The Contribution of Educated Workers to Firms' Efficiency Gains
The Key Role of the Proximity to Frontier

V. Vandenberghe

Discussion Paper 2017-12

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain

The Contribution of Educated Workers to Firms' Efficiency Gains

The Key Role of the Proximity to 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. 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 from the frontier, the more educated workers 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.

1. Introduction

Nelson & Phelps (1966) were the first to suggest that education facilitates the implementation of new technologies and help countries catch up with the leading/frontier country. Grossman & Helpman (1991) went a step further by arguing that advanced economies, in order to achieve efficiency gains, are particularly dependent on their capacity to increase their stock of skilled labour. Then Acemoglu et al. (2003, 2006) showed theoretically that innovation becomes more important than imitation as an economic entity approaches the technological frontier. A few years later, Vandenbussche et al. (2006), Agrion et al. (2009) formalized the idea underpinning this paper. They suggested that the contribution of skilled vs unskilled (or less-skilled) labour to growth should depend on countries' distance/proximity to the technological frontier. And they predicted that highly-educated labour would be particularly efficiency-growth enhancing close to that frontier, under the reasonable assumption that innovation is a relatively more skill-intensive activity than imitation. Vandenbusshe/Aghion and their co-authors subsequently published OECD (Vandenbenbussche 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. The exercise is based on few key considerations. First, we would argue that efficiency growth is, to a large extent, a firm-level phenomenon, and its determinants (education, proximity-to-frontier; the interaction between these two, or any other determinants) should be primarily assessed at that level. In modern economies, where most people work inside firms, educational attainment-related efficiency gains cannot possibly exist at a more aggregate level if they do not show up at firm level. And the same argument can be used about their determinants. Second, many datasets now contain good-quality information about firms' stock of capital, total labour and its breakdown by educational attainment. In addition, many datasets are structured as panels, so they can be used to compute efficiency growth and explore its determinants. Third, a first wave of studies, exploiting these micro data have documented, virtually without exception, enormous and persistent measured (but unexplained) efficiency performance differences across firms, even within narrowly defined industries (Syverson, 2011). When exploiting firm-level data we can thus count on a large variation of the proximity-to-frontier. This is a plus from an econometric point of view. Four, and beyond the econometric considerations, the heterogeneity of firms probably points a division of roles between what Andrews et al. (2015) call "frontier" vs "laggard" firms; with the former probably playing an ever greater role in explaining economy-wide developments.

The results of this paper are essentially twofold. First, firm-level panel data strongly support the idea that educated workers contribute more to efficiency gains when they work inside firms situated closer to the efficiency frontier. Second, that complementarity is significantly stronger for master-educated workers than their bachelor- or upper-secondary-educated peers. These results accord with idea that innovation (i.e. which is presumably how "frontier" firms become more productive) is a relatively more skill-intensive activity 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. A Stage 2 these estimates, in combination with data on firms' workforce educational attainment, are used to assess education's contribution to efficiency growth. A key assumption is that contribution depends on the distance/proximity to the frontier. This framework extends and improves the one developed by Vandenbenbussche 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 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). 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. n 2015). That paper adopts a quantile regression strategy to capture the potentially heterogeneous contribution of education to productivity[1], 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.

While Vandenbenbussche et al. (2006) use off-the-shelf estimates of the USA's TFP level as the efficiencey frontier – and each country's ratio vis-à-vis the USA as an estimate of the distance/proximity to fronrier – we resort to econometric estimates. In what follows, both the frontier and the distance/proxity to it, rest on different econometric methods and specifications of the technology. The first method consist of retrieving via OLS the standard Solow residuals; whereas the second method is based on the stochastic frontier approach (SF hereafter). In both cases, we assume that the production technology is either Cobb-Douglas or Translog (i.e. the quadratic generalisation of the Cobb-Douglas).

i) Estimating firm-level proximity to frontier using OLS-estimated Solow residuals

We first estimate firm-level TFPs as Solow residuals. Considering a Cobb-Douglas technology and allowing for year fixed effects (assuming that that the technology may improve over time and translate into [presumabley upward] shifts of the fontier), we get

$$lnTFP_{it}=lnY_{it} - \hat{\alpha}\ lnL_{it} - \hat{\beta} lnK_{it} - \widehat{\Phi}_t \qquad [1.]$$

These OLS-estimated values of the (log of) TFP – and also those obtained with a Translog specification of the technology – are then used to compute the proximity to frontier (PTF). The key idea is that the firms with the highest (positive) residuals/TFP's form the efficiency frontier and the those with smaller residuals are less effective. More specifically, we define the (log of) PTF as

$$lnPTF_{it}= min\{0, lnTFP_{it} -ln\ \widetilde{TFP}_t\ \} \qquad [2.]$$

with $-\infty \leq lnPTF_{it} \leq 0$; and $0 \leq PTF_{it} \leq 1$

where ln $\widetilde{TFP}$ is the 99th percentile of the overall distribution of TFP/Solow residuals derived from the OLS estimation of [1].

---

1 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).

ii) Estimating firm-level proximity to frontier using stochastic frontier methods

The second method used to estimate firm-level proximity to frontier (PTF) stems from the stochastic production frontier (SF) literature. Aigner, Lovell & Schmidt (1977) and Meeusen & van den Broeck (1977) pioneered the idea of SF models. In a world without error or inefficiencies, the firm i in period t would produce $f(K_i, L_i)$. Stochastic frontier analysis assumes that each firm potentially produces less than it might due to a degree of inefficiency. Specifically, actual output writes $Y_{it}=f(K_{it},L_{it})\xi_{it}$; where $0\leq\xi_{it} \leq 1$ represents the relative efficiency of firm i in t. If $\xi_{it}=1$, the firm is achieving the optimal output with the technology embodied in the production function $f(K_{it},L_{it})$. Its proximity to the productivity frontier is maximal. Andrew et al. (2015) would call it a "frontier" firm. When $\xi_{it} <1$, the firm is not making the most of the inputs K,L and proximity to frontier falls. Andrew et al. (2015) would refer to such a firm as being a "laggard". The other key idea underpinning SF is that production is subject to random shocks v_{it} ; and these shocks should not be confounded with efficiency/proximity to frontier ξ_{it}. Note also that, paralleling the choice made with OLS-estimated Solow residuals [1], we also allow for year fixed effects (i.e. a year-to-year shifts of the frontier due to the technical change). The estimated model is

$$Y_{it} = f(K_{it},L_{it})\xi_{it}e^{\Phi_t v_{it}} \quad [3.]$$

Taking the natural log of both sides of [3] leads to

$$lnY_{it} = lnf(K_{it}, L_{it})+ \Phi_t+ ln\xi_{it}+v_{it} \quad [4.]$$

where the (log of) the proximity to frontier is $lnPTF_{it}\equiv ln(\hat{\xi}_{it})$ with $-\infty\leq lnPTF_{it} \leq 0$ and $0\leq PTF_{it}\leq 1$ as $0< \hat{\xi}_{it} \leq 1$

