The Contribution of Educated Workers to Firms' Efficiency Gains.

The Key Role of Proximity to the 'Local' Frontier

V. Vandenberghe[¥]

Abstract

Vandenbussche et al (2006), Aghion et al. (2009) posit and show that when economies operate close to the technical frontier, their ability to generate efficiency gains rests on the contribution of workers with advanced forms of education (i.e. those who attended tertiary education). The main originality of this empirical paper is to revisit and improve the analysis of that assumption in the context of firms located in advanced economics, assuming that something that has been verified for OECD countries or US states is likely to be observed also at a much more desegregated level. To that purpose, we analyse a rich panel of Belgian firm-level data, covering the 2008-14 period. In the first step, we concentrate on properly estimating each firm's distance/proximity to frontier using stochastic frontier methods. Step 2 consists in regressing each firm's efficiency growth rate on [1] the share of workers by education attainment [2] its (initial) distance/proximity to the frontier and [3] (the main variable of interest here) the interaction between [1] & [2], whose sign provides a direct test of the Vandenbussche/Aghion assumption. The main result of the paper supports the idea that the closer the firms are to what amounts to a local frontier; the more educated workers -- and in particular those with a master's degree -- matter for efficiency gains.

Keywords: Efficiency growth, Highly-educated workers, Frontier Firms, Proximity to frontier.

JEL Classification: J24, I20, E24, O30, O40

[¥] Université catholique de Louvain (UCL), IRES, 3 place Montesquieu, B-1348 Louvain-la-Neuve (Belgium), email: vincent.vandenberghe@uclouvain.be.

Introduction

Productivity plays a key role not only in the prosperity of countries but also in the success of firms. In practice, its level varies enormously across entities. Those forming the frontier cohabit with "laggards" (Andrews et al. 2015). In dynamic terms, productivity growth depends on the ability of the frontier entities to lift the frontier further up, but also -- and perhaps more significantly -- on the capacity of all the other entities to catch up and converge towards the frontier. Not surprisingly, there is a large economic literature on productivity convergence between countries (Barro & Sala-i-Martin, 1997; Aghion & Howitt 1992; Howitt 2000), and an emerging literature on convergence between firms (Bartelsman et al., 2008; Andrews et al. 2015).

The theoretical construct underlying the process of convergence is that of knowledge spillovers emanating from the most productive, or frontier, technology. To the extent that knowledge is non-rival and not fully appropriable, entities below the frontier can potentially improve performance by learning from the best. The learning process may be easy at first, but then proves more complicated beyond a certain threshold. What is more, there might exist various factors facilitating/impeding that learning process, whose importance may increase the closer one tries to match the performance of the best. Nelson & Phelps (1966) were the first to suggest that education facilitates the implementation of new technologies and help countries catch-up with the leading country. They argued that the importance of human capital rises with the innovative content of the tasks performed or with the extent to which it is necessary to follow and to understand new technological developments, implement more complex procedures…. They write: "*At the bottom of this scale are functions that are highly routinized.…In the other direction on this scale we have, for example, innovative functions which demand keeping abreast of improving technology*" (p. 69). Acemoglu et al. (2006) showed theoretically that innovation should be more important than adoption/replication as an economic entity approaches the technological frontier. Vandenbussche et al. (2006), Aghion et al. (2009) combined the two above ideas, and suggested that the contribution of skilled vs unskilled labour to efficiency should depend on countries' distance/proximity to the technological frontier. They predicted that highly-educated labour would be particularly efficiency-growth enhancing close to that frontier, under the assumption that *i)* at the frontier or close to it innovation matters more than adoption/replication (Acemoglu et al. 2006) *ii)* innovation is a relatively more skill-intensive activity than adoption/replication (Nelson & Phelps, 1966) and *iii)* that these skills are typically those people acquire within tertiary education, in particular the most advanced forms of tertiary education

(masters…). Vandenbusshe/Aghion and their co-authors subsequently published OECD (Vandenbussche et al., 2006) - and US state-level (Aghion et al., 2009) empirical evidence in support of this assumption.

The key purpose of this paper is to use firm-level data to revisit and improve the analysis of the advanced-education/proximity-to-frontier complementarity assumption. As far as we know, this is something that has almost never been done before. One exception is the paper by Bartelsman et al. (2015), who use German and Dutch firm-level data, but implement a slightly different econometric strategy than ours. Belgian data used here not only contain good-quality information about firms' stock of capital, total labour and its breakdown by educational attainment, they are also structured as panels, so they can be used to compute efficiency growth and explore its determinants. When exploiting firm-level data, on the condition that one takes good care of the presence of "noise", we can count on a large variation of the proximity-to-frontier. This is a plus from an econometric point of view. Also, the econometric literature on the estimation of efficiency frontiers has developed methods that are particularly suited to the analysis of firm-level data and for distinguishing efficiency performance and statistical noise. The results presented hereafter all rest on distance/proximity[i]-to-efficiency frontier delivered by *stochastic frontier methods* (more on these, and their relevance in Section 2).

But independently of the nature and quality of the data used, this paper is based on few key economic considerations. First, the study of aggregate (i.e. country)-level performance outcomes should, whenever feasible, be linked to that of firm performance within countries. In other words, macro (ie. country)-level regularities may have micro/firm-level foundations (Bartelsman et al., 2013; Andrews et al. 2015). In modern economies, where most people work inside firms efficiency gains cannot possibly exist at a more aggregate level (i.e. countries or states) if they do not show up in firms. And the same argument can be used about their determinants. In particular, if across countries that differ technologically, different types of human capital play different roles in both imitation/adoption vs. innovation activities, then that distinction should also show up when considering heterogenous firms inside a given country. Second, the existing empirical literature on the role of frontier (and education) in fostering convergence tends to overlook the important heterogeneity that exists locally (ie. within a given country). The complexity of the technology used across firms varies a lot, even when these are located very close to each other or belong to the same industry. In other words, within countries -- and perhaps also within industries in these countries --- firms differ considerably as to their distance to the efficiency frontier. Very traditional and simple

forms of production cohabit with top-notch, innovation-driven activities. Studies, exploiting country-specific firms micro data have documented, virtually without exception, enormous and persistent measured efficiency performance differences (Syverson, 2011). The magnitudes involved are striking. Syverson reports that for the USA, within four-digit SIC industries in the manufacturing sector, the average difference in logged total factor productivity (TFP) between an industry's 90th and 10th percentile plants is 0.651. This corresponds to a TFP ratio [that will turn out to correspond to distance/proximity to frontier hereafter] of $e^{0.651} = 1.92$.

The main result of the paper is that of robust econometric evidence in support of the idea that skilled workers contribute more to efficiency gains when they work inside "frontier" firms. Their contribution to efficiency growth among "laggards" is much smaller. We are also able to show the complementarity between proximity-to-frontier (PTF) and education is stronger when the latter corresponds to master's-educated workers compared to bachelor's- or upper-secondary-educated ones. Also, these results are in line with those of Bartelsman et al. (2015) who focused on German and Dutch firm-level evidence, and used quantile regression techniques. The tentative conclusion is that akin frontier countries, firms approaching the frontier inside advanced economies like Belgium, depend on advanced forms of education to achieve efficiency gains; presumably because these gains require more than imitation or replication.

The rest of the paper is organized as follows. In Section 2, we expose our methodological choices regarding the estimation of distance/proximity-to-frontier (Stage 1) and its subsequent use to assess the role of education (i.e. Stage 2). That section also spells out our Stage-2 econometric models. Results of Stage 1 and other data used in Stage-2 analysis are presented in Section 3. Section 4 contains the Stage-2 key econometric results, while Section 5 concludes.

2. Methodology

In Stage 1, we estimate each firm's distance/proximity to the efficiency frontier using stochastic frontier methods (SF hereafter) (Aigner et al., 1977; Meeusen & van den Broeck, 1977). These are methods estimating production frontier $y_{it}=lnf(K_{it},L_{it}) + \varepsilon_{it}$ parametrically. Specifications for $f(K_{it},L_{it})$ can vary (more on this below). The real interest lies in the fact ST takes into account potential measurement errors and other sources of statistical noise. Efficiency is no longer associated to the full Solow residual ε_i but only part of it. SF posit the existence of two terms: a symmetric random error term υ_i and an (in)efficiency term $\mu_{it} \geq 0$. In other words, $y_{it}=lnf(K_{it},L_{it}) + \upsilon_{it} - \mu_{it}$. The frontier is

4

called "stochastic" because output values are bounded from above by the stochastic (i.e. random) variable $lnf(K_{it}, L_{it}) + v_{it}$. Much of the stochastic frontier analysis is directed towards the prediction of the (in)efficiency term μ_{it}. Note that our interest for SF also derives from the fact that - unlike Vandenbussche et al., 2006 and Aghion et al., 2009 - the data we use are much more disaggregated; they consist of firm-level records. And these are known for being particularly prone to measurement errors/reporting inconsistencies. In short, resorting to SF might be the right thing to do to generate realistic estimates of firms' true distance(proximity) to the (in)efficiency to frontier.

