, maximum cumulated duration of successive temporary work agency contracts. The summary indicator is on a scale from 0 (least restrictions) to 6 (most restrictions). This is a country-level variable, covering years from 2003 to 2008 (source: OECD (2012)).

Shareholder Rights. Unweighted sum of 3 sub-indicators. Sub-indicator 1 equals 1 if the sale of more than 50% of the company's assets requires approval of the general meeting, equals 0.5 if the sale of more than 80%

of the assets requires approval, equals 0 otherwise. Sub-indicator 2 equals 1 if shareholders who hold 1% or less of the capital can put an item on the agenda, equals 0.75 if there is a hurdle of more than 1% but not more than 3%; equals 0.5 if there is a hurdle of more than 3% but not more than 5%, equals 0.25 if there is a hurdle of more than 5% but not more than 10%, equals 0 otherwise. Sub-indicator 3 equals 1 if every shareholder can file a claim against a resolution by the general meeting, equals 0.5 if there is a threshold of 10% voting rights, equals 0 if this kind of shareholder action does not exist. This is a country-level variable, covering years from 2003 to 2005 (source: Siems *et al.* (2009)).

Independent Directors. Equals 1 if at least half of the board members must be independent, equals 0.5 if 25% of them must be independent, equals 0 otherwise. This is a country-level variable, covering years from 2003 to 2005 (source: Siems *et al.* (2009)).

Table 10 provides a synthetic description of the variables used in the analysis.

B Simultaneity bias

In Section 3 we stressed the fact that the presence of different technologies should be allowed for in productivity analysis. On the one hand, by estimating technology-specific coefficients, this removed, or at least softened, the associated bias in the estimates. On the other hand, however, using mixture models prevented us from controlling for other important sources of bias which the literature in this field has been highlighting. In particular, we have been forced to neglect the fact that information on actual *tfp*, although unknown to the econometrician, is to some extent included in the information set of the firm when the decision concerning the amount of inputs is made. If this is the case, our production function parameters, which are obtained through WLS, are biased due to the correlation between the regressors and the error term.

To tackle this problem, commonly referred to as “simultaneity”, in this Section we re-estimate the production function coefficients following the semi-parametric approach suggested by Olley and Pakes (1996). Moreover, as the Olley-Pakes method assumes that labour is a fully variable input, we use the more realistic correction suggested by Ackerberg *et al.* (2006) and adopt the acronym OPACF to refer to the resulting estimation procedure. For more details on the Olley-Pakes routine the reader is redirected to Del Gatto *et al.* (2010). Our estimation strictly follows the description in their Section 5.2.1. Although not being a “proper” robustness check, this exercise can give a sense of the difference in terms of estimated coefficients and *tfp*, respect to our mixture regressions.

As a first comparison, we report in Table 11 the coefficients estimated, for the whole sample, through the OPACF method, and compare them with those obtained through a simple OLS estimation. The latter estimation is equivalent to a one-group mixture regression.

A first information from this table is that, as expected, OLS estimates are always in the middle, compared to the mixture-based coefficients obtained in the two-groups case (see Table 2). This is just a consequence of the fact that $\hat{\beta}_n = \sum_{s=1}^{S} \hat{\beta}_{n,s} \frac{\Theta_s}{\sum_{s=1}^{S} \Theta_s}$

A second information is that no clear pattern emerges form the comparison of the OPACF with the OLS coefficients, neither in terms of RSC, nor in terms of capital/labour coefficients ratio. This might be interpreted as evidence of the fact that the simultaneity bias does not affect our results in a given direction.

As a further check, we re-compute firms' *tfp* using OPACF within the two groups, as identified by the mixture analysis. While at first sight this might seem a good strategy, it presents several drawbacks. First, it is not correct to reduce the mixture regressions to a first step aimed at identifying the groups, as firms' belonging to one or the other group is endogenous to the production function estimation. In other words, firms that fall in one group might fall in the other group in an ideal "simultaneity free" mixture regression. Second, the Olley-Pakes procedure requires information on firms' investments, and this dramatically reflects onto the number of observations, which shrinks by even 2/3, so making impossible to replicate the benchmark regressions. This notwithstanding, we re-compute firms' *tfp* by technology groups, using the Olley-Pakes coefficients reported in Table 11. Interestingly enough, the correlation with the mixture-based *tfp* is around 0.7 for both technologies.

Table 1: Amadeus dataset, size and sectoral composition after data cleaning.

