Over-, Required, and Undereducation: Consequences on the Bottom Lines of Firms

Stephan Kampelmann[1] — Benoît Mahy[2] — François Rycx[3] — Guillaume Vermeylen[4]

Abstract. We provide first evidence regarding the direct effect of over-, required, and undereducation on the bottom lines of firms across work environments. We use detailed Belgian linked employer–employee panel data, rely on the methodological approach pioneered by Hellerstein *et al.* (1999), and estimate dynamic panel data models at the firm level. Our findings show an 'inverted L' profitability profile: undereducation is associated with lower profits, whereas higher levels of required and overeducation are correlated with positive economic rents of roughly the same magnitude. The size of these effects is amplified in firms experiencing economic uncertainty or operating in high-tech/knowledge sectors.

1. Introduction

Advanced economies are facing a constant rise in workers' level of education. The share of tertiary educated workers in total employment of EU28 countries increased from around 22 per cent in 2000 to more than 33 per cent in 2016 (Eurostat, 2018). This trend suggests that educational requirements for employment are on the rise, that there is excess demand for tertiary education, and/or that firms employing more educated workers will improve their production techniques to take advantage of those additional skills (Figueiredo *et al.*, 2015; McGuinness, 2006; Quintini, 2011; Sattinger and Hartog, 2013). However, if these assumptions are not satisfied then workers with tertiary education may end up in jobs for which they are overeducated.[1] Moreover, in periods of high unemployment, there may be some crowding out, that is, namely a process by which people with tertiary education accept jobs that could be occupied by less educated ones. According to the OECD (2013), the proportion of overeducated workers[2] among advanced economies stands at around 20 per cent and almost one out of six workers is recorded as undereducated.[3] Undereducation

---

We are most grateful to Statistics Belgium for giving us access to the data. Financial support from IWEPS (Walloon Region) is kindly acknowledged.

[1]Université libre de Bruxelles, Faculty of Architecture La Cambre-Horta, Chair Circular Economy and Urban Metabolism, Laboratory Urbanism Infrastructure and Ecology, Brussels, Belgium.

[2]University of Mons (UMONS), Warocqué School of Business and Economics, humanOrg and DULBEA, Mons, Belgium.

[3]Université libre de Bruxelles, SBS-EM (CEB and DULBEA), IRES, humanOrg, GLO and IZA, Brussels, Belgium.
E-mail: frycx@ulb.ac.be

[4]University of Mons (UMONS), Warocqué School of Business and Economics, humanOrg, CEB and DULBEA, Mons, Belgium.

LABOUR 34 (1) 80–112 (2020) DOI: 10.1111/labr.12152

JEL J21, J24

may notably result from labour shortages (i.e. bottleneck vacancies) and technologically induced changes in job content and complexity.

Our paper provides empirical evidence on two unresolved questions that are linked to this development. First, we estimate how upgrading the firm's job mix towards jobs requiring higher levels of education affects profitability (i.e. productivity-wage gaps) at the firm level.[4] Put differently, we examine whether rising educational requirements pays off for businesses. Second, we provide evidence on what happens to productivity-wage gaps when firms hire workers who do not have the required level of education for the job (i.e. workers that are either over- or undereducated). Unlike much of the earlier literature (still essentially focused on workers' wages, job satisfaction, and related attitudes and behaviours), our econometric estimates are based on direct measures of labour productivity and wage costs. They are also robust to a range of measurement issues, such as time-invariant labour heterogeneity and firm characteristics. To do so, we use detailed Belgian linked employer–employee panel data, rely on the methodological approach pioneered by Hellerstein *et al.* (1999), and estimate dynamic panel data models at the firm level.

Building on the literature regarding educational mismatch (Baert and Verhaest, 2014; Hartog, 2000; Mavromaras and McGuinness, 2012; Sanchez-Sanchez and McGuinness, 2015; Sellami *et al.*, 2018; Verhaest and Omey, 2009, 2012; Verhaest and Van der Velden, 2013), in which direct measures of productivity are still rare (Grunau, 2016; Kampelmann and Rycx, 2012; Mahy *et al.*, 2015) and direct measures of *gaps* between labour productivity and wage costs (i.e. profits) have so far not been estimated, our paper shows that up- and downward deviations from the required education in a job do not have the same consequences: all models suggest that the extent of undereducation is in general harmful for firm profits, whereas higher levels of required and overeducation generate positive economic rents of roughly the same magnitude. Finally, we show that the level of required education — and deviations from this level — are more consequential in some parts of the economy than in others. More precisely, our results suggest that they exert a smaller impact in sectors characterized by low uncertainty but a stronger effect in high-tech/knowledge industries.

The remainder of the paper is organized as follows. A review of the literature is presented in Section 2. Sections 3 and 4 respectively describe our methodology and data set. Econometric results are presented in Section 5. The final section discusses the results and concludes.

2. Literature review

2.1. Educational mismatch and productivity-wage gaps

Although labour economists have produced an abundant literature on the link between educational credentials and wages (Becker, 1975; Card, 1999), it becomes increasingly problematic that most empirical studies have relied on the assumption that these credentials accurately reflect labour productivity. Indeed, the recent literature on wage discrimination has applied methodological advances — notably an approach pioneered by Hellerstein *et al.* (1999) — to extensive matched employer–employee data sets that provide direct measures of labour productivity. This allowed to re-examine the nexus between worker characteristics (such as age, gender, and training), labour productivity, wages, and profits (Cardoso *et al.*, 2011; Damiani *et al.*, 2016; Devicienti *et al.*, 2018; Konings and Vanormelingen, 2015; Nielen and Schiersch, 2014; van Ours and Stoeldraijer, 2010, 2011).

Instead of merely assuming that wage differentials between educational groups mirror differences in productivity, which has long been the modus operandi of empirical research within the Mincer framework (Lemieux, 2006; Pereira and Martins, 2004), Current literature estimates how the observed productivity differences between educational groups compare with observed differences in labour costs. If the productivity effect of additional schooling is higher (lower) than the effect on labour costs, the combined result of these two effects increases (decreases) the firm's profits in the form of positive (negative) rents. Only in the case in which productivity effects are exactly offset by changes in labour costs the conventional human capital hypothesis remains warranted.

The direct measurement of productivity and wage effects can be applied to elucidate the impact of over-, required, and undereducation on the bottom lines of firms which is at the heart of the rapidly expanding literature on educational mismatch (Baert and Verhaest, 2014; Baert *et al.*, 2013; Figueiredo *et al.*, 2015; Mavromaras and McGuinness, 2012; Sattinger and Hartog, 2013; Verhaest and Omey, 2009, 2012; Verhaest and Van der Velden, 2013).

The required level of education reflects the typical staffing practices of firms in each type of job. It is likely that educational requirements are mostly driven by the level of skills that are necessary to perform the tasks of the job adequately. This also explains why there is substantial variation in requirements across different jobs, with some jobs being more demanding and non-routine than others. Over longer time horizons there may also be variations in educational requirements within jobs. In addition, firms can raise the average education of their workforce without exceeding prevailing requirements in each job, for instance when they change their job mix by replacing jobs with low requirements with jobs for which the typical level of required education is higher.

Employing workers with credentials below prevailing requirements appears to hamper productivity, whereas the effect of overeducation is not clear-cut (Grunau, 2016; Kampelmann and Rycx, 2012; Quintini, 2011). Regarding direct measures of wage and labour cost effects, most existing studies such as the meta-analyses from Groot and Maassen van den Brink (2000) and Leuven and Oosterbeek (2011) are consistent in that they suggest that the wage returns to educational credentials beyond the required level in a given job are in general positive, whereas the estimated positive returns to overeducation turn out to be lower with IV compared with OLS estimations (Leuven and Oosterbeek (2011)). From a more theoretical point of view, Sattinger and Hartog (2013) support the idea that employers pay higher wages to workers possessing more education than required; this is because overeducated workers might have higher reservation wages when they search on a labour market with frictions where the optimal allocation between actual and required education cannot be achieved. Moreover, wages rise even more when the educational requirement that corresponds to the firm's job mix increases. The literature also suggests that wage returns to undereducation are negative but not always significant (Battu *et al.*, 1999; Duncan and Hoffman, 1981; Galasi, 2008; Rumberger, 1987; Sicherman, 1991).

This paper overcomes the complete absence of studies that pull the productivity and labour cost effects together: we use linked employer–employee data to estimate how changes in educational requirements and deviations from these requirements affect productivity, labour cost, and productivity-wage gaps (i.e. profits). Indeed, partial analyses of only wage effects or only productivity effects are much less relevant from the perspective of the firm. For instance, the profitability of the latter does not suffer from higher labour costs per se, but only when these higher costs are not at least offset by hikes in productivity.

Rational profit-seeking firms change their labour force composition only if these changes generate positive productivity-wage gaps.

2.2. Theoretical explanations for productivity-wage gaps

Turning to the theoretical mechanisms that potentially drive the effects of educational mismatch on firm profits boils down to asking under which circumstances the educational composition of a firm's workforce has a different effect on labour costs than on productivity. As the focus of this paper lies on the empirical estimation of productivity-wage gaps related to over-, required, and undereducation, we do not reiterate the numerous mechanisms that have been put forward in the large theoretical literature in this field. Following the literature overview in Kampelmann and Rycx (2011), these mechanisms can be divided into a) theories based on efficiency and individual rationality (e.g. when over- or undereducated workers differ from workers with the required level of education in terms of quasi-fixed costs or firm-specific skills, or when efficiency considerations lead firms to compress the wage structure of firms so as to avoid shirking or demotivation — see, e.g., Cardoso (2010) and McGuinness (2006)); and b) institutionalist theories (e.g. when monopsony power, market regulations, wage norms, or collective bargaining is associated with positive or negative rents that differ for over-, required, and undereducated workers — see, e.g., Konings and Vanormelingen (2015) and Quintini (2011)).

Because empirical productivity-wage gaps are likely to confound several of these mechanisms, it is complicated to formulate a priori hypotheses regarding our estimates. We can, however, hypothesize that labour market institutions will have an impact on the sign and size of productivity-wage gaps. For instance, collective bargaining institutions in EU countries like Austria, Denmark, Finland, France, Germany, Italy, and Belgium have been associated with wage rigidities, i.e., wages that do not adjust automatically and swiftly to productivity due to institutional forces. Among institutions associated with wage rigidities are binding statutory minimum wages and combinations of sectoral minima and high collective bargaining coverage, which serve in effect as a functional equivalent of a binding statutory minimum wage (Garnero *et al.*, 2015a,b). A plausible hypothesis is then that firms in these countries receive positive rents when they employ overeducated workers, whereas undereducated workers are likely to be associated with negative rents. Put differently, overeducation (undereducation) is expected to have a stronger positive (negative) effect on productivity than on wages. For the specific case of the Belgian private sector economy, an additional source for productivity-wage gaps in terms of educational mismatch is the strong role of occupational categories for wage setting in this country (Kampelmann and Rycx, 2011). If occupational categories 'fix' wages within relatively rigid intervals, we expect a stronger positive (negative) effect on profitability from over (under) education.

This being said, this prediction might be attenuated by the fact that collective bargaining agreements also typically foresee binding rules for wage differentiation according to educational credentials within occupational categories, meaning that some overeducated workers will automatically receive higher wages compared with the average colleague in the same job. In this case, a share of any positive rent is converted into higher wages, whereas negative rents are absorbed by lower wages because the wages of undereducated workers will generally lie below the average wage in their occupational category. Furthermore, overeducation might not necessarily improve firm productivity. According to human capital theory, i) education (as well as formal training and informal work experience) develops skills that make workers more productive, and ii) wage differentials reflect differences in productivity. Consequently, results showing that overeducated workers get a wage premium with respect to their adequately educated colleagues in similar jobs suggest that the former are more productive than the latter (Battu *et al.*, 1999; Duncan and Hoffman, 1981; Galasi, 2008; McGuinness and Sloane, 2011; Rumberger, 1987; Sicherman, 1991). However, another strand of the literature examines the impact of educational mismatch on job satisfaction and related characteristics. The standard hypothesis is that overeducated workers, as a result of frustration, are less satisfied, have more health problems, and higher rates of shirking, absenteeism, and turnover than their adequately educated colleagues. This hypothesis is supported by most empirical findings comparing overeducated workers with their adequately educated former classmates, i.e., individuals with the same attained education (Leuven and Oosterbeek, 2011; McGuinness, 2006; Verhaest and Omey, 2009). Results are somewhat less clear-cut when comparing overeducated workers with their adequately educated colleagues doing the same job (which is obviously the most relevant approach from a firm perspective). Anyway, this strand of the literature suggests that overeducation might actually be harmful for firm productivity, and accordingly be detrimental for firm profitability.[5]

2.3. Impact of the firm's environment

There are compelling theoretical arguments according to which the extent of productivity-wage gaps associated with educational mismatch will differ across sectors and work environments. One obvious reason for this is the segmentation of labour markets. Firms requiring a workforce with higher skills typically face supply constraints that are absent in low-skilled environments, notably driven by technological change and the failure of the skill supply to keep up with changes in demand (Broecke, 2016); the impact of such supply constraints comes in addition to the impact of institutional factors that can also create excess wage inequality between skills (Broecke *et al.*, 2017). Deming and Noray (2018) point to labour supply shortage for Science, Technology, Engineering and Math (STEM) jobs due to technological change. These supply constrains should decrease rents and benefit wages, independently of whether the former arise for over-, required, or undereducated workers. On the other hand, in high-skilled sectors competences may be firm specific so that some employers are in a position to extract economic rents from their monopsony power. Higher market frictions might also increase firms' rents in higher skilled sectors, for instance when these firms are better able to differentiate themselves from competitors when hiring workers with heterogeneous preferences (Manning, 2011).