A Stage 2 these estimates, in combination with data on firms' workforce educational attainment, are used to assess education's contribution to efficiency growth (i.e. the propensity of firms to get closer to the frontier over time). And the key assumption we want to be tested is whether education's contribution depends on the distance/proximity to the frontier. This framework extends and improves the one developed by Vandenbussche et al. (2006); the main difference being that it is applied to firm-level data with *i)* a lot more dispersion in terms of distance/proximity to frontier than with country-level comparisons and *ii)* a more detailed description of the educational attainment of the workforce (more on this in Section 3). The paper also pays attention to the risk of reverse causality/simultaneity – a form of endogeneity – when generating stage 2 results (more on this in Section 3). This amount to assuming that μ_{it} are partially anticipated by firms and influence labour input decisions, in particular their educational mix. For instance, firms experiencing positive efficiency developments, may respond by expanding their pool of educated workers. On the other hand, negative developments could lead to recruitment freeze and a reduction of the share of educated workers. This could cause Stage 2 OLS estimates at to exaggerate the true contribution of education to efficiency gains. when estimating contribution of education to efficiency gains. Various methods have been proposed to tackle that endogeneity problem. Most rest on the exploitation of the panel dimension of data, and, according to their approaches, it is possible to group them in two families: instrumental Variables (IV) and Control Function (CF). In this paper, we implement both. IV methods are well known and basically consist of instrumenting endogenous variables (i.e. education shares) with their lagged values, on the assumption that these are less likely to be correlated with the efficiency developments. CF is less known. Here, we implement the Ackerberg et al. (2015) version of CF which, at its core, is a method that consists of using the consumption of various intermediates (electricity, supply and services) to *proxy* unobserved efficiency developments. Finally, our paper also differs from the only other paper exploring the role of the distance to the efficiency frontier using firm-level evidence (Bartelsman et al., 2015). That paper adopts a quantile regression strategy to capture the potentially

heterogeneous contribution of education to productivity[ii], while we use a two-stage approach, with an independent measure of distance/proximity to frontier.

2.1. Stage 1 - Estimating each firms' proximity to the efficiency frontier using stochastic frontier methods.

While Vandenbussche et al. (2006) use off-the-shelf estimates of the USA's TFP level as the efficiency frontier – and each country's ratio vis-à-vis the USA as an estimate of the distance/proximity to frontier – we resort to econometric estimates derived from the analysis of firm-level data. What is more, the method we use to estimate firm-level proximity to frontier (PTF hereafter) stems from the stochastic production frontier (SF) literature. That choice is to a large extent justified by one of the weaknesses of firm-level data in comparison with country-level data: the presence of noise and measurement error. To large extent, Aigner, Lovell & Schmidt (1977) and Meeusen & van den Broeck (1977) pioneered the idea of SF models with that issue in mind. For long econometricians had been estimating *average* production functions. Aigner et al. wanted to reconcile econometrics with the more theoretical definition of a firm-level production function; i.e. the maximal output that can be achieve given input bundles and fixed technology. They posited that production (or the way it is measured by economists) is *i)* subject to random shocks v_{it} : weather, unpredictable variations in machine or labour performance, or measurement errors; and these shocks should not be confounded with *ii)* (in)efficiency μ_{it} : the result of factor under the firm's control, such as technical and economic (in)efficiency, the will or effort of the producer and his employees. In a world without shocks/errors or inefficiencies, the firm i in period t in possession of capital and labour inputs K_{it}, L_{it} would produce $f(K_{it}, L_{it})$. But more realistically, output is affected by both and writes $Y_{it}=f(K_{it},L_{it})e^{\mu_{it}+v_{it}}$ or in logs

$$lnY_{it} = lnf(K_{it}, L_{it})+v_{it} - \mu_{it} \qquad\qquad [1.]$$

where $\mu_{it} \geq 0$ represents the technical inefficiency of distance to the frontier, or equivalently

$$lnY_{it} = lnf(K_{it}, L_{it})+v_{it} + lnPTF_{it} \qquad\qquad [2.]$$

where $lnPTF_{it}$ with $lnPTF_{it} \leq 0$ and $PTF_{it} \equiv e^{-\mu_{it}}$ $0\leq PTF_{it}\leq 1$, will be interpreted as the proximity to frontier or technical efficiency: the ratio of observed output to the corresponding stochastic frontier output: $PTF_{it} =Y_{it}/f(K_{it},L_{it})e^{v_{it}}$. If proximity(distance) is maximal(minimal) $PTF_{it}=1$; $\mu_{it} =0$, the firm

is achieving the optimal output with the technology embodied in the production function $f(K_{it}, L_{it})e^{v_{it}}$. Andrew et al. (2015) would call it a "frontier" firm. When $PTF_{it} < 1; \mu_{it} \geq 0$, the firm is not making the most of the inputs K, L and proximity (distance) to frontier falls(rises). Andrew et al. (2015) would refer to such a firm as being a "laggard".

Of course, estimation is complicated by the presence of two error terms in [1]. The "noise" term v_{it}, is assumed to be randomly normally [symmetric] distributed $N(0, \sigma_v)$, while the μ_{it} term is assumed half-normally [asymmetrically] $N^+(0, \sigma_\mu)$ distributed. For the sum of a symmetric normal random variable and a truncated normal random variable [dropping t], the log-likelihood function is

$$lnL = \sum_{i=1}^{N} \{1/ \ln(2/\pi) - \ln(\sigma_S) + \ln\Phi(-\varepsilon_i\lambda/\sigma_S) - \varepsilon_i^2 2 \sigma_S^2\} \qquad [3.]$$

where $\varepsilon_i = v_i - \mu_i = lnY_i - lnf(K_i, L_i); \quad \sigma_S = \sigma_\mu + \sigma_v ; \quad \lambda = \sigma_\mu/\sigma_v$ and $\Phi(\ )$ is the cumulative distribution function of the normal distribution.

To obtain estimation for μ_i one can use the mean of the conditional distribution $f(\mu_i|\varepsilon_i)$

$$E(\mu_i|\varepsilon_i) = \mu^*_i + \sigma^* \{\phi(-\mu^*_i/\sigma^*)/\Phi(\mu^*_i/\sigma^*)\} \qquad [4.]$$

Then the proximity to frontier will be estimated by

$$\widehat{PTF} = E(e^{-u_i}|\varepsilon_i) = \{[1 - \Phi(\sigma^* - \mu^*_i/\sigma^*)]/[1 - \Phi(-\mu^*_i/\sigma^*)]\} e^{\mu^*_i + .5 \sigma^{*2}} \qquad [5.]$$

where $\mu^*_i = -\varepsilon_i \sigma_\mu^2/\sigma_v^2 = -\varepsilon_i \lambda^2 ; \quad \sigma^* = \sigma_\mu \sigma_v/\sigma_S$

Estimated parameter λ in the above likelihood expression is convenient because it is an indicator of the relative variability of the two sources of random error that distinguish firms from one another. In particular $\lambda^2 = 0$ implies $\sigma_\mu^2 = 0$ and/or $\sigma_v^2 = \infty$ i.e. that the symmetric term (random shocks/measurement errors) dominates in the determination of ε_i. In Eq.(5), $\mu^*_i = 0$ and $\sigma^* = 0$ which lead to $\widehat{PTF} = 1$ (i.e. maximal proximity/efficiency).

Turning to the econometric version of the model for the frontier eq. [6], note also the addition of year fixed effects Ψ i.e. year-to-year shifts of the frontier capturing the existence of technological progress, as it is commonly done in the SF literature. More significantly, note the fact that we allow for the frontier to be specific to the sector/industry j to which the firm belongs.

7

$$lnY^j_{it} = lnf(K^j_{it}, L^j_{it}) + \Psi^j_t + lnPTF^j_{it} + v^j_{it} \qquad [6.]$$

Note finally that $f(K^j_{it}, L^j_{it})$ can be specified in various ways. Hereafter, we will speficied it as a Cobb-Douglas, or - to relax the parametric restrictions inherent to the latter - as a Translog.[iii]

2.2. Stage 2 - Analysing the determinants of efficiency growth

Stage 2 focuses on the propensity of firms to become more effective over time in the sense that they get closer to the Stage-1-estimated frontier. Our focus is not productivity growth, but only on one of its components. Productivity growth is combined effect of *i)* a shift of the frontier (technical change). and *ii)* a movement of the economy towards the frontier (efficiency growth). What is more, our aim is to assess the role of the educational attainment of the workforce in achieving these efficiency gains, and more specifically the extent to which that contribution depends on the initial distance/proximity to that frontier. The model we estimate generalises the one used by Vandenbussche et al. (2006) and Aghion et al. (2009)

$$y_{it} \equiv \widehat{ptf}_{it} - \widehat{ptf}_{it-1} = \alpha + \beta\, ptf_{it-1}$$
$$+ \gamma_1 SEC_{it-1} + \gamma_2 BACH_{it-1} + \gamma_3 MAST_{it-1} +$$
$$\eta_1\, SEC_{it-1}\, \widehat{ptf}_{it-1} + \eta_2 BACH_{it-1}\, \widehat{ptf}_{it-1} + \eta_3\, MAST_{it-1}\, \widehat{ptf}_{it-1}$$
$$+ \theta F_{it} + \varepsilon_{it} \qquad [7.]$$

