Assessing human capital theory using firm-level evidence. The Missing Link

V. Vandenberghe[**]

Abstract

Most economists with an interest in human capital have neglected the level of the firm to study the education-productivity-wage nexus. This is suprising, because, in contrast to individual-level data used by Mincer-inspired studies, firm-level data offer a direct measure of productivity. Therefore the impact human capital/education on productivity no longer needs to be inferred from variation of wages/remunerations under the (questionable) assumption that wage differences must reflect productivity differences. Also firm-level data contain wage data and can thus be used to assess the degree of alignment of wage on marginal labour productvity. Finally, firm-level data permit testing the presence of human/education spillover ie the fact that workers with a certain educational attainment may influence the productivity of others. To address this range of issues this paper estimates labour-quality index of production function à la Hellerstein-Neumark using a rich, firm-level, Belgian panel database that contains information on productivity, labour cost and the workforce's educational attainment. It addresss unobserved heterogeneity bias using fixed effect methods and simultaneity bias using both Ackerberg, Caves & Frazer (2006) (ACF) and system-generalized method of moments (S-GMM) methods (Blundell & Bond, 1998) where lagged values of labour inputs are used as instruments. Results suggest that human capital, in particular larger shares of Masters-educated workers inside firms, translate into significantly higher firm-level labour productivity, and that labour costs are relatively well aligned on education-driven labour productivity differences. In other words, we find evidence that the Mincerian relationship between education and individual wages is driven by a strong positive link between education and firm-level productivity. But, in contrast to what many human capital theorists predict, we find not evidence of spillovers across education categories.

Keywords: Education, Human capital, Firm-Level Productivity and Labour Cost

JEL Classification: J24, E24, C51

[**] Université catholique de Louvain (UCL), IRES, 3 place Montesquieu, B-1348 Louvain-la-Neuve (Belgium), email: vincent.vandenberghe@uclouvain.be.
Funding for this research was provided by the Belgian Federal Government - SPP Politique scientifique, programme "Société et Avenir", project Employment Wage Discrimination and Poverty (EDIPO), Research contractTA/00/46.

[1] Productivity, as most economists conceive it, should not be amalgamated with individual capabilities, either physical or (more related to what is discussed here) cognitive ones. There is evidence, stemming from the International Adult Literacy Surveys (IALS) for example, that individuals who have completed more years of education and possess college or university degree achieve better in terms of literacy or other aspect of cognitive performance.

1. Introduction

There exists substantial evidence, based on the analysis of individual data, that general education (schooling) increases wages. Card (1999) for instance, summarizes various Mincer-inspired studies and concludes that the impact of a year of schooling on wages is about 10%. Similar results exist for Belgium (de la Croix & Vandenberghe, 2004) and many other member countries of the Organization for Economic Co-Operation and Development (OECD). These results generally interpreted as a validation of Becker's human capital theory where more educated individuals are more productive (and thus better paid, assuming market remunerate production factors according to their marginal productivity). The puzzling element of that approach is that labour productivity is never measured or estimated. It is inferred from variation of wages/remunerations under the assumption that wage differences must reflect productivity differences.

Some macroeconomists, analysing country-level time series, also support the idea that the continuous expansion of education has contributed positively to revenue per head (Krueger & Lindahl, 2001), or production per worker (Mankiw et al, 1992). But at that level, identification of the proper contribution of education is complicated by the difficulty to separate - using cross-country data over long time periods - the causal effect of education of income, from the wealth-driven surge of the demand for education, in particular of access to tertiary education.

This paper is based on few key considerations. First, jointly investigating the relationship between productivity, wages and workforce composition (e.g. its educational attainment) − which amounts to bridging industrial organisation and labour economics − is a promising research agenda. Second, productivity is, in essence, a firm-level phenomenon and should be primarily assessed at that level. In modern economies, where most people work inside firms, education-related productivity gains cannot possibly exist at the individual-level (as highlighted in Mincer-type analyses) if they do not show up at the firm level. Productivity is probably intrinsically determined by the (heterogeneous) ability of firms to successfully aggregate individual productivities, in conjunction with other factors of production (capital...). A similar reasoning applies to countries: the benefits of human capital should show clearly in the performance of firms, if they are to emerge at a more aggregate level. We thus argue that a study of the relationship between education, productivity and remuneration requires analysing data at the level of the firm. Individual workers' productivity is hardly ever observed.[1] By contrast, many datasets now contain good-quality information about what firms are able to produce (e.g. firm value added). Similarly, the alignment of productivity and pay at the individual level is hard to assess. But it can be evaluated with firm-level aggregates, conditional on adoption of an adequate analytical framework, as we will show in Section 2. Workers' characteristics (e.g. their educational attainment) can be aggregated at the firm level and introduced into firm-level equations in order to explore how they influence productivity and pay/remuneration.

Surprisingly, most economists with an interest in human capital have neglected the level of the firm to study the education-productivity-pay nexus. Other characteristics of the workforce, like gender or age have, by comparison, been much more investigated at the level of the firm by industrial or labour economists (Hellerstein et al., 1999, Aubert & Crépon, 2003; Hellerstein & Neumark, 2007; Vandenberghe, 2011a,b, Vandenberghe, 2012; Rigo, Vandenberghe & Waltenberg, 2012; Dostie, 2011; van Ours & Stoeldraijer, 2011).

At present, the small literature based on firm-level evidence provides some suggestive evidence of the link between education, productivity and pay at the level of firms. Examples are Hægeland & Klette (1999); Haltiwanger et al. (1999). Other notable papers examining a similar question are Galindo-Rueda & Haskel (2005), Prskawetz et al. (2007) and Turcotte & Rennison (2004). The general consensus in this strand of research is that more educated workers are also more productive. They further conclude that there is an alignment of marginal benefit (productivity) and marginal cost (wage).

But, despite offering plausible and intuitive results, many of the above studies essentially rely on cross-sectional evidence and most of them do not tackle the two crucial aspects of the endogeneity problem affecting the estimation of production and wage functions (Griliches & Mairesse, 1995): *i)* heterogeneity bias (unobserved time-invariant determinants of firms' productivity that may be

[1] Productivity, as most economists conceive it, should not be amalgamated with individual capabilities, either physical or (more related to what is discussed here) cognitive ones. There is evidence, stemming from the International Adult Literacy Surveys (IALS) for example, that individuals who have completed more years of education and possess college or university degree achieve better in terms of literacy or other aspect of cognitive performance.

correlated to the workforce structure[2]) and *ii)* simultaneity bias (endogeneity in input choice, in the short-run, that includes the workforce mix of the firm[3]). While we know that labour productivity is highly heterogeneous across firms (Syverson, 2011)[4], only Haltiwanger et al. (1999) control for firm level-unobservables using firm-fixed effects. The problem of simultaneity has also generally been overlooked. Certain short-term productivity shocks affecting the choice of labour inputs, can be anticipated by the firms and influence their employment decision and thus the workforce mix. Yet these shocks and the resulting decisions by firms' manager are unobservable by the econometrician. Hægeland & Klette (1999) try to solve this problem by proxying productivity shocks with intermediate goods, but their methodology inspired by Levinsohn & Petrin (2003) suffers from serious identification issues due to collinearity between labour and intermediate goods (Ackerberg, Caves & Frazer, 2006) (more on this in Section 2).

Our aim here is to provide a methodologically solid investigation into the connection between a key measure of firm performance: labour productivity (i.e. value added per worker) and the composition of firms' workforce in terms of educational attainment, with a particular focus on tertiary education.[5] The latter choice echoes the, now rather dominant, view that in advanced economies like Belgium, productivity gains are driven by the expansion of tertiary education.[6] We exploit longitudinal firm-level Belgian data (edited by Bel-first[7]). The latest release of this data set contains longitudinal information for a sizeable sample of 7,600 firms located in Belgium for the period 1998-2017, on key

---

[2] For instance, the age of the plant/establishment may affect productivity and simultaneously be correlated with the educational attainment of the workers; less-educated workers being overrepresented in older ones.

[3] For instance, the simultaneity of a negative productivity shock (due to the loss of a major contract) and workforce becoming less educated stemming from a recruitment freeze, causing reverse causality: from productivity to education.

[4] The evidence with firm panel data is that fixed effects capture a large proportion (>50%) of the total productivity variation. Another illustration of the same idea is that published studies have documented, virtually without exception, enormous and persistent measured (but unexplained) productivity differences across firms, even within narrowly defined industries.