2.3.1 Knowledge-intensive environments. Being educated beyond (below) the required norm might increase (decrease) labour productivity to a larger extent in knowledge-intensive environments. Knowledge and education are central in the creation of value. Hiring highly competent employees is presented by Wu (2015) as a way to boost the learning ability of a firm and to solve complex problems that are specific to knowledge-intensive firms. Tohmo (2015) also suggests that employees who possess and provide know-how and creativity play a leading role in knowledge-intensive production and innovation systems. In turn, managers from high-tech firms might consider education as more important in value creation and be thus also more inclined to propose higher wages when hiring overeducated workers. Fares and Yuen (2003) notably point out that Canadian university graduates' workers are better paid in the case of research and development-intensive industries.

Overeducated workers might also experience higher productivity when firms evolve in knowledge-intensive environments as these firms are more innovative and adaptive, as suggested for instance by Billon *et al.* (2017) in the case of innovative European companies producing knowledge-intensive services. Nelson and Phelps (1966) develop the notion of adaptability and consider the role of education in terms of ability to innovate and to adapt to new technology. Moreover, Hanushek *et al.* (2017) observe that returns to skills across PIAAC countries are consistent with the assumption that the ability of higher skilled workers to adapt to economic change is often considered as a prime source of economic value of such skills.

Including dynamic effects further supports the idea of a stronger impact in knowledge-intensive environments. Given that these environments are constantly innovating and upgrading, today's overeducated workers might be tomorrow's adequately educated workers. Moreover, their productivity might well be enhanced as these overeducated workers are not only currently beyond the prevailing norm but also already able to adapt to the next wave of innovation. In this regard, Krueger and Kumar (2004) assume that higher (and over) education lowers the risks of suffering from productivity losses in the case of technological shocks. In these knowledge-intensive environments with their ever-changing and challenging conditions, firms might be inclined to devote more latitude to adaptable overeducated workers for the sake of their future competitiveness, favouring thereby overeducated workers' capabilities to develop their creativity and innovation skills and thus their productivity. For instance, Stinebrickner *et al.* (2018) estimate that learning-by-doing is only favoured in the specific case of high-skilled tasks.

For these additional reasons, high-tech firms could be more inclined to propose higher wages to potentially more adaptable overeducated workers. Additional pay for overeducated more adaptable workers is also supported by Aghion *et al.* (2003) model, which suggests that more adaptable workers should benefit from additional labour demand coming from their ability to transfer their (additional) acquired knowledge to new machines.

To sum up, preceding arguments suggest that the positive (negative) impact of over (under) education on both productivity and wages should be enhanced in knowledge-intensive environments, whereas the positive impact on productivity-wage gaps of overeducated workers might be higher in the case of more rigid labour markets such as the Belgian one.

2.3.2 Uncertain (risky) environments. High risk and uncertainty mean that a firm may fare better by having some sort of additional organizational buffer, or extra resources that it can use in case of trouble. Indeed, a firm facing uncertainty in the form of an immediate risk of bankruptcy might turn to overeducated employees to come up with unforeseen contingency measures, emergency reactions, new projects, or products. If overeducation can in normal times be seen as 'organizational slack', it can be a valuable resource in times of danger. Conversely, undereducated workers might be less helpful to deal with unforeseen obstacles and instead become an additional burden.

High risk and uncertainty often imply that firms have to undergo regular organizational changes to stay competitive. Examining organizational change, Schaefer (1998) looks at different approaches and models (including those of Hannan and Freeman (1989)) to understand how organizational barriers are related to influence activities of employees. Influence activities are carried out by employees to affect the distributive results of organizational decisions (Meyer *et al.*, 1992). Schaefer (1998) points out that employees try to affect organizational change in their favour when the change is considered, decided, and implemented. While these approaches cannot be extended automatically to our analysis, they suggest that overeducated workers should favour improving productivity coming from adequate changes as they could expect that these changes will boost their career prospects and working conditions in the firm. In their assessment of influence activities, Meyer *et al.* (1992) suggest that the prospect of decline and consequent layoffs in one part of a multi-unit firm creates additional influence costs that arise when managers and staff of the threatened unit attempt to protect their jobs. It is likely that overeducated workers have a higher probability to be hired in other units of the same firm following a closure or bankruptcy of a unit; compared with their undereducated co-workers, they have therefore less economic incentives to engage in influence activities aimed at preventing productivity-enhancing organizational change.

Next, Prendergast (2002) suggests that economic uncertainty leads to a closer match between wages and productivity because 'firms delegate decision-making power more in uncertain environments but offer output-based contracts in order to constrain the possibility that they use their discretion in harmful ways' (Prendergast, 2002; : 1079). In other words, the possibility for more (less) productive workers such as over (under) educated workers to receive higher wages will be higher in more uncertain environments. This idea is corroborated by Foss and Laursen (2005), who show that in high-skilled jobs increased productivity is associated with higher wages instead of extra profits. Barth *et al.* (2008) also suggest a close tie between productivity and wages for jobs that require high autonomy, which is another characteristic of uncertain or high-tech environments. This implies that overeducated workers, which Barth *et al.* (2008) find to be more productive in such contexts, reap the benefits of their productivity in the form of higher wages.

Assuming that managers in risky environments could incorporate this additional productivity argument in their wage-setting decisions, they should be more inclined to offer relatively higher wages to overeducated workers. Thus, all in all, our assumptions are that the positive (negative) impacts of over- (under-)education on productivity, wages and profits should be higher when firms evolve in more uncertain (risky) environments.

3. Estimation framework

3.1. Baseline specification

The literature offers three ways to measure the required level of education for a job and the incidence of educational mismatch (see Hartog (2000) for a discussion). The first one, called the objective measure or job analysis approach is based on the evaluation by professional analysts of the level and type of education that is required for a specific job. The American Dictionary of Occupational Titles (DOT) is an example of such an approach. The second approach, called subjective or self-assessment approach, requires the employee/employer to determine the type and level of formal education that is associated with the achievement of the tasks in a given job. This measurement thus rests on employee and/or employer surveys. The third approach, called empirical or realized matches approach, derives the required level of education for a job from what workers in the corresponding job or occupation usually have attained. The required education is then generally computed on the basis of the mode of the education in a given occupation.

Each measure has its own advantages and weaknesses (for a detailed discussion see, e.g., McGuinness (2006)). For instance, the job analysis approach is appealing because it is based on clear definitions and explicit measurement instructions. However, because of the cost and difficulty of this exercise, classifications based on job analysis are only published from time to time. Moreover, given that technological progress is likely to affect rapidly the content and complexity of jobs, classifications become fast outdated. Another point is that there is no such classification for Belgium and the use of a foreign one would probably create important measurement errors given that job classifications and requirements vary across countries. The second approach, based on workers' self-assessment, interestingly relies on local up-to-date information. However, it suffers from the fact that it is not based on rigorous instructions. In particular, respondents may overstate the requirements of their own job. It also typically leads to a downward biased proportion of undereducated workers. Finally, the realized matches approach has the advantage that it can be implemented easily and on a regular basis. Moreover, in contrast to the self-assessment approach, it enables to estimate more easily how much a worker is over- or undereducated. However, as Hartog (2000) argues, the realized matches method reflects the equilibrium between actual supply and demand. Put differently, the required education is likely to be influenced by the extent of over- and undereducation.

Overall, it is impossible to say that one measure is strictly better than the others and in practice the choice of a measure is often dictated by data availability (Verhaest and Van der Velden, 2013). Given the feature of ours, we use realized matches in this paper. Thus, we compute the required level of education for a given job by taking the mode of workers' years of education within detailed ISCO 3-digit occupations (113 categories) for each of the 12 years in our sample, yielding as much as 1,356 occupation-year norms[6] and used this information to define our three main explanatory variables.[7]

The first explanatory variable in our model is the extent of required education in firm j at year t. This variable is obtained by summing up the years of required education associated with the different jobs i in firm j at time t ($R_{i,j,t}$). For example, if a firm employs five workers, namely a managing director, an administration professional, a general office clerk, a shop salesperson, and a cashier, than the extent of required education in that firm will correspond to the sum of the required years of education associated to these five different jobs. The required years of education for a cashier (i.e. the mode of years of education in the ISCO category 523 across the entire private sector at time t) might be for instance equal to 9 (which corresponds to a lower secondary education degree) while those for an administration professional (ISCO category 242) may be equal to 16 (which corresponds to a master's degree). For a given number of employees, firms can increase (decrease) the total amount of required years of education by upgrading (downgrading) the firm's job mix towards jobs with higher (lower) educational requirements. In our example, upgrading may notably result from the replacement of the general office clerk by a department manager (whose required years of education is higher).

The second explanatory variable is the extent of overeducation. It is calculated by summing up the years of education in firm j and year t that exceed required levels. Put differently, we make the sum over all workers i employed in firm j at time t of $O_{i,j,t}$, with $O_{i,j,t} = Attained\ education_{i,j,t} — R_{i,j,t}$ if > 0 and 0 otherwise (where $Attained\ education_{i,j,t}$ is the number of years of schooling attained by worker i and $R_{i,j,t}$ the required years of education associated to workers' i job). For example, a worker with 10 years of attained education working in a job whose required level equals 8 years of education will add 2 years of 'overeducation'. Firms who place workers in occupations whose educational requirements they exceed will therefore increase the extent of overeducation in their workforce.

The third explanatory variable reflects undereducation. It sums up the schooling that falls short of educational requirements, i.e., $U_{i,j,t} = R_{i,j,t}$ — *Attained education*$_{i,j,t}$ if > 0 and 0 otherwise. For example, a worker with 10 years of education in a job with an educational requirement of 12 years adds 2 years of 'undereducation'. The more workers of this type a firm employs, the higher will be the extent of undereducation in its workforce.

To render these sums comparable across firms, the years of required, over-, and undereducation ($R_{i,j,t}$, $O_{i,j,t}, U_{i,j,t}$) are divided by $m_{j,t}$, which is the number of workers employed in firm j at year t.

