IJM
35,8

Received 8 June 2012
Revised 15 January 2013
24 May 2013
Accepted 9 August 2013

Assessing education's contribution to productivity using firm-level evidence

Lara Lebedinski and Vincent Vandenberghe
School of Economics, IRES-Université catholique de Louvain, Louvain-la-Neuve, Belgium

Abstract

Purpose – There is plenty of individual-level evidence, based on the estimation of Mincerian equations, showing that better-educated individuals earn more. This is usually interpreted as a proof that education raises labour productivity. Some macroeconomists, analysing cross-country time series, also support the idea that the continuous expansion of education has contributed positively to growth. Surprisingly, most economists with an interest in human capital have neglected the level of the firm to study the education-productivity-wage nexus. And the few published works considering firm-level evidence are lacking a proper strategy to cope with the endogeneity problem inherent to the estimation production and wage functions. The purpose of this paper is to aim at providing estimates of the causal effect of education on productivity and wage labour costs.

Design/methodology/approach – This paper taps into a rich, firm-level, Belgian panel database that contains information on productivity, labour cost and the workforce's educational attainment to deliver estimates of the causal effect of education on productivity and wage/labour costs. Therefore, it exclusively resorts to within firm changes to deal with time-invariant heterogeneity bias. What is more, it addresses the risk of simultaneity bias (endogeneity of firms' education-mix choices in the short run) using the structural approach suggested by Ackerberg *et al.* (2006), alongside more traditional system-GMM methods (Blundell and Bond, 1998) where lagged values of labour inputs are used as instruments.

Findings – Results suggest that human capital, in particular larger shares of university-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, the authors find evidence that the Mincerian relationship between education and individual wages is driven by a strong positive link between education and firm-level productivity.

Originality/value – 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 been much more investigated at the level of the firm by industrial or labour economists (Hellerstein *et al.*, 1999; Aubert and Crépon, 2003; Hellerstein and Neumark, 2007; Vandenberghe, 2011a, b, 2012; Rigo *et al.*, 2012; Dostie, 2011; van Ours and 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 and Klette (1999); Haltiwanger *et al.* (1999). Other notable papers examining a similar question are Galindo-Rueda and Haskel (2005), Prskawetz *et al.* (2007) and Turcotte and Whewell Rennison (2004). 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 and Mairesse, 1995): first, heterogeneity bias (unobserved time-invariant determinants of firms' productivity that may be correlated to the workforce structure) and second, simultaneity bias (endogeneity in input choice, in the short-run, that includes the workforce mix of the firm). While the authors know that labour productivity is highly heterogeneous across firms (Syverson, 2011), 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

JEL Classifications — J24, E24, C51

International Journal of Manpower
Vol. 35 No. 8, 2014
pp. 1116-1139

0143-7720
DOI 10.1108/IJM-06-2012-0090

employment decision and thus the workforce mix. Yet these shocks and the resulting decisions by firms' manager are unobservable by the econometrician. Hægeland and Klette (1999) try to solve this problem by proxying productivity shocks with intermediate goods, but their methodology inspired by Levinsohn and Petrin (2003) suffers from serious identification issues due to collinearity between labour and intermediate goods (Ackerberg *et al.*, 2006).

Keywords Heterogeneity, Education, Endogeneity, Firm-level productivity, Labour cost, Simultaneity

Paper type Research paper

1. Introduction

There exists substantial evidence, based on the analysis of individual data, that general education (schooling) increases wages. Card (1999) for instance, summarises various Mincer-inspired studies and concludes that the impact of a year of schooling on wages is about 10 per cent. Similar results exist for Belgium (de la Croix and 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 and 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, etc.). 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 data sets 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 and Crépon, 2003; Hellerstein and Neumark, 2007; Vandenberghe, 2011a, b, 2012; Rigo *et al.*, 2012; Dostie, 2011; van Ours and 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 and Klette (1999); Haltiwanger *et al.* (1999). Other notable papers examining a similar question are Galindo-Rueda and Haskel (2005), Prskawetz *et al.* (2007) and Turcotte and Whewell 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 and Mairesse, 1995): first, heterogeneity bias (unobserved time-invariant determinants of firms' productivity that may be correlated to the workforce structure[2]) and second, 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 and Klette (1999) try to solve this problem by proxying productivity shocks with intermediate goods, but their methodology inspired by Levinsohn and Petrin (2003) suffers from serious identification issues due to collinearity between labour and intermediate goods (Ackerberg *et al.*, 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 9,970 firms located in Belgium for the period 2002-2011, on key 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 and 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 and Bond, 1991; Aubert and Crépon, 2003; or van Ours and 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-generalised 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 and Pakes (1996), Levinsohn and Petrin (2003) (OP, LP henceforth), and more recently by Ackerberg *et al.* (2006) (ACF henceforth), are somewhat more structural in nature. They consist of using observed intermediate input decisions (i.e. purchases of raw materials, services, electricity, etc.) 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 and Crépon, 2003, 2007; van Ours and Stoeldraijer, 2011; Cataldi *et al.*, 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 university 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 23 per cent (FE-ACF) to 42 per cent (S-GMM) more productive than a worker with a primary education attainment or less. Workers with a two-year college degree or only secondary school appear to be 3.4 per cent[11] (FE-ACF) to 18.5 per cent more productive as primary school graduates. Simultaneously, the labour cost premium associated to workers university degree is 17.3 per cent (FE-ACF) to 43.8 per cent (S-GMM), and 5 per cent[12] (FE-ACF) to 12.4 per cent (S-GMM) for those with a two-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 organised 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 data set. Section 4 contains the econometric results and Section 5 our main conclusions.

2. Methodology

In order to estimate education-productivity profiles, following most authors in this area, we consider a Cobb-Douglas production function (Hellerstein *et al.*, 1999; Aubert and Crépon, 2003, 2007; Dostie, 2011; van Ours and Stoeldraijer, 2011):

$$\ln\left(Y_{it}/L_{it}\right) = \ln A + \alpha \ln QL_{it} + \beta \ln K_{it} + \gamma F_{it} - \ln L_{it} \tag{1}$$

where Y_{it}/L_{it} is the average value added (Y_{it}) per unit of labour (productivity hereafter) in firm i at time t, QL_{it} is an aggregation of different types of workers, K_{it} is the stock of capital and L_{it} is total labour. F_{it} contains (calendar) year × sector dummies and other firm-level controls (more on this below and in Section 3).