[5] Which, in the Belgian context, comprises 2-3year College/Bachelor programmes and 4-5year university/Masters programmes. Both are clearly tertiary (post-secondary) forms of education. The Bachelor programmes are more vocational in essence and would train future high-level technicians of people in charge of the middle managements, whereas the Masters programmes aimed at delivering a more general/academic training, for people who will end up occupying top/managerial positions or professions delivering professional services (doctors, lawyers, architects....).

[6] Macroeconomists like Wolff & Gittleman (1993) show that for the upper-income group of countries (that comprises Belgium) - among which there is much more cross-county variation in tertiary education than in primary or secondary education attainment - tertiary education is the only statistically significant variable. On the other hand, for the poor countries primary education is statistically significant, while differences in tertiary education are not.

[7] https://belfirst.bvdinfo.com

outcomes and costs of the businesses, as well as the educational attainment of their workers.

In this paper, a key objective is to find robust causal evidence of a positive impact on labour productivity per worker of larger shares of better-educated workers (i.e. those with tertiary-education attainment). We follow the methodology pioneered by Hellerstein & Neumark (1999) (HN henceforth).[8] It consists of estimating Cobb-Douglas production (or labour cost functions) that explicitly account for labour heterogeneity. Applied to firm-level data, this methodology presents three main advantages. First, it delivers productivity/labour cost differences across education groups that can immediately be compared and tested. Second, it measures and tests for the presence of market-wide benefits of education for business. Third, the comparison of the estimated productivity and labour cost coefficients helps to assess the link between education-driven productivity gains and education-driven remuneration increments, at the core of a human-capital model. To be precise, we adopt of a fully linearised Cobb-Douglas specification. That implies assuming perfect substitutability of labour inputs and also of labour and capital[9], but it allows us to estimate fixed effect models (FE hereafter) and thus controlling for interfirm unobserved heterogeneity. In one of our preferred models, we will relax the perfect substitutability assumption (the details will be explained in Section 2).

Recent developments of HN's methodology have tried to improve the estimation of the production function by the adoption of alternative techniques to deal with the risk of simultaneity bias[10]. One set of techniques follows the dynamic panel literature (Arellano & Bond, 1991; Aubert & Crépon, 2003; or van Ours & Stoeldraijer, 2011), which basically consists of using lagged values of (first-differenced) labour inputs as instrumental variables. Its more advanced incarnation is the one proposed by Blundell and Bond (1998) called system-generalized method of moments (S-GMM hereafter). Note about this stream of research that first differences are good at purging fixed effects and thus at coping with unobserved heterogeneity. A second set of techniques, initially advocated by Olley & Pakes (1996), Levinsohn & Petrin (2003) (OP, LP henceforth), and more recently by Ackerberg, Caves & Fraser (2006) (ACF henceforth), are somewhat more structural in nature. They

---

8 The key idea of HN is to estimate a production function (or a labour-cost function), with heterogeneous labour input, where different types (e.g. men/women, more educated/less educated) diverge in terms of marginal productivity.

9 There is a relatively abundant literature on skill-capital substitution suggesting that capital and skilled labor are more complementary as inputs than are capital and unskilled labor (see capital-skill complementarity hypothesis by Griliches (1969)). Duffy, Papageorgiou & Perez (2004) provide the cross-country evidence for capital-skill complementarity.

10 See Ackerberg, Caves & Frazer (2006) for a recent review.

consist of using observed intermediate input decisions (i.e. purchases of raw materials, services, electricity...) to "control" for (or proxy) unobserved short-term productivity shocks.

In this paper, we apply these two strategies that are aimed at coping with simultaneity. Following many authors in this area (Aubert & Crépon, 2003, 2007; van Ours & Stoeldraijer, 2011; Cataldi, Kampelmann & Rycx, 2011), we first implement the ACF approach which primarily consists of using intermediate inputs to control for short-term simultaneity bias. Note that we innovate within this stream, as we combine the ACF intermediate-good approach with fixed effects (FE), to better account for both simultaneity and firm heterogeneity (FE-ACF henceforth). More on this in Section 2, we also estimate the relevant parameters of our model using "internal" instruments (i.e lagged values of endogenous labour inputs) (S-GMM henceforth). .

Our main results indicate that the marginal productivity of workers with a Masters' degree is significantly larger than that of workers with primary education attainment or less. In particular, our preferred specifications controlling for endogeneity and firm heterogeneity (S-GMM, FE-ACF) shows that a worker with a university degree is 18% (FE-ACF) to 27% (S-GMM) more productive than a worker with a primary education attainment or less. Workers with a 2-year college degree or only secondary school appear to be 3.3%[11] (FE-ACF) to 10% more productive as primary school graduates. Simultaneously, the labour cost premium associated to workers university degree is 18.7% (FE-ACF) to 29.3% (S-GMM), and 8%[12] (FE-ACF) to 3% (S-GMM) % for those with a 2-year college degree. Workers with only secondary school appear to be not more productive/expensive than workers with a primary school attainment. Hence, we interpret our results as supportive of the alignment of labour costs on productivity, and thus a validation of the Mincerian assumption.

The rest of the paper is organized as follows. In Section 2, our methodological choices regarding the estimation of the production and labour cost functions are detailed. Section 3 is devoted to an exposition of the dataset. Section 4 contains the econometric results and Section 5 our main conclusions.

---

11 But the underlying coefficients are not statistically significant as the 10% threshold

12 Again, the underlying coefficients are not statistically significant as the 10% threshold

2. Methodology

2.1) Algebraic framework

Our methodology owes a lot to Hellerstein & Neumark (1995) but also to the literature on skill diversity (Duffy et al., 2004) and the one on the degree of complementarity/substitutability between skills (educated labour) and capital. The present work echos the way the HN framework is used to analyse the relationship between age, productivity and wage (Vandenberghe et al., 2013) .[13] The common point of all these papers is that they to two-stage production functions. A stage one, these production functions write as standard Cobb-Douglas

[1.] $Y_{it}=A_t K_{it}^{\alpha} QL_{it}^{\beta}$

where Y_{it} is the value added (productivity hereafter) in firm *i* at time *t*, and K_{it} is the stock of capital. The variable that reflects labour diversity (here in terms of educational attainment) is what HN call *the quality of labour index* QL_{it}. It forms the second stage of the production technology.

[2.] $QL_{it} = [\sum_j \mu_j L_{ijt}^{\rho}]^{1/\rho}; j=1...n$

with Lj the number of workers of type *j* (here with a certain educational attainment) in firm *i* at time *t,* and μ_j their contribution to productivity.

Unlike HN or Vandenberghe (2013), we allow for the possibility that workers of various types are not perfectly substitutable ($\rho \neq 1$). This amounts to defining QL_i as a CES,

Injecting [2] into [1] leads to

[3.] $Y_{it}= A_t K_{it}^{\alpha} [\sum_j \mu_j L_{ijt}^{\rho}]^{\beta/\rho}; j=1...n$

Note now that [3] can can be rewritten by dividing(multiplying) all the labour terms by $\mu_r L_{it}^{\rho}$ where *r* is the labour reference type (ex: workers with primary education)

[4.] $Y_{it}= A_t \mu_r^{\beta/\rho} K_{it}^{\alpha} L_{it}^{\beta} [S_{irt}^{\rho} + \sum_{j\neq r} \lambda_{jr} S_{ijt}^{\rho}]^{\beta/\rho} ; j=1...n$

---

[13] Which is relatively more developed than the literature on education diversity, and better connected the standard economic theory of production. It has.

$S_{itj}=L_{itj}/L_{it}$ are the labour labour shares (e.g ecucation categories) and $\lambda_{jr}\equiv\mu_{j}/\mu_r$ is the parameter reflecting the (relative) contribution of type j labour to productivity. The latter expression can be log-linearized, leading to

[5.] $y_{it}=B_t + \alpha\, k_{it} + \beta l_{it} + \beta/\rho\; ln[S_{irt}^{\rho} + \sum_{j\neq r} \lambda_{jr}\; S_{ijt}^{\rho}]\; ; j=1\ldots n$

where $B_t\equiv ln(A_t\; \mu_r^{\beta/\rho});\; y_{it}\equiv lnY_{it};\; k_{it}\equiv lnK_{it};\; l_{it}\equiv lnL_{it}$

Before turning to econometrics, it is worth highlighting a couple of features of expression [5].