To examine the impact of these explanatory variables on firm productivity and wage costs, we use two specifications aggregated at firm level. More precisely, we estimate the following equations:

$$\text{In}\left(\frac{VA_{j,t}}{m_{j,t}}\right) = \beta_0 + \beta_1\left(\text{In}\left[\frac{VA_{j,t-1}}{m_{j,t-1}}\right]\right) + \left[\left(\beta_2\sum_{i=1}^{m_{j,t}}\frac{O_{i,j,t}}{m_{j,t}}\right) + \left(\beta_3\sum_{i=1}^{m_{j,t}}\frac{R_{i,j,t}}{m_{j,t}}\right) + \left(\beta_4\sum_{i=1}^{m_{j,t}}\frac{U_{i,j,t}}{m_{j,t}}\right)\right] + X_{j,t}\beta_5 + Z_{j,t}\beta_6 + \gamma_t + \upsilon_{j,t} \tag{1}$$

$$\text{In}\left(\frac{w_{j,t}}{m_{j,t}}\right) = \beta_0^* + \beta_1^*\left(\text{In}\left[\frac{w_{j,t-1}}{m_{j,t-1}}\right]\right) + \left[\left(\beta_2^*\sum_{i=1}^{m_{j,t}}\frac{O_{i,j,t}}{m_{j,t}}\right) + \left(\beta_3^*\sum_{i=1}^{m_{j,t}}\frac{R_{i,j,t}}{m_{j,t}}\right) + \left(\beta_4^*\sum_{i=1}^{m_{j,t}}\frac{U_{i,j,t}}{m_{j,t}}\right)\right] + X_{j,t}\beta_5^* + Z_{j,t}\beta_6^* + \gamma_t^* + \upsilon_{j,t}^* \tag{2}$$

where $VA_{j,t}$ and $w_{j,t}$ are the productivity and labour cost of firm j at year t, respectively. The dependent variable in equation [1] is therefore the logarithm of the value added (at factor costs) per worker. The dependent variable in equation [2] is the logarithm of the labour cost per worker. It is obtained by dividing the firm's total wage bill (including fixed and variable pay components, in kind benefits, employer-funded extra-legal advantages (related, e.g., to health, early retirement, or pension), hiring and firing costs, and payroll taxes) by the total number of workers. $X_{j,t}$ is a vector including aggregated characteristics of workers in firm j at year t: the share of the workforce that has at least 10 years of tenure, the fractions of workers respectively younger than 30 and older than 49, and the shares of women, blue-collar, and part-time workers. $Z_{j,t}$ is a vector containing firm j characteristics at year t: the sectorial affiliation (eight dummies), the age and size (number of workers) of the firm, and the level of wage bargaining (one dummy). γ_t is a set of 11 year dummies and $\upsilon_{j,t}$ is the error term.[8]

This approach therefore examines the relationship between average years of over-, required, and undereducation within firms, on the one hand, and the productivity and labour cost per worker, on the other hand, while controlling for year dummies and a range of worker and firm characteristics.[9] Moreover, given that equations [1] and [2] are estimated on the same samples with identical control variables, the parameters for productivity, and wage costs can be compared and conclusions can be drawn on how over-, required, and undereducation affect firms' productivity-wage gaps. Put differently, parameters enable us to highlight whether over-, required, and undereducation are beneficial or harmful for firms' productivity, and whether and how the gains or losses associated with over-, required, and undereducation are shared with workers (in terms of higher or lower wages). This two-equation approach, pioneered by Hellerstein *et al.* (1999), is now standard in the literature on the productivity and wage effects of labour heterogeneity, notably in terms of age, gender, and employment contracts (Cardoso *et al.*, 2011; Damiani *et al.*, 2016; Garnero *et al.*, 2014; Giuliano *et al.*, 2017; Hellerstein and Neumark, 2004; Konings and Vanormelingen, 2015; Mahlberg *et al.*, 2013; Nielen and Schiersch, 2014). To our knowledge, it has never been used to examine the nexus between educational mismatch, productivity, and wages.

The coefficients associated with over-, required, and undereducation can be interpreted as follows. A positive (negative) estimate for β_2 and β_2^*, respectively, means that a firm that increases the education of its workforce beyond the required level in each job category will increase (decrease) its productivity (equation 1) and wage cost (equation 2). Conversely, a positive (negative) estimate for β_4 and β_4^*, respectively, can be interpreted as suggesting positive (negative) productivity/wage effects that arise if a firm hires more workers with education below the required level in the jobs it has.

The coefficients β_3 and β_3^* can reflect two alternative phenomena: they can capture the effect of: (i) an increase in the required level of education for the different job categories in a firm or (ii) that the firm has changed its job mix towards jobs with higher educational requirements. An example of the first case is when the required educational attainment in a given job category increases, for instance because the technology involved to carry out the job has become more complex. The computerization of many secretarial tasks could have led to such an increase in required levels of education in clerical jobs. In this case, firms can increase the amount of education of their workforce *without necessarily changing the job mix in the firm*. An example of the second case is when firms modify their job mix by replacing jobs with relatively low educational requirements (such as unskilled manual jobs) with jobs that require more education (such as managerial or supervisory jobs). For instance, a firm that moves from a production technology based on unskilled manual labour to an automated process with skilled machine supervisors is likely to increase the required level of education in the jobs it has *even if the required level of education in each job category has not changed*. In both cases, we interpret positive estimates of β_3 and β_3^* as suggesting a positive association between the level of required education that corresponds to the firm's job mix and the firm's productivity/wages.

The inclusion of the lagged dependent variable among the regressors renders the model dynamic and aims to account for the persistency in firm-level productivity and wage costs.[10] It is also likely to improve the identification of the parameters of interest (even though the coefficients on the lagged dependent variables are not a central issue in the analysis). Indeed, as illustrated by Bond (2002), the use of a dynamic model is necessary to obtain consistent results when estimating a production function with serially correlated productivity shocks and explanatory variables that are correlated with these shocks. Although serial correlation of productivity shocks may arise if the effects of demand shocks are only partially captured by the industry-specific control variables (Hempell, 2005), the responsiveness of input factors to productivity shocks may be explained by an endogeneity issue (see below).

3.2. Firm environments

In light of the literature review, we test whether the impact of educational mismatch differs according to the characteristics of the environment in which the firm evolves. To this end, we estimate equations [1] and [2] separately for different clusters of firms and compare the corresponding coefficients across clusters.

First, the technological environment is investigated using a taxonomy developed by Eurostat (2012), the HT/KIS nomenclature. This nomenclature supplies the NACE 2- or 3-digits codes indicating whether a firm is classified as high-tech/knowledge and others as low-tech/knowledge. It covers industrial and service-oriented firms. Manufactures are aggregated according to the technological intensity based on the R&D expenditure, i.e., high technology and medium-high technology in the first group, medium-low technology and low technology in the second; services are aggregated according to the share of tertiary educated persons, i.e., knowledge-intensive services in the first group, and less knowledge-intensive services in the second. So the group of firms considered as high-tech/knowledge intensive belongs to sectors that are high or medium-high tech/knowledge intensive (HT/KIS), whereas the second group of firms considered as low-tech/knowledge intensive belongs to sectors that are medium-low or low-tech/less knowledge intensive (non-HT/KIS).

Second, the economic uncertainty of the firm's environment is evaluated through an indicator proposed by Mahy *et al.* (2011). It uses the mean rate of bankruptcy at the NACE 3-digit level that is also supplied by Statistics Belgium. The first group gathers firms that belong to sectors registering a higher mean rate of bankruptcy than the average of the whole sample, whereas the second gathers firms belonging to sectors registering lower bankruptcy levels.

3.3. Estimation techniques

Equations [1] and [2] have been estimated with three different methods: pooled ordinary least squares (OLS), a fixed effects (FE) model, and the generalized method of moments (GMM) estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998). The OLS estimator with standard errors robust to heteroscedasticity and serial correlation is based on the cross-section variability between firms and the longitudinal variability within firms over time. However, this OLS estimator suffers from a potential heterogeneity bias because firm productivity can be related to firm-specific, time-invariant characteristics that are not measured in micro-level surveys (e.g. an advantageous location, firm-specific assets such as patent ownership, or other firm idiosyncrasies).

One way to remove unobserved firm characteristics that remain unchanged during the observation period is to estimate a FE model. However, neither pooled OLS nor the FE estimator address the potential endogeneity of our explanatory variables.[11] Yet, there might be some cyclical 'crowding out', namely a process by which highly educated workers take the jobs that could be occupied by less educated ones during recessions, because of excess labour supply (Dolado *et al.*, 2000). This assumption suggests that mean years of overeducation (undereducation) within firms may increase (decrease) as a result of a lower labour productivity (and vice versa). We have performed a direct endogeneity test on the over-, required, and undereducation variables in our models and indeed reject the null hypothesis that our main variables of interest can actually be treated as exogeneous.[12] To control for this endogeneity issue, in addition to state dependence of firm productivity/wages and the presence of firm fixed effects, we estimate equations [1] and [2] with the dynamic system GMM (GMM-SYS) estimator.[13]

The GMM-SYS approach boils down to simultaneously estimating a system of two equations (respectively in level and in first differences) and relying on internal instruments to control for endogeneity. More precisely, education variables are instrumented by their lagged levels in the differenced equation and by their lagged differences in the level equation.[14] The implicit assumption is that differences (levels) in (of) productivity/wages in one period, although possibly correlated with contemporaneous differences (levels) in (of) education variables, are uncorrelated with lagged levels (differences) of the latter. Moreover, differences (levels) in (of) education variables are assumed to be reasonably correlated with their past levels (differences).

Interestingly, the inclusion of the lagged dependent variable in the OLS, fixed effects, and GMM-SYS specifications provides an ad hoc test for the appropriateness of the latter. As outlined by Roodman (2009), this test consists in checking whether the regression coefficient on the lagged dependent variables obtained with system GMM fall between the OLS and fixed effects estimates.

4. Data and descriptive statistics

Our empirical analysis is based on a combination of two large data sets covering all years from 1999 to 2010. The first is the Structure of Earnings Survey (SES) that is carried out by Statistics Belgium. It covers all firms operating in Belgium with more than 10 workers and activities within sections B to N of the NACE Rev. 2 nomenclature. This survey gathers information on firms' characteristics (e.g. sector, number of workers, level of collective wage bargaining) as well as information on workers' characteristics (e.g. age, education, tenure, gross earnings, paid hours, sex, occupation).[15] However, the SES does not provide financial information. It has thus been merged with a firm-level survey, namely the Structure of Business Survey (SBS), also carried out by Statistics Belgium. This survey provides financial information (e.g. the firm-level value added, wage cost, and gross operating surplus per worker). The coverage of the SBS differs from that of the SES in that it does not cover the whole financial sector (NACE K) but only Other financial intermediation and Activities auxiliary to financial intermediation. The merger of the two data sets has been realized by Statistics Belgium using the firms' social security numbers.[16]

Information in the SES refers to the month of October of each year, whereas data in the SBS are measured over entire calendar years, i.e., from January to December. To avoid running a regression where information on the dependent variable (collected for the entire year) precedes the recording of the explanatory variables (collected in October), all explanatory variables in equation [1] have been lagged by 1 year. This way, information on education is recorded in October in year t and used to explain firm-level productivity during the calendar year $t + 1$. The imperfect synchronization of the SBS and SES data might introduce some fuzziness into our estimates as we cannot exclude the occurrence of external events influencing productivity in the intermediate period. This concern could only be completely eliminated if we had firm-level information on education for the entire calendar year. This being said, even if this information were available, there are also arguments for using non-synchronized information on education variables: it is difficult to conceive how changes in educational composition could generate *immediate* effects, and potential productivity effects are thus more likely to occur after a certain adjustment period. The slightly non-synchronized use of SBS and SES is therefore arguably the best option in light of data availability and productivity dynamics.

Our preferred estimator (GMM-SYS) requires firm information on at least two consecutive years. Given that sampling percentages of firms in our data set increase with the size of the latter (see footnote 13), medium-sized and large firms are thus over-represented in our econometric investigation. Note that workers and firms for which data are missing or inaccurate have been excluded.[17] In addition, in order to guarantee that the level of required education is computed on the basis of a sufficient volume of data, we dropped occupations at ISCO 3-digit level with less than 10 observations.[18] We also eliminated firms with less than 10 observations, because the use of average values at firm level requires a suitable number of observations.[19] Finally, we dropped a very small number of firms with more than one NACE 3-digit code (i.e. firms with more than one activity) over the considered period in order to get only one NACE 3-digit code per firm at aggregate level. Our final sample covering the period 1999–2010 consists of an unbalanced panel of 12,290 firm-year-observations and is representative of all medium-sized and large firms in the Belgian private sector,[20] with the exception of large parts of the financial sector (NACE K) and the electricity, gas, and water supply industry (NACE D + E).

Descriptive statistics of selected variables are presented in Table 1. They show that the annual firm-level value added per worker represents on average 91,876 EUR and that workers' mean annual labour cost is evaluated at 48,666 EUR.[21] The mean number of required years of education at the firm level equals 12.01, whereas the proportions of over- and undereducated workers are respectively around 20 and 28 per cent. The average years of over- and undereducation within firms are respectively equal to 0.53 and 0.94. Moreover, we find that around 28 per cent of employees within firms are women, 52 per cent are blue-collars, 61 per cent are prime-age workers (i.e. between 30 and 49 years old), 37 per cent have at least 10 years of tenure, and 16 per cent are part-time workers (i.e. work less than 30 hours per week). Firms have an average of 250 employees and are essentially concentrated in the following sectors: manufacturing (53 per cent); wholesale and retail trade, repair of motor vehicles and motorcycles (15 per cent); real estate activities, professional scientific and technical activities, administrative and support service activities (13 per cent); and construction (9 per cent).

5. Estimation results

5.1. Baseline specification

Given aforementioned econometric issues associated with pooled OLS and FE estimates, we directly focus on results obtained with the GMM-SYS estimator.,[2223] Table 2 reports the impact of education variables on per worker averages of added value and labour costs. We first examine the consistency of our estimates by applying Hansen's (1982) and Arellano and Bond's (1991) tests. We do not reject the null hypothesis of valid instruments and of no second-order autocorrelation.