SECTOR	**SECTOR CODE**	**N. OBSERVATIONS**	**N. FIRMS**
Chemical products (including pharmaceuticals)	CH	13578	3366
Rubber and plastic products	RP	10274	2714
Other non-metallic products	ONM	7711	2021
Basic metals	BM	2867	812
Metal products (except machinary and equipment)	ME	8371	2293
Electronic, electrical and optical products	EL	13380	3701
Machinary and equipment	MA	16248	4477
Motor vehicles and trailers	MV	3604	1025
Other transport equipment	OTR	1308	384
TOTAL	-	77341	20793

Table 2: Mixture regression, 2 technology clusters - $S = 2$ (dep. var.: *Added Value*).

		Cluster 1			Cluster 2		
SECTOR		**K**	**L**	**const**	**K**	**L**	**const**
CH	**coeffs**	0,353***	0,454***	-0,536***	0,183***	0,775***	0,050***
	st.err.	(0,039)	(0,039)	(0,040)	(0,005)	(0,006)	(0,106)
	n.obs		**553**			**13301**	
RP	**coeffs**	0,477***	0,426***	-0,410***	0,153***	0,797***	0,134***
	st.err.	(0,020)	(0,020)	(0,021)	(0,006)	(0,006)	(0,004)
	n.obs		**1422**			**9106**	
ONM	**coeffs**	0,550***	0,290***	-0,380***	0,260***	0,700***	0,110***
	st.err.	(0,026)	(0,247)	(0,027)	(0,007)	(0,008)	(0,004)
	n.obs		**944**			**6960**	
BM	**coeffs**	0,537***	0,322***	-0,574***	0,249***	0,640***	0,256***
	st.err.	(0,037)	(0,039)	(0,044)	(0,013)	(0,012)	(0,006)
	n.obs		**671**			**2297**	
ME	**coeffs**	0,479***	0,243***	-0,673***	0,160***	0,706***	0,314***
	st.err.	(0,019)	(0,021)	(0,022)	(0,006)	(0,006)	(0,004)
	n.obs		**2098**			**6525**	
EL	**coeffs**	0,453	0,348***	-0,385***	0,040***	0,909***	0,155***
	st.err.	(0,348)	(0,018)	(0,017)	(0,007)	(0,008)	(0,004)
	n.obs		**2132**			**11652**	
MA	**coeffs**	0,433***	0,312***	-0,516***	0,071***	0,880***	0,142***
	st.err.	(0,018)	(0,017)	(0,016)	(0,004)	(0,005)	(0,003)
	n.obs		**2182**			**14535**	
MV	**coeffs**	0,493***	0,513***	-0,362***	0,169***	0,751***	0,204***
	st.err.	(0,025)	(0,027)	(0,026)	(0,013)	(0,013)	(0,007)
	n.obs		**980**			**2723**	
OTR	**coeffs**	0,364**	0,574***	-0,554***	0,106**	0,795***	0,172***
	st.err.	(0,095)	(0,092)	(0,056)	(0,018)	(0,018)	(0,008)
	n.obs		**232**			**1120**	

***,**,* significant values at 99, 95, 90%

Table 3: Technology adoption effects, baseline regression results.

	1 **Coeff.** **(Std.Err.)**	**2** **Coeff.** **(Std.Err.)**	**3** **Coeff.** **(Std.Err.)**	**4** **Coeff.** **(Std.Err.)**	**5** **Coeff.** **(Std.Err.)**
Firm Intangibles	0.074***	0.091***	0.121***	-	-
	(0,015)	(0,018)	(0,030)		
Firm Intangibles$_{t-1}$	-	-	-	0.143***	-
				(0,036)	
Firm Intangibles (instrumented)	-	-	-	-	0.045***
					(0,009)
Firm Age	0.015***	0.012***	0.022***	0.020***	0.006***
	(0,002)	(0,003)	(0,004)	(0,005)	(0,001)
Listed Firm	-1.799***	-1.844***	-1.518**	-1.147**	-0.464**
	(0,306)	(0,400)	(0,624)	(0,741)	(0,196)
Firm Size	1.063***	1.143***	0.761***	0.803***	0.227***
	(0,042)	(0,049)	(0,070)	(0,078)	(0,018)
Regional R&D	-	0.001***	0.001***	0.002***	0.001***
		(0,000)	(0,000)	(0,000)	(0,000)
Region Accessibility	-	0.000*	0.000	0.000*	0.000**
		(0,000)	(0,000)	(0,000)	(0,000)
Neighbouring Regions R&D	-	-20.287	17.183	-19.110	-7.889
		(29,430)	(43,230)	(48,880)	(13,015)
Labour Cost	-		0.939***	0.978***	0.309***
			(0,324)	(0,323)	(0,115)
Country EPL	-	-	-1.266	-0.915	-0.235
			(1,418)	(1,447)	(0,542)
Shareholder Rights	-	-	-4.265***	-4.382***	-1.254**
			(1,355)	(1,143)	(0,497)
Independent Directors	-	-	3.947***	5.206***	1.414***
			(1,425)	(1,482)	(0,544)
Constant	-1.369***	-14.777***	-17.021*	-18.788**	-6.150*
	(0,580)	(0,685)	(8,889)	(9,058)	(3,260)
Firm effects	yes	yes	yes	yes	no
Country effects	yes	yes	yes	yes	yes
Sector effects	yes	yes	yes	yes	yes
Year effects	yes	yes	yes	yes	yes
Number of obs.	59728	38657	16191	15142	15121