- First, (relative) marginal labour productivities can easily be retrieved from [5]. Ignoring time indices for a moment, we have indeed that

[6.] $\frac{\partial Y}{\partial L_{ir}}=A\; K_i^{\alpha}\, \beta\; [\sum_j \mu_j L_{ij}^{\rho}]^{\,\beta/\rho-1}\mu_r\, (L_{ir})^{\rho-1}$

and thus

[7.] $\frac{\partial Y}{\partial L_{ij}} / \frac{\partial Y}{\partial L_{ir}} = \lambda_{jr}\; (L_{ij}/L_{ir})^{\rho-1}$

As long as $\rho\neq 1$, (ie. labour types are not perfect substitutes), relative marginal productivities depend on the firm-specific labour mix and not just λ_{jr}'s. and this is where the idea of (education) spillovers has its roots. Another way of capuring the same idea is to consider the 2nd order cross derivatives of the labour qualily index [2]. It is easy to show (ignoring temporarily t indices for simplicity of exposure) that $\frac{\partial^2 QL_i}{\partial L_{ir}\partial L_{ih}}=(1-\rho)/\rho^2\; [\sum_j \mu_j L_{ij}^{\rho}]^{\,(1-2\rho)/\rho}\mu_r\, (L_{ir})^{\rho-1}\mu_h\, (L_{ih})^{\rho-1}$; for any pair of labour types $r\neq h$. And the latter expression is different from 0 only if $\rho\neq 1$.

- Second, if $\rho=1$ [5] becomes

[8.] $y_{it}=B_t^{\$} + \alpha\, k_{it} + \beta l_{it} + \beta\; ln[S_{irt}+ \sum_{j\neq r} \lambda_{jr}\; S_{ijt}]\; ; j=1\ldots n$

where $B_t^{\$}\equiv ln(A_t\; \mu_r^{\beta});\; y_{it}\equiv lnY_{it};\; k_{it}\equiv lnK_{it};\; l_{it}\equiv lnL_{it}$. What is more, as labour shares sum up to 1, $S_{irt}=1-\sum_{j\neq r}S_{ijt}$

Thus the production function can be writhen

[9.] $y_{it}= B_t^{\$}+ \alpha k_{it} + \beta l_{it} + \beta ln[1+\sum_{j\neq r} (\lambda_{jr} -1)S_{ijt}]$; $j=1\ldots n$

Since $ln(1+x)\approx x$, for small values of x, HN suggest adopting the following log-linear specification:

[10.] $y_{it}= B_t^{\$} + \alpha k_{it} + \beta l_{it} +\sum_{j\neq r} \eta_{jr} S_{ijt}$; $j=1\ldots n$

where $\eta_j= \beta(\lambda_{jr}-1)$

and $B_t^{\$} \equiv ln(A_t\ \mu_r^{\beta})$; $y_{it}\equiv lnY_{it}$; $k_{it}\equiv lnK_{it}$; $l_{it}\equiv lnL_{it}$; $=\mu_j/\mu_r$

Note that [10], being loglinear its coefficients can be interpreted as those of the canonical Mincer equation. In other words S_{ijt} coefficients can be directly interpreted as the percentage change in the firm's average labour productivity of a 1 unit (here 100 percentage points) change of the considered type of workers' share among the employees of the firm. Note also that, strictly speaking, in order to obtain a type j worker's relative marginal productivity, (*i.e.* λ_{jr}), coefficients η_j have to be divided by β, and 1 needs to be added to the result.[14]

Following, van Ours & Stoeldraijer (2011) and Vandenberghe (2013), we propose adopting a similar specification as the of [5] or [10] for the labour cost/wage function.

[11.] $w_{it}= B^w{}_t^{\$} + \alpha^w k_{it} + \beta^w l_{it} = \beta^w/\rho^w\ ln[S_{irt}{}^{\rho^w} + \sum_{j\neq r} \lambda^w{}_{jr}\ \ S_{ijt}{}^{\rho^w}]$; $j=1\ldots n$

where $B^w{}_t \equiv ln(A^w{}_t\ \mu^w{}_r{}^{\beta/\rho})$; $w_{it}\equiv lnW_{it}$; $k_{it}\equiv lnK_{it}$; $l_{it}\equiv lnL_{it}$ and $\lambda^w{}_j =\mu^w{}_j/\mu^w{}_r$ the relative remuneration of type j workers with respect to the reference type

Like in the productivity equations [5], [10] coefficients η_j^w capture the sensitivity of laobur cost to changes of the educational structure of the labour force (S_{ijt}). Note that they do not indicate the actual wage distribution within the firms for different categories of workers, but a hypothetical wage distribution that depends on the variation of the proportion of different workers categories and its

---

[14] Does all this matter in practice? Our experience with firm-level data suggests values for β ranging from 0.6 to 0.8 (these values are in line with what most authors estimate for the share of labour in firms' output/added value). This means that $\lambda_j -1$ are larger (in absolute value) than η_j..

correlation with the firm's average wage level.

The key hypothesis test of this paper can now be easily formulated. If more education/human capital leads to more firm-level productivity, one should verify that η_{jr} rise with educational attainment. And if Belgian labour markets rewards human capital/education at its marginal produce (i.e. as assumed in the standard human-capital model), then one should observe that $\lambda_{jr} \approx \lambda^{w}_{jr}$. Any negative (or positive), statistically significant, difference between these lambdas can be interpreted as a quantitative measure of the violation of the human-capital prediction of alignment. In the same vein, the comparison of the lambdas also informs about the legitimacy of works done by economists (Mincer, 1974; and many others) who simply assume that relative remuneration must reflect underlying productivity differences.

For economists with an interest in labour, in particular labour demand, the sign and the magnitude of the gap between the lambdas can also be interpreted as a measure of the financial incentives firms have to modify the educational mix of their workforce. A sizeable positive difference between the productivity and the labour cost lambdas for workers with a Masters' degree means that firms can improve their gross surplus (the difference between value added and total labour cost) by increasing the share of employees with university degree in their overall workforce. Conversely, a large negative difference should rather entice them to reduce that share.

2.2) Econometric framework and identification

The econometric equivalent of the two-stage production function (ie CES-within Cobb-Douglas) can be noted

[12.] $Y_{it} = A_0 K_{it}^{\alpha} QL_{it}^{\beta} e^{\tau + \varepsilon it}$

or in logs, using the developpement exposed above

[13.] $y_{it} = B_t + \alpha \ lnK_{it} + \beta lnL_{it} + \beta/\rho \ ln[S_{irt}^{\rho} + \sum_{j \neq r} \lambda_{jr} S_{ijt}^{\rho}] + \varepsilon_{it}; \ j=1 \ldots n$

where $B_t \equiv ln(A_0 \mu_r^{\beta/\rho}) + \tau t; \ y_{it} \equiv lnY_{it}; \ k_{it} \equiv lnK_{it}; \ l_{it} \equiv lnL_{it}$. The above model is non linear, so it requires

NLSQ techniques to be estimated. But, from an econometric point of view, what primariliy matters is the structure of the error term. We will assume here that it comprises three elements:

[14.] $\varepsilon_{it} = \theta_i + \omega_{it} + \sigma_{it}$

where: $cov(\theta_i, S_{ij,t}) \neq 0$, $cov(\omega_{it}, S_{ij,t}) \neq 0$ $E(\sigma_{it})=0$

In other words the NLSQ sample-error term may consist of *i)* an unobservable firm fixed effect θ_i; *ii)* a short-term shock ω_{it} whose evolution corresponds to a first-order Markov chain, and is observed by the firm (but not by the econometrician) and (partially) anticipated by the firm, and, *iii)* a purely random shock σ_{it}. Parameter θ_i. in [13] represents firm-specific characteristics that are unobservable but driving average productivity. For example the vintage of capital in use, firm-specific managerial skills, location-driven comparative advantages[15], product lifecycle heterogeneity. And some of these might be correlated with the educationial structure of the firm's workforce, biasing NLSQ results. Less educated workers for instance might be overrepresented among plants built a long time ago using older/less productive technology. However, the panel structure of our data allows for the estimation of NLSQ with firm fixed effects controlling for that sort of bias (more on this here after).