The variable that reflects the heterogeneity of the workforce is the quality of labour index QL_{it}. Let L_{ikt} be labour input of type k (e.g. primary school, secondary school, two-year college and university workers' contribution) in firm i at time t, and μ_{ik} type k labour productivity. We assume that various types of labour are perfectly substitutable[13] with different marginal products. As each type of labour k is assumed to be an input in quality of labour aggregate, the latter can be specified as:

$$QL_{it} = \mu_{i0}L_{i0t} + \mu_{i1}L_{i1t} + \mu_{i2}L_{i2t} + \mu_{i3}L_{i3t} \tag{2}$$

0, being the workers with at most a primary education attainment; 1, the workers with (at most) an upper-secondary education attainment; 2, those with (at most) a two-year college attainment; and 3, those with a four-year university degree. Where $L_{it} \equiv \sum_k L_{ikt}$ is the total labour input in the firm, μ_{i0} the marginal productivity of the reference category of labour (here primary school worker) and μ_{ik} that of the other types of workers.

If we further assume that a worker has the same marginal product across firms, we can drop subscript i from the marginal productivity coefficients. After taking logarithms and doing some rearrangements Equation (2) becomes:

$$\ln QL_{it} = \ln\mu_0 + \ln L_{it} + \ln\left(1 + \sum_{k>0}(\lambda_k - 1)P_{ikt}\right) \tag{3}$$

where $\lambda_k \equiv \mu_k/\mu_0$ is the relative productivity of type k worker and $P_{ikt} \equiv L_{ikt}/L_{it}$ the proportion/share of type k workers over the total number of workers in firm i.

Using the approximation that $\ln(1+x) \approx x$, (3) can be simplified as:

$$\ln QL_{it} = \ln\mu_0 + \ln L_{it} + \sum_{k>0}(\lambda_k - 1)P_{ikt} \tag{4}$$

And the production function becomes:

$$\ln\left(Y_{it}/L_{it}\right) = \ln A + \alpha\left[\ln\mu_0 + \ln L_{it} + \sum_{k>0}(\lambda_k - 1)P_{ikt}\right] + \beta \ln K_{it} + \gamma F_{it} - \ln L_{it} \tag{5}$$

or equivalently:

$$\ln\left(Y_{it}/L_{it}\right) = B + (\alpha - 1)l_{it} + \eta_1 P_{i1t} + \ldots \eta_N P_{iNt} + \beta k_{it} + \gamma F_{it} \tag{6}$$

where:

$$B = \ln A + \alpha \ln\mu_0$$

$$\eta_h = \alpha(\lambda_k - 1);\ \lambda_k = \mu_k/\mu_0; \quad k = 1\ldots N$$

$$l_{it} = \ln L_{it}; k_{it} = \ln K_{it}$$

which can be estimated econometrically.

Note first that (6), being loglinear in P_{ikt}, has coefficients that can be directly interpreted as the percentage change in the firm's average labour productivity of a one 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 k worker's relative marginal productivity, (i.e. λ_k), coefficients η_k have to be divided by α, and 1 needs to be added to the result.

A similar approach can be applied to a firm's average labour cost. If we assume that firms operating in the same labour market pay the same wages to the same category of workers, we can drop subscript i from the remuneration coefficient π[14]. Let π_k stand for the (not directly observed) remuneration[15] of type k workers ($k=0$ primary school being reference type). Then the labour cost per hour becomes:

$$W_{it}/L_{it} = \sum\nolimits_k \pi_k L_{ikt} \Big/ L_{it} = \left[\pi_0 L_{it} + \sum\nolimits_{k>0} (\pi_k - \pi_0) L_{ikt}\right] \Big/ L_{it} + \gamma^w F_{it} \tag{8}$$

where F_{it} still contains (calendar) year × sector dummies and other firm-level controls. Taking the logarithm, using again $\log(1+x) \approx x$, we can approximate this by:

$$\ln\left(W_{it}/L_{it}\right) = \ln\pi_0 + \sum\nolimits_{k>0} (\lambda_k^w - 1) P_{ikt} + \gamma^w F_{it} \tag{9}$$

Where the Greek letter $\lambda_k^w \equiv \pi_k/\pi_0$ denotes the relative remuneration of type k workers ($k>0$) with respect to the ($k=0$) reference group, and $P_{ikt}=L_{ikt}/L_{it}$ is again the proportion/share of type k labour our over the total labour input in firm i.

The logarithm of the labour cost unit of labour finally becomes:

$$\ln\left(W_{it}/L_{it}\right) = B_w + \Phi_1 P_{i1t} + \ldots \Phi_N P_{iNt} + \gamma^w F_{it} \tag{10}$$

where:

$$B_w = \ln \pi_0$$

$$\Phi_k = \left(\lambda_w^k - 1\right); \lambda_w^k = \mu_k/\mu_0; k = 1\ldots N$$

Like in the productivity Equation (6) coefficients Φ_k capture the sensitivity to changes of the educational structure of the labour force (P_{ikt}). 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 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 $\lambda_{i3,0} > \lambda_{i2,0} > \lambda_{i1,0} > 1$. And if Belgian labour markets rewards human capital at its marginal produce (i.e. as assumed in the standard human-capital model), then one should observe that $\lambda_{i3,0} \approx \lambda_{i3,0}^w$, $\lambda_{i2,0} \approx \lambda_{i2,0}^w$ and $\lambda_{i1,0} \approx \lambda_{i1,0}^w$. 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. Our preferred models do not enable us to test statistically the alignment of marginal productivity and marginal cost; thus our interpretation of productivity – cost equalisation can be only suggestive. 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 a labour market policy interest, 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 university degree ($\lambda_{i3,1} > \lambda^{w}_{i3,1}$) 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 ($\lambda_{i3,1} < \lambda^{w}_{i3,1}$) should rather entice them to reduce that share.