Further results indicate that current productivity is positively and significantly related to its value in the previous year.[24] As regards educational mismatch variables, estimates show that the levels of required and overeducation both exert a significant and positive impact on productivity, whereas the level of undereducation affects productivity negatively. More precisely, an increase in the average required level of education within a firm by 1 year of schooling is associated with a 2.3 per cent increase in the value added of the subsequent year. The same increase is estimated for each additional year of overeducation. Regarding undereducation, a 1.1 per cent decrease in productivity is expected when average undereducation in the firm increases by 1 year.

Table 1. Descriptive statistics of selected variables, 1999–2010

Variables	**Mean**	**Std. Dev.**
Annual value added per worker[a] (€)	91,876	612,545
Annual labour cost per worker[a] (€)	48,666	24,163
Required education (years)	12.01	1.36
Overeducation		
Percentage of workers	20.35	22.08
Years	0.53	0.62
Undereducation		
Percentage of workers	27.72	26.40
Years	0.94	1.03
Workers with 10 years or more of tenure (%)	37.44	23.75
Women (%)	28.33	24.64
Blue-collar workers[b] (%)	52.39	35.22
Share of workers < 30 years	21.78	14.73
Share of workers between 30 and 49 years	60.89	13.79
Share of workers > 49 years	17.33	12.51
Part-time (%)	16.40	17.55
Firm size (number of workers)	250.19	448.32
Firm-level collective agreement (%)	28.69	45.08
Sector (%)		
Mining and quarrying (B)	0.68	
Manufacturing (C)	52.66	
Electricity, gas, steam, and air conditioning supply; Water supply, sewerage, waste management, and remediation activities (D + E)	0.63	
Construction (F)	8.85	
Wholesale and retail trade, repair of motor vehicles and motorcycles (G)	15.28	
Accommodation and food services activities (I)	1.81	
Transport and storage; Information and communication (H + J)	5.88	
Financial and insurance activities (K)	1.59	
Real estate activities; Professional, scientific, and technical activities; Administrative and support serviceactivities (L + M + N)	12.60	
Number of firm-year observations	12,290	

Notes: [a] At 2004 constant prices.
[b] The distinction between blue- and white-collar workers is based on the International Standard Classification of Occupations (ISCO-08). Workers belonging to groups 1–5 are considered to be white-collar workers (1: Managers; 2: Professionals; 3: Technicians and associate professionals; 4: Clerical support workers; 5: Services and sales workers), and those from groups 7–9 are considered to be blue-collar workers (7: Craft and related trades workers; 8: Plant and machine operators and assemblers; 9: Elementary occupations).

Turning to the impact of educational mismatch on labour costs, our GMM-SYS estimates suggest that lagged labour costs positively and significantly influence their current value. Concerning educational variables, it appears that if a firm employs workers in jobs with higher educational requirements it can expect rising labour costs: an increase in the average level of required education by 1 year leads to an increase in labour costs by 1 per cent in the following year. Overeducation is also estimated to have a positive and significant impact on labour costs (an additional year of overeducation is associated with a 2.3 per cent increase in labour costs). Finally, our results suggest that undereducation has no significant influence on average labour costs.

It should be noted that our results for labour costs are not directly comparable to standard results in the educational mismatch literature. Indeed, most previous studies

Table 2. Educational mismatch, productivity, and labour costs

	GMM-SYS	
Estimator/Dependent variables	**Value added per worker (ln)[d]**	**Labour cost per worker (ln)[e]**
Value added per worker (1 year lagged, in ln)	0.623*** (0.043)	
Labour cost per worker (1 year lagged, in ln)		0.706*** (0.059)
Required education (1 year lagged, in years)	0.023*** (0.005)	0.010* (0.006)
Overeducation (1 year lagged, in years)	0.023*** (0.007)	0.023* (0.013)
Undereducation (1 year lagged, in years)	−0.011** (0.004)	−0.006 (0.007)
Worker characteristics[a]	YES	YES
Firm characteristics[b]	YES	YES
Year dummies (11)	YES	YES
Sig. model (*p*-value)	0.000	0.000
Hansen statistic	483.7	538.81
p-value	0.41	0.20
Arellano–Bond statistic (AR2)[c]	1.47	0.81
p-value	0.14	0.42
Number of firm-year observations	12,290	12,290

Notes: Robust standard errors are reported between brackets.
***, **, *Significant at respectively 1%, 5%, and 10% levels.
[a] Share of the workforce that: (i) has at least 10 years of tenure, and (ii) is younger than 30 and older than 49 years, respectively. The shares of women, blue-collar, and part-time workers are also included.
[b] Sectorial affiliation (8 dummies), number of workers, age of the firm, and level of wage bargaining (1 dummy).
[c] AR2 displays the test for second-order autocorrelation in the first-differenced errors.
[d] First and second lags of explanatory variables are used as instruments in the GMM specification, excluding time dummies.
[e] Second and third lags of explanatory variables are used as instruments in the GMM specification, excluding time dummies.

adopt the perspective of individual employees, use wages as a dependent variable, and find that the wage returns of overeducation are lower than the returns to the level of education that is required for a given job (McGuinness, 2006). In contrast, we look at the effects of educational mismatch from the perspective of the firm, which means that the total labour cost is the more relevant outcome variable. As a consequence, we think that our finding that increases in overeducation have a larger effect on labour costs than upgrading to jobs with higher educational requirements is not necessarily at odds with the literature. First, a standard *t*-test for equality of regression coefficients shows that the difference between the labour cost coefficients of required and overeducation is actually not statistically significant. Second, there are theoretical and empirical arguments for why the impact of overeducation on labour costs could be stronger than on wages. For instance, overeducated workers exhibit higher turnover than their adequately educated peers (McGuinness, 2006) — a difference that could generate extra (labour) adjustment costs borne by the employer.[25] Third, even if wages represent the lion share of labour costs, the latter are nevertheless substantially higher, especially in a country like Belgium with very high (and progressive) payroll taxes, including employer social security contributions. It is therefore conceivable that overeducation has both a relatively lower effect than required education on wages and a similar effect on labour costs. Fourth, unlike Mincer-type wage regressions, all of our equations include lagged dependent variables among the regressors. Yet, lagged labour costs will tend to capture the human capital profile of workers in the previous period, so that educational variables will capture only the impacts of *changes* in the shares that have occurred since the previous period, i.e., marginal rather than average effects — which is also why we focus our interpretation on the marginal impact of changes in the firm's educational mix rather than the total impact of over-, required, and undereducation.

What about the impact of educational variables on productivity-wage gaps? Given that mean sample values of productivity and wage costs reach respectively 91,876 and 48,666 EUR, GMM-SYS estimates suggest that when mean years of overeducation in a firm increase by 1 year, annual productivity per worker rises on average the year after by 2,113 EUR (i.e. 0.023 * 91,876 EUR) and the annual wage cost per worker by 1,119 EUR (i.e. 0.023 * 48,666 EUR). Put differently, we find that: (i) profitability (i.e. the productivity-wage gap) depends positively on overeducation (i.e. profitability increases by around 2.3 per cent when overeducation increases by 1 year), and (ii) productivity gains associated with overeducation are shared almost equally between wages and profits. Profitability is also found to be fostered by an increase in the (average) required level of education within a firm. GMM-SYS estimates indeed show that the required level of education within a firm has a stronger impact on productivity than on labour costs. In contrast, profits are found to decrease when the firm's workforce accumulates more years of undereducation with respect to required levels. Overall, results suggest that increasing mean years of required (over) education enhances firm's profitability, whereas undereducation is found to be harmful for firm profits. As a robustness test, we re-estimated our model using the log of the firm-level gross operating surplus per worker (i.e. of the difference between firm-level value added and labour costs per worker) as a dependent variable. GMM-SYS estimates, reported in Appendix A, corroborate our conclusions.

5.2. *Interactions with the firm's environment*

We now examine to what extent our benchmark results for the private sector change when we account for the different environments in which firms operate.

5.2.1. Technology/knowledge intensity. We first estimate whether the effects of educational mismatch on productivity and labour costs depend on the technological/knowledge intensity of a firm's environment. For this purpose, we divide the data set into two subsamples according to the HT/KIS nomenclature (see Section 3.2 for more details) and run separate regressions on 3,888 firm-year observations of high-tech/knowledge-intensive firms and 8,402 firm-year observations of low-tech/knowledge-intensive firms.[26]

The reliability of GMM estimates is assessed through Arellano and Bond (1991) and Hansen (1982) tests. As shown in Table 3, we can neither reject the null hypothesis of valid instruments nor the null hypothesis of no autocorrelation in all four regressions.

According to the productivity coefficients presented in Table 3 (second and third columns), the effects of the educational variables are larger when the firm operates in an environment with greater knowledge intensity. Increasing the level of required education by 1 year is expected to increase firm's productivity by 2.5 per cent in a low-tech environment, but by 3.4 per cent in a high-tech environment.[27] A 1-year increase in the level of overeducation is expected to increase firm's productivity by 5 per cent in a high-tech environment, which is 2 percentage points higher compared with a low-tech one. Increasing the level of undereducation by 1 year is expected to decrease firm productivity by 1.3 per cent in a low-tech environment, but has no significant effect in high-tech firms.

Table 3. Educational mismatch, productivity, and labour costs according to the technological/knowledge intensity

	GMM-SYS			
	Value added per worker (ln)		**Labour cost per worker (ln)**	
Estimator/Dependent variables:	**Low-Tech/ Knowledge intensive[d]**	**High-Tech/ Knowledge intensive[d]**	**Low-Tech/ Knowledge intensive[e]**	**High-Tech/ Knowledge intensive[f]**
Value added per worker	0.535***	0.758***		
(1 year lagged, in ln)	(0.074)	(0.037)		
Labour cost per worker			0.611***	0.756***
(1 year lagged, in ln)			(0.075)	(0.054)
Required education	0.025***	0.034**	0.006	0.025***
(1 year lagged, in years)	(0.007)	(0.014)	(0.007)	(0.009)
Overeducation	0.030***	0.050*	0.018	0.030*
(1 year lagged, in years)	(0.008)	(0.028)	(0.011)	(0.016)
Undereducation	−0.013***	0.004	−0.003	−0.019
(1 year lagged, in years)	(0.005)	(0.015)	(0.007)	(0.012)
Worker characteristics[a]	YES	YES	YES	YES
Firm characteristics[b]	YES	YES	YES	YES
Year dummies (11)	YES	YES	YES	YES
Sig. model (*p*-value)	0.000	0.000	0.000	0.000
Hansen statistic	596.0	337.1	505.5	453.5
p-value	0.49	0.43	0.20	0.32
Arellano–Bond statistic (AR2)[c]	1.63	0.42	0.66	0.48
p-value	0.103	0.68	0.51	0.63
Number of firm-year observations	8,402	3,888	8,402	3,887

Notes: Robust standard errors are reported between brackets.
***, **, *: significant at respectively 1%, 5%, and 10% levels.
[a] Share of the workforce that: (i) has at least 10 years of tenure, and (ii) is younger than 30 and older than 49 years, respectively. The share of women, blue-collar, and part-time workers are also included.
[b] Sectorial affiliation (8 dummies), number of workers, age of the firm, and level of wage bargaining (1 dummy).
[c] AR2 displays the test for second-order autocorrelation in the first-differenced errors.
[d] First and third lags of explanatory variables are used as instruments in the GMM specification, excluding time dummies.
[e] Second and third lags of explanatory variables are used as instruments in the GMM specification, excluding time dummies.
[f] First and second lags of explanatory variables are used as instruments in the GMM specification, excluding time dummies.

Regarding labour costs, estimates in the fourth and fifth columns of Table 3 show that educational variables are only correlated with higher labour costs in high-tech/knowledge environments. In this subsample, increasing the level of required education by 1 year is expected to increase labour costs in the subsequent year by 2.5 per cent. A 1-year increase in the mean years of overeducation is expected to increase labour costs by 3 per cent.

What are the implications for productivity-wage gaps? Taken together, estimates suggest that gains (losses) in productivity are not entirely offset by higher (lower) labour costs. Moreover, extra profits related to levels of required and overeducation are found to be somewhat higher in high-tech/knowledge firms. Finally, we observe that increasing the level of undereducation is expected to decrease profitability among low-tech/knowledge firms.