The other challenge is to go around the simultaneity bias (Griliches & Mairesse, 1995). The economics underlying that concern is intuitive. In the short run, firms could be confronted to productivity deviations, ω_{it}; say, a lower turnover, itself the consequence of a missed sales opportunity. Contrary to the econometrician, firms may know about ω_{it}. An anticipated downturn could translate into selective recruitment freeze or layoffs, with less educated workers more at risk of not being recruited or laid-off. Negative spells, would then create a negative correlation between rising shares of better-educated workers and productivity, thereby leading to underestimated estimates of their productivity (when resorting to NLSQ or even NLSQ-FE estimates).

i) Firm heterogeneity as fixed effects (NLSQ-FE)

Duffy et al. (2004) show that fixed effects can be eliminated from a non-linear model like [12] by estimating its growth-equivalent (over T periods)

---

[15] Motorway/airport in the vicinity of logistics companies for instance.

[15.] $y_{it}- y_{it-T} = \tau T + \alpha(k_{it}-k_{it-T}) + \beta(l_{it}-l_{it-T}) + \beta/\rho \, ln[S_{irt}^{\rho} + \sum_{j\neq r} \lambda_{jr} S_{ijt}^{\rho}] / [S_{irt-T}^{\rho} + \sum_{j\neq r} \lambda_{jr} S_{ijt-T}^{\rho}]$

$+ \omega_{it}- \omega_{it-T} + \sigma_{it}- \sigma_{it-T}$

Note that the new constant term stems form T-differencing $B_t \equiv ln(A_t \, \mu_r^{\beta/\rho})$ under the assumption that $A_t = A_0 e^{\tau t}$ *;* where A_{i0} is the starting value of TFP and τ its the constant/annual rate of growth.

ii) Simultaneity bias (ACF-FE and S-GMM)

There are several methods to account for the presence of the endogeneity bias. But as far as we know, they all imply (log) linearising our equations. This means assuming perfect susbsitutability between labour types (in other words the absence of human capital/education spillover at firm level). Of course the validity of that assumption can be assessed using NLSQ/NLSQ-FE. We will comment more on this in the result section.

The first method we implement corresponds to the structural approach initiated Ackerberg, Caves & Frazer (2006). The essence of the ACF approach is to use some function of a firm's material inputs (for instance raw materials, electricity, purchase of services...) in the estimation of unobserved productivity. ACF argue that firms can swiftly (and also at a relatively low cost) respond to productivity developments, ω_{it}, by adapting the volume of the intermediate inputs they buy on the market. ACF argue that there is some solid and intuitive identification idea in the LP paper, but they claim that their two-stage estimation procedure delivers poor estimates of the labour coefficients due and propose an improved version of it.

Simplifying our notations to make them alike those used by ACF, productivity per hour equation becomes:

[16.] $y_{it} = B_t + \alpha \, k_{it} + \varphi \, ql_{it} + \varepsilon_{it}$

with the HN labour quality index (or vector of labour inputs) equal to:

[17.] $\varphi \, ql_{it} \equiv \beta l_{it} + \sum_{j\neq r} \eta_{jr} S_{ijt} + \boldsymbol{\varepsilon_{it}}, \;\; j=1 \ldots n$

and the error term $\varepsilon_{it} = \theta_i + \omega_{it} + \sigma_{it}$

Note that unlike AC, we explicitly assume that ε_{it} contains a proper fixed effect (here our reference hereafter to ACF-FE estimemates). But like ACF, we assume that firms' (observable) demand for intermediate inputs (int_{it}) is a function of the unobserved terms as well as (log of) capital, and the quality of labour index (in logs) ql_{it} and its components[16]:

[18.] $int_{it} = f_t(\omega_{it}, k_{it}, ql_{it})$

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 $\theta_i + \omega_{it}$ as a function of int_{it} , k_{it}, ql_{it}, and introduced into the production function: $\omega_{it} \equiv f_t^{-1}(int_{it}, k_{it}, ql_{it}),$ leading to

[19.] $y_{it} = B_t + \alpha\, k_{it} + \varphi\, ql_{it} + \theta_i + f_t^{-1}(int_{it}, k_{it}, ql_{it}) + \sigma_{it}$

In practice, how are the parameter forming vector $\varphi \equiv (\beta, \eta_{1r},, \eta_{2r}, \ldots \eta_{nr})$ and α estimated? The ACF algorithm consists of two stages. We argue that only stage one needs to be adapted to account for fixed effects θ_i, by resorting to first-differences. In stage one, like ACF, we regress average productivity (i.e value added per fte worker) on a composite term Φ_{it} that comprises a constant, a 3^{rd} order polynomial expansion in $int_{it}, k_{it}, ql_{it.}$ (and the fixed effect θ_i). This leads to

[20.] $y_{it} = \Phi_{it}\,(int_{it}, ql_{it}, k_{it}) + \sigma_{it}$

Note that $\varphi \equiv (\beta, \eta_{1r},, \eta_{2r}, \ldots \eta_{nr})$ and α are clearly not identified yet, implying the need of a second stage.[17] Note in particular that Φ_{it} encompasses $f_t^{-1}(.)$ proxying ω_{it}. The point made by ACF is that this first-stage regression delivers an unbiased estimate of the composite term Φ_{it}^{hat}; i.e productivity net of the purely random term σ_{it}. We go a step further and also get rid of θ_i by resorting to first-differences when estimating equation [20]. The resulting FD-estimated coefficients - provided they

[16] Note, as an aside that LP unrealistically assume that the demand of intermediate goods is not influenced by that of labour inputs. Consider the situation where ql_{it} is chosen at $t\text{-}b$ ($0<b<1$) and int_{it} is chosen at t. Since ql_{it} is chosen before int_{it}, a profit-maximizing (or cost-minimizing) optimal choice of int_{it} will generally directly depend on ql_{it} (Ackerberg, et al., 2006) creating strong collinearity problems.

[17] Note in particular that the non identification of vector φ (ie. labour input coefficients) in the first stage is one of the main differences between ACF and LP.

are applied to variables in levels - deliver an unbiased prediction of Φ_{it}^{hat}. Specifically, Φ_{it}^{hat}, net of the noise term σ_{it}. and firm-fixed effects, is calculated as $\Phi_{it}^{hat} = (\upsilon_{a1})^{FD}\, int_{it} + (\upsilon_{a2})^{FD}\, int^2_{it} + \ldots + (\upsilon_{b1})^{FD}\, k_{it} + \ldots + (\upsilon_{c1})^{FD} ql_{it} + \ldots + (\upsilon_{d1})^{FD}\, int_{it} k_{it} \ldots$, where $(\upsilon_{a1})^{FD}$, $(\upsilon_{a2})^{FD} \ldots$ represent the first-differenced coefficient estimates on the polynomial terms.

As the second step of ACF we basically argue that it is unaffected by the modifications discussed above. Key is the idea that one can generate implied values for $_{it}$ using first-stage estimates Φ_{it}^{hat} and candidate[18] values for the coefficients φ, α:[19]

[21.] $\omega_{it} = \Phi_{it}^{hat} - ql_{it}\,\varphi - \alpha k_{it}$

ACF assume further that the evolution of ω_{it} follows a first-order Markov process

[22.] $\omega_{it} = E[\omega_{it} | \omega_{it-1}] - \xi_{it}$

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}.

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

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 φ, α. Like ACF, we would also argue that capital in period t was determined at period $t-1$ (or earlier). The economics behind this is that it may take a full period for new capital to be ordered and put to use. Since k_{it} is actually decided upon $t-1$, $t-2\ldots$, it must be uncorrelated with the implied innovation terms ξ_{it}:

[24.] $E[\xi_{it} | k_{it}] = 0$

---

18 OLS estimates for example.

19 Equation (16) appears as a Cobb-Douglas. However, its key component is Φ_{it}^{hat} that consists of 3^{rd} order polynomial expansion in int_{it}, k_{it}, ql_{it} where the latter variables are systematically interacted, implicitly allowing for imperfect substitutability across labour types and with capital.