Remember that in all the econometric specifications of productivity (6) and wage estimation (10) we include the vector of controls F_{it} containing year × sector[16] dummies. This allows for systematic and proportional productivity variation among firms along this dimension. This assumption can be seen to expand the model by controlling for year and sector-specific productivity shocks or trends, labour quality and intensity of efficiency wages differentials across sectors and other sources of systematic productivity differentials (Hellerstein *et al.*, 1999). More importantly, since the data set we use does not contain sector price deflators, the introduction of these dummies can control for asymmetric year-to-year variation in the price of firms' outputs at the sector level. An extension along the same dimensions is made with respect to the labour cost equation. Of course, the assumption of segmented labour markets, implemented by adding linearly to the labour cost equation the set of year/sector dummies, is valid as long as there is proportional variation in wages by age group along those dimensions. Detailed discussion of all firm-level controls included in F_{it} will be presented in the data section below.

But, as to a proper identification of the causal links, the main challenge consists of dealing with the various constituents of the residual ε_{it} of Equation (9)[17]. We assume that the latter has a structure that comprises three elements:

$$\varepsilon_{it} = \tau_{it} + \theta_i + \sigma_{it} \tag{11}$$

where $\text{cov}(\theta_i, P_{ik,t}) \neq 0$, $\text{cov}(\omega_{it}, P_{ik,t}) \neq 0$, $E(\sigma_{it}) = 0$.

In words, the ordinary least squares (OLS) sample-error term potentially consists of an unobservable firm fixed effect θ_i; a short-term shock τ_{it} (whose evolution may correspond to a first-order Markov chain), and is observed by the firm (but not by the econometrician) and (partially) anticipated by the firm, and 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, or the overall stock of human capital[18], firm-specific managerial skills, location-driven comparative advantages […][19]. And these might be correlated with the education structure of the firm's workforce, biasing OLS results. More educated workers for instance might be underrepresented among plants built a long time ago, that use older technology. However, the panel structure of our data allows for the estimation of models that eliminate fixed effects (FE). The results from FE can be interpreted as follows: a group (e.g. primary school, secondary school, college or university graduates) is estimated to be more (less) productive than another group if, within firms, a increase of that group's share in the overall workforce translates into productivity gains (loss).

This said, the greatest econometric challenge is to go around the simultaneity bias (Griliches and 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} (and similarly about its short-term dynamics). An anticipated downturn could translate into a recruitment freeze, and possibly also, into a multiplication of layoffs. A recruitment freeze affects presumably younger and more educated workers, and translates into falling share of more educated workers during negative spells, creating a positive correlation between more educated workers' share and productivity, thereby leading to overestimated estimates of their productivity (when resorting to OLS or even FD estimates). By contrast, if firms primarily layoff older less-educated workers when confronted with adverse demand shocks, we would expect the correlation between more educated workforce and productivity to be negative, leading to an underestimation of more educated workers' productivity with OLS or FD.

To account for the presence of this endogeneity bias, we first adopt the structural approach initiated by Olley and Pakes (1996), Levinsohn and Petrin (2003) and more recently by Ackerberg *et al.* (2006). The essence of the OP approach is to use some function of a firm's investment to control for (proxy) time-varying unobserved productivity, ω_{it}. The drawback of this method is that only observations with positive investment levels can be used in the estimation. Many firms indeed report no investment in short panels. LP overcome this problem by using material inputs (for instance raw materials, electricity, etc.) instead of investment in the estimation of unobserved productivity. They 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:

$$\ln\left(Y_{it}/L_{it}\right) = B + \varphi ql_{it} + \beta k_{it} + \gamma F_{it} + \varepsilon_{it} \tag{12a}$$

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

$$\varphi ql_{it} \equiv \alpha l_{it} + \eta_1 P_{i1t} + \eta_2 P_{i2t} + \ldots \eta_N P_{iNt} \tag{12b}$$

and the ACF error term:

$$\varepsilon_{it} = \omega_{it} + \sigma_{it} \tag{12c}$$

with $\omega_{it} \equiv \tau_{it} + \theta_{i,}$ meaning that unlike AC, we explicitly assume that ω_{it} contains a proper fixed effect.

Like ACF, we assume that firms' (observable) demand for intermediate inputs (int_{it}) is a function of the time-varying unobserved term ω_{it} (which comprises the fixed effect) as well as (log of) capital, and the quality of labour index (in logs) ql_{it} and its components[20]:

$$int_{it} = f_t(\tau_{it} + \theta_i, k_{it}, ql_{it}) \tag{13}$$

ACF further assume that this function f_t is monotonic in ω_{it} (i.e. $\tau_{it} + \theta_i$) and its other determinants, meaning that it can be inverted to deliver an expression of $\tau_{it} + \theta_i$ as a function of int_{it}, k_{it}, ql_{it}, and introduced into the production function:

$$\ln(Y_{it}/L_{it}) = B + \varphi ql_{it} + \beta k_{it} + \gamma F_{it} + \tau_{it} + \theta_i + \sigma_{it} \tag{14a}$$

with $\tau_{it} + \theta_i \equiv f_t^{-1}(int_{it}, k_{it}, ql_{it})$, leading to:

$$\ln(Y_{it}/L_{it}) = B + \varphi ql_{it} + \beta k_{it} + \gamma F_{it} + f_t^{-1}(int_{it}, k_{it}, ql_{it}) + \sigma_{it} \tag{14b}$$

In practice, how are the parameters φ, β and γ estimated? The ACF algorithm consists of two stages. We argue that only stage one needs to be adapted to account for fixed effects, by resorting to first-differences. In stage one, like ACF, we regress average productivity (i.e value added) on a composite term Φ_t that comprises a constant, a third-order polynomial expansion in int_{it}, k_{it}, ql_{it}, and our vector of controls added linearly. This leads to:

$$\ln(Y_{it}/L_{it}) = \Phi_{it}(int_{it}, ql_{it}, k_{it}, F_{it}) + \sigma_{it} \tag{15}$$

Note that φ, β and γ are clearly not identified yet, implying the need of a second stage[21]. Note in particular that Φ_t encompasses $f_t^{-1}(.)$ proxying $\omega_{it}+\theta_i$. 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 (17). The resulting FD-estimated coefficients – provided they are applied to variables in levels – deliver an unbiased prediction of Φ_{it}^{hat}. Specifically, Φ_{it}^{hat}, net of the noise term and firm-fixed effects, is calculated as $\Phi_{it}^{hat} = (v_{a1})^{FD}\, int_{it} + (v_{a2})^{FD}\, int_{it}^2 + \ldots + (v_{b1})^{FD}\, k_{it} + \ldots + (v_{c1})^{FD} ql_{it} + \ldots + (v_{d1})^{FD}\, int_{it} k_{it} \ldots$, where $(v_{a1})^{FD}$, $(v_{a2})^{FD}$ … represent the first-differenced coefficient estimates on the polynomial terms.