These findings are supported by our sensitivity test using the log of the gross operating surplus per worker as a dependent variable (see Appendix A).[28]

5.2.2 Economic uncertainty. We now present evidence on whether the educational composition of the firm affects productivity and labour costs differently in more uncertain environments. The first regressions are run on a subsample of 4,685 firm-year observations of firms operating in a more uncertain economic context (for definitions see Section 3.2); the second regressions are based on the subsample of 7,605 firm-year observations of firms evolving in a less uncertain environment.[29]

All regressions pass the Arellano–Bond and Hansen tests. Results in columns two and three of Table 4 suggest that the impact of educational variables on productivity is bigger (in absolute value) when the firm operates in a more uncertain economic context. More precisely, increasing the level of required (over) education by 1 year is expected to increase labour productivity by 2.4 per cent (2.2 per cent) in a more stable environment, whereas it is expected to increase productivity by 3.1 per cent (4 per cent) in a more uncertain environment.[30] Concerning undereducation, a 1-year increase is expected to decrease productivity by 0.9 per cent in a context of low uncertainty, compared with a 1.2 per cent decrease in a more uncertain environment.

Labour costs coefficients presented in the fourth and fifth column of Table 4 suggest stronger effects in a more uncertain environment. Increasing the level of required education by 1 year is expected to raise labour costs by 1.2 per cent and 2.1 per cent under low and high uncertainty, respectively. Concerning overeducation, a 1-year increase in the mean years of overeducation is expected to increase labour costs by 4.1 per cent in a more uncertain economic context, whereas the effect is not significant in the other subsample. Finally, the extent of undereducation is not significant in both subsamples.

Overall, these results imply that required and overeducation improve firms' productivity-wage gaps in both contexts. However, the gains appear to be larger for firms operating in a more uncertain environment. As regards undereducation, estimates show that this variable is slightly more detrimental for profitability in a more uncertain environment. This conclusion is confirmed by our robustness test using a direct measure of profitability as a dependent variable (see Appendix A).

6. Discussion and conclusion

Using a large linked employer–employee panel data set covering the Belgian private sector over the period 1999–2010, this paper provides first evidence regarding the direct impact of educational mismatch on firm productivity, labour costs, and productivity-wage gaps (i.e. profitability). It therefore fills a gap in the literature on over-, required, and undereducation as existing studies have not been able to address frontally the question whether productivity effects associated with over- and undereducation are offset by corresponding changes in labour costs. Moreover, the paper is the first to assess how the impact of educational mismatch on profitability differs according to firms' economic environment.

Our findings — based on the GMM-SYS estimator and controlling for a large set of covariates, simultaneity issues, time-invariant unobserved firm characteristics, and dynamics in the adjustment process of productivity and wages — suggest that educational mismatch has a stronger impact on firm productivity than on labour costs. This gives rise to a profitability profile in the form of an inverted L: at the firm level, undereducation is

Table 4. Educational mismatch, productivity, and labour costs according to the uncertainty of the economic environment

	GMM-SYS			
	Value added per worker (ln)		**Labour cost per worker (ln)**	
Estimator/Dependent variables	**Less uncertain**[d]	**More uncertain**[e]	**Less uncertain**[d]	**More uncertain**[e]
Value added per worker (1 year lagged, in ln)	0.693*** (0.054)	0.637*** (0.070)		
Labour cost per worker (1 year lagged, in ln)			0.690*** (0.082)	0.749*** (0.053)
Required education (1 year lagged, in years)	0.024*** (0.007)	0.031*** (0.009)	0.012* (0.007)	0.021*** (0.007)
Overeducation (1 year lagged, in years)	0.022*** (0.009)	0.040*** (0.012)	0.012 (0.013)	0.041*** (0.011)
Undereducation (1 year lagged, in years)	−0.009* (0.006)	−0.012* (0.007)	−0.010 (0.007)	−0.005 (0.010)
Worker characteristics[a]	YES	YES	YES	YES
Firm characteristics[b]	YES	YES	YES	YES
Year dummies (11)	YES	YES	YES	YES
Sig. model (*p*-value)	0.000	0.000	0.000	0.000
Hansen statistic	679.0	457.0	498.1	467.5
p-value	0.19	0.20	0.41	0.46
Arellano–Bond statistic (AR2)[c]	1.27	1.55	0.66	1.50
p-value	0.20	0.12	0.51	0.13
Number of firm-year observations	7,605	4,685	7,605	4,684

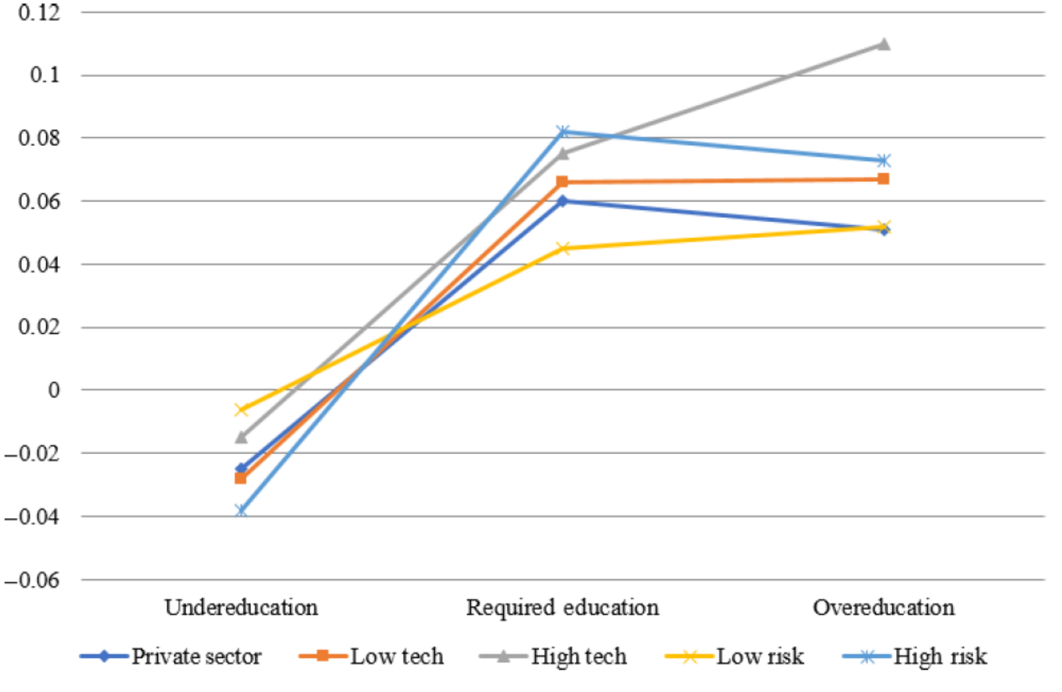
Notes: Robust standard errors are reported between brackets.
***, **, *: significant at respectively 1%, 5%, and 10% levels.
[a] Share of the workforce that: (i) has at least 10 years of tenure, and (ii) is younger than 30 and older than 49 years, respectively. The share of women, blue-collar, and part-time workers are also included.
[b] Sectorial affiliation (8 dummies), number of workers, age of the firm, and level of wage bargaining (1 dummy).
[c] AR2 displays the test for second-order autocorrelation in the first-differenced errors.
[d] First and third lags of explanatory variables are used as instruments in the GMM specification, excluding time dummies.
[e] Second and third lags of explanatory variables are used as instruments in the GMM specification, excluding time dummies.

associated with a negative impact on profits, whereas higher levels of required and overeducation are correlated with positive economic rents of roughly the same magnitude (see Figure 1).[31] This inverted L shape stems from the fact that the positive impact of required and overeducation on productivity is only partly offset by relatively higher labour costs. In other words, both upgrading the required level of education in a firm and hiring workers with credentials beyond prevailing requirements are associated with positive rents. In contrast, the lower productivity of undereducated workers is not associated with significantly lower wages, so that the combined effect on profits is negative.

These findings are consistent with theoretical expectations in light of Belgium's relatively compressed wage structure, a result of labour institutions such as strong centralized collective bargaining. All other things being equal, they could also be expected in other EU countries whose collective bargaining translates into more rigid wage setting, like for instance in France, Germany, or Italy. Our findings suggest that wage compression limits wage decreases for undereducated workers more than they cap wage increases for overeducated workers — a result that could be due to binding wage floors like Belgium's national and sectoral minimum wages (Garnero *et al.*, 2014). The increased profitability associated with upgrading the firm's job mix and/or hiring more overeducated workers can also be the result of profit-maximizing rent extraction by the employers. It should be noted that the policy implications of these alternative explanations are somewhat different: if rigidities in wage bargaining institutions drive our results, then labour market policies should be made more flexible to bring labour costs closer to productivity; in contrast, if the higher profitability is seen as unfair rent extraction at the expense of employees, this calls for stronger wage bargaining institutions so that highly educated workers can capture more of the economic rent that their employment appears to entail.

We further show that the profitability profile of educational mismatch varies with respect to the environment in which the firm operates. Although the subsample of low-tech/knowledge firms displays a similar inverted L shape than the baseline regression, the profile is steeper in high-tech/knowledge firms: for them, hiring beyond educational requirement is even more attractive as overeducated workers have such a high positive effect on productivity that overeducation easily offsets the associated hike in labour costs (see Figure 1). This is further evidence for wage compression due to the above-mentioned labour institutions, as higher productivity among overeducated workers in high-tech/knowledge firms or riskier environments does not translate into proportionate additional labour costs.

Figure 1. Comparison of profitability profiles of educational variables (based on Appendix A).



Comparing the profitability profiles of educational mismatch in firms facing low and high uncertainty confirms theoretical expectations: the profile of firms in relatively stable environments is more flat, meaning that the effects of under- and overeducation are slightly less consequential compared with the baseline profile. In particular, undereducation is not associated with negative rents, arguably because firms in stable environments are able to avoid unprofitable mismatches. In uncertain environments, the inversed L profile is stretched, with all effects being larger compared with the baseline regression. This suggests that uncertainty amplifies the effects of educational mismatch on profitability rather than changing their sign.

To conclude, the results presented in this paper underline important caveats of relying on the conventional human capital hypothesis assuming that wage differentiation in terms of educational credentials reflects productivity differences. First, increasing the amount of jobs with higher educational requirements appears to be associated with productivity gains that surpass hikes in labour costs. In Belgium, the resulting rents are captured by firms in the form of higher profits.[32] Second, although the human capital hypothesis correctly predicts higher productivity for overeducated workers, hiring beyond educational requirements is also found to be profitable, especially in high-tech/knowledge and more uncertain environments. This supports the idea, supported by Nelson and Phelps (1966) and Bulmahn and Kräkel (2002) among others, according to which higher, and by extension, more educated workers are more adaptable, more reactive, and thus more valuable in moving technological and more unstable contexts. Finally, our estimates show that firms employing undereducated workers are not only less productive but also less profitable. Given that undereducation is a sizable phenomenon in all advanced economies (Quintini, 2011), this is an alarming result. It notably calls for more initiatives to tackle bottleneck vacancies (i.e. labour shortages) and to ensure that workers' skills and knowledge remain up-to-da3te.[33]

Appendix A. Educational mismatch and profitability (GMM-SYS estimates)

	Profit per worker (ln)[a]				
Dependent variable	**Overall sample[e]**	**Low-Tech/ Knowledge Intensive[e]**	**High-Tech/ Knowledge intensive[f]**	**Less uncertain[g]**	**More uncertain[e]**
Profit per worker	0.413***	0.417***	0.599***	0.539***	0.424***
(1 year lagged, in ln)	(0.029)	(0.035)	(0.039)	(0.033)	(0.048)
Required education	0.060***	0.066***	0.075***	0.045**	0.082***
(1 year lagged, in years)	(0.017)	(0.019)	(0.025)	(0.019)	(0.023)
Overeducation	0.051**	0.067**	0.110***	0.052*	0.073*
(1 year lagged, in years)	(0.026)	(0.030)	(0.042)	(0.030)	(0.039)
Undereducation	−0.025*	−0.028*	−0.015	−0.006	−0.038*
(1 year lagged, in years)	(0.014)	(0.016)	(0.030)	(0.016)	(0.023)
Worker characteristics[b]	YES	YES	YES	YES	YES
Firm characteristics[c]	YES	YES	YES	YES	YES
Year dummies (11)	YES	YES	YES	YES	YES
Sig. model (*p*-value)	0.000	0.000	0.000	0.000	0.000
Hansen statistic	500.72	492.1	626.7	508.9	458.0
p-value	0.22	0.32	0.11	0.21	0.56

Appendix A. Continued

Dependent variable	Profit per worker (ln)[a] Overall sample[e]	Low-Tech/ Knowledge Intensive[e]	High-Tech/ Knowledge intensive[f]	Less uncertain[g]	More uncertain[e]
Arellano–Bond statistic (AR2)[d]	1.39	2.46	0.33	0.85	1.87
p-value	0.17	0.01	0.74	0.39	0.06
Number of firm-year observations	12,290	8,402	3,888	7,605	4,685

Notes: Robust standard errors are reported between brackets.