The STATA frontier command that we used is capable of delivering estimates of the second component of the residual i.e. $ln\xi_{it}\equiv lnPTF_{it}$; and that for any log-linear specification of $f(K_{it},L_{it})$. Like with the Solow residual method, we retain the Cobb-Douglas and the Translog. Note finally that the crucial two-component residual at the heart of a SF model; v_{it}, is assumed to be randomly $N(0,\sigma v)$ distributed over the observations, while the PTF_{it} are assumed independently half-normally $N^+(0,\sigma PTF)$ distributed.

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 first model we estimate generalises the one used by Vandenbussche et al. (2006) and Aghion et al. (2009)

$$g_{it} \equiv ln\widehat{PTF}_{it} - ln\widehat{PTF}_{it-1} = \alpha + \beta 1 n\widehat{PTF}_{it-1} + \gamma_{second} S^{second}_{it-1} + \gamma_{bach} S^{bach}_{it-1}. + \gamma_{mast} S^{mast}_{it-1} + \eta_{second}\, S^{second}_{it-1} * ln\widehat{PTF}_{it-1} + \eta_{bach}\, S^{bach}_{it-1} * ln\widehat{PTF}_{it-1} + \eta_{mast}\, S^{mast}_{it-1} * ln\widehat{PTF}_{it-1} + \theta F_{it} + \varepsilon_{it} \quad [5.]$$

The efficiency growth achieved by firm i between period t and $t-1$ is the change in the proximity to frontier i.e. $g_{it} \equiv ln\widehat{PTF}_{it} - ln\widehat{PTF}_{it-1}$. The initial share of the workers with more than primary education is S^{j}_{it-1}.; with j= *secondary, bachelor, master* and least educated workers being the reference (more on the definition of education below). The (initial) proximity to the frontier is $ln\widehat{PTF}_{it-1}$. The main variables of interest are the S^{j}_{it-1} interacted with the (log of) proximity to frontier $ln\widehat{PTF}_{it-1}$. The sign of their coefficients η provide a direct test of the Vandenbussche/Aghion et al. assumption. Positive and statistically significant η*'s* will be interpreted as evidence that better-educated labour has a larger efficiency growth-enhancing effect closer to the technological frontier. Note that $ln\widehat{PTF}_{it-1}$ is a negative number (i.e. log of a ratio <1) with a maximum of 0. That means that coefficients γ_j 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 [5], one can thus assess whether the efficiency growth enhancing effect closer to the efficiency frontier is larger for, say, master-educated workers compared to bachelor -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[2]/year fixed effects (that among other things control for output price inflation differences),

[2] NAICS 2-digit

plus the annual change of the share of workers by educational attainment ($\Delta S^{second.}{}_{it};\Delta S^{bach.}{}_{it};\Delta S^{mast.}{}_{it}$). 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 *t-1* and not *t* educational attainment as a predictor of *t-1* to *t* efficiency growth. 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 econometric terms, this means simultaneity bias or reverse causality. In more economic ones, it corresponds the situation where efficiency gains pushes firms to recruit more educated workers. To limit that risk, the strategy is to prioritise *t-1* (or further lagged) values of educational attainment as (causal) explanation of firm-level efficiency growth.

We also extend on Vandenbussche/Aghion's work by estimating a much more flexible version of equation [5]

$$g_{it} \equiv ln\widehat{PTF}_{it} - ln\widehat{PTF}_{it-1} = \Phi(1n\widehat{PTF}_{it-1}, S^{second}{}_{it-1}, S^{bach}{}_{it-1}, S^{master}{}_{it-1}) \qquad [6.]$$

Here the key idea is to specify Φ as a 3rd-order polynomial expansion in its components; including all the terms combining PTF and educational shares. That model remains linear, and can be estimated by OLS. But unlike [5], none of its coefficients has straightforward economic interpretation. This difficulty can be overcome by resorting to post-estimation STATA margins command using the dydx($S^j{}_{it-1}$) over($qln\widehat{PTF}_{it-1}$) options; where $qln\widehat{PTF}_{it-1}$ is the quantile (here decile) of the overall proximity-to-frontier distribution. STATA margins dydx($S^j{}_{it-1}$) retains the estimated coefficients and the variables corresponding to the first derivative of the 3rd-order polynomial expansion vis-à-vis education share $S^j{}_{it-1}$. Adding over($qln\widehat{PTF}_{it-1}$) means it then computes the average fitted marginal effect of $S^j{}_{it-1}$ using the observed values of the explanatory variables; and that separately for each quantile (decile) of the proximity to frontier.