As an aside, also note the presence in Φ_{it} of a third-order polynomial expansion in (inter alia) ql_{it}. and its components, namely lnh_{it}, $lnhwk_{it}$, P_{i1}, P_{i2} *and* P_{i3} and capital. To this point, the production function (a Cobb-Douglas) has been specified so that workers of different types have different marginal products but are perfectly substitutable. Because this specification may be too restrictive, we should also consider evidence from estimates of a production function in which worker types between themselves and with capital are imperfect rather than perfect substitutes. Resorting to a translog specification is what Hellerstein *et al.* (1999) did in their seminal paper. But the ACF first-stage equation above (17) consists of regressing the log of productivity on a third-order polynomial that contains interaction terms between the various labour input variables and capital. When we report ACF and FE-ACF estimates below, one should thus bear in mind that we have gone part-way toward doing what Hellerstein *et al.* (1999) do when estimating translog production function to allow for imperfect substitutability across labour types and with capital. This feature will be called up when commenting the results in Section 4.

Returning to the ACF procedure, we basically argue that their second stage is unaffected by the modifications discussed above. Key is the idea that one can generate implied values for τ_{it} using first-stage estimates Φ_{it}^{hat} and candidate[22] values for the coefficients φ, β, γ[23]:

$$\tau_{it} = \Phi_{it}^{hat} - ql_{it}\varphi - \beta k_{it} - \gamma F_{it} \tag{16}$$

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

$$\tau_{it} = E[\tau_{it}|\tau_{it-1}] - \xi_{it} \tag{17}$$

That assumption simply amounts to saying that the realisation of τ_{it} depends on some function $g(.)$ (known by the firm) of $t-1$ realisation and an (unknown) innovation term ξ_{it}:

$$\tau_{it} = g(\tau_{it-1}) + \xi_{it} \tag{18}$$

By regressing non-parametrically (implied) τ_{it} on (implied) τ_{it-1}, τ_{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, β and γ. We would argue that residuals ξ_{it} are orthogonal to our controls F_{it}:

$$E[\xi_{it}|F_{it}] = 0 \tag{19a}$$

Analogous to 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 $t-1$, $t-2$ …, it must be uncorrelated with the implied innovation terms ξ_{it}:

$$E[\xi_{it}|k_{it}] = 0 \tag{19b}$$

Labour inputs observed in t are probably also chosen sometime before, although after capital – say in $t-b$, 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 $t-b-1$ (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:

$$E[\xi_{it}|ql_{it-1}, ql_{it-2}\ldots] = 0 \tag{19c}$$

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

$$E[\xi_{it}|\ln l_{it-1}, \ln l_{it-2}\ldots] = 0 \tag{20a}$$

$$E[\xi_{it}|P_{i1t-1}, P_{i1t-2}\ldots] = 0 \tag{20b}$$

$$E[\xi_{it}|P_{i2t-1}, P_{i2t-2}\ldots] = 0 \tag{20c}$$

$$E[\xi_{it}|P_{i3t-1}, P_{i3t-2}\ldots] = 0 \tag{20d}$$

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 and Crépon, 2003, 2007; van Ours and Stoeldraijer, 2011; Cataldi *et al.*, 2011)[24]. 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 and 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).

3. Data description

We are in possession of a large unbalanced panel of around 73,794 firm-year observations corresponding to the situation of about 9,970 firms, from all sectors forming the Belgian private economy, in the period 2002-2011. These firms are largely documented in terms of sector (SIC[25]), size, capital used, labour cost levels[26] and productivity (value added). These observations come from the Bel-first database[27], that most for-profit firms located in Belgium must feed tocomply with the legal prescriptions. All the firms occur at least four times in the panel; the maximum being ten times. This seems to be a reasonable time span as most economists would a priori consider that a proper assessment of how education/human capital affects productivity requires a medium-term perspective, meaning that firms' performance need to be observed over a certain number of years for human capital's beneficial contribution to production to become visible.

Descriptive statistics, forming this large sample are reported in Table I. Of prime interest in this paper is the breakdown by educational attainment. Table I shows that, during the observed period (2002-2011), about 73 per cent of the workforce of private for-profit firms located in Belgium have still, at most, an upper secondary school degree. Workers with a two-year college degree represented 19 per cent of the total workforce. Slightly < 8 per cent consisted of individuals with a (four-year) university degree. This means a mere 27 per cent of workers with a tertiary education background; clearly less than the percentage among the current generation of school leavers[28]. This discrepancy logically reflects the lower propensity of older generations to stay on beyond secondary education, and complete a tertiary education programme.

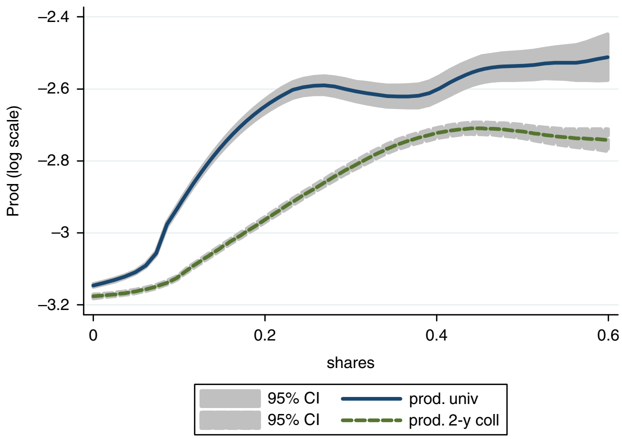
Labour costs used in this paper, which were measured independently of value added, include the value of all monetary compensations paid to the total labour force (both full- and part-time, permanent and temporary), including social security contributions paid by the employers, throughout the year. In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

Figure 1 displays how the (log of) productivity per hour (value added per hour) evolves with the share of university- and two-year college-educated workers for the period 2002-2011. These stylised facts suggest that, in the Belgian private economy, the productivity regularly rises with human capital, in particular between the 5 per cent and the 20 per cent range. Productivity seems to plateau for the share of two-year college workers above the 40 per cent threshold. At this stage any deductions can hardly be regarded as conclusive.