Dependent variable: firm-level *Group H*
Estimation method: LOGIT
***,**,* significant values at 99, 95, 90%

Table 4: *tfp* effects, baseline regression results.

	1 **Coeff.** **(Std.Err.)**	**2** **Coeff.** **(Std.Err.)**	**3** **Coeff.** **(Std.Err.)**	**4** **Coeff.** **(Std.Err.)**	**5** **Coeff.** **(Std.Err.)**
Firm Intangibles	0,011*** (0,001)	0,013*** (0,001)	0,012*** (0,001)	-	-
Firm Intangibles t-1	-	-	-	0,013*** (0,001)	-
Firm Intangibles (instrumented)		-	-	-	0,013*** (0,001)
Firm Age	-0,000 (0,000)	-0,000 (0,000)	-0,000 (0,000)	0,000 (0,000)	0,000 (0,000)
Listed Firm	-0,126*** (0,016)	-0,124*** (0,020)	-0,114*** (0,025)	-0,121*** (0,026)	-0,121*** (0,026)
Firm Size	0,103*** (0,002)	0,099*** (0,002)	0,077*** (0,002)	0,079*** (0,003)	0,079*** (0,003)
Regional R&D	-	0,000*** (0,000)	0,000*** (0,000)	0,000*** (0,000)	0,000*** (0,000)
Region Accessibility	-	0,000*** (0,000)	0,000** (0,000)	0,000** (0,000)	0,000** (0,000)
Neighbouring Regions R&D	-	1,200 (1,343)	0,228 (1,687)	-0,268 (1,770)	-0,313 (1,765)
Labour Cost	-	-	0,041*** (0,010)	0,045*** (0,009)	0,044*** (0,009)
Country EPL	-	-	0,423*** (0,055)	0,457*** (0,055)	0,451*** (0,055)
Shareholder Rights	-	-	-0,365*** (0,038)	-0,375*** (0,038)	-0,380*** (0,038)
Independent Directors	-	-	0,529*** (0,044)	0,540*** (0,045)	0,541*** (0,045)
Constant	-1,309*** (0,029)	-1,261*** (0,032)	-2,121*** (0,271)	-2,298*** (0,274)	-2,254*** (0,276)
Firm effects	yes	yes	yes	yes	yes
Country effects	yes	yes	yes	yes	yes
Sector effects	yes	yes	yes	yes	yes
Year effects	yes	yes	yes	yes	yes
Number of obs.	59728	38657	16191	15242	15121

Dependent variable: firm-level *tfp*
Estimation method: GLS
***,**,* significant values at 99, 95, 90%

Table 5: Robustness, mixture regression with clustering over sub-periods (dep.var: *Value Added*).