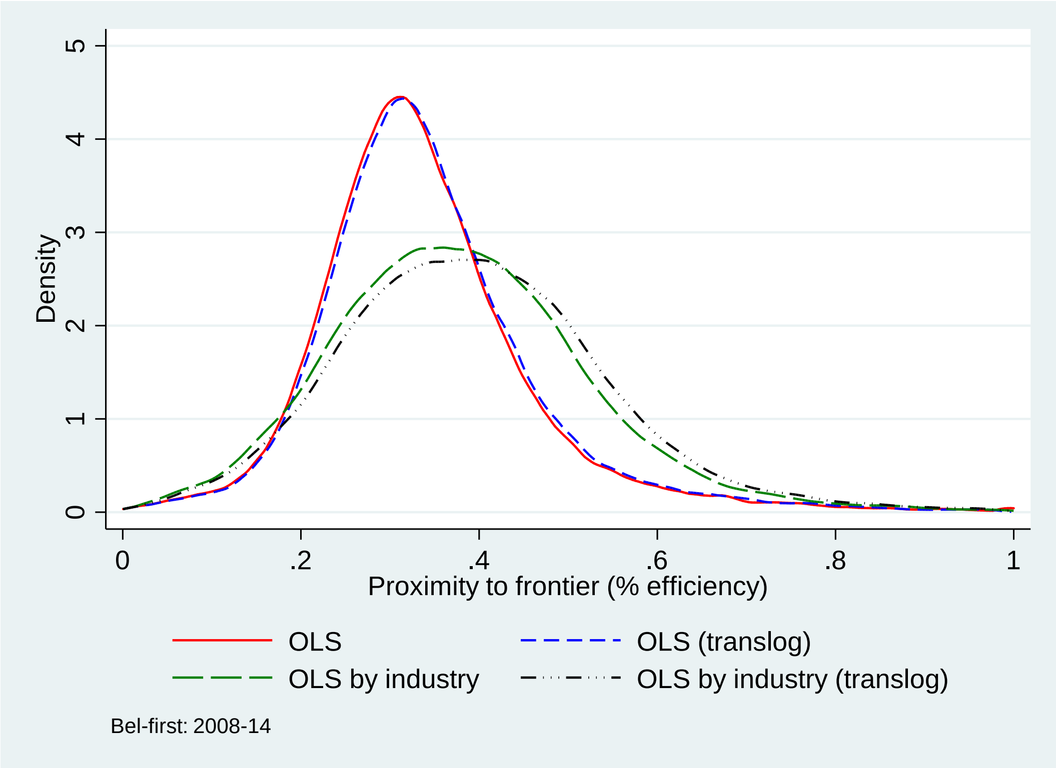
3. Stage-1 results and data description

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 50,525 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, and total numbers of hours worked. Note that Bel-first contains the breakdown of the total workforce 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 degrees hereafter), and [4] 4-5-year university attainment (masters hereafter).

Figure 1 reports Stage 1-estimated PTFs computed with OLS/Solow-residuals. On display are the distributions obtained when using both Cobb-Douglas and Translog specifications. The two distributions are very similar. Figure 1 also displays the distribution of PTF when estimation is carried out industry by industry.[3] Quite logically, the latter distribution shifts to the right, as many firms tend to be closer to the "local" frontier (the one characterizing their industry) than to the "global" frontier. Figure 3 compares PFT estimates obtained with OLS vs SF. The main difference rests is the magnitude of the estimated distance/proximity to frontier. The SF-estimated PTF is on average, significantly smaller, essentially because, by construction, that method allows for a random component which comes in deduction to the overall residual. However, Figure 3 shows that firm-level estimates based on these two methods are highly correlated, meaning that the position of a particular firm in the overall efficiency distribution is not fundamentally dependant on the method used to estimate frontier and distance/proximity to frontier.

[3] NAICS 2-digit

Figure 1 – OLS-Solow-residuals approach to estimation of proximity to frontier. Density distribution. Cobb-Douglas vs Translog specification. Polled data or analysis by industry



Source: Belfirst 2008-2014

Figure 2 – OLS- vs Stochastic Frontier(SF)-estimated proximity to frontier. Comparison of density distributions. Year 2014

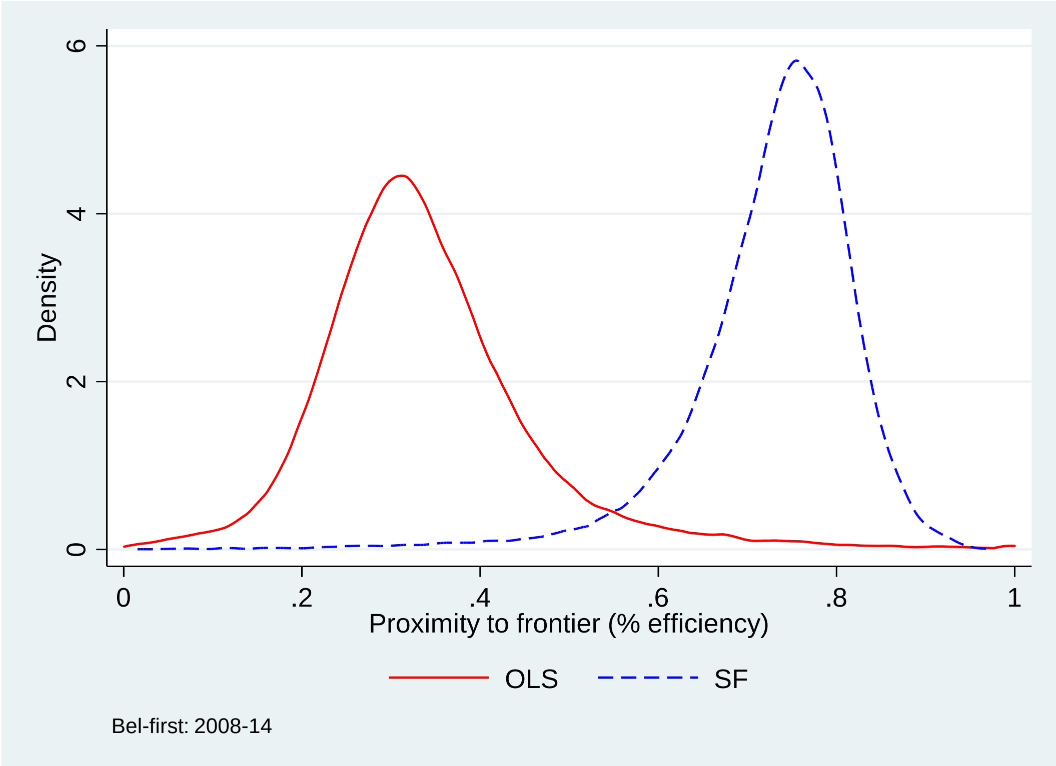
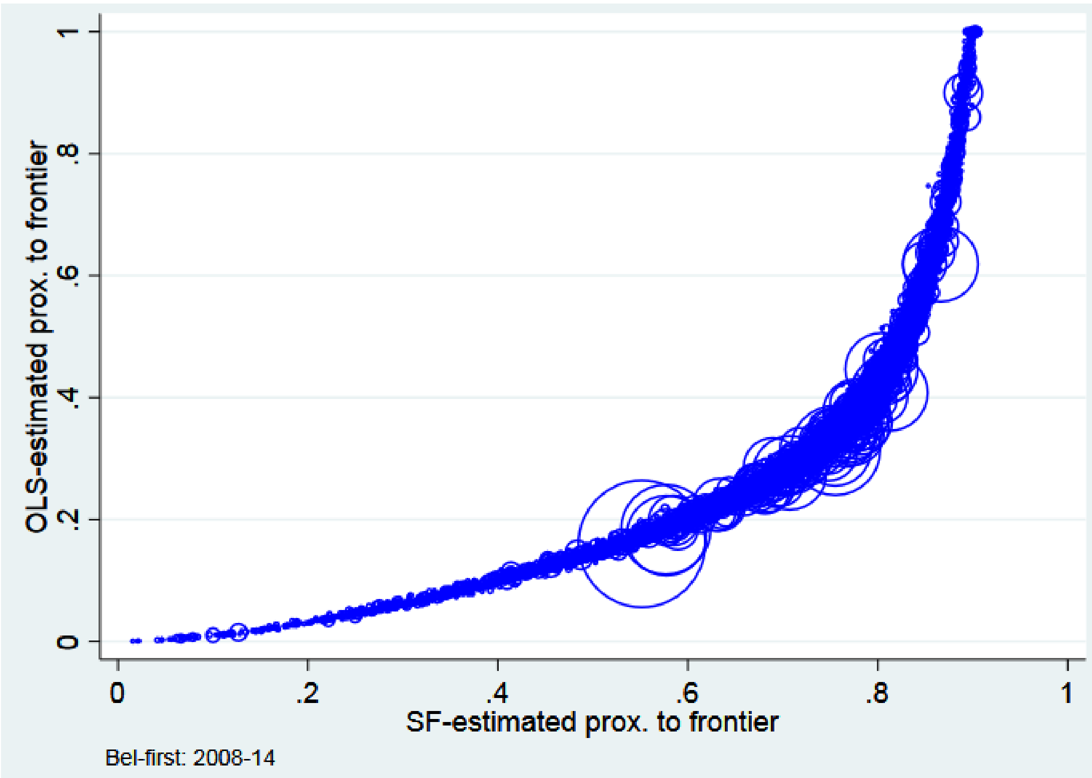


