

CEBRIG Working Paper



CEBRIG
CENTRE EMILE BERNHEIM
DE RECHERCHE INTERDISCIPLINAIRE EN GESTION



Solvay Brussels School
Economics & Management

Does Over-education Raise Productivity and Wages Equally ? The Moderating Role of Workers' Origin and Immigrants' Background

Valentine Jacobs, François Rycx, Mélanie Volral

We provide first evidence of the impact of over-education, among natives and immigrants, on firm-level productivity and wages. We use Belgian linked panel data and rely on the methodology from Hellerstein et al. (1999) to estimate ORU (over-, required, and under-education) equations aggregated at the firm level. Our results show that the over-education wage premium is higher for natives than for immigrants. However, since the differential in productivity gains associated with over-education between natives and immigrants outweighs the corresponding wage premium differential, we conclude – based on OLS and dynamic GMM-SYS estimates – that over-educated native workers are in fact underpaid to a greater extent than their over-educated immigrant counterparts. This conclusion is refined by sensitivity analyses, when testing the role of immigrants' background (e.g. region of birth, immigrant generation, age at arrival in the host country, tenure).

Keywords Immigrants, Over-Education, Productivity, Wages, Linked Panel Data, Belgium

JEL Classifications J24, J71

Acknowledgments The authors are most grateful to Statistics Belgium for giving them access to the data. Financial support from the Belgian Federal Science Office (BELSPO, IMMILAB project) is also kindly acknowledged.

CEBRIG Working Paper N°22-003

February 2022

Does Over-education Raise Productivity and Wages Equally ? The Moderating Role of Workers' Origin and Immigrants' Background

Valentine Jacobs

*Université de Mons (Soci&ter) and
Université libre de Bruxelles (CEBRIG and DULBEA)*

François Rycx

Université libre de Bruxelles (CEBRIG and DULBEA), GLO, IRES, IZA, Soci&ter

Mélanie Volral

Université de Mons (Soci&ter) and DULBEA

Abstract

We provide first evidence of the impact of over-education, among natives and immigrants, on firm-level productivity and wages. We use Belgian linked panel data and rely on the methodology from Hellerstein *et al.* (1999) to estimate ORU (over-, required, and under-education) equations aggregated at the firm level. Our results show that the over-education wage premium is higher for natives than for immigrants. However, since the differential in productivity gains associated with over-education between natives and immigrants outweighs the corresponding wage premium differential, we conclude – based on OLS and dynamic GMM-SYS estimates – that over-educated native workers are in fact underpaid to a greater extent than their over-educated immigrant counterparts. This conclusion is refined by sensitivity analyses, when testing the role of immigrants' background (*e.g.* region of birth, immigrant generation, age at arrival in the host country, tenure).

KEYWORDS: Immigrants, over-education, productivity, wages, linked panel data, Belgium
JEL classification: J24, J71

Corresponding author:

François Rycx
Université libre de Bruxelles
50 avenue F. Roosevelt
CP 114/03 B-1050 Brussels
Belgium
E-mail: Francois.Rycx@ulb.be

Acknowledgments: The authors are most grateful to Statistics Belgium for giving them access to the data. Financial support from the Belgian Federal Science Office (BELSPO, IMMILAB project) is also kindly acknowledged.

1 Introduction

Over-education has become a major challenge for societies over time. A worker is considered as over-educated if her/his attained level of education is higher than the level of education required for her/his job. According to Eurostat, in 2020, 21.5% of workers were affected by this phenomenon in the EU27 countries (Eurostat, 2021). Given the magnitude of this figure, “many observers point to: i) the failure of the education system in providing youth with the skills required at work, and ii) the inability of labour markets to sort many workers to suitable jobs” (OECD, 2011: 193). Moreover, McGuinness (2006) emphasizes that educational mismatch may be costly for the economy as a whole (*e.g.* as a result of wasted tax revenues due to the financing of excessive levels of education), for firms (*e.g.* due to a loss in efficiency if over-educated workers are less productive than their adequately educated colleagues) but also for individuals (*e.g.* if over-educated workers earn a lower wage and/or have a lower level of job satisfaction than their former classmates employed in jobs matching their level of education).