		Sub-period 2003-2005						Sub-period 2006-2009					
		Cluster 1			Cluster 2			Cluster 1			Cluster 2		
SECTOR		**K**	**L**	**const**	**K**	**L**	**const**	**K**	**L**	**const**	**K**	**L**	**const**
CH	**coeffs**	0,361***	0,509***	-0,469***	0,184***	0,777***	0,088***	0,247***	0,639***	-0,357***	0,169***	0,772***	0,196***
	st.err.	(0,029)	(0,031)	(0,021)	(0,008)	(0,008)	(0,006)	(0,017)	(0,017)	(0,011)	(0,008)	(0,008)	(0,006)
	n.obs		**991**			**5568**			**2298**			**4997**	
RP	**coeffs**	0,502***	0,451***	-0,575***	0,172***	0,781***	0,116***	0,367***	0,531***	-0,444***	0,153***	0,793***	0,238***
	st.err.	(0,026)	(0,025)	(0,023)	(0,007)	(0,008)	(0,005)	(0,020)	(0,018)	(0,015)	(0,008)	(0,008)	(0,005)
	n.obs		**778**			**4028**			**1690**			**4032**	
ONM	**coeffs**	0,586***	0,357***	-0,573***	0,307***	0,652***	0,130***	0,498***	0,392***	-0,467***	0,252***	0,708***	0,175***
	st.err.	(0,027)	(0,026)	(0,024)	(0,009)	(0,009)	(0,006)	(0,024)	(0,029)	(0,024)	(0,010)	(0,011)	(0,007)
	n.obs		**593**			**3208**			**946**			**3157**	
BM	**coeffs**	0,526***	0,423***	-0,904***	0,235***	0,661***	0,323***	0,439***	0,438***	-0,826***	0,269***	0,613***	0,189***
	st.err.	(0,052)	(0,057)	(0,027)	(0,016)	(0,016)	(0,008)	(0,041)	(0,044)	(0,031)	(0,015)	(0,016)	(0,008)
	n.obs		**231**			**892**			**424**			**1421**	
ME	**coeffs**	0,502***	0,314***	0,313***	0,157***	0,712***	0,255***	0,355***	0,497***	-0,871***	0,154***	0,704***	0,341***
	st.err.	(0,028)	(0,032)	(0,004)	(0,009)	(0,008)	(0,005)	(0,021)	(0,028)	(0,020)	(0,007)	(0,007)	(0,005)
	n.obs		**698**			**2726**			**1350**			**3849**	
EL	**coeffs**	0,485***	0,404***	-0,620***	0,096***	0,841***	0,138***	0,373***	0,460***	-0,436***	0,101***	0,833***	0,288***
	st.err.	(0,024)	(0,021)	(0,018)	(0,010)	(0,011)	(0,007)	(0,017)	(0,015)	(0,012)	(0,009)	(0,009)	(0,006)
	n.obs		**1334**			**4883**			**2559**			**5008**	
MA	**coeffs**	0,466***	0,349***	-0,751***	0,097***	0,864***	0,097***	0,378***	0,445***	-0,519***	0,082***	0,856***	0,208***
	st.err.	(0,027)	(0,026)	(0,022)	(0,006)	(0,006)	(0,004)	(0,018)	(0,020)	(0,014)	(0,006)	(0,006)	(0,003)
	n.obs		**976**			**6605**			**1924**			**7212**	
MV	**coeffs**	0,422***	0,583***	-0,540***	0,194***	0,747***	0,169***	0,460***	0,551***	-0,414***	0,175***	0,748***	0,254***
	st.err.	(0,036)	(0,041)	(0,029)	(0,016)	(0,016)	(0,008)	(0,028)	(0,026)	(0,018)	(0,015)	(0,015)	(0,007)
	n.obs		**375**			**1159**			**728**			**1441**	
OTR	**coeffs**	0,440**	0,512**	-0,719***	0,105**	0,800***	0,132***	0,147	0,739**	-0,609***	0,074**	0,823**	0,226***
	st.err.	(0,145)	(0,140)	(0,071)	(0,022)	(0,022)	(0,010)	(0,086)	(0,100)	(0,051)	(0,025)	(0,024)	(0,012)
	n.obs		**85**			**558**			**190**			**519**	

Standard errors among brackets
***,**,* significant values at 99, 95, 90%

Table 6: Robustness, re-run estimation with clustering over sub-periods.

	Technology adoption effects	***tfp* effects**
	Coeff.	**Coeff.**
Firm Intangibles (instrumented)	0.058***	0.012***
Firm Age	0.005***	0.000
Listed Firm	-0.528***	-0.115***
Sales	0.265***	0.080***
Region Accessibility	0.000*	0.000*
Neighbouring Regions R&D	16.105	-1.131
Regional R&D	0.001***	0.000**
Labour Cost	-0.037	0.055***
Country EPL	-0.178	0.358***
Shareholder Rights	-0.181	-0.488***
Independent Directors	-0.075	0.690***
Constant	-2.159***	-2.131***
Firm effects	not	yes
Country effects	yes	yes
Sector effects	yes	yes
Year effects	yes	yes
Number of obs.	15363	15363
Dependent variable:	*Group H*	*tfp*
Estimation method:	Logit	GLS

***,**,* significant values at 99, 95, 90%

Table 7: Robustness, flexible number of clusters - LR test.