Figure 3- OLS- vs SF-estimated proximity to frontier- Scatter plot. Year 2014



Circle size represents the number of employee in the firm
Source: Belfirst 2008-2014

Tables 1 to 3 describe the key variables used at stage 2, where we estimate equations [5] & [6]. 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 1 reports the annual growth rates whereas Table 3 the 6-year equivalent. Table 2 describes the main stage 2 explanatory variables. It shows that secondary-educated workers represent about 59% of the total workforce, while bachelor-educated and master-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 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 tha we just keep one period when we use the 6-year growth rate as dependent variable (Table 3).

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

	Cobb-Douglas		Translog		Cobb-Douglas by Industry*	
	OLS[$]	SF[$$]	OLS	SF	OLS	SF
2009	-0.0021	0.0002	-0.0013	0.0009	-0.0013	0.0001
2010	-0.0006	0.0020	0.0006	0.0032	-0.0002	0.0023
2011	0.0023	0.0010	0.0025	0.0013	0.0022	0.0012
2012	-0.0003	-0.0007	-0.0005	-0.0006	-0.0006	-0.0008
2013	-0.0010	0.0003	-0.0014	0.0004	-0.0011	-0.0000
2014	-0.0103	-0.0029	-0.0104	-0.0035	-0.0088	-0.0027
Average	**-0.0019**	**-0.0000**	**-0.0017**	**0.0003**	**-0.0016**	**0.0000**
N	242,025					

[$]: annual growth rates of OLS-/Solow-residual based estimates of proximity to frontier. Assuming either a Cobb-Douglas or a Translog specification, we first estimate firm-level Solow residuals. These are then used to compute the proximity to frontier as $lnPTF=min\{0, lnTFP_{it} -ln\ \widetilde{TFP}_t\}$ where $ln\ \widetilde{TFP}_t$ is the 99th percentile of the overall distribution of TFP.
[$$]: annual growth rates of Stochastic-frontier (SF) based estimates of proximity to frontier. SF assumes that each firm potentially produces less than it might due to a degree of inefficiency. Specifically, actual output writes $Y_{it}=f(K_{it},L_{it})\xi_{it}$; where $0\leq\xi_{it}\leq1$ represents the proximity to frontier and $F()$ can be specified as Cobb-Douglas or Translog. If $\xi_{it}=1$, the firm is achieving the optimal output with the technology embodied in the production function $f(K_{it},L_{it})$, and proximity to the productivity frontier is maximal. When $\xi_{it}<1$, the firm is not making the most of the inputs K,L and proximity to frontier falls.
*:NAICS 2-digit
Source: Belfirst 2008-2014

Table 2: Descriptive statistics- main regressors

	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 educated workers	Master 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 3: Descriptive statistics- proximity to frontier growth (6-year)

	Cobb-Douglas		Translog		Cobb-Douglas by Industry*	
	OLS	SF	OLS	SF	OLS	SF
2008-2014	0.0029	0.0053	0.0035	0.0079	0.0045	0.0057
N	39,901					

Source: Belfirst 2008-2014
*:NAICS 2-digit

4. Stage-2 econometric results

Tables 4 and Figure 5 present a first series of stage-2 results, corresponding to the estimation of equations [5]. They are divided in three parts corresponding to the Cobb-Douglas and Translog specifications of the efficiency frontier; and also the Cobb-Douglas specification estimated industry by industry.[4] In all cases, the key coefficient corresponds to the educationX*PTF* interaction variable (η). 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 that this is always the case. And it tends to be true for all types of educated workers, but with strong evidence [e.g. col. 2 of Table 4] that complementarity with PTF is significantly stronger for master-educated workers (η^{mast}=0.142) compared to bachelor-educated (η^{bach}=0.045) or secondary-educated workers (η^{second}=0.035). At the bottom of Table 4, we report the results of hypothesis test that η*'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, OLS vs SF, with all industries pooled or not).