Focusing on the wage consequences of over-education, two main findings are obtained in the literature. On the one hand, estimates based on the ORU (over-, required, and under-education) approach (Duncan and Hoffman, 1981) show that over-educated workers receive a wage premium in comparison with their adequately educated colleagues in similar jobs. This result suggests that the additional human capital of the over-educated is not totally unproductive and is therefore partly reflected in their earnings. On the other hand, estimates based on the dummy specification (Verdugo and Verdugo, 1989) indicate that over-educated workers face a wage penalty compared to their former classmates employed in jobs matching their education level. This result suggests that over-educated workers underuse their skills and therefore suffer from a wage penalty compared to what they could have earned in a job matching their education level. A complementary explanation is that over-educated workers may be less satisfied with their jobs, which could have a negative impact on their productivity and earnings. To sum up, empirical evidence emphasizes that over-educated workers generally earn a wage premium in comparison with their adequately educated colleagues doing the same job, but a wage penalty compared to their former classmates employed in jobs matching their level of education.

In recent years, the phenomenon of over-education has also been studied through the prism of workers' origin. Indeed, the situation of immigrant workers on the labour market is a major concern in many countries, especially when these workers originate from developing countries. This is notably illustrated by the fact that first-generation immigrants (*i.e.* foreign-born people) face a significant employment penalty (*e.g.* Brinbaum, 2018; Piton and Rycx,

2021) and are found to be paid less than their native counterparts, even when differences in productive characteristics are taken into account (*e.g.* Chiswick, 1978; Borjas, 1985; Sanromá *et al.*, 2015a; Kampelmann and Rycx, 2016; Basilio *et al.*, 2017; Fays *et al.*, 2021). Numerous studies have also focused on the integration of these immigrants' children (*i.e.* second-generation immigrants) into the labour market. As Card (2005) points out, since second-generation immigrants are born, educated, and socialized in the country of residence, their relative success or failure is often considered as the ultimate benchmark for integration. However, most of the empirical literature supports the segmented assimilation theory, stating that second-generation immigrants might still experience high levels of discrimination and downward assimilation (*e.g.* Portes and Rumbaut, 2001).¹ More precisely, estimates generally show divergent intergenerational mobility patterns between different ethnic groups, with a lower likelihood for children of immigrants from poorer countries to outperform their parents (*e.g.* Meurs *et al.*, 2006; Zorlu and Hartog, 2012; Brinbaum, 2018; Piton and Rycx, 2021).

With regard to educational mismatch, studies first show that first-generation immigrants are more likely to be over-educated than the native-born (*e.g.* Wald and Fang, 2008; Nielsen, 2011; Joonas *et al.*, 2014; Jacobs *et al.*, 2021a). As far as we know, only Falcke *et al.* (2020) have been able to compare the likelihood of over-education between natives and second-generation immigrants. The authors show that the greater likelihood of being over-educated persists among the second-generation, compared to natives. Several studies have also examined the impact of over-education on wages according to workers' origin (*e.g.* Battu and Sloane, 2004; Chiswick and Miller, 2008, 2010; Wald and Fang, 2008; Nielsen, 2011; Joonas *et al.*, 2014; Schwientek, 2016; Jacobs *et al.*, 2021b; Maani and Wen, 2021). Their results show that: i) the wage premium received by over-educated workers in comparison with their adequately educated colleagues tends to be smaller for first-generation immigrants than for the native-born, and ii) the wage penalty suffered by over-educated workers in comparison with their former classmates employed in jobs matching their education level appears to be greater for first-generation immigrants than for native-born workers. To our knowledge, no study has yet investigated whether the impact of over-education on wages is different for first- and second-generation immigrants and, thereby, whether the integration of children of immigrants appears to be more successful than that of their parents in this respect.