SECTOR	**LR (S=1)**	**LR (S=2)**	**LR (S=3)**	**LR (S=4)**
CH	0.05	0.05	**0.25**	
RP	0.02	0.02	**0.38**	
ONM	0.02	0.02	**0.71**	
BM	0.02	**0.08**		
ME	0.02	0.04	0.02	**0.29**
EL	0.02	0.04	**0.58**	
MA	0.02	0.04	0.04	**0.52**
MV	0.02	0.04	0.04	**0.57**
OTR	0.02	0.04	**0.54**	

LR = log-likelihood ratio test values for H_0 : numb. of segments = S, H_1 : numb. of segments > S
Boostrapped test computed over 25 replications of the basic model that had 100 replications
Decisive tests in boldface. $S = 3$ in sector MA based on insignificant β_k in the $S = 4$ specification.

Table 8: Robustness, flexible number of clusters - mixture regression output. (dep. var.: *Added Value*).

		Cluster 1			Cluster 2			Cluster 3			Cluster 4		
		K	**L**	**const**	**K**	**L**	**const**	**K**	**L**	**const**	**K**	**L**	**const**
CH	**coeffs**	0,381***	0,467***	-0,571***	0,272***	0,781***	0,016***	0,155***	0,771***	0,056***			
	st.err.	(0,041)	(0,041)	(0,045)	(0,026)	(0,025)	(0,017)	(0,014)	(0,011)	(0,009)			
	n.obs		**573**			**696**			**12585**				
RP	**coeffs**	0,495***	0,413***	-0,405***	0,413**	0,436**	-1,522***	0,154***	0,796***	0,149***			
	st.err.	(0,018)	(0,018)	(0,021)	(0,191)	(0,212)	(0,252)	(0,006)	(0,007)	(0,004)			
	n.obs		**926**			**1842**			**7760**				
ONM	**coeffs**	0,549***	0,308***	-0,381***	0,297***	0,662***	0,157***	0,097***	0,896***	-0,036**			
	st.err.	(0,024)	(0,025)	(0,026)	(0,010)	(0,011)	(0,011)	(0,021)	(0,024)	(0,016)			
	n.obs		**1130**			**6410**			**364**				
BM	**coeffs**	0,537***	0,322***	-0,574***	0,249***	0,640***	0,256***						
	st.err.	(0,037)	(0,039)	(0,044)	(0,013)	(0,012)	(0,006)						
	n.obs		**671**			**2297**							
ME	**coeffs**	0,479***	0,247***	-0,672***	0,245***	0,534***	0,326***	0,134***	0,758***	0,272***	0,124***	0,738***	,738***
	st.err.	(0,019)	(0,021)	(0,021)	(0,023)	(0,034)	(0,020)	(0,011)	(0,014)	(0,011)	(0,021)	(0,020)	(0,037)
	n.obs		**2180**			**112**			**5301**			**1030**	
EL	**coeffs**	0,456***	0,339***	-0,383***	0,406*	0,397*	-1,783***	0,040***	0,908***	0,167***			
	st.err.	(0,017)	(0,016)	(0,019)	(0,200)	(0,180)	(0,292)	(0,008)	(0,009)	(0,004)			
	n.obs		**2140**			**2364**			**9280**				
MA	**coeffs**	0,452**	0,239	-1,593***	0,433***	0,315***	-0,509***	0,070***	0,882***	0,129***	0,017	0,930***	0,648***
	st.err.	(0,138)	(0,172)	(0,240)	(0,016)	(0,016)	(0,016)	(0,004)	(0,005)	(0,003)	(0,012)	(0,011)	(0,013)
	n.obs		**350**			**2070**			**9058**			**5239**	
MV	**coeffs**	0,516***	0,552***	-0,440***	0,379***	0,391***	-0,013	0,122***	0,831***	0,181***	0,070**	0,806***	0,502***
	st.err.	(-0,030)	(-0,033)	(-0,036)	(0,041)	(0,039)	(0,044)	(0,019)	(0,021)	(0,012)	(0,026)	(0,027)	(0,023)
	n.obs		**773**			**401**			**2160**			**369**	
OTR	**coeffs**	0,368**	0,574***	-0,553***	0,116***	0,796***	0,121***	0,018	0,855***	0,462***			
	st.err.	(0,09)	(0,089)	(0,053)	(0,018)	(0,019)	(0,013)	(0,030)	(0,029)	(0,022)			
	n.obs		**241**			**928**			**183**				

Standard errors among brackets
***,**,* significant values at 99, 95, 90%

Table 9: Robustness, flexible number of clusters - re-run *tfp* estimation.