***, **, *Significant at respectively 1%, 5%, and 10% levels.

[a] The profit per worker corresponds to the gross operating surplus per worker, i.e., the difference between the value added (at factor costs) per worker and the labour cost per worker.

[b] Share of the workforce that: (i) has at least 10 years of tenure, and (ii) is younger than 30 and older than 49 years, respectively. The shares of women, blue-collar, and part-time workers are also included.

[c] Sectorial affiliation (8 dummies), number of workers, age of the firm, and level of wage bargaining (1 dummy).

[d] AR2 displays the test for second-order autocorrelation in the first-differenced errors.

[e] First and second lags of explanatory variables are used as instruments in the GMM specification, excluding time dummies.

[f] Second and third lags of explanatory variables are used as instruments in the GMM specification, excluding time dummies.

[g] First and third lags of explanatory variables are used as instruments in the GMM specification, excluding time dummies.

Appendix B. Descriptive statistics of selected variables according to technology/knowledge intensity, 1999–2010

Variables	Low-tech/knowledge intensive Mean	Std. Dev.	High-tech/knowledge intensive Mean	Std. Dev.
Annual value added per worker[a] (€)	87,925	733,064	100,416	157,160
Annual labour cost per worker[a] (€)	45,325	19,559	55,887	30,704
Required education (years)	11.73	1.26	12.61	1.36
Overeducation				
Percentage of workers	19.06	22.03	23.14	21.95
Years	0.51	0.64	0.57	0.59
Undereducation				
Percentage of workers	29.55	27.28	23.77	23.91
Years	1.03	1.09	0.75	0.87
Workers with 10 years or more of tenure (%)	38.71	22.77	34.69	25.53
Women (%)	26.09	23.76	33.17	25.81
Blue-collar workers[b] (%)	58.96	33.00	38.21	35.71
Share of workers < 30 years	21.15	14.09	23.15	15.92
Share of workers between 30 and 49 years	60.85	13.54	60.96	14.33
Share of workers > 49 years	18.00	12.37	15.89	12.69
Part-time (%)	17.29	17.34	14.49	17.86

Appendix B. Continued

Variables	Low-tech/knowledge intensive		High-tech/knowledge intensive	
	Mean	Std. Dev.	Mean	Std. Dev.
Firm size (number of workers)	211.94	310.60	332.85	645.72
Firm-level collective agreement (%)	27.15	44.30	32.01	46.55
Sector (%)				
Mining and quarrying (B)	0.99		0.00	
Manufacturing (C)	51.86		54.40	
Electricity, gas, steam, and air conditioning supply; Water supply, sewerage, waste management, and remediation activities (D + E)	0.93		0.00	
Construction (F)	12.95		0.00	
Wholesale and retail trade, repair of motor vehicles and motorcycles (G)	22.35		0.00	
Accommodation and food services activities (I)	2.65		0.00	
Transport and storage; Information and communication (H + J)	7.03		3.42	
Financial and insurance activities (K)	0.00		5.02	
Real estate activities, Professional, scientific, and technical activities; Administrative and support service activities (L + M + N)	1.24		37.16	
Number of firm-year observations	8,402			3,888

Notes: [a] At 2004 constant prices.
[b] The distinction between blue- and white-collar workers is based on the International Standard Classification of Occupations (ISCO-08). Workers belonging to groups 1–5 are considered to be white-collar workers (1: Managers; 2: Professionals; 3: Technicians and associate professionals; 4: Clerical support workers; 5: Services and sales workers), and those from groups 7–9 are considered to be blue-collar workers (7: Craft and related trades workers; 8: Plant and machine operators and assemblers; 9: Elementary occupations).

Appendix C. Educational mismatch and profitability according to the technological/knowledge intensity: sensitivity analysis

Estimator/ Dependent variable:	GMM-SYS/Profit per worker (ln)			
	Low-Tech/Less Knowledge Intensive[d]	Medium-Low-Tech/Less Knowledge Intensive[d]	Medium-High-Tech/ Knowledge Intensive[d]	High-Tech/ Knowledge Intensive[e]
Profit per worker	0.468***	0.446***	0.664***	0.599***
(1 year lagged, in ln)	(0.040)	(0.042)	(0.039)	(0.039)
Required education	0.054**	0.105***	0.091**	0.075***
(1 year lagged, in years)	(0.026)	(0.028)	(0.024)	(0.025)
Overeducation	0.063**	0.095***	0.098*	0.110***
(1 year lagged, in years)	(0.029)	(0.035)	(0.040)	(0.042)
Undereducation	−0.036*	−0.036	−0.009	−0.015
(1 year lagged, in years)	(0.020)	(0.023)	(0.030)	(0.030)

Appendix C. Continued

Estimator/ Dependent variable:	GMM-SYS/Profit per worker (ln)			
	Low-Tech/Less Knowledge Intensive[d]	**Medium-Low-Tech/Less Knowledge Intensive[d]**	**Medium-High-Tech/ Knowledge Intensive[d]**	**High-Tech/ Knowledge Intensive[e]**
Worker characteristics[a]	YES	YES	YES	YES
Firm characteristics[b]	YES	YES	YES	YES
Year dummies (11)	YES	YES	YES	YES
Sig. model (p-value)	0.000	0.000	0.000	0.000
Hansen statistic	413.5	444.92	441.97	626.7
p-value	0.36	0.45	0.08	0.11
Arellano–Bond statistic (AR2)[c]	2.53	1.70	0.98	0.33
p-value	0.11	0.09	0.33	0.74
Number of firm-year observations	5,203	4,528	3,489	3,888

Notes: Robust standard errors are reported between brackets.
***, **, *: significant at respectively 1%, 5%, and 10% levels.
[a] Share of the workforce that: (i) has at least 10 years of tenure, and (ii) is younger than 30 and older than 49 years, respectively. The share of women, blue-collar, and part-time workers are also included.
[b] Sectorial affiliation (8 dummies), number of workers, age of the firm, and level of wage bargaining (1 dummy).
[c] AR2 displays the test for second-order autocorrelation in the first-differenced errors.
[d] First and second lags of explanatory variables are used as instruments in the GMM specification, excluding time dummies.
[e] First and third lags of explanatory variables are used as instruments in the GMM specification, excluding time dummies.

Appendix D. Descriptive statistics of selected variables according to economic uncertainty, 1999–2010

Variables	Less uncertain		More uncertain	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Annual value added per worker[a] (€)	103,636	775,970	72,788	94,023
Annual labour cost per worker[a] (€)	49,189	22,204	47,817	27,024
Required education (years)	12.09	1.22	11.87	1.55
Overeducation				
Percentage of workers	19.78	20.83	21.29	23.96
Years	0.51	0.58	0.56	0.68
Undereducation				
Percentage of workers	28.80	26.31	25.97	26.46
Years	0.98	1.02	0.87	1.05
Workers with 10 years or more of tenure (%)	41.03	24.16	31.62	21.86
Women (%)	26.92	23.70	30.62	25.94
Blue-collar workers[b] (%)	56.23	32.90	46.16	37.89
Share of workers < 30 years	20.94	14.30	23.14	15.30

Appendix D. Continued

Variables	Less uncertain Mean	Less uncertain Std. Dev.	More uncertain Mean	More uncertain Std. Dev.
Share of workers between 30 and 49 years	61.50	13.55	59.92	14.13
Share of workers > 49 years	17.56	12.45	16.94	12.61
Part-time (%)	15.41	15.54	18.01	20.30
Firm size (number of workers)	274.20	512.44	211.22	313.97
Firm-level collective agreement (%)	34.75	47.51	18.85	38.88
Sector (%)				
Mining and quarrying (B)	1.10		0.00	
Manufacturing (C)	77.50		12.34	
Electricity, gas, steam, and air conditioning supply; Water supply, sewerage, waste management, and remediation activities (D + E)	1.03		0.00	
Construction (F)	0.79		21.94	
Wholesale and retail trade, repair of motor vehicles and motorcycles (G)	2.85		35.45	
Accommodation and food services activities (I)	0.99		3.16	
Transport and storage; Information and communication (H + J)	4.16		8.69	
Financial and insurance activities (K)	2.18		0.62	
Real estate activities; Professional, scientific, and technical activities; Administrative and support service activities (L + M + N)	9.40		17.80	
Number of firm-year observations		7,605		4,685

Notes: [a] At 2004 constant prices.
[b] The distinction between blue- and white-collar workers is based on the International Standard Classification of Occupations (ISCO-08). Workers belonging to groups 1–5 are considered to be white-collar workers (1: Managers; 2: Professionals; 3: Technicians and associate professionals; 4: Clerical support workers; 5: Services and sales workers), and those from groups 7–9 are considered to be blue-collar workers (7: Craft and related trades workers; 8: Plant and machine operators and assemblers; 9: Elementary occupations).

Notes

[1] In this paper, we address only peripherally the issue of the performance of occupational training or schooling over time. Indeed, the productivity of a specific occupational training or schooling obtained at time t could be lower or higher to the same degree obtained at $t + 1$. Overall, the empirical literature suggests that the productivity associated with a given educational level tends to diminish rather than to increase over time; this might notably give rise to a phenomenon of 'horizontal mismatch' (Domadenik *et al.*, 2013; Pecoraro, 2016).

[2] Developed by Freeman (1976), educational mismatch (or simply over- and undereducation) refers to the difference between the worker's attained level of education and the education required in her job.

[3] Results for Belgium, reported in the same study, are slightly lower as regards overeducation (15 per cent) and about the same for undereducation.

[4] By definition, the gap between productivity (i.e. value added at factors costs) and wage costs corresponds to the gross operating surplus.

[5] There is still very little *direct* evidence on whether overeducation is beneficial or detrimental for firm productivity (Grunau, 2016; Quintini, 2011). However, results of Kampelmann and Rycx (2012) for the Belgian private sector suggest that the net effect is significant and positive. The authors do not exclude that, for a given job, educational mismatch may lead to less job satisfaction and worse

correlated workers' attitudes and behaviours. However, they conclude that productivity gains associated with overeducation are on average larger then potential losses. Whether these productivity gains translate into higher profitability remains an open question. Moreover, results are likely to vary across working environments. For instance, it could be assumed that an overeducated worker with a PhD in statistics might add value to a bank's trading room or to a high-tech company facing rapidly changing market conditions. In contrast, the same person employed in a traditional retail store or cleaning company could be less profitable than his adequately educated colleagues doing the same job, due to frustration and worse productivity-related characteristics.

[6]The workers' educational attainment is available in seven categories in our data set. This information, reported by firms' human resource departments (on the basis of their registers), has been transformed in years of education. To this end, we applied the following rule: (1) primary education: 6 years of education; (2) lower secondary education: 9 years of education; (3–4) general, technical, and artistic upper secondary education: 12 years of education; (5) higher non-university education, short: 14 years of education; (6) university and non-university education, long: 16 years of education; (7) post-graduate education: 17 years of education.

[7]As a robustness test, we computed the required education for a given job by taking the mode of workers' years of education within detailed ISCO 3-digit occupations (113 categories) *and* industries (NACE 1-digit nomenclature) for each of the 12 years in our sample. Moreover, we also tested the sensitivity of our estimates controlling for the birth cohort of workers, that is, computing the required education separately among young and older workers. Results, available on request, support our conclusions.

[8]The control variables that have been included in our regressions are in line with extant literature (for a review of the set of covariates that should be included in this type of analysis see, e.g., Göbel and Zwick, 2009). As highlighted by Mahlberg *et al.* (2013: 10): 'by including a rather broad set of independent variables, we account for heterogeneity among firms, in order to mitigate the bias that could be caused by omitted variables'.

[9]Note that: $\frac{1}{m_{j,t}}\left(\sum_{i=1}^{m_{j,t}} O_{i,j,t} + \sum_{i=1}^{m_{j,t}} R_{i,j,t} - \sum_{i=1}^{m_{j,t}} U_{i,j,t}\right) = \frac{1}{m_{j,t}}\sum_{i=1}^{m_{j,t}} Attained_{i,j,t}$, i.e., the sum of the average years of over- and required education minus the average years of undereducation in firm j at time t is equal to the average years of education attained by the workers employed in firm j at time t.

[10]The assumption of persistent productivity both at the industry and firm level is strongly supported in the literature (see, e.g., Bartelsman and Doms, 2000). Researchers 'documented, virtually without exception, enormous and persistent measured productivity differences across producers, even within narrowly defined industries' (Syverson, 2011: 326). Large parts of these productivity differences are still hard to explain. The persistence of wage costs is also highlighted in the literature (see, e.g., Fuss and Wintr, 2009; Heckel *et al.*, 2008). Wage stickiness is notably the outcome of labour market institutions, adjustment costs and efficiency wages' motives.