To assess the impact of over-education on wages, empirical studies rely on either a Mincer-type equation (Mincer, 1974), an ORU specification (Duncan and Hoffman, 1981), a

¹ This theory is generally opposed to the classical assimilation perspective, which suggests that ethnic differences should be reduced over time and across generations for all origin groups (*e.g.* Alba and Nee, 2003).

dummy approach (Verdugo and Verdugo, 1989), or an Oaxaca-Blinder (1973) decomposition using detailed information at the worker level. These approaches, especially the last one, known as residual methods, decompose the difference in mean wages between over- and adequately educated workers into explained and unexplained components. The share of the wage gap that cannot be explained by the differences in observable characteristics is then interpreted as the wage penalty (wage premium) incurred by over-educated workers compared to their well-matched former classmates (well-matched colleagues). As highlighted by Bartolucci (2014: 1167), “the residual approach(es) can be understood as a comparison of wages and productivity where the latter is approximated using a function of observable characteristics”. However, if *unobservable* characteristics vary between over- and adequately educated workers and are correlated with productivity, then it is likely that the wage penalty (wage premium) associated with over-education will be biased.

The availability of linked employer-employee data offers a response to this potential weakness of the residual approaches (Bartolucci, 2014). If worker-level data are not rich enough to account for differences in productivity between over- and adequately educated workers, one possibility is to directly estimate a wage equation and a production function at the firm level. The wage equation estimates the impact of over-education (*i.e.* the incidence of over-educated workers within firms) on the average wage, whereas the production function yields this impact on the average marginal product. Estimating both equations using the same samples and identical control variables allows to compare the parameters for marginal wage and products associated with over-education, and to draw conclusions on how over-education affects productivity, wages, and productivity-wage gaps. This method has been developed by Hellerstein *et al.* (1999) and is now standard in the literature on the effects of firm heterogeneity² on wages and productivity (*e.g.* Hellerstein and Neumark, 2004; Cataldi *et al.*, 2011; van Ours and Stoeldraijer, 2011; Vandenberghe, 2013; Garnero *et al.*, 2014; Konings and Vanormelingen, 2015; Giuliano *et al.*, 2017; Garnero *et al.*, 2020). To our knowledge, the study by Kampelmann *et al.* (2020) is the only one so far to use this method in order to provide evidence of how over-education (*i.e.* the incidence of over-educated workers within firms) affects both wages³ and productivity⁴. On the basis of detailed Belgian linked panel data and an

² Firm heterogeneity has notably been assessed by the diversity of the workforce within firms in terms of age, gender, occupations, or employment contracts.

³ More precisely, the authors focus on wage costs.

⁴ A few other papers have relied on firm-level data to investigate whether over-education is beneficial or harmful to productivity (*e.g.* Quintini, 2011; McGowan and Andrews, 2015; Grunau, 2016). For the Belgian private sector, the results of Kampelmann and Rycx (2012a) and Mahy *et al.* (2015) suggest that the effect is significant and positive.

ORU specification aggregated at the firm level, the authors find that over-education has a positive effect on productivity and that this effect is only partly offset by higher wage costs. In other words, the authors find that, from firms' perspective, over-education is associated with positive economic rents (*i.e.* productivity-wage gaps). From workers' perspective, the authors' results suggest that over-educated workers are more productive than their well-matched colleagues and that this productivity differential is only partially reflected in the wage premium earned by over-educated workers. Accordingly, the authors conclude that, in view of their productivity, over-educated workers are actually underpaid compared to their well-matched colleagues. This result is important because it changes the interpretation of the over-education wage premium, as the latter only partially reflects the productivity increase generated by over-educated workers. Moreover, it questions the literature on the wage effects of over-education by origin. Current evidence shows that the over-education wage premium is higher for natives than for immigrants, suggesting that immigrants are disadvantaged and potentially discriminated against. However, it is not clear whether this conclusion would still hold if the productivity differences between over-educated natives and immigrants were properly taken into account. In this respect, starting from the study by Kampelmann *et al.* (2020), a key question to investigate is whether over-educated natives and immigrants, considering their productivity, are both underpaid compared to their well-matched colleagues and, if so, whether this underpayment is more pronounced among over-educated immigrants than among over-educated natives. To the best of our knowledge, this issue has never been studied before.