	1 **Coeff.**	**2** **Coeff.**	**3** **Coeff.**	**4** **Coeff.**	**5** **Coeff.**
Firm Intangibles	0.011***	0.012***	0.012***		
Firm Intangibles t-1				0.012***	
Firm Intangibles (instrumented)					0.012***
Firm Age	0.000	0.000	0.000	0.000	0.000
Listed Firm	-0.165***	-0.128***	-0.139***	-0.159***	-0.160***
Sales	0.116***	0.123***	0.092***	0.095***	0.096***
Region Accessibility		0.000	0.000	0.000	0.000
Neighbouring Regions R&D		-0.301	0.263	-0.96	-0.993
Regional R&D		0.000	0.000**	0.000*	0.000*
Labour Cost			0.051***	0.056***	0.057***
Country EPL			0.380***	0.382***	0.380***
Shareholder Rights			-0.081	-0.086	-0.093
Independent Directors			0.207**	0.226**	0.228***
Constant	-1.258***	-1.259***	-2.732***	-2.826***	-2.839***
Firm effects	yes	yes	yes	yes	yes
Country effects	yes	yes	yes	yes	yes
Sector effects	yes	yes	yes	yes	yes
Year effects	yes	yes	yes	yes	yes
Number of obs.	52141	34455	14372	13678	13662

Dependent variable: firm-level *tfp*
Estimation method: GLS
***,**,* significant values at 99, 95, 90%

Table 10: Data sources and definitions.

VARIABLE	**STAT. UNIT**	**SOURCE**	**TIME SPAN**	**DESCRIPTION**
Added value	Firm	Amadeus	2003-2009	Log of added value.
Labour input	Firm	Amadeus	2003-2009	Log of total number of employees.
Capital input	Firm	Amadeus	2003-2009	Log of tangible assets.
Firm intangibles	Firm	Amadeus	2003-2009	Log of intangible to tangible assets ratio.
Firm age	Firm	Amadeus	2003-2009	Age of the firm (years).
Listed firm	Firm	Amadeus	2003-2009	Dummy variable (1 = the firm is listed in the stock market, 0 = otherwise).
Sales	Firm	Amadeus	2003-2009	Log of net sales.
Regional R&D	Region	Eurostat (2012)	2003-2009	Total intramural R&D expenditure
Region accessibility	Region	Espon	2001, 2006	Multi-modal potential accessibility, std.
Neighbouring regions R&D	Region	Eurostat (2012)	2003-2009	Local clusters indicator of R&D expenses of the neighbouring regions.
Labour cost	Country	Eurostat (2012)	2003-2008	Hourly labour costs, manufacturing.
Country EPL	Country	OECD (2012).	2003-2008	Index of employment protection legislation.
Shareholder rights	Country	Siems *et al.* (2009).	2003-2005	Index of minority shareholder power to intervene in the general meeting.
Independent directors	Country	Siems *et al.* (2009).	2003-2005	Extent at which independent directors must be present in the board by law.

Table 11: Robustness, production function coefficients.

		OPACF			OLS			OPACF - group H			OPACF - group L		
SECTOR		β_K	β_L	RSC	β_K	β_L	RSC	β_K	β_L	RSC	β_K	β_L	RSC
CH	**coeffs**	0.15	0.57	0.72	0.22	0.73	0.95	0.14	0.51	0.65	0.10	0.85	0.95
	st.err.	0.043	0.084	-	0.006	0.006	-	0.038	0.060	-	0.358	0.146	-
RP	**coeffs**	0.21	0.34	0.55	0.28	0.66	0.94	0.14	0.65	0.80	0.27	0.63	0.89
	st.err.	0.053	0.066	-	0.007	0.007	-	0.035	0.063	-	0.114	0.203	-
ONM	**coeffs**	0.22	0.75	0.97	0.40	0.55	0.95	0.20	0.79	0.98	0.31	0.71	1.02
	st.err.	0.068	0.090	-	0.008	0.008	-	0.039	0.055	-	0.244	0.138	-
BM	**coeffs**	0.49	0.76	1.25	0.47	0.41	0.88	0.13	0.79	0.91	0.48	0.49	0.96
	st.err.	0.148	0.247	-	0.018	0.018	-	0.087	0.108	-	0.171	0.134	-
ME	**coeffs**	0.31	0.32	0.63	0.36	0.47	0.83	0.07	0.74	0.82	0.40	0.58	0.98
	st.err.	0.102	0.148	-	0.010	0.010	-	0.044	0.069	-	0.089	0.117	-
EL	**coeffs**	0.21	0.51	0.71	0.22	0.69	0.90	0.03	0.81	0.84	0.35	0.32	0.67
	st.err.	0.057	0.085	-	0.007	0.007	-	0.038	0.055	-	0.068	0.070	-
MA	**coeffs**	0.17	0.80	0.97	0.20	0.72	0.92	0.05	0.79	0.83	0.28	0.66	0.94
	st.err.	0.039	0.074	-	0.006	0.006	-	0.031	0.050	-	0.060	0.046	-
MV	**coeffs**	0.10	0.52	0.62	0.36	0.58	0.94	0.09	0.74	0.83	0.31	0.67	0.98
	st.err.	0.110	0.147	-	0.014	0.014	-	0.083	0.100	-	0.103	0.097	-
OTR	**coeffs**	0.30	0.89	1.19	0.20	0.71	0.91	-0.08	0.78	0.70	-0.05	0.63	0.58
	st.err.	0.237	0.263	-	0.028	0.027	-	0.121	0.143	-	0.468	0.344	-