The efficiency growth achieved by firm *i* between period *t* and *t-1* is the change in the proximity to frontier i.e. $y_{it} \equiv \widehat{ptf}_{it} - \widehat{ptf}_{it-1}$; with $\widehat{ptf}_{it} \equiv ln\widehat{PTF}_{it-1}$. The initial shares of the workers with, a secondary, bachelor's and master's attainment are, respectively SEC_{it-1}; $BACH_{it-t}$; $MAST_{it-1}$ with the least educated workers being the reference (more on the definition of education below). The (initial) (log of) the proximity to the frontier is $\widehat{ptf}_{it-1}$. The main variables of interest are the education shares interacted with the (log of) proximity to frontier $\widehat{ptf}_{it-1}$. The sign of their coefficients $\eta's$ provide a direct test of the Vandenbussche/Aghion et al. assumption. Positive and statistically significant $\eta's$ will be interpreted as evidence that better-educated labour has a larger efficiency growth-enhancing effect closer to the technological frontier. Note that ptf_{it-1} is a negative number (i.e. log of a ratio <1) with a maximum of 0. That means that coefficients $\gamma's$ capture education's contribution at the frontier. Compare to Vandenbussche/Aghion, the main advantage of the above model is that it reflects the heterogeneity of educational attainment and allows for *i)* variable contribution of education to

efficiency growth and *ii)* variable interaction with the proximity to frontier. When estimating [7], one can thus assess whether the efficiency growth enhancing effect closer to the efficiency frontier is larger for, say, master's-educated workers compared to bachelor's-educated workers or secondary-educated workers. Vector F_{it} represents the list of control variables. There will be more on this in Section 3. But note already that it systematically comprises industry[iv]/year fixed effects (that among other things control for output price inflation differences), plus the annual change of the share of workers by educational attainment (ΔSEC_{it};$\Delta BACH_{it}$;$\Delta MAST_{it}$). Why distinguishing the contribution of *t-1* stock of education from it short-term variation? Fundamentally, for the same reason that led Vandenbussche/Aghion to retain educational attainment in *t-1* and not in *t* as a predictor of *t-1* to *t* efficiency growth: the problem of reverse causality/simultaneity. Educational attainment in *t* (or the variation of educational attainment between *t-1* and *t*) may be – at least partially – caused by firms' efficiency gains rather than the opposite. In econometrics this problem is refered to as simultaneity bias or reverse causality. To limit that risk, the strategy is to prioritise *t-1* values of educational attainment (or further lags – more on this when we will implement IV of Control function estimation of [7.]) as (causal) explanation of firm-level efficiency growth).

3. Data description & Stage-1 results

Our data come from the Bel-first database. We have extracted a large (unbalanced) panel of 261,935 firm-year observations corresponding to the situation of 39,901 firms, from all industries forming the Belgian private economy, in the period 2008-2014 (6 consecutive years). These firms are largely documented in terms of output (value added), capital used[v], sums dedicated to the purchase of intermediates and services,[vi] and total numbers of hours worked. Note that Bel-first contains the breakdown of the total workforce -- and this is a relatively unique feature for firm-level data -- into 4 levels of educational attainment; [1] at most a primary education attainment, [2] at most upper-secondary education attainment, [3] with a 2 to 3-year college attainment (bachelor's degrees hereafter), and [4] 5-year university attainment (master's hereafter).

Table 1 displays some descriptive statistics about the variable used at Stage 1. Figure 1 reports Stage 1-estimated PTFs computed with SF vs OLS/Solow-residuals. It's sole purpose is to illustrate who much SF makes a difference in estimating PTF. On display are the distributions obtained when using the Cobb-Douglas specification (all sectors/industries pooled). The shift to the right with SF (firms appear much closer to the frontier) is the direct consequence of allowing the frontier to be stochastic

i.e. to comprise a purely random term v_{it}. Figure 2 also displays the distribution of PTF when estimation is carried out industry by industry[vii] (as is the case for all the results presented hereafter). Quite logically, the latter distribution shifts further to the right, as many firms tend to be closer to the industry frontier (the one characterizing their industry) than to the national frontier.

[insert Figures 1 to 2 about here]

Tables 2 to 4 describe the key variables used at Stage 2, where we estimate equation [7]. First, the dependant variable: i.e. the proximity-to-frontier growth (PTF) rate reflecting the propensity of firms to become more effective over time. Table 2 reports the annual growth rates, whereas Table 3 the 6-year equivalent. Table 3 describes the main Stage-2 explanatory variables. It shows that secondary-educated workers represent about 59% of the total workforce, while bachelor's-educated and master's-educated workers count for about 13.5% and 5% of the total respectively. Note, in the last column, the information about the (average) age of firms (i.e. # years elapsed since incorporation) that we systematically include in our list of controls alongside industry (NAICS 2 digits) interacted with year, and the province[viii] where the firm is located. Note finally that at Stage 2, we lose one period (i.e. 2008) as the dependent variable is the annual PTF growth rate. And that we just keep one period when we use the 6-year growth rate as dependent variable (Table 4).

[insert Tables 1 to 4 about here]

4. Stage-2 econometric results

4. 1. Main results

Table 5 and Figure 3 present a first series of Stage-2 results, corresponding to the estimation of equations [4]. They are divided in four parts corresponding to the Cobb-Douglas and Translog specifications of the SF-estimated efficiency frontier; and results when the model is estimated using a unique vs industry-by-industry frontier.[ix] In all cases, the key coefficient corresponds to the education x *PTF* interaction variable ($\eta's$). If it is true that educated labour has a higher efficiency growth-enhancing effect closer to the technological frontier, then the estimated coefficient for that variable should be positive and statistically significant. We verify this, but with strong evidence [e.g. col. 2 of Table 5, or Figure 3] that complementarity with PTF is significantly stronger for master's-

educated workers (η^{mast}=0.162 for th Cobb-Douglas specification, 0.177 for the Translog) compared to bachelor's-educated or secondary-educated workers for which we have coefficient ranging from 0 to 0.048 and not always statistically significant. At the bottom of Table 5, we report the results of hypothesis test that $\eta's$ taken pairwise are equal. While one cannot reject the possibility that $\eta^{bach}=\eta^{second}$, there is little doubt that $\eta^{mast}>\eta^{bach}$ or that $\eta^{mast}>\eta^{second.}$. These results hold whatever the way we estimate PTF (i.e. Cobb-Douglas vs Translog specification, with all industries pooled or not).

Remember that, in Table 5, reported coefficients for the education attainment inform about the latter's contribution to productivity growth at the frontier.[x] A .079 value thus suggests that a 0 to 100% rise of the share of master's-educated workers in those firms leads to a 7.9 %-point rise of the annual efficiency growth rate. More realistically, a 10%-point rise of the share of master's-educated workers inside frontier firms adds 0.79 %-point to that rate.

The other variables present is the model have the expected sign. The closer firms initially (i.e. in *t-1*) are to the frontier the lower their efficiency growth during the subsequent period. This is supportive of the standard idea of convergence over time: firms with lower efficiency levels (the "laggards") tend to catch up with the frontier firms, via dissemination or imitation of technological or managerial of best practices.

4.2. Robustness checks

We checked the robustness of our empirical results in two directions. First, we use 6-year efficiency growth instead of 1-year growth. Results are reported in Table 6 and are qualitatively very similar to those on display in Table 5. Second, and more significantly, we address the problem of reverse causality/simultaneity bias. To that end, we first estimate the relevant parameters of our model using only "internal" instruments and the generalize method of moments (IV-GMM here after). The essence of this strategy is to use lagged values of endogenous variables (here lags 1 to 3) as instruments for the endogenous variables. An alternative to IV-GMM that seems promising and relevant is to adopt the more structural approach initiated by Olley & Pakes (1998) and further developed by Ackerberg et al. (2015) (ACF-GMM, hereby). The essence of these approaches is to use some function of a firm's consumption of intermediates to control for (proxy) time-varying efficiency developments that are (partially) anticipated by the firms and correlated with the observed level of education (see Appendix for a full exposition of the ACF 2 steps estimator).