The first objective of the present paper is therefore to analyse how the incidence of over-education, among natives and immigrants, respectively, affects productivity, wages, and productivity-wages gap at the firm level.⁵ Put differently, using an ORU specification aggregated at the firm level in combination with the method developed by Hellerstein *et al.* (1999), we aim to test whether the wage premia obtained by over-educated native and immigrant workers are aligned with their productivity differentials, so as to draw conclusions on the potential underpayment of over-educated workers compared to their well-matched

⁵ As this paper focuses primarily on the empirical estimation of productivity-wage gaps related to over-education according to workers' origin, we choose not to reiterate the numerous mechanisms put forward in the large theoretical literature that do not take workers' origin into account. Following the literature overview in Kampelmann and Rycx (2012b), these mechanisms can be divided into: a) theories based on efficiency and individual rationality (*e.g.* when over-educated 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 their wage structure so as to avoid shirking or demotivation (see *e.g.* McGuinness, 2006; Cardoso, 2010); and b) institutionalist theories (*e.g.* when monopsony power, market regulations, wage norms, or collective bargaining are associated with positive or negative rents that differ between over-educated and well-matched workers (see *e.g.* Quintini, 2011; Konings and Vanormelingen, 2015).

colleagues and on the magnitude of this underpayment according to workers' origin. To this end, we rely on detailed Belgian matched employer-employee panel data for the years 1999 to 2016. Our dataset includes information on the characteristics of workers (*e.g.* gross hourly wage, age, gender, education, tenure, working time, occupation, country of birth of the workers and their parents, and age at arrival in the host country for first-generation immigrants) and firms (*e.g.* value added, sector of activity, number of employees, and level of collective agreement). As it contains detailed information on the country of birth of both workers and their parents, it is particularly well suited to identify natives (*i.e.* individuals born in Belgium from parents both born in Belgium) and immigrants (*i.e.* individuals either born abroad, referred to as "first-generation immigrants", or born in Belgium with at least one parent born abroad, referred to as "second-generation immigrants"). Our final sample, covering 21,086 firm-year-observations from 5,704 firms, enables us to estimate productivity and wage equations at the firm level while controlling for a large range of covariates and, in a robustness test section, also for firm-level fixed effects and potential endogeneity issues.

We further aim to contribute to the existing literature by examining the role of different moderating variables, reflecting key characteristics of immigrants, in the relationship between over-education, productivity, and wages. For this purpose, we first examine the role of immigrants' region of origin. Whereas the region of origin for first-generation immigrant workers corresponds to their region of birth, we need to take into account the region of birth of the workers' parents when dealing with second-generation immigrants. Considering immigrants as a homogenous group hides significant disparities in labour market performance. Across Europe, we can indeed distinguish two groups of immigrants: on the one hand, people originating from developed countries, whose employment rate and earnings are very close to those of natives in all countries; and, on the other hand, immigrants originating from developing countries, whose access to employment is much more problematic and whose earnings are significantly lower on average. Belgium is no exception in this respect (FPS Employment and Unia, 2017; Fays *et al.*, 2021; Jacobs *et al.*, 2021). Accordingly, immigrants are classified according to whether they originate from a developed or a developing country, based on the their region of birth and that of their parents. Our first sensitivity test aims to determine whether the impact of over-education on wages and productivity varies when focusing on immigrants originating from developed and developing countries, respectively. Factors such as the imperfect transferability of human capital and discrimination have been identified as causing a higher incidence of over-education as well as a lower over-education wage premium among immigrants (*e.g.* Chiswick and Miller, 2009; Aleksynska and Tritah, 2013; Matano *et al.*, 2015).

As these factors are likely to be less detrimental for immigrants from developed countries, due to their greater proximity with natives in terms of culture, education systems, languages, and/or economic development (Sanromá *et al.*, 2015b), the returns to over-education are expected to be greater in terms of wages and productivity for them than for immigrants from developing countries. The effect on productivity-wage gaps, however, is less clear a priori.

Our second sensitivity test aims to examine the moderating role of immigrants' generation. As mentioned earlier, considering that second-generation immigrants are born, educated, and socialized in the same country as natives, they should be, on average, better integrated than foreign-born workers. Therefore, from a theoretical perspective, we expect the effects of over-education on wages and productivity to be greater for second-generation immigrants than for first-generation immigrants and to potentially converge with those estimated for natives. This assumption remains yet to be tested empirically. Another open question that we investigate is whether the misalignment between the productivity and wage effects of over-education (*i.e.* the potential underpayment of over-educated workers) is more pronounced for first-generation than for second-generation immigrants.