Variable	Obs.	Mean	SD	Min.	Max.
Value added per hour (log)	73,794	−3.011	0.603	−10.315	4.150
Labour cost per hour (log)	73,794	−3.467	0.375	−17.476	4.052
Number of workers (log)	73,794	3.751	1.148	0.693	10.287
Number of workers	73,794	113	481	1	29,344
Capital (th. €) (log)	73,794	7.841	1.814	0	17.437
Share of workers with at most a primary education attainment	73,794	0.152	0.276	0	1
Share of workers with at most a secondary education attainment	73,794	0.582	0.341	0	1
Share of workers with a two-year college attainment	73,794	0.188	0.231	0	1
Share of workers with a university attainment	73,794	0.078	0.156	0	1
Share of women	73,794	0.285	0.241	0	1
Number of hours worked annually per employee (log)	73,794	7.369	0.132	6.215	8.510
Share of workers with open-ended contracts	73,794	0.963	0.096	0	1

Source: Bel-first panel, our calculus

Table I.
Bel-first – basic descriptive statistics (unbalanced panel 2002-2011)



Note: Panel 2002-2011 (95 per cent confidence intervals)

Figure 1.
Productivity per hour according to share of university or two-year college-educated workers

Figure 2 is essentially stylised facts that do 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 firm fixed effects and other controls will allow us to draw more substantiated conclusions.

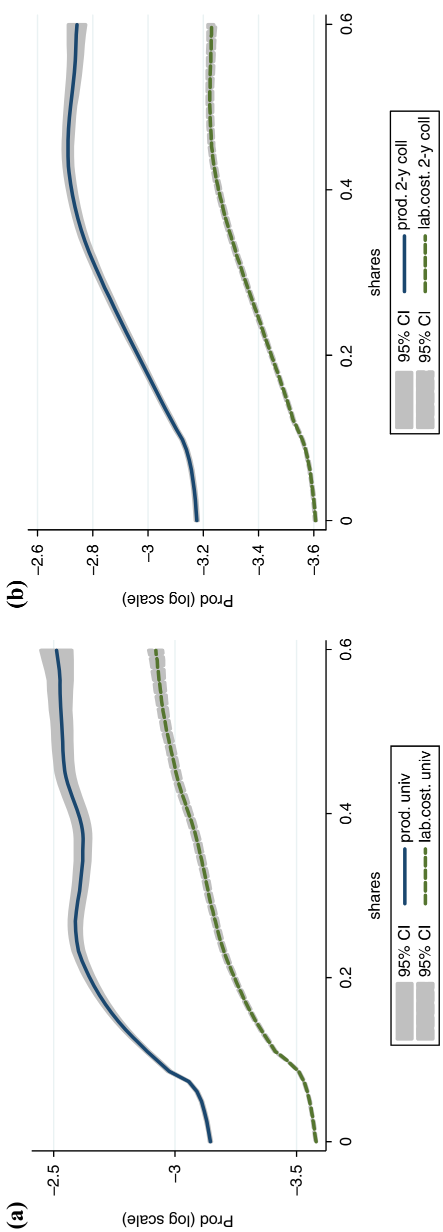
Curves on display correspond to a local polynomial smooth of y (i.e. log of average productivity,) on x (i.e. share of university- or two-year-college-educated workers). A kernel function of the Epanechnikov type is used to calculate the weighted local polynomial estimate.

Curves on display correspond to a local polynomial smooth of y (i.e. log of productivity per worker or log of labour cost per workers) on x (i.e. share of university- or two-year-college-educated workers). A kernel function of the Epanechnikov type is used to calculate the weighted local polynomial estimate.

Remember that all our regressions contain a vector of control F_{it} with year/sector interaction dummies. Additionally, F_{it} contains the share of women to control for gender based productivity differences. Our list of controls comprises also the share of workers with an open-ended contract (vs those with a temporary contract). These are individuals who may possess more firm-specific human capital, acquired via on-the-job learning, and have developed a degree of attachment to their employer that could positively affect productivity. These determinants of firm performance are conceptually different from general human capital acquired at school, college or university. Hence, we think it is worth trying to isolate their specific contribution.

Another possibility to better understand the data is to examine the evolution of the educational mix of the workforce over the observed period of time (2002 to 2011). Note that firms in our sample have experienced a marked rise of their share of better-educated workers (Table II). On average, their share of two-year-college-educated workers has climbed from 17.9 per cent to 19.2 per cent; their share of university-educated employees from 7.4 per cent to 8 per cent.

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



Notes: Panel 2002-2011 (95 per cent confidence intervals). (a) Share of university-educated workers; (b) share of 2-year college educated workers

Figure 2.
Productivity per hour and labour cost per hour

Table II.
Bel-first – unbalanced panel sample of 9,970 firms followed between 2002 and 2011

Year	Number of observations	Value added per hour (log)	Labour cost per hour (log)	Number of workers (log)	Number of workers	Capital (th. €) (log)	Share of workers with a two-year college attainment	Share of workers with a university attainment	Share of women	Share of workers with open-ended contracts	Number of hours worked annually per employee (log)
2002	5,166	−3.139	−3.593	3.906	129	7.788	0.179	0.074	0.270	0.970	7.392
2003	5,439	−3.095	−3.564	3.877	121	7.833	0.182	0.076	0.273	0.969	7.389
2004	5,676	−3.052	−3.533	3.861	123	7.870	0.184	0.078	0.277	0.969	7.400
2005	5,832	−3.013	−3.508	3.864	126	7.958	0.191	0.0805	0.277	0.970	7.389
2006	5,979	−2.970	−3.470	3.884	127	8.041	0.193	0.084	0.283	0.967	7.371
2007	6,316	−2.934	−3.443	3.869	128	8.111	0.198	0.085	0.289	0.965	7.373
2008	9,867	−3.019	−3.467	3.648	104	7.693	0.175	0.069	0.287	0.959	7.370
2009	9,767	−3.014	−3.421	3.651	101	7.758	0.189	0.079	0.288	0.961	7.344
2010	9,957	−2.979	−3.414	3.644	103	7.797	0.192	0.081	0.290	0.959	7.352
2011	9,795	−2.971	−3.397	3.627	101	7.767	0.192	0.080	0.294	0.956	7.355