Results for IV-GMM are reported in Table 7. We reproduce the key results of Table 5, namely that of strong and statistically significant complementarity with PTF for master's-educated workers (η^{mast}=0.149 for th Cobb-Douglas specification, 0.181 for the Translog) compared to bachelor's-educated or secondary-educated workers for which we have coefficient ranging from 0 to 0.096 but not statistically significant. At the bottom of Table 7, we report the results of hypothesis test that $\eta's$ taken pairwise are equal. And again, there is little doubt that $\eta^{mast} > \eta^{second.}$. That $\eta^{mast} > \eta^{bach}$ or $\eta^{bach} > \eta^{sec}$ is less clear cut. Note that the bottom of Table 7 reports that usual statitics assessing the relevance of our IV-GMM estimator. The first test determines whether endogenous regressors in the model are in fact exogenous. Reported GMM-C statistics is high and statistically different than zero, meaning that there are good reasons to believe that our regressors are endogenous and that IV-GMM should be preferred to OLS. Next, we have statistics that measure the relevance of our instruments. For them to be a valid they must be sufficiently correlated with the included endogenous regressors. Moreover, what really matters IV estimation is whether the component of a regressor that is orthogonal to the other regressors can be explained by the component of the predicted value of that regressor that is orthogonal to the predicted values of the other regressors in the model. Shea's partial R^2 statistic measures this correlation. Because the bias of instrumental-variables estimators increases as more instruments are used, we report Shea's *adjusted* partial R^2 statistics, as they make a degrees-of-freedom adjustment for the number of instruments, analogous to the adjusted R^2 measure used in OLS regression. Although what constitutes a "low" value for Shea's partial R^2 depends on the specifics of the model being fit and the data used, our results suggest that our instruments are good at predicting endogenous regressors. Finally, in addition to the requirement that instrumental variables be correlated with the endogenous regressors, the instruments must also be uncorrelated with the structural error term. If (as is our case) the model is overidentified, meaning that the number of additional instruments exceeds the number of endogenous regressors, then we can test whether the instruments are uncorrelated with (i.e. orthogonal to) the error term. Results of that test are reported at the very end of Table 6. The null hypothesis is that all instruments are uncorrelated with the sample-equivalent of the error term. And in our case it appears very likely (prob>.8). We thus verify the second condition for our instruments to be considered as valid.

Our last set of econometric results are reported in Table 8 and correspond to the implementation of the ACF 2-step estimator (see Appendix or Vandenberghe et. al. 2013). They also allow us to reproduce the key OLS results of Tables 5, 6 or those obtained with IV-GMM and reported in Table 7. We confirm that there is a statistically significant complementarity between master's-educated workers and firms' proximity to frontier PTF. We find estimated form η^{mast}=0.266 for th Cobb-

Douglas specification and 0.146 for the Translog. The corresponding coefficient for bachelor's-educated or secondary-educated workers are systematically smaller and not statistically significant. Suggesting that they play a much less decisive role in generating efficiency gains closer to the efficiency frontier.

[insert Tables 4 to 6 about here]
[insert Figure 4 about here]

4. Conclusion

There is plenty of individual-level evidence, based on the estimation of Mincerian equations, showing that better-educated individuals earn more, presumably because they are more productive. Many macroeconomists (Mankiw et al, 1992), analysing cross-country time series, also support the idea that the continuous expansion of education has contributed positively to growth by raising productivity.[xi] There is also a relatively vast empirical literature on the effects of human capital on firms' productivity. Using matched employer-employee data sets, Vandenberghe & Lebedinski (2014) for Belgium, Turcotte & Rennison (2004) for Canada, Fox & Smeets (2011) for Denmark, Abowd et al. (1999) for France, Galindo-Rueda & Haskel (2005) for the UK, Haltiwanger et al. (2007) for the United States, and Van Biesebroeck (2011) for Zimbabwe all find positive effects of workers' skills on firm/plant productivity. But this paper tries to go way beyond simply showing that education matters for firms' productivity.

The key idea here is to assess the role of different levels of educational attainment in fostering efficiency at different distances of the technological frontier; and to do that using firm-level panel evidence. This is something that, to our knowledge, has rarely been done before. Vandenbussche et al (2006), Aghion et al. (2009) using much more aggregate data have shown that when economies operate close to the productivity frontier, their ability to generate productivity growth rests on advanced forms of education. The idea underpinning the present paper is that something that has been observed for OECD countries (Vandenbussche et al, 2006) or US states (Aghion et al., 2009) should also be visible at much more disaggregated level; that of firms. To that purpose, we analyse a rich panel of firms forming the private-for-profit Belgian economy, covering the 2008-14 period.

The results of this paper are essentially twofold. First, the paper shows that the 'national' frontier – and within a particular country the industry frontier - plays a role in fostering convergence, and that advanced forms of education play a complementarity role at that level. That result can be visualised on Figure 4, left-hand axis. Their contribution to efficiency growth among "laggards" is much smaller (on Figure 4, thick solid lines are negatively slopped when going from right to left). Our Belgian firm-level panel data strongly support the idea that workers with advanced types of education contribute more to efficiency gains when they work inside firms situated closer to the national/industry efficiency frontier. So far, the empirical literature had emphasised this result in reference to the international/global frontier. The fact that we are able to replicate that results with a much more "local" frontier might come as a surprise to people who believe that small advanced economies like Belgium – and the industries inside it – are relatively homogeneous in terms of efficiency and educational attainment of the workforce. It is probably less to people who have been working with witin country/firm-level data, and have quantified the (important) heterogeneity that exists at the level.

The second key result is that complementarity is significantly stronger for master's-educated workers than their bachelor's- or upper-secondary-educated peers. Our results suggest that master's/ISCED[xii] 7 clearly dominate bachelor's/ISCED 6 degrees when it come to fostering efficiency gains close to the frontier. While the existing empirical literature on distance-to-frontier and education concludes that it is crucial to distinguish between the two margins that are primary/secondary versus tertiary educational attainment [with only the latter type of education being a significant source of convergence closer to the international frontier], we conclude that [within the context of a relatively advanced economy like Belgium] it might also be important to distinguish bachelor's vs master's. These two results accord with idea that innovation (i.e. which is presumably how firms improve performance when they start approaching the frontier) is a relatively more skill-intensive activity than imitation or replication and that only. The second one also suggests that, in the context of relatively advanced economies like Belgium, only workers who have been exposed to a master's/ISCED 7 education within tertiary education possess those skills. This result may have (educational) policy implications.

Note finally that our results raise the question of the optimal allocation of workers with master's across firms and industries. Should policies aimed at boosting efficiency growth foster the mobility of more educated workers towards "frontier" entities? Moving more workers with master's to firms closer to the efficiency frontier would have a large impact on productivity. At the 10th decile of the

proximity-to-frontier distribution, we estimate that a 10%-point increase in the share of these workers [a big increment that would double the current share] could raise annual efficiency growth by up to 5.5% points (Figure 4, right-hand figure). The size of the effect declines the further the firm is from the technological frontier. Around the 5th decile of the proximity-to-frontier distribution, the equivalent gains falls below 2%-points.

Is there evidence that these individuals do not spontaneously concentrate in firms close to the frontier, where they are more productive? This evidently calls for future research investigating labour (re)allocation mechanisms. But Figure 4 contains some preliminary elements of answer about the adequacy of workers' current allocation. The dashed blue line (common to the 3 sub-figures) corresponds to the overall distribution of the workforce along the proximity-to-frontier (PTF) axis. The fact that it is not horizontal (and equal to 0.1) reflects the unequal distribution of firm size along the PTF distribution.[xiii] The solid thin lines inform about the distribution of each category of educated workers (secondary-, bachelor's- or master's-educated). When the solid thin line is above the dashed one, educated workers are overrepresented and vice-versa. The good news, in the case of Belgium, is that master's-educated workers are clearly overrepresented among "frontier" firms; which is precisely where they contribute the most to efficiency gains. Still, many of them work in "laggard" firms.

5. References

Abowd, J. M., F. Kramarz and D.N. Margolis (1999), 'High wage workers and high wage firms', *Econometrica*, 67(2), 251–334.

Acemoglu, D., P. Aghion and F. Zilibotti (2006), 'Distance to Frontier, Selection, and Economic Growth', *Journal of the European Economic Association*, 4, 37-74.

Ackerberg; D.A., K. Caves & G. Frazer (2015), 'Identification Properties of Recent Production Function Estimators', *Econometrica*, 83, 2411-2451,

Aghion, P., Boustan, L., Hoxby, C., & Vandenbussche, J. (2009), 'The Causal Impact of Education on Economic Growth: Evidence from United States', in Romer, D. & Wolfers, J. (Eds.), *Brookings Papers on Economic Activity: Conference Draft,* Brookings, Washington, DC.

Aghion, P., and P. Howitt (1992), 'AModel of GrowthThrough CreativeDestruction', *Econometrica*, 60, 323–351

Aigner, D. J., C. A. K. Lovell, and P. Schmidt (1977), 'Formulation and estimation of stochastic frontier production function models', *Journal of Econometrics,* 6, 21–37.

Andrews, D., C. Criscuolo and P. Gal (2015), 'Frontier firms, technology diffusion and public policy: micro evidence from OECD countries', *OECD Productivity Working Papers* 2.

Bartelsman, E. J & Haskel, J. & M. Ralf (2008), 'Distance to Which Frontier? Evidence on Productivity Convergence from International Firm-level Data', *CEPR Discussion Papers* 7032.

Bartelsman, E. & S. Dobbelaere & B. Peters (2015), 'Allocation of human capital and innovation at the frontier: firm-level evidence on Germany and the Netherland's', *Industrial and Corporate Change*, 24(5), 875-949.

Barro, Robert, and Xavier Sala-i-Martin (1997),'Technological Diffusion, Convergence and Growth', *Journal of Economic Growth*, 2, 1–26.