Labour inputs observed in t are probably also chosen sometime before, although after capital – say in $t\text{-}b$, with $0<b<1$. As a consequence, ql_{it} will be correlated with at least part of the productivity innovation ξ_{it}. On the other hand, assuming *lagged* labour inputs were chosen at time $t\text{-}b\text{-}1$ (or earlier), ql_{it-1}, ql_{it-2}… should be uncorrelated with the innovation terms ξ_{it}. This gives us the third (vector) of moment conditions needed for identification of φ:

[25.] $E[\xi_{it}|\ ql_{it-1}, ql_{it-2}...]=0$

or more explicitly, given the composite nature of ql_{it}, we have:

[26.] $E[\xi_{it}|\ lnl_{it-1}, lnl_{it-2}...]=0$
[27.] $E[\xi_{it}|\ S_{i1t-1}, S_{i1t-2}...]=0$
[28.] $E[\xi_{it}|\ S_{i2t-1}, S_{i2t-2}...]=0$
…

An alternative to the structural approach is to estimate the relevant parameters of our model using only "internal" instruments. The essence of this strategy is to use lagged values of endogenous labour inputs as instruments for the endogenous (first-differenced) labour inputs (Aubert & Crépon, 2003, 2007; van Ours & Stoeldraijer, 2011; Cataldi, Kampelmann & Rycx, 2011).[20] First differences are good at purging fixed effects and thus at coping with unobserved heterogeneity terms θ_i. But (lagged) variables in level, although they might be orthogonal to the short-term shock ω_{it}, tend to prove poor predictors of first differences (i.e. they are weak instruments). Blundell & Bond (1998) then proposed an improved estimator called system-GMM (S-GMM hereafter) that uses extra moment conditions. S-GMM consists of a system of two equations estimated simultaneously. One corresponds to the above-mentioned first-difference equation, where the instruments are the (lagged) labour inputs in level. The second equation consists of using regressors in level, with (lagged) first-differenced of the endogenous variables as instruments. S-GMM estimator has become the estimator of choice in many applied panel data settings. We use it here to cope with simultaneity/endogeneity of the labour inputs (i.e. both the overall level of labour and the share of education group).

---

[20] The other key feature of these methods is that they are based on the Generalized Method of Moments (GMM), known for being more robust than 2SLS to the presence of heteroskedasticity (see appendix in Arellano, 2003).

3. Data description

We are in possession of a large unbalanced panel of around 69,148 firm-year observations (Table 1) corresponding to the situation of a 7,673 firms (Table 4) and 150,000 workers on average per year of observatin[21], from all sectors forming the Belgian private economy, in the period 1998– 2012. These firms are largely documented in terms of sector (NACE2), size, capital used, labour cost levels[22] and productivity (value added per full-time equivelent worker). These observations come from the Bel-first database,[23] that most for-profit firms located in Belgium must feed to comply with the legal prescriptions. All the firms occur at least 3 times in the panel; the maximum being 16 times (Table 4).

Descriptive statistics, forming this large sample are reported in Table 1. Labour costs used in this paper, which were measured independently of value added, include the value of all monetary compensations paid to the total labour force (both full- and part-time, permanent and temporary), including social security contributions paid by the employers, throughout the year. In the r part of Table 1, one also sees that labour costs (overall labour costs per hour) is logically inferior to productivity (value-added per hour).

Of prime interest in this paper is the breakdown by educational attainment and also the consequence of changing (risng) educational attainment of the workerforce. Table 2 (last line) shows that, during the observed period (1997-2012), about 74% of the workforce of private for-profit firms located in Belgium have still, at most, an upper secondary school degree. Workers with a 2-year college/bachelor degree represented 17% of the total workforce. Slightly less than 9% consisted of individuals with a (4-5year) Masters' degree. This means a mere 26% of workers with a tertiary education background; clearly less than the percentage among the current generation of school leavers.[24] This discrepancy logically reflects the higher propensity of younger generations to stay on beyond secondary education, and complete a tertiary education programme. And the impact of this trend is clearly visible in Table 2. Between 1997 and 2012, the share of people with at most an upper secondary

---

21 About 5% of the total workforce of the Belgian private economy.

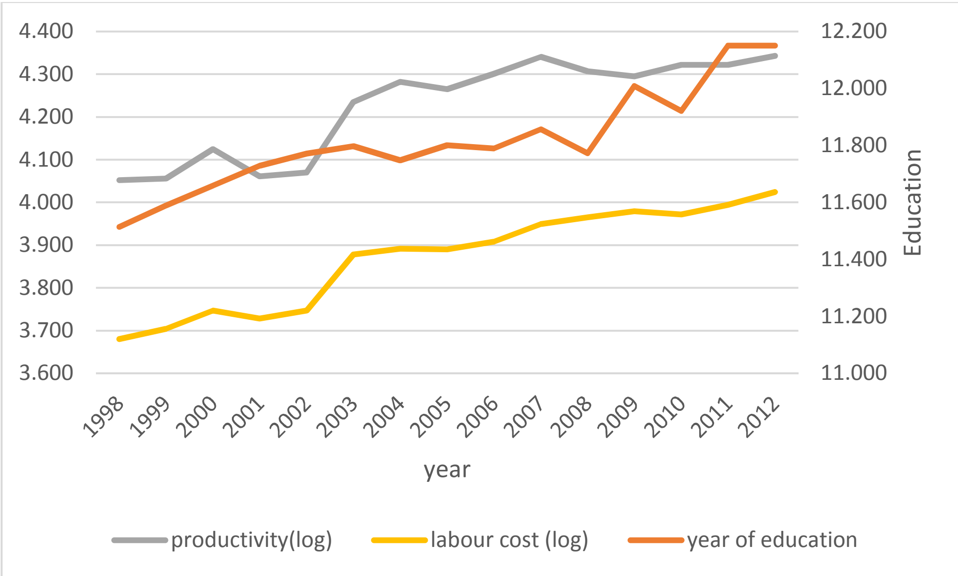
22 Labour costs used in this paper, which were measured independently of value added, include the value of all monetary compensations paid to the total labour force (both full- and part-time, permanent and temporary), including social security contributions paid by the employers, throughout the year. The summary statistics of the variables in the data set are presented in Table 1.

23 http://www.bvdinfo.com/Products/Company-Information/National/Bel-First.aspx

24 Statistics Belgium estimates that they now represent between 35 and 40% of a cohort.

degree has declined markedly whereas the one of workers with a tertiary degree (Bachelor or Masters) has risen.

Figure 1 – Education, productivity and labour cost: 1998-2012



Source: Bel-first. Note Average year of education is defined at firm-level using the following durations Primary=6y; Secondary=12y; Bachelor=15y; Master=17y

Table 1- Descriptive statistics: Belgian private economy 1998-2012

	Prod. th.Eur (log)	Labour cost th.Eur (log)	Number of fte workers (log)	Capital in th.Eur (log)	Workforce (shares by educational attainment) Primary	Sec (upper)	Bachelor	Master	Average years of educ($)	Raw materials, consumables & other goods in th.Eur per fte worker
Mean	4.285	3.920	6.315	10.764	0.186	0.552	0.175	0.090	11.868	4.502
St dev	0.648	0.410	1.820	2.565	0.270	0.293	0.181	0.133	2.195	2.026
N	69,148									

Source: Bel-first. Weights=firm-level number of fte workers

$ Average year of education is defined at firm-level using the following durations Primary=6y; Secondary=12y; Bachelor=15y; Master=17y

Table 2-Evolution of the composition of the workforce by educational attainment: Belgian private economy 1997-2012

	Workforce (shares by educational attainment)				
	Primary	Secondary (upper)	Bachelor	Master	Average years of educ($)
1998	0.211	0.550	0.162	0.083	11.639
1999	0.222	0.539	0.161	0.083	11.579
2000	0.216	0.540	0.167	0.083	11.653
2001	0.201	0.550	0.169	0.085	11.748
2002	0.189	0.571	0.162	0.082	11.764
2003	0.192	0.560	0.167	0.085	11.798
2004	0.192	0.555	0.174	0.089	11.805
2005	0.198	0.546	0.170	0.093	11.765
2006	0.199	0.550	0.169	0.089	11.784
2007	0.189	0.556	0.170	0.090	11.826
2008	0.184	0.557	0.170	0.089	11.852
2009	0.171	0.549	0.186	0.095	12.005
2010	0.185	0.537	0.186	0.093	11.914
2011	0.152	0.559	0.190	0.099	12.153
2012	0.154	0.556	0.189	0.102	12.152
Total	0.186	0.552	0.175	0.090	11.868
N	69,148				

Source: Bel-first. Weights=firm-level number of fte workers
$ Average year of education is defined at firm-level using the following durations
Primary=6y; Secondary=12y; Bachelor=15y; Master=17y

Figure 1 displays, at the agregate level, how the (log of) productivity and labour cost per worker and labour have evolved with the average number of years of education during the period 1998 – 2012. These stylised facts suggest that, in the Belgian private economy, the productivity and labour cost regularly rise with human capital. But this is essentially a stylized fact that does not control for the important difference in the way workers with different educational background distribute across sectors that may dramatically differ in terms of productivity and labour cost for reasons that are independent from the educational structure of their workforces. Only adequate econometric analysis, with sector and/or firm fixed effects and other controls will allow us to draw more substantiated conclusions.