For first-generation immigrants, we further test whether the results differ between those who arrived in Belgium before (or at) the age of 18 and those who arrived at a later age. Since foreign-born immigrants who arrived in Belgium before (or at) the age of 18 have spent at least the last years of their schooling (and sometimes all of their schooling) in the host country, they are likely to have accumulated more country-specific human capital. They should, on average, have fewer problems with the recognition of their qualifications, have a better command of the host country's languages, be better informed about the functioning of the country and its institutions, and have a wider and more effective network, all of which are likely to favor their integration into the labour market (Mattoo *et al.*, 2008; Nielsen, 2011). Therefore, we expect the effects of over-education on wages and productivity to be greater among first-generation immigrants who arrived in Belgium before (or at) the age of 18 than among those who arrived at a later age. In addition to testing this hypothesis, we also investigate whether this 18-year threshold affects the productivity-wage gaps associated with over-education, in order to examine whether the potential underpayment of over-educated foreign-born workers varies according to whether they arrived in Belgium before or after their majority.

Finally, we investigate the moderating role of tenure (*i.e.* the number of years an employee has been working for her/his current employer). Tenure has been shown to foster workers' wages and productivity (Flabbi and Ichino, 2001; Gagliardi *et al.*, 2021), and some studies suggest that it also affects workers' likelihood of being over-educated, for all workers

as well as according to the worker's origin (Jacobs *et al.*, 2021a). Hence, this variable is also likely to influence the over-education wage premium (and productivity differential) for natives and immigrants. Concerning the role of tenure in the relationship between over-education and productivity, the intuition is as follows. According to human capital theory (Becker, 1964), we expect the productivity of over-educated workers to increase with tenure, possibly in a curvilinear way (Gagliardi *et al.*, 2021). This assumption is supported by the notions of firm-specific human capital (Becker, 1964) and tacit knowledge (Polanyi, 1966). The question of whether the productivity gains associated with tenure among over-educated workers are greater for natives or for immigrants remains open and particularly depends on the steepness of the workers' relative learning curves. If over-educated immigrants have a lower quality of human capital (at a given level of education) than their native counterparts at the time of hiring (for example, because they have less command of the languages used at work and/or less experience of doing business in the host country) but are able to catch up quickly, then their productivity should increase faster than that of their native counterparts as years of service increase. However, the outcome could also be the opposite if over-educated immigrants fail to make up for their initial knowledge deficit with increased work experience. As for the potential influence of tenure in the nexus between over-education and wages according to workers' origin, if the smaller wage premium earned by immigrants is driven by statistical discrimination (Phelps, 1972), we might expect the differences in premia according to workers' origin to decrease as years of tenure increase.⁶ However, if the differences in wage premia are due to other phenomena (such as taste-based discrimination or unobserved differences in actual skills), then the effect of tenure is more likely to be non-significant or at least quite limited. Finally, as regards the impact of tenure on the magnitude of the productivity-wage gaps associated with over-education by workers' origin (*i.e.* the extent of the potential underpayment of over-educated native and immigrant workers), given the ambivalence of the arguments described above, theoretical predictions remain uncertain, and this point thus also requires an empirical analysis.

The remainder of this paper is organised as follows. Section 2 describes our methodology, and Section 3 presents our dataset. Our econometric results are presented in Section 4, and Section 5 finally concludes.

⁶ Statistical discrimination theory indeed postulates that employers' asymmetric information about the productivity of immigrants should decrease as the duration of the employment within a firm increases and that, consequently, immigrants' wages should better reflect their true productivity as the length of their service increases.

2 Methodology

2.1 Benchmark specification

The literature proposes three main approaches to measure educational mismatch: the job analysis (JA), realized matches (RM), and worker self-assessment (WA) methods. The JA approach is an objective measurement based on analysts' criteria to determine the education requirement for a job to be compared with workers' educational attainments; the RM is a statistical approach that compares workers' educational attainments with those of workers in the same occupation using the mean (Verdugo and Verdugo, 1989) or modal value (Kiker *et al.*, 1997) as reference; and the WA is a subjective measurement based on surveys in which workers are asked to evaluate the level of education required to do their jobs.