Remember that, in Table 4, reported coefficients γ_j for the education attainment (S^j_{it-1}) inform about the latter's contribution to productivity growth at the frontier.[5] A .232 value thus suggests that a 0 to 100% rise of the share of master-educated workers in those firms leads to a 23.2%-point rise of the annual efficiency growth rate. More realistically, a 10%-point rise of the share of master-educated workers inside frontier firms adds 2.32 %-points to that rate.

4 NAICS 2-digits

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

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.

Figures 5 and 6 display the results stemming for the estimation of equation [6]: a generalized and very flexible version of the Vandenbussche/Aghion model. The plotted lines inform about the (average) predicted marginal contribution of the different types of educated workers to annual efficiency gain. Note that the expected marginal contribution combines *i)* the direct contribution of education (i.e. the equivalent of the γ's in eq. [5]) and *ii)* the one that depends on PTF (paralleling the η's in eq. [5]). Considering the highest/10th decile of the PTF distribution, it is clear that master-educated workers contribute more at the margin than bachelor- or secondary educated workers. What is key however is the propensity of these marginal gains to go down when turning to the lower deciles of the PTF distribution. On both Figure 5 and Figure 6, we verify that the relationship is negatively sloped for master-educated and bachelor-educated workers; but much less so for secondary-educated workers. This accords with the idea of complementarity between advanced forms of education (i.e. master and bachelor) and proximity to frontier when it comes to fostering efficiency gains. Note that both Figures 5 and 6 suggest that education [whatever the type] plays no statistically significant role in raising the efficiency performance of firms forming the 1st decile of the PTF distribution.

Finally, we check 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 5 and are qualitatively very similar to those on display in Table 4. Second, and more significantly, we use Instrumental Variable estimate (IV) in order to assess the risk of endogeneity/simultaneity discussed in Section 2.2. To implement IV, we use as instruments the three educational shares present in [5] by their lagged 2 to 4 values. Results are reported in Table 6, and again, are qualitatively very similar to those reported in Table 4.

Table 4 - The contribution to 1-year efficiency growth of educated workers, according to the proximity to frontier. Log linear specification [5]

	Cobb-Douglas		Translog		Cobb-Douglas frontier by industry [$]	
	OLS	SF	OLS	SF	OLS	SF
$1n\widehat{PTF}_{it-1}$ [β]	-0.378***	-0.451***	-0.380***	-0.454***	-0.365***	-0.451***
	(0.010)	(0.017)	(0.011)	(0.017)	(0.008)	(0.017)
$S^{second.}_{it-1}$ [γ^{second}]	0.036**	0.010	0.036*	0.012	0.018*	0.007
	(0.014)	(0.006)	(0.014)	(0.007)	(0.009)	(0.005)
$S^{bach.}_{it-1}$ [γ^{bach}]	0.102***	0.025**	0.102***	0.039***	0.072***	0.012
	(0.021)	(0.010)	(0.021)	(0.010)	(0.017)	(0.009)
S^{mast}_{it-1} [γ^{mast}]	0.232***	0.087***	0.217***	0.101***	0.189***	0.079***
	(0.026)	(0.011)	(0.026)	(0.011)	(0.024)	(0.012)
$1n\widehat{PTF}_{it-1}X\ S^{second.}_{it-1}$ [η^{second}]	0.035**	0.038	0.036**	0.043*	0.022*	0.032
	(0.012)	(0.020)	(0.013)	(0.020)	(0.009)	(0.019)
$1n\widehat{PTF}_{it-1}X\ S^{bach.}_{it-1}$ [η^{bach}]	0.045*	0.025	0.045*	0.058	0.025	-0.002
	(0.020)	(0.032)	(0.021)	(0.032)	(0.016)	(0.031)
$1n\widehat{PTF}_{it-1}X\ S^{mast.}_{it-1}$ [η^{mast}]	0.142***	0.219***	0.128***	0.240***	0.083***	0.162***
	(0.026)	(0.038)	(0.027)	(0.036)	(0.022)	(0.039)
Controls	yearXindustry(NAICS 2-digit), province, year of incorporation, change of educational mix ($\Delta S^{second.}_{it}$; $\Delta S^{bach.}_{it}$; $\Delta S^{mast.}_{it}$)					
Nobs	242,025					
R^2	0.1858	0.1845	0.1882	0.1911	0.2214	0.2254
prob $\eta^{mast}=\eta^{bach}$	0.0033	0.0298	0.0138	0.0571	0.0001	0.0006
prob $\eta^{mast}=\eta^{second}$	0.0000	0.0038	0.0003	0.0055	0.0000	0.0004
prob $\eta^{bach}=\eta^{second}$	0.5961	0.8469	0.6490	0.5800	0.6733	0.2303

Estimates obtained using Bel-first firm-level data (2008-2014)
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
$: NAICS 2-digit

Table 5 - The contribution to 6-year efficiency growth of educated workers, according to the proximity to frontier. Log linear specification [5]

	Cobb-Douglas frontier		Translog frontier		Cobb-Douglas frontier by Industry $^{\$}$	
	OLS	SF	OLS	SF	OLS	SF
$1n\widehat{PTF}_{it-1}$ $[\beta]$	-0.666***	-0.719***	-0.668***	-0.715***	-0.647***	-0.697***
	(0.0176)	(0.0239)	(0.0179)	(0.0233)	(0.0153)	(0.0267)
$S^{second.}_{it-1}$ $[\gamma^{second}]$	0.020	0.004	0.019	0.002	-0.003	-0.005
	(0.0243)	(0.0092)	(0.0244)	(0.0098)	(0.0173)	(0.0086)
$S^{bach.}_{it-1}$ $[\gamma^{bach}]$	0.134***	0.034*	0.140***	0.050**	0.082**	0.006
	(0.0373)	(0.0145)	(0.0368)	(0.0157)	(0.0312)	(0.0148)
S^{mast}_{it-1} $[\gamma^{mast}]$	0.353***	0.113***	0.332***	0.126***	0.313***	0.103***
	(0.0547)	(0.0212)	(0.0537)	(0.0225)	(0.0518)	(0.0222)
$1n\widehat{PTF}_{it-1} X\, S^{second.}_{it-1}$ $[\eta^{second}]$	0.025	0.025	0.025	0.019	0.004	-0.005
	(0.0211)	(0.0287)	(0.0215)	(0.0282)	(0.0174)	(0.0305)
$1n\widehat{PTF}_{it-1} X\, S^{bach.}_{it-1}$ $[\eta^{bach}]$	0.035	0.023	0.037	0.056	-0.002	-0.042
	(0.0347)	(0.0474)	(0.0346)	(0.0470)	(0.0279)	(0.0480)
$1n\widehat{PTF}_{it-1} X\, S^{mast.}_{it-1}$ $[\eta^{mast}]$	0.155**	0.208**	0.130*	0.209**	0.088*	0.132*
	(0.0541)	(0.0714)	(0.0533)	(0.0681)	(0.0458)	(0.0695)
Controls	Industry(NAICS 2-digit), province, year of incorporation, change of educational mix ($\Delta S^{second.}_{it}$; $\Delta S^{bach.}_{it}$; $\Delta S^{mast.}_{it}$)					
Nobs	39,265					
R^2	0.40	0.44	0.41	0.44	0.40	0.44
prob $\eta^{mast} = \eta^{bach}$	0.0544	0.0304	0.1309	0.0619	0.0837	0.0317
prob $\eta^{mast} = \eta^{second}$	0.0136	0.0086	0.0415	0.0041	0.0603	0.0402
prop $\eta^{ach} = \eta^{second}$	0.7788	0.9734	0.6991	0.4086	0.8226	0.3903