Intermediate inputs play a key role in our analysis, as they are central to one of our strategies to overcome the simultaneity/ endogeneity bias (see Section 2 ACF/FE-ACF models). Our measure is a direct one. It is the value (in th. EUR per full-time-equivalent worker) of raw materials, consumables and other goods and services consumed or used up as inputs in production by firms.

Finally, it is clear from Table 3 that there has been a rise in the number of firms included in the panel between 1997 and 2012. This reflects the history of Bel-first's way of collecting data on educational attainment. Until 2007, the only information reported was about the educational attainment of workers leaving or joining the workforce. After 2008, it became mandatory for all firms to communicate the information about the educational attainment of their workforce. What we have done is to reconstruct the pre-2008 times series using the information about in- and out-flows.

Table 3-Number of firms in the Bel-first sample

	Number of firms
1998	2,553
1999	2,711
2000	2,867
2001	3,045
2002	3,493
2003	2,942
2004	3,892
2005	4,045
2006	4,227
2007	4,395
2008	5,219
2009	7,659
2010	7,543
2011	7,462
2012	7,095
Total	69,148

Source: Bel-first

Table 4-Number of firms by spell duration

Duration of spell (years)	Number of firms
3	8
4	20
5	46
6	499
7	511
8	562
9	644
10	1,041
11	737
12	206
13	204
14	306
15	569
16	2,320
Total	7,673

Source: Bel-first

4. Econometric results

Table 5 summarises the main econometric results. We first estimate the productivity and labour-cost regression with non-linear least squares (NLSQ) (model (1)). To account for firm unobserved heterogeneity we then turn to growth models (with firm fixed effects) that are also estimated with non-linear least squares (NLSQ-FE) (model (2) in Table 5). Table 6 present the models (3) (4) dealing with the risk of simultaneity bias, (ACF-FE & S-GMM). Table 7 contains the results of our preferred models (NSLQ-FE (2a), ACF-FE (3a), S-GMM (4a)) applied to the (large) subsample forming the vast services industry. Notice relative marginal productiviy and labour cost are never direclty delivered by our econometric esimations, but they can be retrieved expost. The estimated retrieved values are reported at the bottom of Tables 5, 6, 7.

The basic NSLQ regression puts forward the presence of a relative increase in productivity of bachelor and university degres. with respect to primary education. Marginal productivity of bachelor degree is estimated to be 1.74 times larger than of workers with a primary education attainment. That of a worker with a Masters degree appears to be 3 times that of the reference group. NLSQ-estimated marginal labour cost convey the idea that bachelor workers cost 1.47 times more to their employer than the reference group, whereas the corresponding ratio for masters'-educated workers is 2.4. Results for secondary education are smaller in magnitude but also supportive of a positive impact on productivty, but not on labour cost.

Turning to the results of the NLSQ-FE/growth model (2) in Table 5, we immediately see at the bottom of th Table that a higher educational attainment translates into lower (marginal) productivity advantages compared with NLSQ in levels. This stems from controlling for firm unobserved heterogeneity (i.e. FE), and it suggests that bette r-educated individuals, in particular those with a masters' background, concentrate in firms that are intrinsically more productive. Holding a secondary degree does not seem to make any statistically significant difference compared with possessing a primary degree. Will will comment more on this result in the conclusion. Those with a bachelor now appear only 1.025 times more productive than the reference group (vs. 1.78 with NLSQ). And workers possessing masters degree appear only 1.06 times more productive (vs. 3 with NLSQ). Similar falls are observed among our estimates of (marginal) labour costs. Mush lower point estimimates when controling for unobserved heterogeneity across firms (or equivalently, when resorting a within firm growth) hints at the existence of assortative matching between workers and firms. Workplaces that are intrinsically more productivie attract better-educated individuals, in line with the skill segregation assumption put forth by Kremer (1993) or Sattinger (1993).

Another important result from NLSQ-FE (2) in Table 2 is that the substitubility parameter ρ is not statistically different from 1, meaning that we find no evidence of human capital spillovers in our data. This result is confirmed when we restimate NLSQ-FE using only firms form the service industry. One can indeed see at the bottom of Table 7 (model 2a) that ρ is very close to one, leading to very high elasticities of substitution (ie. perfect substitutability of education types).

We now turn to our preferred models, that also account for the fact that NLSQ suffers from endogeneity bias. We first combine ACF with firm fixed effects (FE-ACF). Results (model 3, Table 6) show that the relative marginal productivity for secondary and bachelor degree do not reach statistical significance. By contrast, workers with masters still display a significant productivity advantage of 1.18 (18% more) with respect to primary school graduates. Our second preferred model is S-GMM. Estimates, for both relative productivity and labour cost, are somewhat larger than those delivered by FE-ACF. A worker with a master degree appears 1.27 times (27%) more productive than workers with a primary school attainment (vs. 18% with FE-ACF). This could be explained by the fact that S-GMM does not completely evacuate data in level.[25] This said, S-GMM results largely comfort the evidence gathered: more educated workers, in particular master graduates, and to a lesser extent workers with a bachelor degree, contribute more to producitivy growth than the reference category (at most primary school graduates). It is again quite stricking that both models (3) and (4) suggest an absence of producitivy advantage for workers with an upper secondary degree.

Focusing on estimates of (relative) marginal contribution of education to labour cost, we come to a similar conclusion. Broadly speaking, the comparison of productivity and labour cost estimates delivered by our preferred models suggests an overall alignment. Note that both FE-ACF and S-GMM deliver estimates of education-related producitivity (and labour cost) gains that are of *higher magnitude* than those delivered by the models that only control for fixed efffects (model 2). This tentatively suggests that the simultaneity bias the former control for i) exists ii) is biasing estimates

---

[25] As is explained at the end of Section 2,.S-GMM consists of a system of two equations estimated simultaneously. One corresponds to the first-difference equation, where the instruments are the (lagged) labour inputs *in level*. The second equation consists of using regressors *in level*, with (lagged) first-differenced of the endogenous variables as instruments. S-GMM.

downwards. The latter observation is compatible with the idea formulated above that Belgian firms tend to reduce the share of less-educated workers (and thus increase the share of more educated ones) when confronted to negative productivity shocks.

Table 5- NLSQ, NLSQ-FE (growth), coefficients estimates (+ standard errors)

	(1) NLSQ		(2) NLSQ-FE	
	Prod	Labour cost	Prod	Labour cost
α	0.0704***	0.154***	0.0107***	0.0340***
	(46.54)	(42.20)	(4.34)	(12.15)
β	0.936***	0.807***	0.815***	0.697***
	(322.43)	(144.98)	(49.44)	(38.95)
λ_{sec} ;$\lambda^{w}{}_{sec}$	1.052***	0.926***	0.993***	1.006***
	(116.44)	(69.57)	(177.37)	(65.56)
λ_{bach} ;$\lambda^{w}{}_{bach}$	1.049***	1.058***	1.008***	1.026***
	(91.91)	(52.15)	(166.82)	(56.61)
λ_{mast};$\lambda^{w}{}_{mast}$	1.667***	1.546***	1.030***	1.050***
	(59.58)	(34.81)	(104.88)	(36.75)
Controls	Year*nace2 fixed effets		Firm fixed effects	
Nobs	69,148		65,970	
ρ	1.052	0.926	0.993	1.006
Prob ρ =1	0.000	0.000	0.240	0.684
$\sigma \equiv 1/(1- \rho)$	-19.215	13.561	151.883	-160.277
Implied marginal productivities (ref=1= workers with primary education)				
Rmp/mlc sec($)	**1.154**	**0.934**	**0.996**	**1.037**
Rmp/mlc bach	**1.738**	**1.471**	**1.025**	**1.055**
Rmp/mlc mast	**3.038**	**2.306**	**1.059**	**1.080**
Prob $\lambda_{sec}=1$	0.000	0.004	0.177	0.154
Prob $\lambda_{bach=1}$	0.000	0.000	0.002	0.078
Prob $\lambda_{mast=1}$	0.000	0.000	0.004	0.144