Chambers, R. G. (1988), *Applied production analysis: a dual approach*, Cambridge University Press.

de la Fuente, A. & R. Doménech, (2006), 'Human Capital in Growth Regressions: How Much Difference Does Data Quality Make?', *Journal of the European Economic Association*, 4(1),. 1-36.

Fox, J. and V. Smeets (2011), 'Does input quality drive measured differences in firm productivity?', *International Economic Review*, 52(4), 961–989.

Fuss, M., D. McFadden et Y. Mundlak (1978), « A Survey of Functional Forms in the Economic Analysis of Production ». In : *Production Economics: A Dual Approach to Theory and Applications*, North-Holland, under the direction of Melvin Fuss et Daniel McFadden.

Galindo-Rueda, F. and J. Haskel (2005), 'Skills, workforce characteristics and firm-level productivity: evidence from the matched ABI/Employer skills survey', *IZA Discussion Paper* 1542. IZA: Bonn.

Haltiwanger, J., J. Lane and J. Spletzer (2007), 'Wages, productivity, and the dynamic interaction of businesses and workers', *Labour Economics*, 14(3), 575–602.

Howitt, Peter (2000), 'Endogenous Growth and Cross-Country Income Differences', *American Economic Review*, 90, 829–846.

Mankiw, G., D. Romer, and D. Weil (1992), 'A Contribution to the Empirics of Economic Growth', *Quarterly Journal of Economics*, 107(2), 407-437.

Meeusen, W., and J. van den Broeck (1977), 'Productivity estimation from Cobb–Douglas production functions with composed error', *International Economic Review*, 18, 435–444.

Nelson, R. and Phelps, E. (1966), 'Investment in Humans, Technological Diffusion, and Economic Growth', *The American Economic Review*, 56, 69-75.

Sianesi, B. and J. van Reenen (2003),'The returns to education: macroeconomics', *Journal of Economic Surveys*, 17(2), 157–200.

Syverson, C. (2011), 'What Determines Productivity?', *Journal of Economic Literature*, 49(2), 326–365.

Turcotte, J. and L. W. Rennison (2004), 'The link between technology use, human capital, productivity and wages: firm-level evidence', *International Productivity Monitor*, 9(Fall), 25–36.

Van Biesebroeck, J. (2011), 'Wages equal productivity. Fact or fiction? Evidence from SubSaharan Africa', *World Development*, 39(8), 1333–1346.

Vandenberghe, V. and L. Lebedinski, (2014), 'Assessing education's contribution to productivity using firm-level evidence', *International Journal of Manpower*, 35(8), 1116-1139.

Vandenberghe, V., M. Rigo and F. Waltenberg (2013), 'Ageing and Employability. Evidence from Belgian Firm-Level Data', *Journal of Productivity Analysis*, 40(1), 111-136.

Vandenbussche, J., Aghion, P. and Meghir; C. (2006), 'Distance to Frontier, Growth, and the Composition of Human Capital', *Journal of Economic Growth*, 11(2), 97-127.

6. Tables and Figures

Table 1: Descriptive statistics- main regressors used to estimated proximity to frontier (PTF) (Stage 1 equation)

s	Number of workers in full-time equivalent	Total capital [per employee]in th. EUR
p5	6.00	39
p25	16.00	117
p50	34.00	225
p75	74.00	508
p95	325.00	4,102
N	39,901	

Source: Belfirst 2008-2014

Figure 1 –Stochastic Frontier(SF)- vs OLS-estimated proximity to frontier. Comparison of density distributions. Year 2014
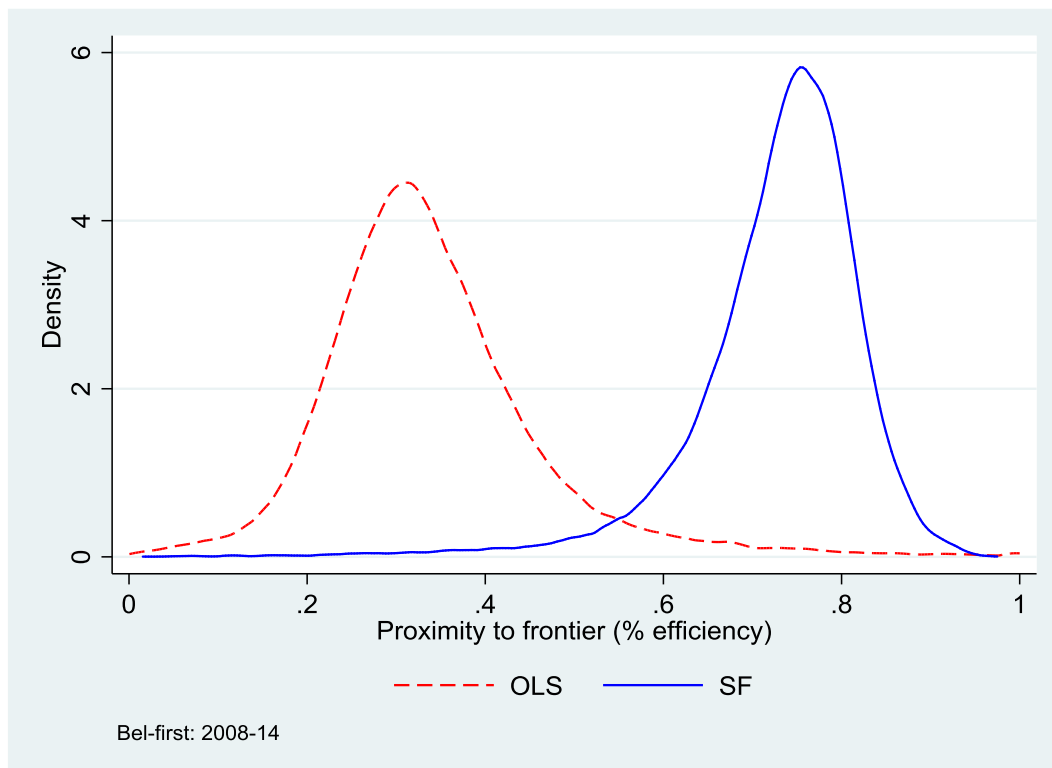
Source: Belfirst 2008-2014
The STATA **frontier** command was use to generate the SF-estimated values for the proximity to frontier.

'

Figure 2- SF-estimated proximity to frontier. Polled vs industry by industry sample- Scatter plot. Year 2014
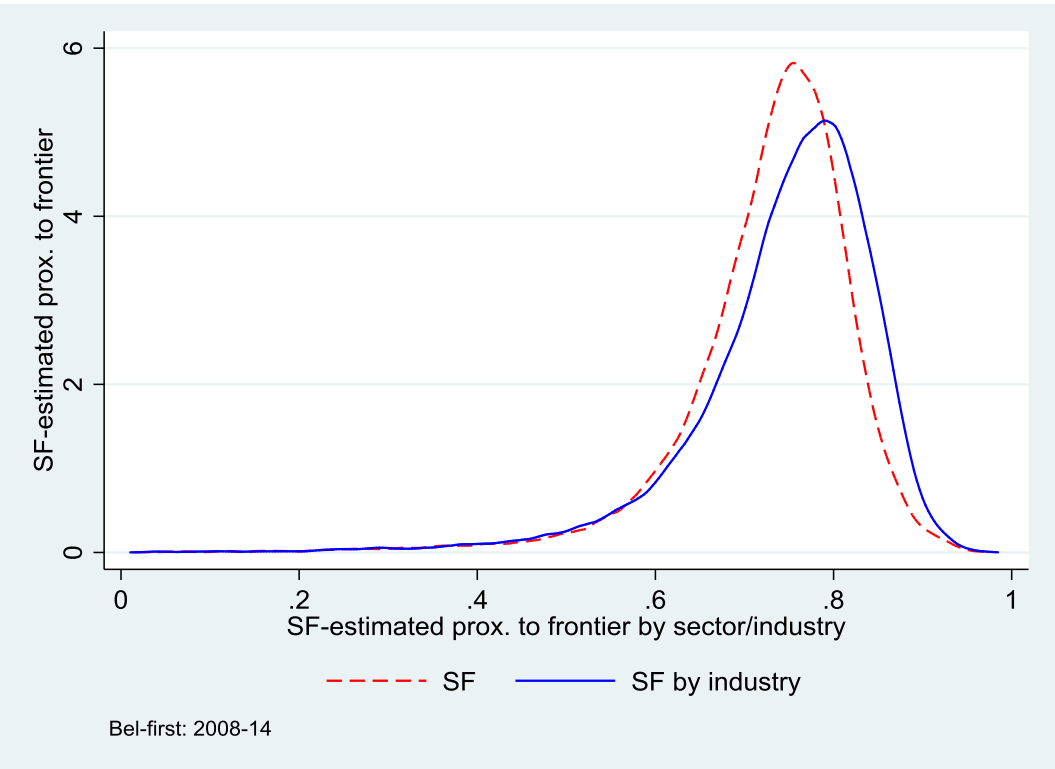
Source: Belfirst 2008-2014
The STATA **frontier** command was use to generate the SF-estimated values for the proximity to frontier.