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

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

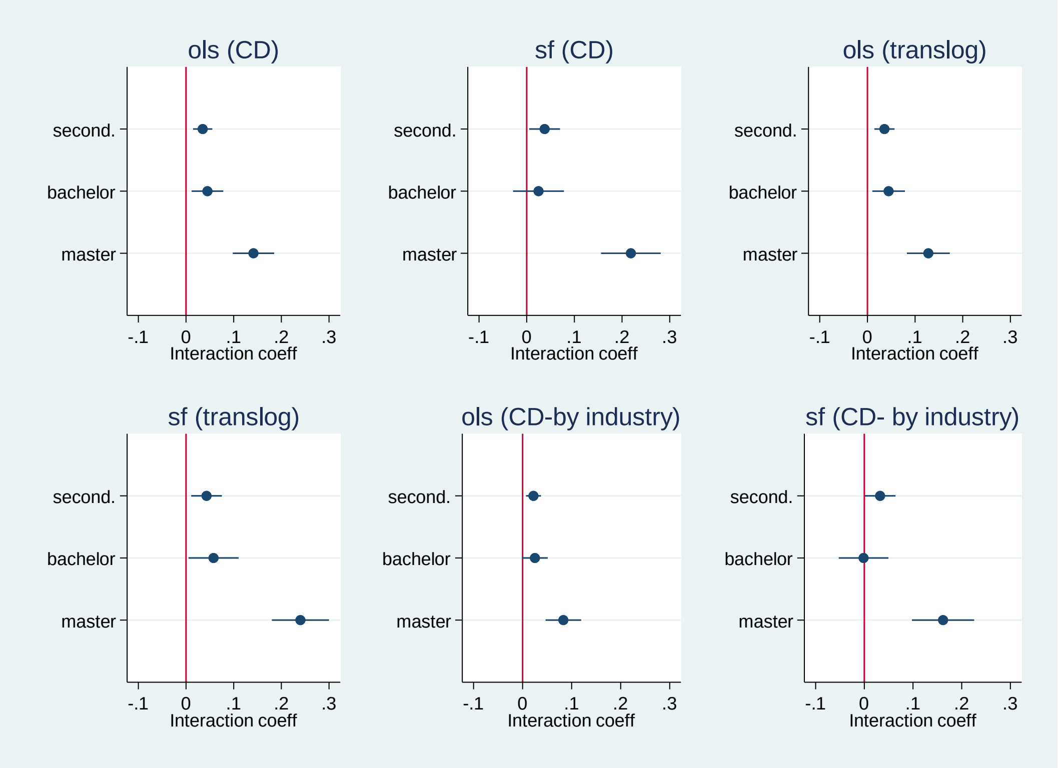
$: NAICS 2-digit

Table 6- IV regression. The contribution to 1-year efficiency growth of educated workers, according to the proximity to frontier. Log linear specification [5] where $S^{j.}_{it-1}$ are instrumented by $S^{j.}_{it-2}; S^{j.}_{it-3}$; $S^{j.}_{it-4}$

	Cobb-Douglas frontier		Translog frontier		Cobb-Douglas frontier by Industry $^{\$}$	
	OLS	SF	OLS	SF	OLS	SF
$1n\widehat{PTF}_{it-1}$ [β]	-0.284***	-0.317***	-0.282***	-0.322***	-0.263***	-0.310***
	(0.0069)	(0.0079)	(0.0069)	(0.0078)	(0.0058)	(0.0076)
$S^{second.}_{it-1}$ [γ^{second}]	0.005	-0.006	0.002	-0.006	-0.008	-0.007*
	(0.0101)	(0.0033)	(0.0099)	(0.0036)	(0.0073)	(0.0029)
$S^{bach.}_{it-1}$ [γ^{bach}]	0.138***	0.038***	0.137***	0.053***	0.091***	0.025***
	(0.0131)	(0.0044)	(0.0130)	(0.0048)	(0.0114)	(0.0043)
S^{mast}_{it-1} [γ^{mast}]	0.212***	0.073***	0.193***	0.081***	0.173***	0.062***
	(0.0158)	(0.0053)	(0.0158)	(0.0059)	(0.0152)	(0.0057)
$1n\widehat{PTF}_{it-1} X\, S^{second.}_{it-1}$ [η^{second}]	0.010	-0.011	0.007	-0.008	-0.003	-0.015
	(0.0084)	(0.0096)	(0.0084)	(0.0095)	(0.0069)	(0.0092)
$1n\widehat{PTF}_{it-1} X\, S^{bach.}_{it-1}$ [η^{bach}]	0.088***	0.080***	0.087***	0.114***	0.047***	0.046***
	(0.0113)	(0.0124)	(0.0113)	(0.0122)	(0.0094)	(0.0121)
$1n\widehat{PTF}_{it-1} X\, S^{mast.}_{it-1}$ [η^{mast}]	0.140***	0.185***	0.120***	0.192***	0.085***	0.125***
	(0.0137)	(0.0141)	(0.0139)	(0.0135)	(0.0119)	(0.0145)
Controls	yearXindustry(NAICS 2-digit), province, year of incorporation, change of educational mix ($\Delta S^{second.}_{it}$; $\Delta S^{bach.}_{it}$; $\Delta S^{mast.}_{it}$)					
Nobs	101,410					
R^2	0.11	0.12	0.12	0.12	0.11	0.12
prob $\eta^{mast} = \eta^{bach}$	0.0028	0.0000	0.0610	0.0000	0.0116	0.0000
prob $\eta^{mast} = \eta^{second}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
prob $\eta^{bach} = \eta^{second}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