Source: Bel-first & Carrefour; Std errors are robust to firm-level clustering

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

($) Rmp_j= mean of $\lambda_j(L_{ij}/L_{ir})^{\rho-1}$ *$Rmlc_j$*= mean of $\lambda^{w}{}_{j}(L_{ij}/L_{ir})^{\rho^{w}-1}$

Table 6 - ACF-FE S-GMM, coefficients estimates (+ standard errors)

	(3) ACF-FE		**(4) S-GMM**	
	Prod	Labour cost	Prod	Labour cost
α	0.161^{***}	0.0982^{***}	0.0886^{***}	0.0370^{***}
	(39.22)	(78.32)	(55.58)	(52.18)
β-1	-0.270^{***}	-0.108^{***}	0.0957^{***}	0.0109^{***}
	(-31.34)	(-31.86)	(17.39)	(4.45)
$\eta_{\text{sec}};\eta^{w}_{\text{sec}}$	0.0199	0.00311	-0.0508^{**}	-0.0874^{***}
	(0.93)	(0.29)	(-2.61)	(-10.07)
$\eta_{\text{bach}};\eta^{w}_{\text{bach}}$	0.0241	0.0730^{***}	0.119^{***}	0.0279^{**}
	(0.90)	(5.38)	(4.97)	(2.62)
$\eta_{\text{mast}};\eta^{w}_{\text{mast}}$	0.133^{*}	0.167^{***}	0.303^{***}	0.297^{***}
	(2.05)	(5.68)	(9.15)	(20.13)
Controls		Firm fixed effects		
Nobs	54,128		54,796	
Rmp/mlc sec($)	**1.027**	**1.003**	**0.954**	**0.914**
Rmp/mlc bach	**1.033**	**1.082**	**1.109**	**1.028**
Rmp/mlc mast	**1.182**	**1.187**	**1.277**	**1.293**
Prob $\lambda_{sec}=1$	0.351	0.773	0.009	0.000
Prob $\lambda_{bach=1}$	0.369	0.000	0.000	0.009
Prob $\lambda_{mast=1}$	0.041	0.000	0.000	0.000

Source: Bel-first & Carrefour; Std errors are robust to firm-level clustering

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

($) Rmp$_j$ $= \lambda_j = \eta_j/\beta + 1$; **Rmlc$_j$** $= \lambda^w_j = \eta^w_j/\beta^w + 1$

Table 7 – NLSQ-FE (growth), ACF-FE S-GMM, coefficients estimates (+ standard errors) . Services only

	(2a) NLSQ-FE			**(3a) ACF-FE**		**(4a) S-GMM**	
	Prod	Labour cost		Prod	Labour cost	Prod	Labour cost
α	0.0379***	0.0120**	α	0.0917***	0.0357***	0.0903***	0.0364***
	(9.65)	(3.25)		(72.88)	(63.52)	(58.44)	(52.88)
β	0.675***	0.798***	β-1	0.0748***	0.0110***	0.0883***	0.0117***
	(33.03)	(38.42)		(18.11)	(5.98)	(16.73)	(4.97)
λ_{sec} ;$\lambda^{w}{}_{sec}$	1.039***	1.005***	η_{sec};$\eta^{\text{w}}{}_{\text{sec}}$	-0.155***	-0.140***	-0.0357	-0.0796***
	(40.51)	(118.21)		(-7.38)	(-15.00)	(-1.88)	(-9.43)
λ_{bach} ;$\lambda^{w}{}_{bach}$	1.055***	1.025***	η_{bach};$\eta^{\text{w}}{}_{\text{bach}}$	0.143***	0.0389***	0.127***	0.0323**
	(30.53)	(84.43)		(5.52)	(3.36)	(5.46)	(3.12)
λ_{mast};$\lambda^{w}{}_{mast}$	1.091***	1.054***	η_{mast};$\eta^{\text{w}}{}_{\text{mast}}$	0.338***	0.311***	0.332***	0.327***
	(18.13)	(43.07)		(9.21)	(18.94)	(10.39)	(22.98)
Controls	Firm fixed effects		Controls	Firm fixed effects			
Nobs	40,808			33,837		33,837	
ρ	0.989	0.995					
Prob ρ =1	0.538	0.465					
$\sigma \equiv 1/(1-\rho)$	88.452	189.256					
Implied marginal productivities (ref=1= workers with primary education)			Implied marginal productivities (ref=1= workers with primary education)				
Rmp/mlc sec($)	**1.018**	**0.995**	***Rmp sec($$)***	**0.856**	**0.861**	**0.967**	**0.921**
Rmp/mlc bach	**1.044**	**1.020**	***Rmp bach***	**1.133**	**1.038**	**1.117**	**1.032**
Rmp/mlc mast	**1.085**	**1.051**	***Rmp mast***	**1.315**	**1.307**	**1.305**	**1.323**
Prob $\lambda_{sec}=1$	0.127	0.543	Prob $\lambda_{sec}=1$	0.000	0.000	0.060	0.000
Prob $\lambda_{bach=1}$	0.111	0.037	Prob $\lambda_{bach=1}$	0.000	0.001	0.000	0.002
Prob $\lambda_{mast=1}$	0.129	0.028	Prob $\lambda_{mast=1}$	0.000	0.000	0.000	0.000

Source: Bel-first & Carrefour; Std errors are robust to firm-level clustering

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; **($)** *Rmp$_j$*= mean of $\lambda_j(L_{ij}/L_{ir})^{\rho-1}$ *Rmlc$_j$*= mean of $\lambda^{w}{}_{j}(L_{ij}/L_{ir})^{\rho^{w}-1}$ **($$)Rmp$_j$** $= \lambda_j = \eta_j/\beta + 1$; **Rmlc$_j$** $= \lambda^{w}{}_{j} = \eta^{w}{}_{j}/\beta^{w} + 1$

5. Conclusions

In this paper, we use firm-level micro data to try to validate the fact that the abundantly-documented relationship between education and wages/labour costs is causally driven by a positive relationship between education and firm-level productivity. The existing empirical literature is primarily based on individual-level wage data, with no direct measure of productivity. In the Mincerian tradition, productivity is assumed to be relected in wage. There is a small literature that exploits firm-level evidence (which contains direct measure of producitivy alongside information on wage or labour cost) provides some suggestive evidence of the link between education, productivity and pay at the level of firms. But, despite offering plausible and intuitive results, it essentially relies on cross-sectional evidence and most of it does not tackle two crucial aspects of the endogeneity of production and wage functions: heterogeneity and simultaneity. We have tried to fill that void using good-quality Belgian data, covering the private economy during the 2000s, analysed with state-of-the-art panel models that control for heterogeneity and simultaneity. Our results are essentially sixfold.

First, marginal productivity of workers with a tertiary education (mainly 4-5 year Masters) is positively associated with firm-level overall labour productivity. Referring to our preferred models that control for firm-level unobserved heterogeneity and simultaneity bias (FE-ACF, S-GMM), a worker with a university degree appears 18% (FE-ACF) to 27% (S-GMM) more productive than workers with a primary school attainment or less. Using Psacharopoulos' (1981) 'shortcut' method to estimating rate of return, and assuming that university graduates have studied during 10 additional years compared with the reference group (workers with at most a primary degree), these figures correspond to rates of return of 1.8 to 2.7% per year of schooling; somewhat below the 5.2% obtained by de la Croix & Vandenberghe (2004) when estimating a Mincerian gross montly wage equation using Belgian data.

For those with a 2-3 year Bachelor degree similar estimates range from 3.3% (FE-ACF)[26] to 11% (S-GMM). Those for individuals with secondary school attainment are not statistically different from zero. Particularly for Masters, we find strong evidence supportive of labour costs' alignment on

---

[26] Not statistically significant.

marginal productivity. In short, the traditional relationship between individual wages and education, highlighted in innumerable estimations of Mincerian equations, could be driven by a positive link between education and the capacity of firms to be more productive. Belgium is generally considered as a country where labour issues - in particular those related to wages and labour cost formation - are highly regulated and determined by centralised tripartite bargaining. Yet, his paper provides evidence that, at the level of the firm, productivity remains a key determinant of pay. The alignment of marginal labour cost on marginal productive that we observe is compatible with the textbook assumption of spot labour markets.

Second, estimated substitubility parameter ρ are not statistically different from 1, meaning that we find no evidence of human capital spillovers in our data. In other words, the productivity of one worker with a given educational attainment is not affected/improved by the presence of less- or better-educated colleagues.