Note: Basic descriptive statistics: mean models). Our measure is a direct one. It is the value (in thousand Euros 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 II that there has been a rise in the number of firms included in the panel between 2002 and 2008. This reflects the history of Bel-first's way of collecting data on educational attainment. Until 2007, reporting that information was optional and most of the (voluntary) respondents where large firms. After 2008, it became mandatory for all firms to communicate the information about the educational attainment of their workforce. The results specific to large-firms present in the panel from 2002 to 2011 are not shown here because they yield little additional insights. Qualitatively, they do not differ, but are available from the authors upon request.

4. Econometric results

Table III summarises the main econometric results. We first estimate the productivity and labour-cost regression with OLS (columns (1) and (2)). To account for firm unobserved heterogeneity we then turn to models with firm fixed effects (columns (3) and (4)). To account for simultaneity bias, we then turn to the structural approach proposed by ACF (columns (5) and (6)). Next are our preferred models, i.e. those presenting the enviable characteristic of dealing with heterogeneity and simultaneity, in an integrated way. Columns (7) and (8) display those delivered by the model that combines FD and the ACF intermediate-goods proxy idea. The last two columns (9) and (10) present results of the system-GMM estimation. All our regressions include year × sector fixed effects. The vector of controls F_{it} comprises the share of women and the share of workers with an open-ended contract. The coefficients in the table should be interpreted with respect to the reference group (ie. workers/employees with at most primary education). Notice that we cannot test the hypothesis that relative marginal productivity equals relative marginal cost, because we estimate separately the regressions.

The basic OLS regression puts forward the presence of a relative increase in productivity of attending a two-year college and university education with respect to primary education. Marginal productivity of a two-year college worker is estimated to be 1.27 times larger than of workers with a primary education attainment. That of a worker with a university degree appears to be 1.52 times that of the reference group. OLS-estimated marginal labour cost convey the idea that two-year college workers cost 1.30 times more to their employer than the reference group, whereas the corresponding ratio for university-educated workers is 1.76. Results for secondary education are not statistically significant. At this stage, we also test for the possibility that the error terms in the productivity and labour cost equations are correlated and a source of bias. We do so by estimating a seemingly unrelated regression (SUR) system. Results are not shown here but are available from the authors upon request. They are very similar the OLS results reported above.

Turning to the results of the FE model, we immediately see at the bottom of column (3) that a higher educational attainment translates into lower (marginal) productivity advantages compared with OLS. This stems from controlling for firm unobserved heterogeneity (i.e. FE), and it suggests that better-educated individuals, in particular those with a university background, concentrate in firms that are intrinsically more productive. Holding a secondary degree still does not seem to make any statistically significant difference compared with possessing a primary degree. Those with two-year college now appear only 1.075 times more productive than the

Table III.
Econometric results

	OLS		Fixed effect (FE)		ACF		FE-ACF		System-GMM à-la Blundell Bond (S-GMM)	
	(1) Productivity per hour (log)	(2) Labour cost per hour (log)	(3) Productivity per hour (log)	(4) Labour cost per hour (log)	(5) Productivity per hour (log)	(6) Labour cost per hour (log)	(7) Productivity per hour (log)	(8) Labour cost per hour (log)	(9) Productivity per hour (log)	(10) Labour cost per hour (log)
$\mu_1; \mu_1^w$ (secondary school)	0.016	0.043***	0.011	0.006	–0.013	0.015	0.013	0.032*	–0.003	–0.005
	(0.012)	(0.008)	(0.012)	(0.007)	(0.019)	(0.010)	(0.015)	(0.018)	(0.012)	(0.010)
$\mu_1; \mu_1^w$ (two-year college)	0.230***	0.305***	0.063***	0.076***	0.139***	0.246***	0.027	0.050	0.153***	0.124
	(0.020)	(0.013)	(0.018)	(0.011)	(0.030)	(0.015)	(0.025)	(0.031)	(0.015)	(0.014)
$\mu_3; \mu_3^w$ (university)	0.446***	0.765***	0.160***	0.177***	0.601***	0.664***	0.184**	0.173**	0.349***	0.438
	(0.038)	(0.023)	(0.039)	(0.021)	(0.087)	(0.042)	(0.076)	(0.077)	(0.021)	(0.019)
Firm fixed effects	No	No	Yes	Yes	No	No	Yes	Yes	No	No
Controls[a]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year × sector fixed effects[b]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	73,794	73,794	73,794	73,794	73,794	73,794	73,794	73,794	73,794	73,794
Number of firms	9,970	9,970	9,970	9,970	9,970	9,970	9,970	9,970	9,970	9,970
R^2	0.236	0.166	0.213	0.067	–	–	–	–		
Implied relative marginal productivity/labour cost (a) (ref = primary or less)										
$\lambda_{i1,0}; \lambda_{i1,0}^w$ (secondary school)	1.019	1.043	1.014	1.006	0.983	1.015	1.017	1.031*	0.996	0.995
$\lambda_{i2,0}; \lambda_{i2,0}^w$ (two-year college)	1.271***	1.305***	1.075***	1.076***	1.169***	1.246***	1.034	1.050	1.185***	1.124***
$\lambda_{i3,0}; \lambda_{i3,0}^w$ (university)	1.526***	1.765***	1.193***	1.178***	1.734***	1.663***	1.234**	1.173**	1.422***	1.438***

Notes: Parameter of the production/labour cost function. Implied relative marginal productivity/labour cost. Panel 2002-2011. [a]Controls: share of women and share of workers with an open-ended contract; [b]SIC1 (No. 9 sectors). ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

reference group (vs 1.27 with OLS). And workers possessing university degree appear only 1.193 times more productive (vs 1.53 with OLS). Similar falls are observed among our estimates of (marginal) labour costs. Workers with a two-year college attainment are 1.076 times more expensive to employ than the reference group, and university graduates are 1.178 times more expensive. Note that these estimates of the relative cost of employing workers with a tertiary education attainment are almost perfectly aligned on estimates of their productivity.