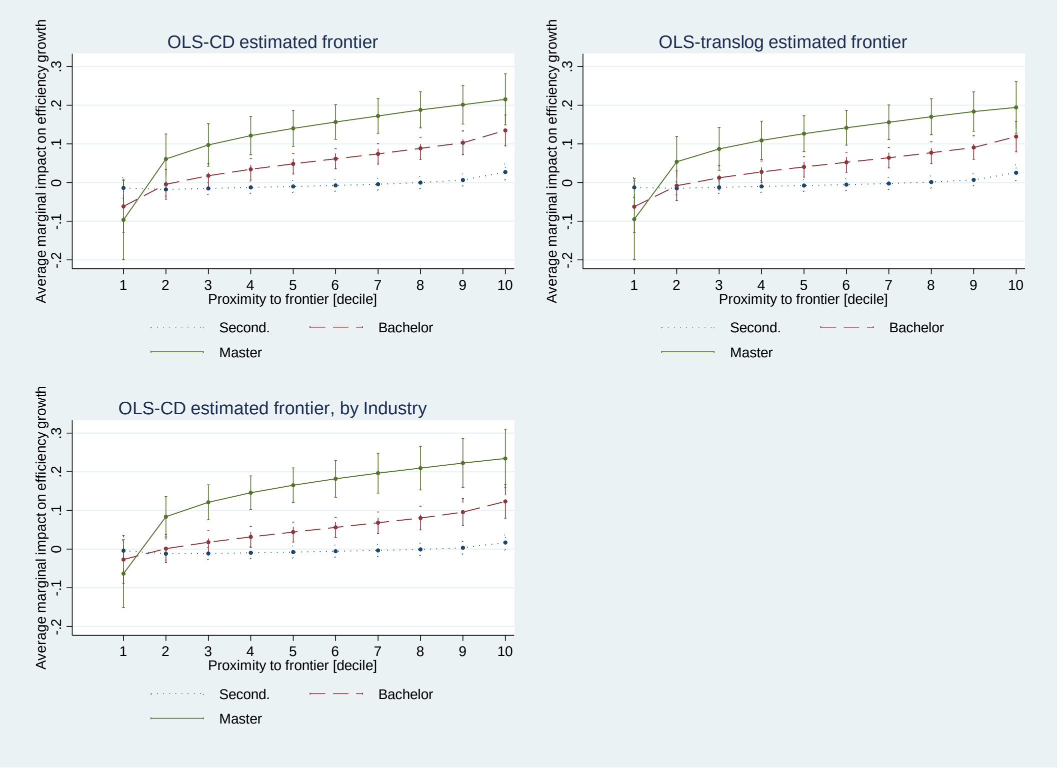
Estimates obtained using Bel-first firm-level data (2008-2014)
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
$: NAICS 2-digit

Figure 4 – The contribution to 1-year efficiency growth of highly-educated workers, according to the proximity to frontier – plot of the interaction education- proximity-to-frontier coefficients of Eq. [5]



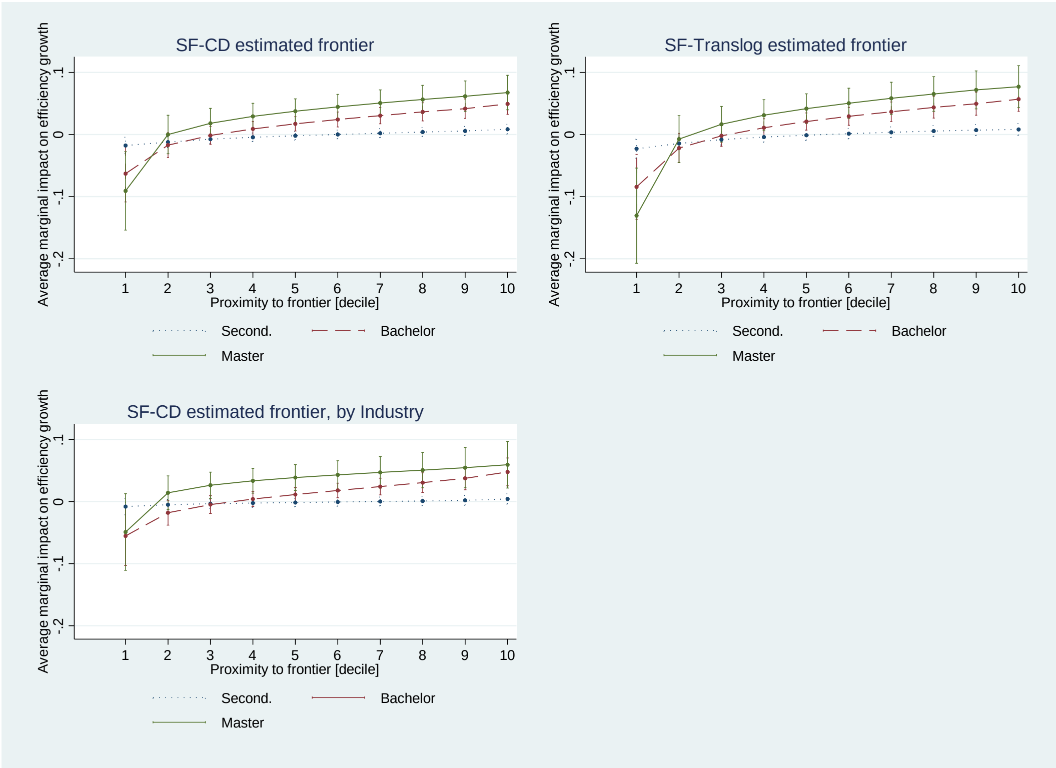
Estimates obtained using Bel-first firm-level data (2008-2014)

Figure 5 - Marginal contribution to 1-year OLS-estimated efficiency growth of educated workers, according to the proximity to frontier. 3rd order polynomial specification [6]. Estimated using **`STATA margins dydx`**



Estimates obtained using Bel-first firm-level data (2008-2014)
`STATA margins dydx`(S^j_{it-1}) retains the estimated coefficients and the variables corresponding to the first derivative of the 3rd-order polynomial expansion vis-à-vis education share S^j_{it-1}. Adding **`over`**(*qln(PTFit-1)*`) means it then computes the average fitted marginal effect of S^j_{it-1} using the observed values of the explanatory variables, separately for each quantile (decile) of the proximity to frontier.