Table 2: Descriptive statistics- proximity to frontier growth rate (1-year)

	2009		2010		2011		2012		2013		2014	
	SF-CD by industry*	SF-Translog by industry	SF-CD by industry*	SF-Translog by industry	SF-CD by industry*	SF-Translog by industry	SF-CD by industry*	SF-Translog by industry	SF-CD by industry*	SF-Translog by industry	SF-CD by industry*	SF-Translog by industry
p25	-0.0321	-0.0398	-0.0332	-0.0412	-0.0301	-0.0381	-0.0304	-0.0394	-0.0297	-0.0374	-0.0308	-0.0388
p50	0.0018	0.0026	-0.0017	-0.0016	0.0005	0.0003	0.0012	0.0010	0.0005	0.0002	-0.0013	-0.0019
p75	0.0339	0.0426	0.0307	0.0393	0.0308	0.0384	0.0312	0.0388	0.0293	0.0368	0.0260	0.0326
N	242,025											

*:NAICS 2-digit
The STATA **frontier** command was use to generate the SF-estimated values for the proximity to frontier.
Source: Belfirst 2008-2014

Table 3: Descriptive statistics- main regressors (Stage 2 equation)

	1-year lagged shares (level)				Shares (1-year difference)				1-year lagged (log of) proximity to frontier (OLS-estimated)	Age of firm
	At most prim. educated workers	Second. educated workers	Bachelor's educated workers	Master's educated workers	Share at most prim. educated workers - 1-year growth	Share second. educated workers - 1-year growth	Share bach. educated workers - 1-year growth	Share mast educated workers - 1-year growth		
2009	0.224	0.600	0.131	0.045	-0.004	-0.006	-0.004	-0.004	-1.148	19.419
2010	0.223	0.598	0.131	0.047	-0.002	-0.001	-0.002	-0.002	-1.148	19.791
2011	0.221	0.597	0.133	0.049	-0.002	-0.001	-0.002	-0.002	-1.148	20.119
2012	0.224	0.592	0.134	0.050	0.002	-0.004	0.002	0.002	-1.144	20.617
2013	0.227	0.586	0.136	0.052	0.002	-0.004	0.002	0.002	-1.141	21.242
2014	0.228	0.577	0.140	0.055	0.003	-0.006	0.003	0.003	-1.133	22.178
Average	**0.224**	**0.592**	**0.134**	**0.050**	**0.000**	**-0.004**	**0.000**	**0.000**	**-1.144**	**20.559**
N	242,025									

Source: Belfirst 2008-2014

Table 4: Descriptive statistics- proximity to frontier growth (6-year)

	2014 SF-Cobb-Douglas frontier, by industry[*]	SF-Translog frontier, by industry
p25	-0.0538	-0.0667
p50	0.0007	0.0006
p75	0.0530	0.0669
N	39,901	

Source: Belfirst 2008-2014

The STATA **frontier** command was use to generate the SF-estimated values for the proximity to frontier.

[*]:NAICS 2-digit

Table 5 - The contribution to 1-year efficiency growth[μ] of educated workers, according to the proximity to frontier. **OLS** *estimates for Stage-2 Eq.[4] (cluster-robust[£] standard errors)*

	Cobb-Douglas SF- by industry[*]	Translog SF- by industry[*]
$\widehat{ptf}_{it-1}$	-0.451[***]	-0.476[***]
	(0.017)	(0.018)
SEC_{it-1}	0.007	0.011
	(0.005)	(0.006)
$BACH_{it-1}$	0.012	0.030[**]
	(0.009)	(0.010)
$MAST_{it-1}$	0.079[***]	0.083[***]
	(0.012)	(0.013)
$\widehat{ptf}_{it-1}.SEC_{it-1}\ [\eta^{second}]$	0.032	0.048[*]
	(0.019)	(0.021)
$\widehat{ptf}_{it-1}.BACH_{it-1}\ [\eta^{bach}]$	-0.002	0.051
	(0.031)	(0.031)
$\widehat{ptf}_{it-1}MAST_{it-1}\ [\eta^{mast}]$	0.162[***]	0.177[***]
	(0.039)	(0.039)
Controls	yearXindustry(NAICS 2-digit), province, year of incorporation, change of educational mix (ΔSEC_{it}; $\Delta BACH_{it}$; $\Delta MAST_{it}$)	yearXindustry(NAICS 2-digit), province, year of incorporation, change of educational mix (ΔSEC_{it}; $\Delta BACH_{it}$; $\Delta MAST_{it}$)
Nobs (firmsXyear)	242,025	242,025
Nfirms	39,901	39,901
R^2	0.2254	0.2311
prob $\eta^{mast}=\eta^{bach}$	0.0006	0.0065
prob $\eta^{mast}=\eta^{second}$	0.0004	0.0004
prob $\eta^{bach}=\eta^{second}$	0.2303	0.9322

Estimates obtained using Bel-first firm-level data (2008-2014)

The STATA **frontier** command was use to generate the SF-estimated values for the proximity to frontier.

[μ]: $\widehat{ptf}_{it}$ - $\widehat{ptf}_{it-1}$

[£]: Standard errors are computed to account for firm-level clustering (correlation of error terms)

[*] $p < 0.05$, [**] $p < 0.01$, [***] $p < 0.001$

[*]: NAICS 2-digit

Figure 3 – The contribution to 1-year efficiency growth[μ] of educated workers, according to the proximity to frontier – plot of the interaction education- proximity-to-frontier OLS-estimated coefficients Eq. [4]. SF estimates (cluster-robust[£] standard errors)

Estimates obtained using Bel-first firm-level data (2008-2014)

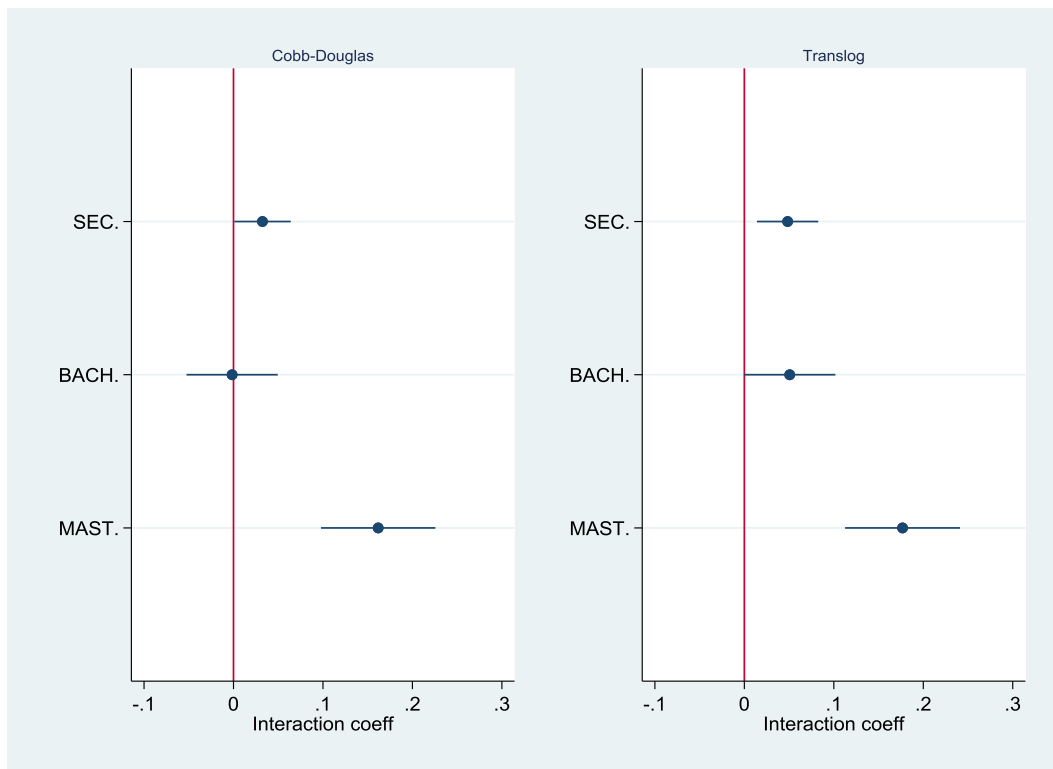The STATA **frontier** command was use to generate the SF-estimated values for the proximity to frontier.