Third, our regressions with firm-fixed effects (FE) estimates of human capital-related productivity gains are (sometimes much) smaller in magnitude than those emerging from regressions without firm FE, but still statistically significant contrary to those obtained by Haltiwanger et al. (1999) who analysed productivity changes within US firms between 1985 and 1996. We interpret this as an indication that the gradual rise of the educational attainment of the workforce, in particular the rise of the number of univeristy/masters graduates[27], is good for the productivity of Belgian firms. At the same time, cross-sectional evidence stemming from standard regressions is conducive to systematic exaggeration of human-capital-related productivity gains. This is because better-educated individuals self-select in, or are selected by, those of the Belgian private firms that are intrinsically more productive; something a priori in line with Kremer's assumption of skill segregation at the level of the firm (Kremer, 1993).

Fourth, when we account for firm-heterogeneity and simultaneity bias with the ACF methodology, we obtain higher point estimates than those delivered by models that only control for FE. We conjecture that, in our setting, the simultaneity bias is significant and negative; meaning that Belgian firms tend to reduce(increase) the share of less-(better) educated workerswhen confronted to negative productivity shocks.

[27] Haltiwanger et al. (1999) used a very loose definition of education levels (low, medium, high).

Fifth, its is quite stricking that all our models (if we excluse the one that controls neither for unoberseved heterogeneity (FE) nor for similulaneity bias) conclude to the absence of producitivy advantage for workers with an upper secondary degree compare to those with a primary or no educational attaiment. One might see this as validation of distance-to-frontier idea of Aghion et al., (2006).[28] Another assumption, that does not exclude the previous one, is that our results capture the very high degree of negative selection affecting people with little or no education record in the context of the Belgian labour market. The employment rate among these individuals is very low by international standards and it possible that only the most productive are visible in our data.

Sixth, in terms of labour demand, estimates delivered by our preferred models (FE-ACF, S-GMM) are supportive of the alignment of marginal productivity on marginal labour cost. This tentatively suggests that private firms located in Belgium face no financial incentives to modify the educational mix of their workforce.

References

Ackerberg, D.A, Caves, K. and Frazer, G. (2006), 'Structural Identification of Production Functions', Department of Economics, Working Paper, UCLA.

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

Aubert, P. and B. Crépon (2007). 'Are older workers less productive. Firm-level evidence on age-productivity and age-wage profiles', mimeo, (French version published in *Economie et Statistique*, No 368).

Aubert. P. and B. Crépon (2003). "La productivité des salariés âgés : une tentative d'estimation". *Economie et Statistique*, No 368, pp. 95-119.

Blundell, R. and S. Bond (1998), Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, *Journal of Econometrics*, 87, pp. 115-143.

Card, D.(1999), The causal effect of education on earnings, : O. Ashenfelter & D. Card (editors),Handbook of Labour Economics, edition 1, 3(30), pp. 1801-1863, Elsevier.

Cataldi, A., S. Kampelmann and F. Rycx (2011), Productivity-Wage Gaps Among Age Groups: Does the ICT Environment Matter?, *De Economist*, 159(2), pp. 193-221.

de la Croix, D. and Vandenberghe, V. (2004), *Human capital as a factor of growth and employment at the regional level. The case of Belgium, Report for the European Commission*, DG for

[28] Aghion and his co-authors develop a theoretical model showing that skilled labor (our Masters) has a higher growth-enhancing effect closer to the technological frontier under the reasonable assumption that innovation is a relatively more skill-intensive activity than imitation. Also, they provide evidence in favor of this prediction using a panel dataset covering 19 OECD countries between 1960 and 2000 and explain why previous empirical research had found no positive relationship between initial schooling level and subsequent growth in rich countries. In particular, they show that in OECD economies it is crucial to isolate the two separate margins of primary/secondary and tertiary education.

Employment and Social Affairs, Brussels.

Dostie, B. (2011), Wages, Productivity and Aging, *De Economist*, 159(2), pp. 139-158.

Duffy J., Ch. Papageorgiou and F. Perez-Sebastian (2004), Capital-Skill Complementarity? Evidence from a Panel of Countries, *The Review of Economics and Statistics*, 86(1), pp. 327-344.

Galindo-Rueda, F. and Haskel, J., 2005, Skills, Workforce Characteristics and Firm-Level Productivity: Evidence from the Matched ABI/Employer Skills Survey, *IZA Discussion Paper,* No. 1542.

Griliches Z. (1969), Capital-skill complementarity, *Review of Economics and Statistics*, 51, pp. 465-468.

Griliches. Z. and J. Mairesse, (1995), Production functions: the search for identification, *NBER working paper, No 5067*, NBER, Ma.

Hægeland, T. and Klette, T. J. (1999), Do Higher Wages Reflect Higher Productivity? Education, Gender and Experience Premiums in a Matched Plant-Worker Data Set. In: *The Creation and Analysis of Employer-Employee Matched Data*, edited by Haltiwanger J., Lane J. , Spletzer J. R., Theeuwes J. and Troske K., Amsterdam: North Holland.

Haltiwanger, J, and J. Lane and J. Spletzer, (1999), "Productivity Differences across Employers: The Roles of Employer Size, Age, and Human Capital", *American Economic Review*, 89(2), pp. 94-98.

Hellerstein, J. and D. Neumark (2007), Production Function and Wage Equation Estimation with Heterogeneous Labour: Evidence from a New Matched Employer-Employee Data Set, in *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, Ernst R. Berndt and Charles R. Hulten (editors).

Hellerstein, J. K. and D. Neumark (1999), ‘Sex, Wages and Productivity: An Empirical Analysis of Israel Firm-level Data’, *International Economic Review*, 40(1), pp. 95-123.

Hellerstein, J.K., D. Neumar, and K. Troske (1999), ‘Wages. Productivity and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations’, *Journal of Labour Economics*, 17(3), pp. 409-446.

Kremer, M. (1993), The O-Ring Theory of Economic Development, *The Quarterly Journal of Economics*, MIT Press, 108(3), pp 551-75.

Krueger, A.B. and M. Lindahl, 2001. "Education for Growth: Why and for Whom?," *Journal of Economic Literature*, 39(4), pp. 1101-1136.

Levinsohn. J. and A. Petrin (2003), Estimating production functions using inputs to control for unobservables, *Review of Economic Studies*, 70 (2), pp. 317-341.

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

Mincer, J. (1974); *Schooling, Experience, and Earnings*, NBER Press, New York.

Olley, G.S, and A. Pakes (1996), ‘The Dynamics of Productivity in the Telecommunications Equipment Industry’, *Econometrica*, 64(6), pp. 1263-1297.

Prskawetz, A., Mahlberg B. and Skirbekk V., 2007, “Firm productivity, and educational structure in Austrian industries” in A. Mason and R. Clark (eds.) *Intergenerational Transfers and the Macroeconomy* , Edward Elgar.

Psacharopoulos, G. (1981) “Returns to Education: An Updated International Comparison”, *Comparative Education*, 17(3), pp 321-341.

Rigo, M., Vandenberghe, V. and Waltenberg. F. (2012), Ageing and Employability. Evidence from Belgian Firm-Level Data (2012), *Journal of Productivity Analysis*, DOI: 10.1007/s11123-012-0297-8.

Sattinger, M. (1993), Assignment Models of the Distribution of Earnings, *Journal of Economic Literature*, 31(2), pp. 831-80.

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

Turcotte J. and Whewell Rennison L., 2004, "The Link between Technology Use, Human Capital, Productivity and Wages: Firm-level Evidence", *International Productivity Monitor*, Centre for the Study of Living Standards, vol. 9, 25-36.

van Ours, J.C. and L, Stoeldraijer (2011), Age, Wage and Productivity in Dutch Manufacturing, *De Economist*, 159(2), pp. 113-137.

Vandenberghe, V. (2011a), Firm-level Evidence on Gender Wage Discrimination in the Belgian Private Economy, *Labour: Review of Labour Economics and Industrial Relations*, 25(3), pp. 330-349

Vandenberghe, V. (2011b), Boosting the employment rate of older men and women. An empirical assessment using Belgian firm-level data on productivity and labour costs, *De Economist*, 159(2), pp. 159-191.

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

Wolff, E.N. and M. Gittleman. (1993), The Role of education in Productivity Convergence: Does Higher Education Matter?, in Szirmai A., Van Ark B., Pilat D., *Explaining Economic Growth*, Elsevier.