OLS also potentially suffers from endogeneity bias. This justifies considering ACF i.e. using intermediate goods to proxy for a plant's unobservable short-term productivity shocks. ACF has the advantage over the more typical FE panel data approach of allowing for time-varying firm effects and allowing for more identifying variation in the other inputs. It is not, however, a complete panacea. We have explained in Section 2 that it is difficult to believe in the existence of a one-to-one relationship between a firm's consumption of intermediates goods and a term ω_{it} that would systematically comprise all the firms' unobservables (shocks + fixed effects). ACF results (columns (5) and (6) in Table III somehow comfort us in our *a priori* scepticism. ACF fails to take us significantly away from OLS, as point estimates are of similar magnitude.

Remember also that ACF – due to the inclusion of interaction terms between the various labour share variables is a way to allow for imperfect substitutability across labour groups and between labour and capital (Hellerstein *et al.*, 1999). We interpret the similarity between our ACF results and those of the OLS-estimated production function as a possible indication that the assumption of perfect substitutability may not be abusive or a major source of distortion of our key estimates.

We now turn to our preferred models. We first combine ACF with firm fixed effects (FE-ACF). Results (columns (7) and (8)) show that the relative marginal productivity for secondary and two-year college do not reach statistical significance. By contrast, university graduates still display a significant productivity advantage of 1.23 (23 per cent more) with respect to primary school graduates. Note that FE and FE-ACF-results first, are fairly similar but second, are much lower than those delivered by ACF which themselves tend to be similar to OLS results. This tentatively suggest that first, the simultaneity bias is not pronounced in the case of Belgian firms but second, that firm unobserved heterogeneity is important and hint at the existence of assortative matching between workers and firms. High-productivity workplaces attract better-educated individuals, in line with the skill segregation assumption put forth by Kremer (1993) or Sattinger (1993).

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 university degree appears 1.42 times (42 per cent) more productive than workers with a primary school attainment (vs 23 per cent with FE-ACF). This could be explained by the fact that S-GMM does not completely evacuate data in level[29]. This said, S-GMM results largely comfort the evidence gathered: more educated workers, in particular university graduates, are more productive than the reference category (at most primary school graduates).

Focusing on estimates of (relative) marginal contribution of education to labour cost, we come to a similar conclusion. The estimated contribution of educated workers to labour cost is positive among the different specifications. In our first preferred specification (ACF-FE) we do not find statistical significance for secondary or two-year college, but we estimate a marginal labour cost of 1.17 times that of the reference group increase for university graduates (1.23 times in terms of productivity). With our other preferred specification (S-GMM), the corresponding estimate is 1.43 (1.42 in terms of productivity). Broadly speaking, the comparison of productivity and labour cost estimates delivered by our preferred models suggests an overall alignment.

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 is causally driven by a positive relationship between education and firm-level productivity. The existing empirical literature contains surprisingly little evidence of a causal relationship supporting this standard assumption of the human capital theory. The small literature that exploits on firm-level evidence 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 fourfold.

First, marginal productivity of workers with a tertiary education 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 23 per cent (FE-ACF) to 42 per cent (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 ten additional years compared with the reference group (workers with at most a primary degree), these figures correspond to rates of return of 2.3 per cent to 4.2 per cent per year of schooling; somewhat below the 5.2 per cent obtained by de la Croix and Vandenberghe (2004) when estimating a Mincerian gross montly wage equation.

For those with a two-year college degree similar estimates range from 3.4 per cent (FE-ACF)[30] to 18.5 per cent (S-GMM). Those for individuals with secondary school attainment are not statistically different from zero. Simultaneously, the labour cost premium of workers with university degree ranges from 17 per cent (FE-ACF) to 43 per cent (S-GMM). For those with a two-year college, it ranges from 5 per cent (FE-ACF)[31] to 12.4 per cent (S-GMM). It is not significant for workers with a secondary education attainment. We interpret these results as supportive of labour costs' alignment on 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, our regressions with firm-fixed effects (FE) estimates of human capital-related productivity gains are 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 university graduates[32], is good for the productivity of Belgian firms. At the same time, cross-sectional evidence stemming from OLS 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).

Third, when we account for firm-heterogeneity and simultaneity bias with the ACF methodology, we obtain similar results to those delivered by standard model with FE. We conjecture that, in our setting, the simultaneity bias is not large.

Fourth, 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.

Notes

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 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.
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 per cent) 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 two-year College programmes and four-year university programmes. Both are clearly tertiary (post-secondary) forms of education. The College programmes are more vocational in essence and would train future high-level technicians of people in charge of the middle managements, whereas the university 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 etc.).
6. Macroeconomists like Wolff and 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

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 labour are more complementary as inputs than are capital and unskilled labour (see capital-skill complementarity hypothesis by Griliches (1969)). Duffy *et al.* (2004) provide the cross-country evidence for capital-skill complementarity.
10. See Ackerberg *et al.* (2006) for a recent review.
11. But the underlying coefficients are not statistically significant as the 10 per cent threshold.
12. Again, the underlying coefficients are not statistically significant as the 10 per cent threshold.
13. We will see later in this section, how this assumption can be relaxed, when we present the econometric models used to identify the key coefficients of this production function.
14. We will see, how, in practice via the inclusion of dummies, this assumption can be relaxed to account for sector/industry wage effects, that must be important given Belgium's tradition of binding sector-level wage bargaining.
15. It should be clear that this HN framework is suitable for analysis, like ours, where only firm-level data on wages or labour costs are available. The reader must be aware that the labour cost function's coefficients obtained hereafter do not indicate the actual wage distribution within the firms but the sensitivity of average labour cost to changes in the educational mix of the workforce.
16. NACE2 level.
17. And its equivalent in Equation (12).
18. At least the part of that stock that is not affected by short-term recruitments and separations.
19. Motorway/airport in the vicinity of logistic firms for instance.
20. 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 $t-b$ $(0<b<1)$ 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.
21. 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.
22. OLS estimates for example.
23. Equation (16) appears as a Cobb-Douglas. However, its key component is Φ_{it}^{hat} that consists of 3rd 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.
24. 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).
25. The Standard Industrial Classification (abbreviated SIC) is a United States government system for classifying industries by a four-digit code.
26. 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 I.