These measurements all have advantages and shortcomings (for a discussion, see *e.g.* Hartog, 2000), and the approach chosen in practice is therefore often driven by data availability. Given the features of our dataset, we rely on the RM approach. We thus define the required level of education for a given job by taking the mode of workers' years of attained education (ISCED: 7 categories)⁷ within each occupation (ISCO 3-digit: 150 categories). As a result, a worker is considered as over- (under-) educated if her/his educational attainment is higher (lower) than that required (*i.e.* observed) for her/his occupation.

To examine the influence of over-education on firm-level productivity and wages, we rely on an ORU (over-, required, under-education) specification that has been aggregated at the firm level. Following common practice, we first decompose the number of years of attained education of each worker i employed in firm j at year t ($E_{i,j,t}^A$) into three parts:

- i) years of required education ($E_{i,j,t}^R$) for her/his job,
- ii) years of attained education exceeding those required for her/his job, *i.e.* years of over-education ($E_{i,j,t}^O = \begin{cases} E_{i,j,t}^A - E_{i,j,t}^R, & \text{if } E_{i,j,t}^A > E_{i,j,t}^R \\ 0, & \text{otherwise} \end{cases}$), and
- iii) years of attained education below those required for her/his job, *i.e.* years of under-education ($E_{i,j,t}^U = \begin{cases} E_{i,j,t}^R - E_{i,j,t}^A, & \text{if } E_{i,j,t}^A < E_{i,j,t}^R \\ 0, & \text{otherwise} \end{cases}$).

⁷ Information on workers' educational attainments, available in 7 categories in our dataset, were reported by firms' HR departments (based on their registers). We converted that information into years of education, applying the following rule: (i) primary education: 6 years of education; (ii) lower secondary education: 9 years; (iii-iv) general, technical and artistic upper secondary education: 12 years; (v) higher university and non-university education, short type: 15 years; (vi) university and non-university education, long type: 17 years; (vii) postgraduate education: 18 years. Given that information on workers' levels of education were provided by firms' HR departments, those levels might be somewhat under-estimated for immigrants. The findings reported in this paper should therefore be considered as a lower bound.

Accordingly, in the case of a worker with a university degree of long type (*i.e.* 17 years of education) employed in a job requiring a university degree of short type (*i.e.* 15 years of education), for example, the years of attained education would be decomposed into 15 years of required education ($E_{i,j,t}^R$), 2 years of over-education ($E_{i,j,t}^O$), and 0 years of under-education ($E_{i,j,t}^U$).⁸

Next, we estimate the following firm-level wage and productivity equations:

$$\begin{aligned} \log(W_{j,t}) = & \beta_0 + \beta_1 \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,N}}{m_{j,t}} \right) + \beta_2 \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,IM}}{m_{j,t}} \right) + \beta_3 \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^R}{m_{j,t}} \right) + \\ & \beta_4 \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,N}}{m_{j,t}} \right) + \beta_5 \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,IM}}{m_{j,t}} \right) + \beta_6 X_{j,t} + \beta_7 Z_{j,t} + \gamma_t + u_{j,t} \end{aligned} \quad (1)$$

$$\begin{aligned} \log(VA_{j,t}) = & \beta_0^* + \beta_1^* \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,N}}{m_{j,t}} \right) + \beta_2^* \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,IM}}{m_{j,t}} \right) + \beta_3^* \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^R}{m_{j,t}} \right) + \\ & \beta_4^* \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,N}}{m_{j,t}} \right) + \beta_5^* \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,IM}}{m_{j,t}} \right) + \beta_6^* X_{j,t} + \beta_7^* Z_{j,t} + \gamma_t^* + u_{j,t}^* \end{aligned} \quad (2)$$

In these equations, $W_{j,t}$ is the average gross hourly wage in firm j at year t , measured by the ratio of the firm's total wage bill (*i.e.* the sum of gross hourly wages, including premia for overtime, weekend or night work, bonuses, and other premia, paid by firm j at year t) over the total number of effective paid hours; $VA_{j,t}$ is the productivity of firm j at year t , measured by the average hourly value added, which corresponds to the total value added computed at factor costs divided by the total number of effective paid hours; $m_{j,t}$ is the number of workers employed in firm j at year t ; the superscripts N and IM refer to natives (*i.e.* individuals born in Belgium with both parents born in Belgium) and immigrants (*i.e.* individuals born abroad, referred to as “first-generation immigrants”, and those born in Belgium with at least one parent born abroad, referred to as “second-generation immigrants”), respectively; $X_{j,t}$ is a vector including aggregated characteristics of workers in firm j at year t : the share of the workforce having at least 10 years of tenure, the percentages of workers younger than 25 and older than 49 years, respectively, and the shares of women, immigrants, part-time, temporary (*i.e.* fixed-term, trainee, and agency), and blue-collar⁹ workers. $Z_{j,t}$ is a vector containing characteristics