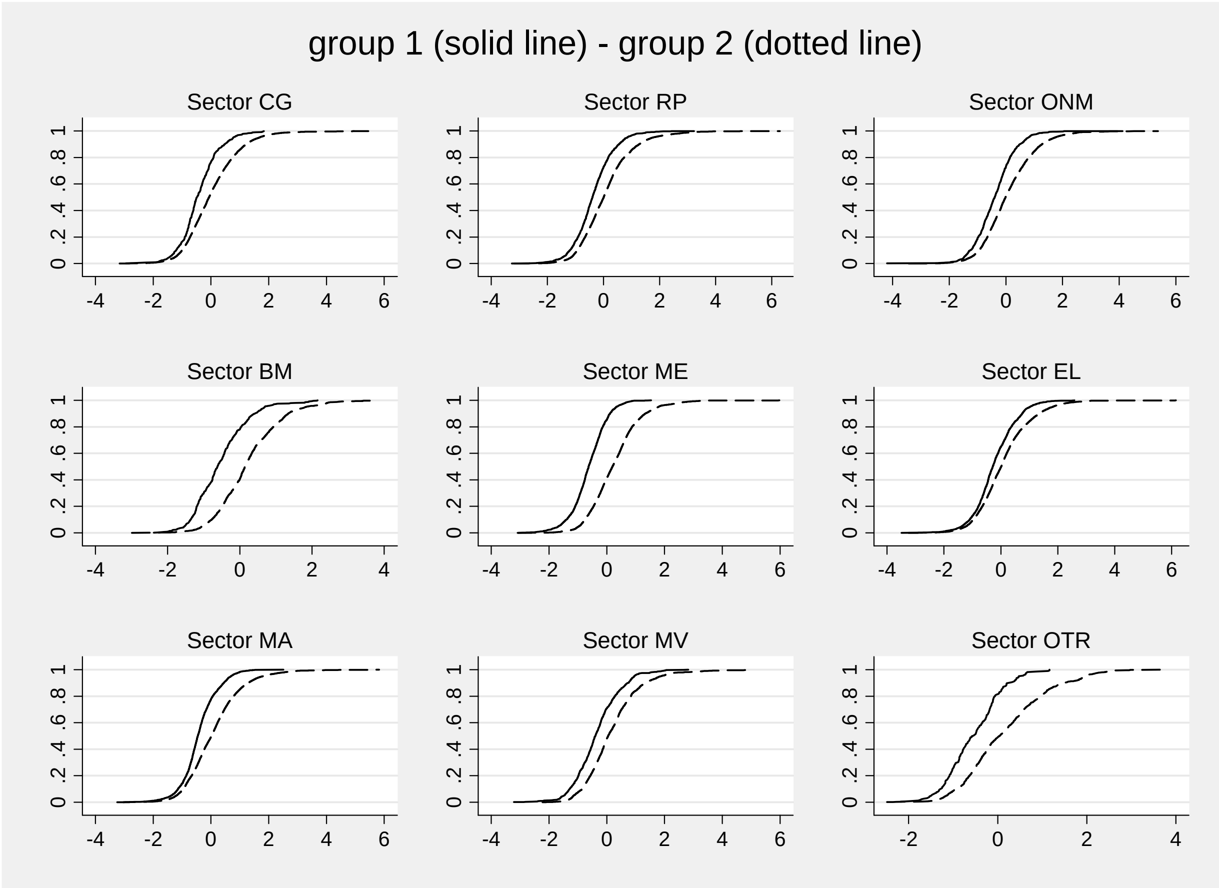

Figure 1: Predicted output, cumulative distribution functions by sector

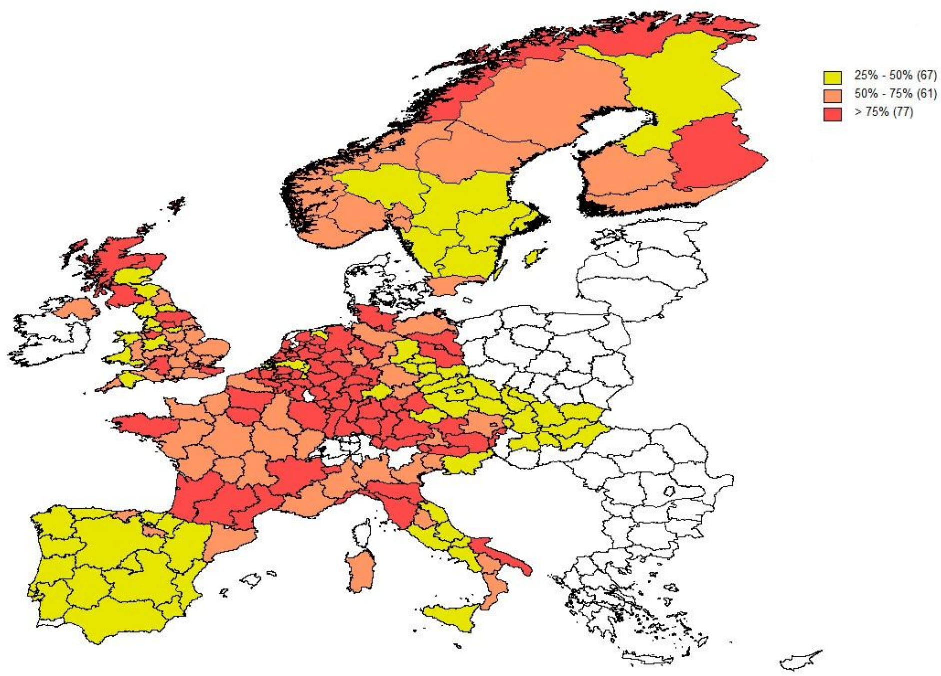