μ: $\widehat{ptf}_{it} - \widehat{ptf}_{it-1}$

*Table 6 - The contribution to 6-year efficiency growth[μ] of educated workers, according to the proximity to frontier. **OLS** estimates for Stage 2 Eq.[4] (cluster-robust[£] standard errors)*

	Cobb-Douglas SF- by industry[*]	Translog SF- by industry*
$\widehat{ptf}_{it-1}$	-0.697***	-0.724***
	(0.0098)	(0.0097)
SEC_{it-1}	-0.004	-0.007*
	(0.0026)	(0.0032)
$BACH_{it-1}$	0.019***	0.018***
	(0.0042)	(0.0052)
$MAST_{it-1}$	0.045***	0.042***
	(0.0069)	(0.0085)
$\widehat{ptf}_{it-1}.SEC_{it-1}\ [\eta^{second}]$	-0.005	-0.006
	(0.0043)	(0.0049)
$\widehat{ptf}_{it-1}.BACH_{it-1}\ [\eta^{bach}]$	0.006	0.030***
	(0.0066)	(0.0077)
$\widehat{ptf}_{it-1}MAST_{it-1}\ [\eta^{mast}]$	0.103***	0.115***
	(0.0091)	(0.0108)
Controls	yearXindustry(NAICS 2-digit), province, year of incorporation, change of educational mix (ΔSEC_{it}; $\Delta BACH_{it}$; $\Delta MAST_{it}$)	yearXindustry(NAICS 2-digit), province, year of incorporation, change of educational mix (ΔSEC_{it}; $\Delta BACH_{it}$; $\Delta MAST_{it}$)
Nobs (firmsXyear)	39,901	39,901
Nfirms	39,901	39,901
R^2	0.44	0.44
prob $\eta^{mast} = \eta^{bach}$	0.0000	0.0000
prob $\eta^{mast} = \eta^{second}$	0.0000	0.0000
prob $\eta^{bach} = \eta^{second}$	0.0086	0.0557

Estimates obtained using Bel-first firm-level data (2008-2014)

The STATA **frontier** command was use to generate the SF-estimated values for the proximity to frontier.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

μ: $\widehat{ptf}_{it} - \widehat{ptf}_{it-6}$

£: Standard errors are computed to account for firm-level clustering (correlation of error terms)

$: NAICS 2-digit

*Table 7- The contribution to 1-year efficiency growth$^{\mu}$ of educated workers, according to the proximity to frontier. **IV-GMM** estimates for Stage 2 Eq.[4] where regressors are instrumented by their lagged 1 to 3 values (cluster-robust$^{£}$ standard errors)*

	Cobb-Douglas SF- by industry*	Translog SF- by industry*
$\widehat{ptf}_{it-1}$	-0.307***	-0.340***
	(0.0337)	(0.0394)
SEC_{it-1}	-0.006	-0.004
	(0.0105)	(0.0139)
$BACH_{it-1}$	0.022	0.043*
	(0.0173)	(0.0197)
$MAST_{it-1}$	0.068***	0.078***
	(0.0204)	(0.0229)
$\widehat{ptf}_{it-1}.SEC_{it-1}\ [\eta^{second}]$	-0.014	0.001
	(0.0384)	(0.0459)
$\widehat{ptf}_{it-1}.BACH_{it-1}\ [\eta^{bach}]$	0.037	0.096
	(0.0577)	(0.0620)
$\widehat{ptf}_{it-1}MAST_{it-1}\ [\eta^{mast}]$	0.149*	0.181*
	(0.0692)	(0.0741)
Controls	yearXindustry (NAICS 2-digit), province, year of incorporation, change of educational mix (ΔSEC_{it}; $\Delta BACH_{it}$; $\Delta MAST_{it}$)	yearXindustry (NAICS 2-digit), province, year of incorporation, change of educational mix (ΔSEC_{it}; $\Delta BACH_{it}$; $\Delta MAST_{it}$)
Nobs (firmsXyear)	102,342	102,342
Nfirms	39,902	39,902
WaldX	1092.6365	1018.7167
prob $\eta^{mast}= \eta^{bach}$	0.2233	0.3770
prob $\eta^{mast}= \eta^{second}$	0.0086	0.0053
prob $\eta^{bach}= \eta^{second}$	0.3266	0.0789
IV-Test of endogeneity of regressors		
GMM C-statistic χ2	17.6231	12.9878
prob_C-stat.=0	0.0072	0.0432
C-stat. df	6	6
IV-First-stage assessment (Shea's Adj. Partial R^2)		
SEC_{it-1}	.68	.69
$BACH_{it-1}$	.68	.68
$MAST_{it-1}$	.75	.75
$\widehat{ptf}_{it-1}.SEC_{it-1}$	.66	.67
$\widehat{ptf}_{it-1}.BACH_{it-1}$	.69	.69
$\widehat{ptf}_{it-1}MAST_{it-1}$	.77	.77
Exogeneity (orthogonality of instruments)		
Hansen	7.7696	6.4926
prob Hansen=0	0.8029	0.8892
Hansen df	12	12

Estimates obtained using Bel-first firm-level data (2008-2014). All regressors with and educational content are instrumetent using lags 1 to 3.

The STATA **frontier** command was use to generate the SF-estimated values for the proximity to frontier.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

μ: $\widehat{ptf}_{it} - \widehat{ptf}_{it-1}$

£: Standard errors are computed to account for firm-level clustering (correlation of error terms).

$: NAICS 2-digit

*Table 8 - The contribution to 1-year efficiency growth$^\mu$ of educated workers, according to the proximity to frontier. **ACF$^\$$-GMM** estimates for Eq.[4] (cluster-robust$^\pounds$ standard errors)*

	Cobb-Douglas SF-by industry[*]	Translog SF-by industry[*]
$\widehat{ptf}_{it-1}$	-0.410[***]	-0.456[***]
	(0.0437)	(0.0559)
SEC_{it-1}	0.001	0.009
	(0.0116)	(0.0197)
$BACH_{it-1}$	0.004	0.020
	(0.0124)	(0.0214)
$MAST_{it-1}$	0.065[*]	0.052
	(0.0263)	(0.0300)
$\widehat{ptf}_{it-1}.SEC_{it-1}\,[\eta^{second}]$	-0.008	0.029
	(0.0520)	(0.0636)
$\widehat{ptf}_{it-1}.BACH_{it-1}\,[\eta^{bach}]$	0.005	0.025
	(0.0426)	(0.0691)
$\widehat{ptf}_{it-1}MAST_{it-1}\,[\eta^{mast}]$	0.226[*]	0.146[*]
	(0.1078)	(0.0612)
Controls	yearXindustry(NAICS 2-digit), province, year of incorporation, change of educational mix (ΔSEC_{it}; $\Delta BACH_{it}$; $\Delta MAST_{it}$)	yearXindustry(NAICS 2-digit), province, year of incorporation, change of educational mix (ΔSEC_{it}; $\Delta BACH_{it}$; $\Delta MAST_{it}$)
Nobs (firmsXyear)	89,556	89,556
Nfirms	39,901	39,901
R^2	0.44	0.44
prob $\eta^{mast}=\eta^{bach}$	0.0689	0.1189
prob $\eta^{mast}=\eta^{second}$	0.0024	0.0041
prob $\eta^{bach}=\eta^{second}$	0.8518	0.9494

Estimates obtained using Bel-first firm-level data (2008-2014)

The STATA **frontier** command was use to generate the SF-estimated values for the proximity to frontier.
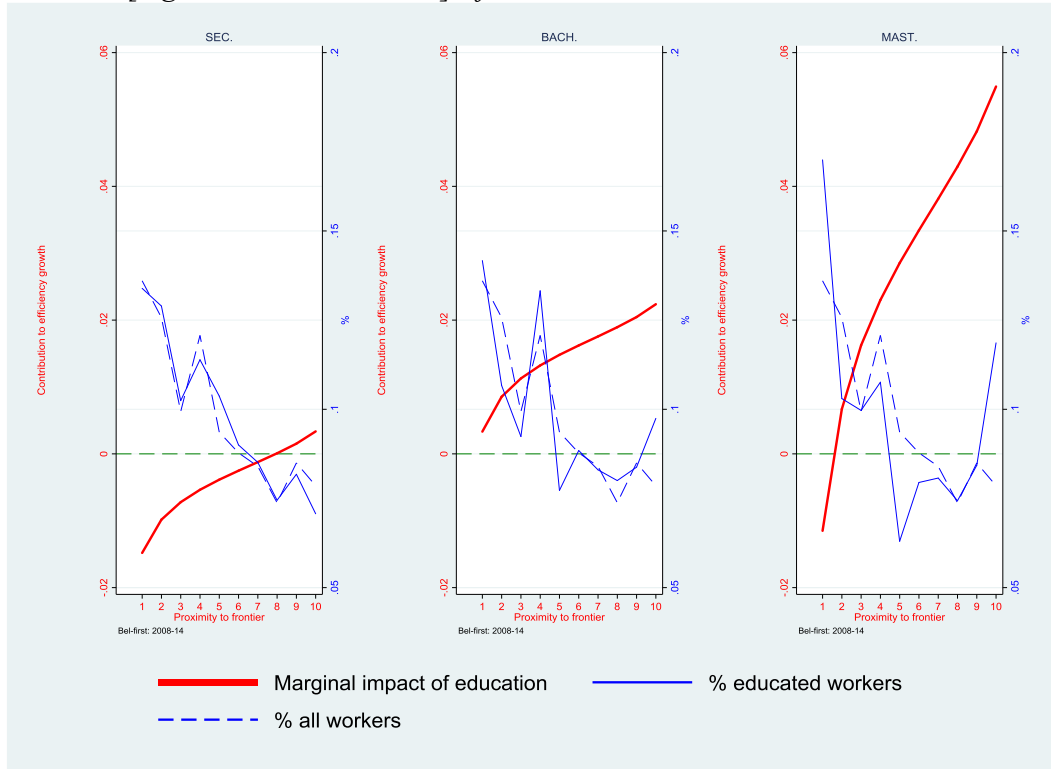
[*] $p < 0.05$, [**] $p < 0.01$, [***] $p < 0.001$

[$\$$]: The essence of this method is to use some function of a firm's consumption of intermediates [here the firm's total purchase of intermediates and services] to control for/proxy time-varying efficiency developments that are (partially) anticipated by the firms and correlated with the observed level of education (see Appendix for a full exposition of the ACF 2-steps estimator).