[11]Expected biases associated with OLS and the relatively poor performance and shortcomings of the FE estimator in the context of firm-level productivity regressions are reviewed in Van Beveren (2012).

[12]We have performed such a test using a 2SLS estimator on an equation in levels in which our variables of interest have been instrumented by their first differences. Both equations (i.e. value added and wage costs) pass standard underidentification and weak identification tests. This means that the endogeneity test for the over-, required, and undereducation variables is valid. This test suggests that for both equations we have to reject the null hypothesis that our main variables of interest can actually be treated as exogenous.

[13]It is standard in the literature to use dynamic panel data methods such as those proposed by Arellano and Bond (1991) to overcome key econometric issues, in particular, lag dependency, firm fixed effects, and endogeneity of input shares. Accordingly, many recent papers rely on dynamic GMM methods to estimate the impact of workforce and job characteristics on productivity and/or labour costs (see, e.g., Göbel and Zwick, 2012, 2013; Ilmakunnas and Ilmakunnas, 2011; Mahlberg *et al.*, 2013; Nielen and Schiersch, 2012, 2014).

[14] Bond and Söderbom (2005) provide a review of the literature regarding the identification of production functions. The authors notably highlight that adjustment costs of labour and capital can justify the use of lagged values (of endogenous variables) as instruments.

[15] The SES is a stratified sample. The stratification criteria refer respectively to the region (NUTS-groups), the principal economic activity (NACE-groups), and the size of the firm. The sample size in each stratum depends on the size of the firm. Sampling percentages of firms are respectively equal to 10, 50, and 100 per cent when the number of workers is between 10 and 50, between 50 and 99, and above 100. Within a firm, sampling percentages of employees also depend on size. Sampling percentages of employees reach respectively 100, 50, 25, 14.3, and 10 per cent when the number of workers is between 10 and 20, between 20 and 50, between 50 and 99, between 100 and 199, and between 200 and 299. Firms employing 300 workers or more have to report information for an absolute number of employees. This number ranges between 30 (for firms with between 300 and 349 workers) and 200 (for firms with 12,000 workers or more). To guarantee that firms report information on a representative sample of their workers, they are asked to follow a specific procedure. First, they have to rank their employees in alphabetical order. Next, Statistics Belgium give them a random letter (e.g. the letter O) from which they have to start when reporting information on their employees (following the alphabetical order of workers' names in their list). If they reach the letter Z and still have to provide information on some of their employees, they have to continue from the letter A in their list. Moreover, firms that employ different categories of workers, namely managers, blue- and/or white-collar workers, have to set up a separate alphabetical list for each of these categories and to report information on a number of workers in these different groups that is proportional to their share in total firm employment. For example, a firm with 300 employees (namely, 60 managers, 180 white-collar workers, and 60 blue-collar workers) will have to report information on 30 workers (namely, 6 managers, 18 white-collar workers, and 6 blue-collar workers). Finally, let us notice that no threshold at the upper limit of wages is found in the SES. To put it differently, wages are not censored. For an extended discussion see Demunter (2000).

[16] Note that the coverage of the SBS is not the same as that of the SES, as it does not cover the entire financial sector (NACE K).

[17] For instance, we eliminate a (very small) number of firms for which the recorded value added was negative.

[18] We did some robustness tests by fixing the threshold at 50 observations. However, given that the number of data points per occupation at the ISCO 3-digit level is quite large, this alternative threshold has little effect on sample size and leaves results (available on request) unaffected.

[19] This restriction is unlikely to affect our results as it leads to a very small drop in sample size.

[20] Larger firms are likely to employ a lower share of overeducated workers because they generally have more sophisticated HRM procedures (notably in terms of recruitment) and a wider range of jobs (Dolton and Silles, 2008). Moreover, the required level of education is probably better defined in bigger firms. As a result, the fact that medium and large firms are over-represented in our sample may underestimate the incidence of overeducation. Yet, caution is required. Indeed, empirical results provided by Karakaya *et al.* (2007) suggest that the impact of firm size on overeducation is very weak in the Belgian private sector. Using matched employer–employee data for 1995, the authors suggest that the likelihood for a worker to be overeducated decreases by only 0.1 per cent *ceteris paribus* if firm size increases by 100 extra workers.

[21] All variables measured in monetary terms have been deflated to constant prices of 2004 by the consumer price index taken from Statistics Belgium.

[22] OLS and FE estimates, not reported here for the sake of conciseness, are available on request.

[23] Detailed dynamic system GMM estimates, including control variables, are also available on request. Interestingly, regression coefficients associated with the covariates are in line with earlier findings. Industry dummies, for instance, are generally significant and they follow a similar pattern than those reported in the literature on inter-industry wage differentials (see, e.g., du Caju *et al.* (2012)). Among the highly productive sectors, we notably find the electricity, gas, and water supply industry (NACE D and E) and financial and insurance activities (NACE K). Not surprisingly, as

shown in du Caju *et al.* (2011), these sectors are also found at the top of the conditional wage distribution. The coefficient on part-time is found to be significantly negative. This corroborates estimates of Devicienti *et al.* (2018) showing that firms employing more part-timers are *ceteris paribus* less productive. An insignificant coefficient for blue-collar workers is also reported in Kampelmann and Rycx (2012). The authors find that occupations play different roles for remuneration and productivity in the Belgian private sector. Although their estimations indicate a significant upward-sloping occupational wage-profile, they cannot reject the hypothesis of a flat productivity profile. Finally, the insignificant coefficient associated with the share of women is in line with Garnero *et al.* (2014). The latter show that women are associated with economic rents.

[24] Overall, this confirms that productivity (and wage costs, see next paragraph) are highly persistent at the firm level. Moreover, GMM coefficients on lagged dependent variables fall systematically between the OLS and FE estimates (available on request). As highlighted by Roodman (2009), these results support the appropriateness of our dynamic GMM-SYS specification.

[25] As highlighted in Section 3.1, our measure of labour costs refers to the firm's total wage bill, including fixed and variable pay components, in kind benefits, employer-funded extra-legal advantages (related, e.g., to health, early retirement, or pension), hiring and firing costs, and payroll taxes. Labour adjustment costs are thus taken into account.

[26] Descriptive statistics of selected variables related to the first environment (i.e. technology/knowledge intensity) are presented in Appendix B. They show that the mean number of required years of education at the firm level equals 11.7 in low-tech/knowledge-intensive firms and 12.6 in a high-tech/knowledge environment, whereas the proportions of over- and undereducated workers are respectively 19.1 and 29.6 per cent in low-tech versus 23.1 and 23.8 per cent in high-tech firms. Low-tech firms present a lower (higher) proportion of over- (under) educated workers compared with their high-tech counterparts. Results also show that the mismatch incidence is more balanced between over- and undereducated workers in high-tech firms than in low-tech ones, where the latter are over-represented. Such statistics therefore seem to provide evidence that firms in a high-tech environment hire overeducated workers to a bigger extent than those behaving in a lower one.

[27] Standard *t*-tests for equality of coefficients (available on request) indicate that differences in regression coefficients (associated with a given educational variable) across economic environments are statistically significant. In contrast, as for our benchmark specification, *t*-tests show that regression coefficients associated with required and overeducation, for a given economic environment, are never statistically different from each other. These comments also apply to results discussed in Section 5.2.2.

[28] They are also robust to sensitivity analysis regarding the HT/KIS indicator: rather than using two groups (high and low level of knowledge intensity), we categorized firms into four levels of technology and knowledge intensity (low-tech/less knowledge intensive, medium-low-tech less knowledge intensive, medium-high-tech knowledge intensive and high-tech/knowledge intensive). As shown in Appendix C, the higher the degree of knowledge intensity, the more overeducation is expected to provide extra profits, suggesting a roughly linear impact of (at least) overeducation.

[29] Descriptive statistics of selected variables related to the second environment (i.e. economic uncertainty) are presented in the Appendix D. They show a higher spread between the percentages of overeducated and undereducated workers in less uncertain economic environments representing respectively 19.8 per cent and 28.8 per cent, to be compared with 21.3 per cent and 26.0 per cent in the case of more uncertain environments. It thus means that if the overall incidence of mismatch is rather close in both environments, firms in a higher uncertainty environment tend to hire overeducated workers to a somewhat bigger extent than those that face lower uncertainty.

[30] Again, *t*-tests for equality of regression coefficients (available on request) indicate that differences in coefficients (associated with a given educational variable) across economic environments are statistically significant.

[31] The vertical axis of Figure 1 describes differences in profits associated with additional years of under-, required, and overeducation among firms in the whole private sector and specific environments (knowledge intensity and high risk). The coefficients presented in this figure correspond to

those provided in Appendix A. For example, the estimate -0.025 for undereducation in the whole private sector means a 2.5 per cent decrease in profits per additional year of undereducation in these firms.

[32] These results echo the estimates of Konings and Vanormelingen (2015). Using Belgian firm-level panel data, the latter show that on-the-job training (i.e. another component of workers' human capital) increases firms' profitability. Put differently, their results indicate that the productivity premium of a trained worker is substantially higher compared with the wage premium, an outcome that appears to be consistent with recent theories explaining work-related training by imperfect competition in the labour market.

[33] For examples of initiatives that could be implemented to reach these goals see European Commission (2014).

References

Aghion P., Howitt P. and Violante G. (2003) 'Wage Inequality and Technological Change: A Nelson-Phelps Approach', in Aghion P., Frydman R., Stiglitz J. and Woodford M. (eds.) *Knowledge, Information, and Expectations in Modern Macroeconomics.* Honor of Edmund S. Phelps, Princeton University Press: 443–61.

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* 58: 277–297.

Arellano M. and Bover O. (1995) 'Another Look at the Instrumental Variable Estimation of Error-Components Models', *Journal of Econometrics* 68: 29–51.

Baert S. and Verhaest D. (2014) 'Unemployment or Overeducation: Which is a Worse Signal to Employers?', IZA Discussion Papers, 8312.

Baert S., Cockx B. and Verhaest D. (2013) 'Overeducation at the Start of the Career: Stepping Stone or Trap?', *Labour Economics* 25: 123–140.

Bartelsman E. and Doms M. (2000) 'Understanding Productivity: Lessons From Longitudinal Microdata', *Quarterly Journal of Economics* 122: 1721–1758.

Barth E., Bratsberg B., Haegeland T. and Raaum O. (2008) 'Who Pays for Performance?', *International Journal of Manpower* 29: 8–29.

Battu H., Seaman P. and Sloane P. (1999) 'Overeducation, Undereducation and the British Labour Market', *Applied Economics* 31: 1437–1453.

Becker G. (1975) *Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education*. Chicago: University of Chicago Press.

Billon M., Marco R. and Lera-Lopez F. (2017) 'Innovation and ICT use in the EU: An Analysis of Regional Drivers', *Empirical Economics* 53: 1083–1108.

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

Bond S. (2002) 'Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice', *Portuguese Economic Journal* 1: 141–162.

Bond S. and Söderbom M. (2005) 'Adjustment Costs and the Identification of Cobb Douglas Production Functions', IFS Working Paper, 05/04.

Broecke S. (2016) 'Do Skills Matter for Wage Inequality?' *IZA World of Labor*, 2016-232.

Broecke S., Quintini G. and Vandeweyer M. (2017) 'Explaining International Differences in Wage Inequality: Skills Matter', *Economics of Education Review* 60: 112–24.

Bulmahn G. and Kräkel M. (2002) 'Overeducated Workers as an Insurance Device', *Labour* 16: 383–402.

du Caju P., Rycx F. and Tojerow I. (2011) 'Inter-Industry Wage Differentials: How Much Does Rent-Sharing Matter?', *Manchester School* 79: 691–717.

du Caju P., Rycx F. and Tojerow I. (2012) 'Wage Structure Effects of International Trade in a Small Open Economy: The Case of Belgium', *Review of World Economics* 148: 297–331.

Card D. (1999) '*The Causal Effect of Education on Earnings*', in *Handbook of Labour Economics* (Eds) O. Ashenfelter and Card D., Vol. 3, North-Holland, Amsterdam.

Cardoso A. (2010) 'Do Firms Compress the Wage Distribution?', in Marsden D. and Rycx F. (eds.) *Wage Structures, Employment Adjustments and Globalisation: Evidence From Linked and Firm-Level Panel Data*. UK: Palgrave Macmillan: 202–218.

Cardoso A., Guimaraes P. and Varejao J. (2011) "Are Older Workers Worthy of Their pay?', *An Empirical Investigation of age-Productivity and age-Wage Nexuses', De Economist* 159: 95–111.