Figure 2: High to low technology firms ratio (Θ_H/Θ_L), regional values.

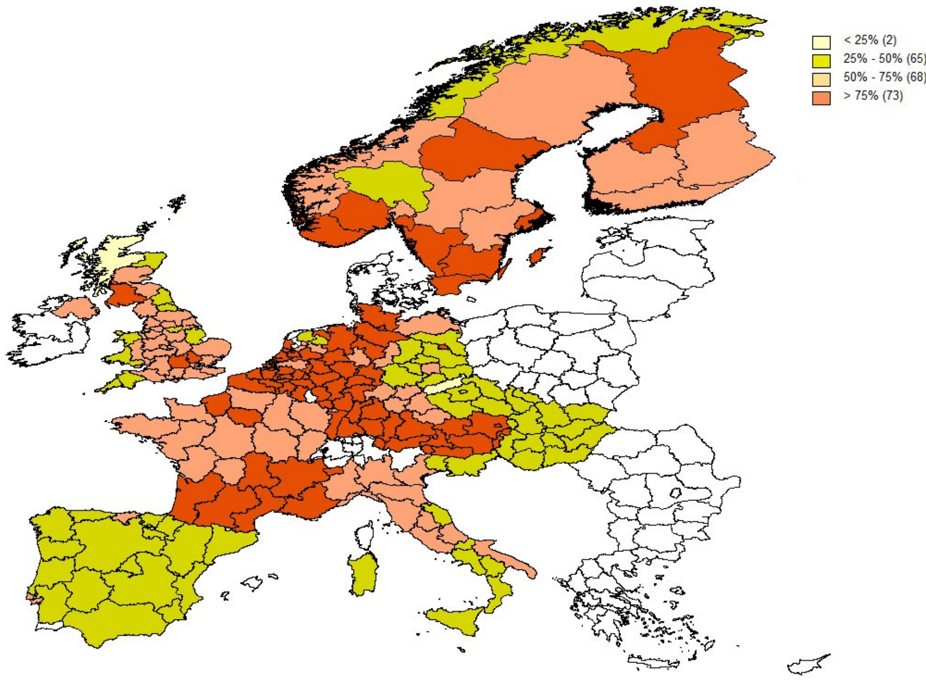

Figure 3: Firms' *tfp* distribution, regional averages.