[μ]: $\widehat{ptf}_{it}$ - $\widehat{ptf}_{it-1}$

[$\pounds$]: Standard errors are computed to account for firm-level clustering (correlation of error terms)

[*]: NAICS 2-digit

Figure 4 – Marginal impact of educated workers[$] on annual efficiency growth [left-hand vertical axis] according to the proximity to frontier[£] decile [horizontal axis] + frequency distribution [right-hand vertical axis] of educated workers vs all workers[$$].

[$] : The left-hand vertical axis measures the impact on annual firm-level efficiency growth of a 100 %-points rise of the share of educated workers. The values are computed using the OLS estimation of eq. [4]- see coefficients reported in Table 4, The closer the firm to the frontier the larger the positive productivity growth and below a certain threshold that contribution is even negative. The right-hand axis (and the two frequency curves) measures the distribution of the workforce according to the distance to the efficiency frontier of the firm they work in.

£:The frontier is industry-specific (ie. NAICS 2-digit) and specified as a Translog. It has been estimated using stochastic frontier methods (SF) exposed in Section 2.

[$$]: The dashed blue line (common to the 3 sub-figures) corresponds to the overall distribution of the workforce along the proximity-to-frontier axis. The solid blue lines inform about the distribution of the different categories of educated workers (secondary-, bachelor's or master's-educated). When the solid line is above the dashed one, it means that the educated workers are overrepresented, and vice-versa.

Estimates obtained using Bel-first firm-level data (2008-2014)

Appendix: the ACF-GMM estimator

Considering a stylised version our Stage 2 model Eq.[4],

$$y_{it} = X_{it}\, \lambda + \varepsilon_{it} \qquad\qquad [8.]$$

with

$$y_{it} \equiv \widehat{ptf}_{it} - \widehat{ptf}_{it\text{-}}$$

$$X_{it} = [\widehat{ptf}_{it\text{-}1},\ SEC_{it\text{-}1};\ BACH_{it\text{-}1};\ MAST_{it\text{-}1};\ SEC_{it\text{-}1}*\widehat{ptf}_{it\text{-}1};\ BACH_{it\text{-}1}*\widehat{ptf}_{it\text{-}}1; AST_{it\text{-}1}*\widehat{ptf}_{it\text{-}1}]$$

$$\lambda \equiv [\alpha, \beta, \gamma_1, \gamma_2, \gamma_3, \eta_1, \eta_2, \eta_{3,}\theta]$$

the logic of the ACF-GMM estimator is to consider that the error term comprise ε_{it} a time-varying unobserved terms (partially known or anticipated by the management of the firm) ω_{it} that is potentially correlated with the other regressors. It might for instance correspond to a (positive/negative) efficiency development influencing their educational mix.

$$\varepsilon_{it} = \omega_{it} + \sigma_{it} \tag{9.}$$

Key with ACF it to assume that firms' (observable) demand for intermediate inputs (int_{it}) is a function of ω_{it} as well as labour inputs (here the shares of workers by educational attainement):

$$int_{it} = f_t(\omega_{it}, X_{it}) \tag{10.}$$

ACF further assume that this function f_t is monotonic in ω_{it} and its other determinants, meaning that it can be inverted to deliver an expression of ω_{it} as a function of int_{it}, X_{it} and introduced into the model

$$y_{it} = X_{it}\lambda + f_t^{-1}(int_{it}, X_{it}) + \sigma_{it} \tag{11.}$$

The ACF algorithm consists of two steps. In step one, outcome variable (here efficiency growth) is regressed on a composite term Φ_{it} that comprises a constant, and 3rd order polynomial expansion in i int_{it}, $X_{it,}$, approximating $f_t^{-1}(int_{it}, X_{it})$

$$y_{it} = \Phi_{it}(int_{it}, X_{it}) + \sigma_{it} \tag{12.}$$

At that point λ is clearly not identified yet. The first-step regression delivers an unbiased estimate of the composite term Φ_{it}^{hat}; i.e outcome net of the purely random term σ_{it}. This is done at step 2. Key is the idea that one can generate implied values for ω_{it} using first-stage estimates Φ_{it}^{hat} and candidate[1] values for the vector of coefficients λ

$$\omega_{it} = \Phi_{it}^{hat} - X_{it}\lambda^{\$} \tag{13.}$$

ACF assume further that the evolution of ω_{it} follows a first-order Markov process
$$\omega_{it} = E[\omega_{it}|\omega_{it-1}] - \xi_{it} \tag{14.}$$

That assumption simply amounts to saying that the realization of ω_{it} depends on some function g(.)

(known by the firm) of *t-1* realisation and an (unknown) innovation term ξ_{it}.

$$\omega_{it} = g(\omega_{it-1}) + \xi_{it} \qquad\qquad [15.]$$

By regressing non-parametrically (implied) ω_{it} on (implied) ω_{it-1}, ω_{it-2}, one gets residuals that correspond to the (implied) ξ_{it} that can form a sample analogue to the orthogonality (or moment) conditions identifying $\lambda \equiv [\alpha, \beta, \gamma_1, \gamma_2, \gamma_3, \eta_1, \eta_2, \eta_3, \theta]$.

$$E[\xi_{it}|X_{it}] = 0 \qquad\qquad [16.]$$

with the possibility, for some of the variables forming X_{it} to resort to *lagged* values, under that assumption that these are less likely to be uncorrelated with the innovation terms ξ_{it}. We followed this recommendation for the variables containing education share, meaning that our moment conditions are

$$E[\xi_{it}|\ SEC_{it-2},\ SEC_{it-3}...] = 0 \qquad\qquad [17.]$$

$$E[\xi_{it}|\ BACH_{it-2},\ BACH_{it-3}...] = 0 \qquad\qquad [18.]$$

$$E[\xi_{it}|\ MAST_{it-2},\ MAST_{it-3}...] = 0 \qquad\qquad [19.]$$

$$E[\xi_{it}|\ SEC_{it-2}\ \widehat{ptf}_{it-2},\ SEC_{it-3}\ \widehat{ptf}_{it-3}\ ...] = 0 \qquad\qquad [20.]$$

$$E[\xi_{it}|\ BACH_{it-2}\ \widehat{ptf}_{it-2},\ BACH_{it-3}\ \widehat{ptf}_{it-3}\ ...] = 0 \qquad\qquad [21.]$$

$$E[\xi_{it}|\ MAST_{it-2}\ \widehat{ptf}_{it-2},\ MAST_{it-3}\ \widehat{ptf}_{it-3}\ ...] = 0 \qquad\qquad [22.]$$

i   A natural (inverse) proxy for distance to frontier is "proximity to frontier".

ii   In the Bertelsman et al. (2015) study, the th quantile return to educated labour corresponds to the marginal change in productivity due to a marginal change in the share of that type of workers conditional on being in a firm belonging to the th quantile of the overall outcome distribution (i.e. the outcome being labour productivity in their case).

iii   The Translog is know for being a flexible functional form for production functions. One of its main advantages is that, unlike the Cobb-Douglas, it does not assume rigid premises such as unitary elasticity of substitution or constant returns to scale. Also Translog functions have been shown to be second-order Taylor approximations of any technology (Fuss, McFadden & Mundlak, 1978; Chambers, 1988).

iv   NAICS 2-digit

v   As to capital, we use the most encompasing estimate available in the dataset (i.e. "total de l'actif"/ total assets). It comprises a firm's current (receivables & stocks) and non-currents assets (both non-financials and financials).

vi   That are necessary to implement the Ackerberg et al. (2015) control/proxy econometric estimation aimed at controling for the risk of reverse causality/simultaneity.

vii   NAICS 2 digits

viii   Belgium counts 11 provinces. These are 50-80 kilometer-wide administrative boundaries. We use province fixed affect as way to capture location-based determinants of efficiency growth (presence of a large city, density of infrastructure…) highlighted by economic geography.

ix   NAICS 2 digits

x   The log of proximity-to-frontier being negative ($-\infty \leq lnPTF_{it} \leq 0$) the positive $\eta's$ mean that we get smaller contribution to annual efficiency gains when distance to frontier rises (proximity to frontier falls).

xi   The role of education as a net contributor to country-level growth is in fact more disputed than its contribution to individuals' fortune (see Sianesi & van Reenen (2003); de la Fuente (2011) for surveys). This is due to the methodological difficulties related to measuring skills and, what is more, modelling and identifying the channels through which skills impact on macroeconomic performance.

xii   The International Standard Classification of Education (ISCED) developed by UNESCO and OECD (for more details see http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf)

xiii   And Figure 7 suggests that larger firms (with more employees) are overrepresented among "frontier" firms (i.e. they form the 6th to 10th decile of the PTF distribution) whereas smaller sizes are predominant among "laggards".