Figure 6 - Marginal contribution to 1-year SF-estimated efficiency growth of educated workers, according to the proximity to frontier. 3rd order polynomial specification [6]. Estimated using **`STATA margins dydx`**



Estimates obtained using Bel-first firm-level data (2008-2014)
`STATA margins dydx`(S^j_{it-1}) retains the estimated coefficients and the variables corresponding to the first derivative of the 3rd-order polynomial expansion vis-à-vis education share S^j_{it-1}. Adding **`over`**(*qln(PTFit-1)*) means it then computes the average fitted marginal effect of S^j_{it-1} using the observed values of the explanatory variables, separately for each quantile (decile) of the proximity to frontier.

5. 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.[6] 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.

6 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.

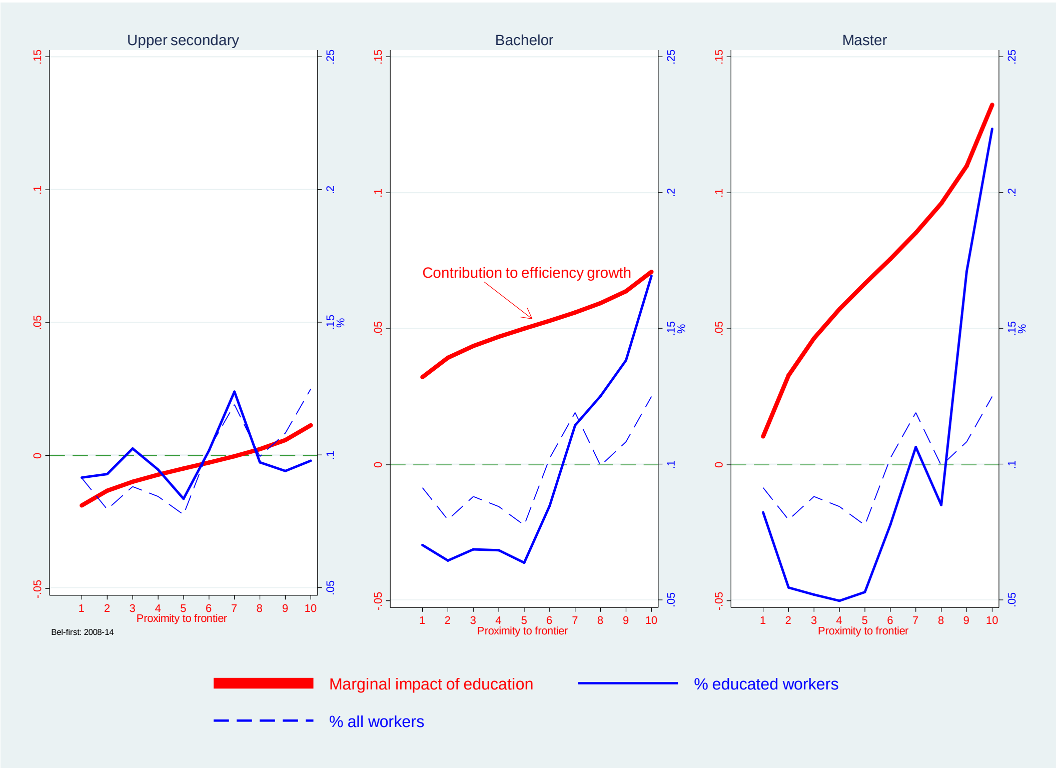
(1999) for France, Galindo-Rueda & Haskel (2005) for the UK, Haltiwanger et al. (2007) for the United States, and Van Biesebroeck (2011) for Zimbabwe all find positive effects of workers' skills on firm/plant productivity. But this paper tries to go way beyond simply showing that education matters for 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 techological 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 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 main result of the paper (that can be visualised on Figure 7, left-hand axis) 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 (on Figure 7, thick solid lines are negatively slopped when going from right to left). We are also able to show the complementarity between proximity-to-frontier (PTF) and education is stronger when the latter corresponds to master-educated workers compared to bachelor- 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, frontier firms insided advanced economies like Belgium, depend on advanced forms of education to achieve efficiency gains; presumably because these gains require more than imitation or replication.

Note finally that our results raise the question of the optimal allocation of educated workers across firms and industries. Should policies aimed at boosting efficiency growth foster the mobility of more educated workers towards "frontier" entities? It there evidence that these individuals do not spontaneously concentrate in these? This evidently calls for future research investigating labour (re)allocation mechanisms. But Figure 7 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.[7] The solid thin lines inform about the distribution of each category of educated workers (secondary-, bachelor- or master-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- and to a lesser extent bachelor-educated workers are clearly overrepresented among "frontier" firms; which is precisedly where they contribute the most to efficiency gaims. Still, many of them work in "laggard" firms.

Figure 7 – Marginal impact of educated workers[$] on annual efficiency growth [left-hand vertical axis] according to the proximity to frontier decile [horizontal axis].[£] Plus distribution [right-hand vertical axis] of educated workers vs all workers[$$]



[$] : of a 1 unit (i.e. 100%-points) rise of the share of educated workers.
£: Estimates derived from the estimation of eq. [5]- see coefficients reported in Table 4, OLS-Cobb-Douglas specification.
[$$] : 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 or master-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)

---

7 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".

6. References

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

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.

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. & 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.

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.

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.

Grossman, G. and H. Elhanan (1991), *Innovation and Growth in the Global Economy*, Cambridge MA, MIT Press.

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

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.

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.

Institut de Recherches Économiques et Sociales
Université catholique de Louvain

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

ISSN 1379-244X D/2017/3082/12