Damiani M., Pompei F. and Ricci A. (2016) 'Performance Related pay, Productivity and Wages in Italy: A Quantile Regression Approach', *International Journal of Manpower* 37: 232–256.

Deming D. and Noray K. (2018) 'STEM Careers and Technological Change', NBER Working Paper, 25065.

Demunter C. (2000) 'Structure and Distribution of Earnings Survey', DGSIE Working Paper.

Devicienti F., Grinza E. and Vannoni D. (2018) 'The Impact of Part-Time Work on Firm Total Factor Productivity: Evidence From Italy', *Industrial and Corporate Change* 27: 321–347.

Dolado J. J., Felgueroso F. and Jimeno J. F. (2000) 'Explaining Youth Labor Market Problems in Spain: Crowding-out, Institutions, or Technology Shifts?', *European Economic Review* 44: 943–956.

Dolton P. and Silles M. (2008) 'The Effects of Over-Education on Earnings in the Graduate Labour Market', *Economics of Education Review* 27: 125–139.

Domadenik P., Farcnik D. and Pastore F. (2013) 'Horizontal Mismatch in the Labour Market of Graduates: The Role of Signalling', IZA Discussion Paper, 7527.

Duncan G. and Hoffman S. (1981) 'The Incidence and Wage Effects of Overeducation', *Economics of Education Review* 1: 75–86.

European Commission (2014) *Mapping and Analyzing Bottleneck Vacancies in EU Labour Markets – Overview Report*. Luxembourg: Publications Office of the European Union.

Eurostat (2012) *High-Tech Statistics – Statistics Explained*. Luxembourg: Eurostat.

Eurostat (2018) *Labour Market Statistics*. Luxembourg: Eurostat.

Fares J. and Yuen T. (2003) 'Technological Change and the Education Premium in Canada: Sectoral Evidence', Bank of CanadaWorking Paper, 2003-18.

Figueiredo H., Rocha V., Biscaia R. and Teixeira P. (2015) 'Gender pay Gaps and the Restructuring of Graduate Labour Markets in Southern Europe', *Cambridge Journal of Economics* 39: 565–598.

Foss N. J. and Laursen K. (2005) 'Performance pay, Delegation and Multitasking Under Uncertainty and Innovativeness: An Empirical Investigation', *Journal of Economic Behavior & Organization* 58: 246–276.

Freeman R. (1976) *The Overeducated American*. New York: Academic Press.

Fuss C. and Wintr L. (2009) 'Rigid Labour Compensation and Flexible Employment? Firm-Level Evidence With Regard to Productivity for Belgium', ECB Working Paper, 1021.

Galasi P. (2008) 'The Effect of Educational Mismatch on Wages for 25 Countries', Budapest Working Papers on the Labour Market, 8.

Garnero A., Kampelmann S. and Rycx F. (2014) 'Part-Time Work, Wages and Productivity: Evidence From Belgian Matched Panel Data', *Industrial and Labor Relations Review* 67: 926–954.

Garnero A., Kampelmann S. and Rycx F. (2015a) 'Minimum Wage Systems and Earnings Inequalities: Does Institutional Diversity Matter?', *European Journal of Industrial Relations* 21: 115–30.

Garnero A., Kampelmann S. and Rycx F. (2015b) "Sharp Teeth or Empty Mouths?', *European Institutional Diversity and the Sector-Level Minimum Wage Bite', British Journal of Industrial Relations* 53: 760–788.

Giuliano R., Kampelmann S., Mahy B. and Rycx F. (2017) "Short Notice, Big Difference ?', *The Effect of Temporary Employment on Firm Competitiveness Across Sectors', British Journal of Industrial Relations* 55: 421–449.

Göbel C. and Zwick T. (2009) ‘Age and Productivity: Evidence From Linked Employer Employee Data’, ZEW Discussion Paper, 09-020.

Göbel C. and Zwick T. (2012) ‘Age and Productivity: Sector Differences’, *De Economist* 160: 35–57.

Göbel C. and Zwick T. (2013) ‘Are Personnel Measures Effective in Increasing Productivity of old Workers?’, *Labour Economics* 22: 80–93.

Groot W. and Maassen van den Brink H. (2000) ‘Overeducation in the Labor Market: A Meta-Analysis’, *Economics of Education Review* 19: 149–158.

Grunau P. (2016) ‘The Impact of Overeducated and Undereducated Workers on Firm-Level Productivity: First Evidence for Germany’, *International Journal of Manpower* 37: 258–283.

Hannan M. T. and Freeman J. (1989) *Organizational Ecology*. Harvard: University Press.

Hansen L. (1982) ‘Large Sample Properties of Generalized Method of Moments Estimators’, *Econometrica* 50: 1029–1054.

Hanushek E., Schwerdt G., Wiederhold S. and Woessmann L. (2017) ‘Coping With Change: International Differences in the Returns to Skills’, *Economics Letters* 153: 15–19.

Hartog J. (2000) ‘Over-Education and Earnings: Where are we, Where Should we go?’, *Economics of Education Review* 19: 131–147.

Heckel T., Le Bihan H. and Montornès J. (2008) ‘Sticky Wages: Evidence From Quarterly Microeconomic Data’, ECB Working Paper, 893.

Hellerstein J. and Neumark D. (2004) ‘Production Function and Wage Equation Estimation With Heterogeneous Labor: Evidence From a new Matched Employer-Employee Data set’, NBER Working Paper, 10365.

Hellerstein J., Neumark D. and Troske K. (1999) ‘Wages, Productivity and Worker Characteristics: Evidence From Plant-Level Production Functions and Wage Equations’, *Journal of Labor Economics* 17: 409–446.

Hempell T. (2005) ‘‘What’s Spurious? What’s Real?’, *Measuring the Productivity Impacts of ICT at the Firm Level’, Empirical Economics* 30: 427–464.

Ilmakunnas P. and Ilmakunnas S. (2011) ‘Diversity at the Workplace: Whom Does it Benefit?’, *De Economist* 159: 223–255.

Kampelmann S. and Rycx F. (2011) ‘‘Are Occupations Paid What They are Worth?’, *An Econometric Study of Occupational Wage Inequality and Productivity’, De Economist* 160: 257–287.

Kampelmann S. and Rycx F. (2012) ‘The Impact of Educational Mismatch on Firm Productivity: Evidence From Linked Panel Data’, *Economics of Education Review* 31: 918–931.

Karakaya G., Plasman R. and Rycx F. (2007) ‘Overeducation on the Belgian Labour Market: Evaluation and Analysis of the Explanatory Factors Through two Types of Approaches’, *Compare* 37: 513–532.

Konings J. and Vanormelingen S. (2015) ‘The Impact of Training on Productivity and Wages: Firm Level Evidence’, *Review of Economics and Statistics* 97: 485–497.

Krueger D. and Kumar K. (2004) ‘Skill-Specific Rather Than General Education: A Reason for US-Europe Growth Differences?’, *Journal of Economic Growth* 9: 167–207.

Lemieux T. (2006) ‘*The ‘Mincer Equation’: Thirty Years After Schooling, Experience, and Earnings*’ in Grossbard S. (ed.) *Jacob Mincer: A Pioneer of Modern Labor Economics*. New York: Springer.

Leuven E. and Oosterbeek H. (2011) ‘*Overeducation and Mismatch in the Labor Market*’ in Hanushek E., Machin M. and Woessman L. (eds.) *Handbook of the Economics of Education (Chapter 3)*. Amsterdam: North-Holland: pp. 283–326.

Mahlberg B., Freund I., Cuaresma J. and Prskawets A. (2013) ‘Ageing Productivity and Wages in Austria’, *Labour Economics* 22: 5–15.

Mahy B., Rycx F. and Volral M. (2011) ‘Wage Dispersion and Firm Productivity in Different Working Environments’, *British Journal of Industrial Relations* 49: 460–485.

Mahy B., Rycx F. and Vermeylen G. (2015) ‘Educational Mismatch and Firm Productivity: Do Skills, Technology and Uncertainty Matter?’, *De Economist* 163: 233–262.

Manning A. (2011) *Imperfect Competition in the Labor Market*, in Ashenfelter O. and Card D. (eds.) *Handbook of Labor Economics*, Volume 4, Part B, 973-1041, Elsevier (North Holland).
Mavromaras K. and McGuinness S. (2012) 'Overskilling Dynamics and Education Pathways', *Economics of Education Review* 31: 619–628.
McGuinness S. (2006) 'Overeducation in the Labour Market', *Journal of Economic Surveys* 20: 387–418.
McGuinness S. and Sloane P. (2011) 'Labour Market Mismatch Among UK Graduates: An Analysis Using REFLEX Data', *Economics of Education Review* 30: 130–145.
Meyer M., Milgrom P. and Roberts J. (1992) 'Organizational Prospects, Influence Costs, and Ownership Changes', *Journal of Economics and Management Strategy* 1: 9–35.
Nelson R. and Phelps E. (1966) 'Investment in Humans, Technological Diffusion and Economic Growth', *American Economic Review* 56: 69–75.
Nielen S. and Schiersch A. (2012) 'Productivity in German Manufacturing Firms: Does Fixed-Term Employment Matter?', Schumpeter Discussion Paper, 4.
Nielen S. and Schiersch A. (2014) 'Temporary Agency Work and Firm Competitiveness: Evidence From German Manufacturing Firms', *Industrial Relations* 53: 365–393.
OECD (2013) *Skills Outlook*. Paris: OECD.
van Ours J. C. and Stoeldraijer L. (2010) 'Age, Wage and Productivity', CESifo Working Paper, 2965.
van Ours J. C. and Stoeldraijer L. (2011) 'Age, Wage and Productivity in Dutch Manufacturing', *De Economist* 159: 113–137.
Pecoraro M. (2016) 'The Incidence and Wage Effects of Overeducation Using the Vertical and Horizontal Mismatch in Skills: Evidence From Switzerland', *International Journal of Manpower* 37: 536–555.
Pereira P. T. and Martins P. S. (2004) 'Returns to Education and Wage Equations', *Applied Economics* 36: 525–531.
Prendergast C. (2002) 'The Tenuous Trade-off Between Risk and Incentives', *Journal of Political Economy* 110: 1071–1102.
Quintini G. (2011) 'Over-Qualified or Under-Skilled: A Review of Existing Literature', OECD Social, Employment and Migration Working Papers, 121.
Roodman D. (2009) 'How to do Xtabond2: An Introduction to Difference and System GMM in Stata', *Stata Journal* 9: 86–136.
Rumberger R. (1987) 'The Impact of Surplus Schooling on Productivity and Earnings', *Journal of Human Resources* 22: 24–50.
Sanchez-Sanchez N. and McGuinness S. (2015) 'Decomposing the Impacts of Overeducation and Overskilling on Earnings: An Analysis Using Reflex Data', *Education Economics* 24: 419–432.
Sattinger M. and Hartog J. (2013) 'Nash Bargaining and the Wage Consequences of Educational Mismatches', *Labour Economics* 23: 50–56.
Schaefer S. (1998) 'Influence Costs, Structural Inertia, and Organisational Change', *Journal of Economics and Management Strategy* 7: 237–263.
Sellami S., Verhaest D. and Van Trier W. (2018) "How to Measure Field-of-Study Mismatch?', *A Comparative Analysis of the Different Methods', LABOUR* 32: 141–173.
Sicherman N. (1991) 'Overeducation in the Labor Market', *Journal of Labor Economics* 9: 101–122.
Stinebrickner T., Stinebrickner R. and Sullivan P. (2018) 'Job Tasks and the Gender Wage gap Among College Graduates', NBER Working Paper, 24790.
Syverson C. (2011) 'What Determines Productivity?', *Journal of Economic Literature* 49: 326–365.
Tohmo T. (2015) 'The Creative Class Revisited: Does the Creative Class Affect the Birth Rate of High-Tech Firms in Nordic Countries?', *Journal of Enterprising Culture* 23: 63–89.
Van Beveren I. (2012) 'Total Factor Productivity Estimation: A Practical Review', *Journal of Economic Surveys* 26: 98–128.
Verhaest D. and Omey E. (2009) 'Objective Over-Education and Worker Well-Being: A Shadow Price Approach', *Journal of Economic Psychology* 30: 469–481.

Verhaest D. and Omey E. (2012) ‘Overeducation, Undereducation and Earnings: Further Evidence on the Role of Ability and Measurement Error Bias’, *Journal of Labor Research* 33: 76–90.

Verhaest D. and Van der Velden R. (2013) ‘Cross-Country Differences in Graduate Overeducation’, *European Sociological Review* 29: 642–653.

Wu Y. (2015) ‘Organizational Structure and Product Choice in Knowledge-Intensive Firms’, *Management Science* 61: 1830–1848.