⁸ Note that, by definition: $E_{i,j,t}^A = E_{i,j,t}^R + E_{i,j,t}^O - E_{i,j,t}^U$, *i.e.* the sum of a worker's years of required and over-education minus her/his years of under-education is equal to her/his years of attained education.

⁹ The distinction between blue-collar and white-collar workers is based on the International Standard Classification of Occupations (ISCO-08). Workers belonging to groups 1-5 (*i.e.* managers, professionals, technicians and associate professionals, clerical support workers, and services and sales workers) are classified as white-collars,

of firm j at year t : its sectorial affiliation (13 categories)¹⁰, its size (*i.e.* the log of the full-time equivalent number of workers), a dummy taking the value 1 if the firm is covered by a firm-level collective agreement and 0 otherwise (*i.e.* if it is covered only by a national and sectoral collective agreement), its age¹¹, the region where it is located (*i.e.* Brussels and Wallonia, respectively; Flanders being the reference category), and the logarithm of its capital stock per worker¹²; γ_t is a set of 17 year dummies; and $u_{j,t}$ is the error term.

The first main explanatory variable in equations (1) and (2) is the extent of required education in firm j at the year t $\left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^R}{m_{j,t}}\right)$. This variable is obtained by summing up the years of required education associated with the different jobs i in firm j at year t ($E_{i,j,t}^R$) and dividing this sum by the total number of workers employed in firm j at year t ($m_{j,t}$). 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, then the mean level of required education in that firm will correspond to the average value of the required years of education associated to with these five different jobs. The required years of education for a cashier (*i.e.* the mode of the years of education in the ISCO category 523 across the entire private sector) may be, for instance, equal to 9 (which corresponds to a lower secondary education degree), whereas those for an administration professional (ISCO category 242) may be equal to 16 (which corresponds to a master's degree). In our example, the replacement of a cashier by an administrative professional will lead to an increase in the mean value of required education in that firm.

The second main explanatory variable is the extent of over-education in firm j at year t , which is computed separately for natives $\left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,N}}{m_{j,t}}\right)$ and for immigrants $\left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,IM}}{m_{j,t}}\right)$. It is calculated by summing up the years of over-education, for natives and immigrants, respectively, and dividing these sums by the total number of workers employed in firm j at year t ($m_{j,t}$). For example, a native (an immigrant) worker with 10 years of attained education working in a job

whereas those in groups 7-9 (*i.e.* craft and related trades workers, plant and machine operators and assemblers, and elementary occupations) are classified as blue-collars.

¹⁰ It covers the following sectors: i) mining and quarrying, ii) manufacturing, iii) electricity, gas, steam and air conditioning supply, iv) water supply, sewerage, waste management and remediation activities, v) construction, vi) wholesale and retail trade, repair of motor vehicles and motorcycles, vii) transportation and storage, viii) accommodation and food service activities, ix) information and communication, x) financial and insurance activities, xi) real estate activities, xii) professional, scientific and technical activities, xiii) administrative and support service activities.

¹¹ Our dataset does not provide direct information of the age of firms, so we have proxied this variable with the length of employment of each firm's most senior employee.

¹² This is estimated through the 'perpetual inventory method' (or PIM, see *e.g.* OECD (2009) for more details). The PIM incorporates the idea that capital stock results from investment flows and corrects for capital depreciation and efficiency losses. Following standard practice, we assume a 5 percent annual rate of depreciation.

requiring 8 years of education will add 2 years of over-education to the group of native (immigrant) workers. Firms who place workers in occupations whose educational requirements they exceed will therefore increase the extent of over-education in their workforce.