27. www.bvdinfo.com/Products/Company-Information/National/Bel-First.aspx
28. Statistics Belgium estimates that they now represent between 35 per cent and 40 per cent of a cohort.
29. 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.
30. Not statistically significant.
31. Not statistically significant.
32. Haltiwanger *et al.* (1999) used a very loose definition of education levels (low, medium, high).

References

Ackerberg, D.A., Caves, K. and Frazer, G. (2006), "Structural identification of production functions", working paper, Department of Economics, UCLA, Los Angeles.

Arellano, M. (2003), "Discrete choices with panel data", *Investigaciones Economicas, Fundación SEPI*, Vol. 27 No. 3, pp. 423-458.

Arellano, M. and Bond, S. (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *Review of Economic Studies*, Vol. 58, pp. 277-297.

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

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

Blundell, R. and Bond, S. (1998), "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, Vol. 87 No. 1, pp. 115-143.

Card, D. (1999), "The causal effect of education on earnings", in Ashenfelter, O. and Card, D. (Eds), *Handbook of Labour Economics*, 1st ed., Vol. 3, Elsevier, Amsterdam, pp. 1801-1863.

Cataldi, A., Kampelmann, S. and Rycx, F. (2011), "Productivity-wage gaps among age groups: does the ict environment matter?", *De Economist*, Vol. 159 No. 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 Employment and Social Affairs, Brussels.

Dostie, B. (2011), "Wages, productivity and aging", *De Economist*, Vol. 159 No. 2, pp. 139-158.

Duffy, J., Papageorgiou, C. and Perez-Sebastian, F. (2004), "Capital-skill complementarity? Evidence from a panel of countries", *The Review of Economics and Statistics*, Vol. 86 No. 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", Discussion Paper No. 1542, IZA, Bonn.

Griliches, Z. (1969), "Capital-skill complementarity", *Review of Economics and Statistics*, Vol. 51 No. 4, pp. 465-468.

Griliches, Z. and Mairesse, J. (1995), "Production functions: the search for identification", Working Paper No. 5067, NBER, Cambridge, 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 Haltiwanger, J., Lane, J., Spletzer, J.R., Theeuwes, J. and Troske, K. (Eds), *The Creation and Analysis of Employer-Employee Matched Data*, Emerald, North Holland, Amsterdam.

Haltiwanger, J., Lane, J. and Spletzer, J. (1999), "Productivity differences across employers: the roles of employer size, age, and human capital", *American Economic Review*, Vol. 89 No. 2, pp. 94-98.

Hellerstein, J. and Neumark, D. (2007), "Production function and wage equation estimation with heterogeneous labour: evidence from a new matched employer-employee data set", in Berndt, E.R. and Hulten, C.R. (Eds), *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, NBER, Cambridge, MA, pp. 31-71.

Hellerstein, J.K. and Neumark, D. (1999), "Sex, wages and productivity: an empirical analysis of Israel firm-level data", *International Economic Review*, Vol. 40 No. 1, pp. 95-123.

Hellerstein, J.K., Neumar, D. and Troske, K. (1999), "Wages. Productivity and worker characteristics: evidence from plant-level production functions and wage equations", *Journal of Labour Economics*, Vol. 17 No. 3, pp. 409-446.

Kremer, M. (1993), "The o-ring theory of economic development", *The Quarterly Journal of Economics*, Vol. 108 No. 3, pp. 551-575.

Krueger, A.B. and Lindahl, M. (2001), "Education for growth: why and for whom?", *Journal of Economic Literature*, Vol. 39 No. 4, pp. 1101-1136.

Levinsohn, J. and Petrin, A. (2003), "Estimating production functions using inputs to control for unobservables", *Review of Economic Studies*, Vol. 70 No. 2, pp. 317-341.

Mankiw, G., Romer, D. and Weil, D. (1992), "A contribution to the empirics of economic growth", *Quarterly Journal of Economics*, Vol. 107 No. 2, pp. 407-437.

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

Olley, G.S. and Pakes, A. (1996), "The dynamics of productivity in the telecommunications equipment industry", *Econometrica*, Vol. 64 No. 6, pp. 1263-1297.

Prskawetz, A., Mahlberg, B. and Skirbekk, V. (2007), "Firm productivity, and educational structure in Austrian industries", in Mason, A. and Clark, R. (Eds), *Intergenerational Transfers and the Macroeconomy*, Edward Elgar, Northampton, MA, pp. 38-66.

Psacharopoulos, G. (1981), "Returns to education: an updated international comparison", *Comparative Education*, Vol. 17 No. 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*, Vol. 40 No. 1, pp. 111-136

Sattinger, M. (1993), "Assignment models of the distribution of earnings", *Journal of Economic Literature*, Vol. 31 No. 2, pp. 831-880.

Syverson, C. (2011), "What determines productivity?", *Journal of Economic Literature*, Vol. 49 No. 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, Fall, pp. 25-36.

van Ours, J.C. and Stoeldraijer, L. (2011), "Age, wage and productivity in dutch manufacturing", *De Economist*, Vol. 159 No. 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*, Vol. 25 No. 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*, Vol. 159 No. 2, pp. 159-191.

Vandenberghe, V. (2012), "Are firms willing to employ a greying and feminizing workforce?", *Labour Economics*, doi:10.1016/j.bbr.2011.03.031.

Wolff, E.N. and Gittleman, M. (1993), "The role of education in productivity convergence: does higher education matter?", in Szirmai, A., Van Ark, B. and Pilat, D. (Eds), *Explaining Economic Growth*, Elsevier, Amsterdam.

Corresponding author
Professor Vincent Vandenberghe can be contacted at: vincent.vandenberghe@uclouvain.be