The third variable related to education is the extent of under-education in firm j at year t , which is computed separately for natives $\left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,N}}{m_{j,t}}\right)$ and for immigrants $\left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,IM}}{m_{j,t}}\right)$.¹³ It is calculated by summing up the years of under-education, for natives and immigrants respectively, and dividing these sums by the total number of workers employed in firm j at year t ($m_{j,t}$). For example, a native (an immigrant) worker with 10 years of attained education working in a job requiring 12 years of education adds 2 years of under-education to the group of native (immigrant) workers. The more a firm employs workers of this type, the higher the extent of under-education in its workforce.¹⁴

Equations (1) and (2) enable us to investigate how the extent of over- (and under-) education among natives and immigrants, respectively, affect productivity and wages within firms, conditional on mean years of required education, year dummies, and a wide range of worker and firm characteristics. Moreover, given that equations (1) and (2) are estimated on the same samples with identical control variables, we can compare the parameters for wages and productivity and draw conclusions on how over-, required, and under-education affect firms' productivity-wage gaps. This method has been developed by Hellerstein *et al.* (1999) and is now standard in the literature on the effects of firm heterogeneity on wage and productivity (*e.g.* Hellerstein and Neumark, 2004; Cataldi *et al.*, 2011; van Ours and Stoeldraijer, 2011; Vandenberghe, 2013; Garnero *et al.*, 2014; Konings and Vanormelingen, 2015; Giuliano *et al.*, 2017; Garnero *et al.*, 2020; Kampelmann *et al.*, 2020). To our knowledge, it has never been used to examine the nexus between over-education, wages, and productivity according to workers' origin.

The parameters in equations (1) and (2) associated with mean years of over-education can be interpreted as follows. Positive estimates for β_1 (β_2) and β_1^* (β_2^*) imply that increasing the extent of over-education among natives (immigrants) is associated with an improvement in firms' average wages/productivity. In contrast, negative estimates for β_1 (β_2) and β_1^* (β_2^*), mean

¹³ The focus of our paper is on the effects of over-education. Therefore, the results regarding under-education will be neither reported nor discussed in the empirical section of this manuscript. Our estimates for under-education are available on request.

¹⁴ Note that: $\frac{1}{m_{j,t}} \left(\sum_{i=1}^{m_{j,t}} E_{i,j,t}^{O,N} + \sum_{i=1}^{m_{j,t}} E_{i,j,t}^{O,IM} + \sum_{i=1}^{m_{j,t}} E_{i,j,t}^R - \sum_{i=1}^{m_{j,t}} E_{i,j,t}^{U,N} - \sum_{i=1}^{m_{j,t}} E_{i,j,t}^{U,IM} \right) = \frac{1}{m_{j,t}} \sum_{i=1}^{m_{j,t}} E_{i,j,t}^A$, *i.e.* the sum of the average years of over-education (among natives and immigrants, respectively), years of required education, and the average years of under-education (among natives and immigrants, respectively) in firm j at year t is equal to the average years of education attained by the workers employed in firm j at year t .

that wages and firm productivity would be negatively affected. Since the Hellerstein *et al.* (1999) approach ensures the comparability between the parameters of the two equations, it further allows us to assess the effect of over-education on productivity-wage gaps. If the estimated value of β_1^* (β_2^*) is higher than that of β_1 (β_2), then the effect of over-education among natives (immigrants) on productivity is larger than that on wages, and over-education is therefore associated with a positive gap between productivity and wages. This would imply that over-educated native (immigrant) workers, given their productivity differential, are underpaid compared to their well-matched colleagues. Conversely, we would conclude that over-educated native (immigrant) workers are overpaid if β_1^* (β_2^*) is smaller than β_1 (β_2).

2.2 Moderating factors

Our paper further aims to investigate the role of various moderating variables, reflective of immigrants' background, in the relationship between over-education, productivity, and wages. More precisely, we respectively examine the role of immigrants': a) region of origin; b) generation, c) age at arrival in the host country, and d) years of tenure.

In order to determine immigrants' region of origin, we followed common practice: for first-generation immigrants, we took their region of birth into account; for second-generation immigrants, we considered their parents' region of birth.¹⁵ Using information on the country of birth of both the workers and their parents, as well as the United Nations' classification of countries according to their economic development (UNCTAD, 2020), we split workers into three groups: i) natives (*i.e.* workers born in Belgium with both parents born in Belgium); ii) immigrants from developed countries (*i.e.* workers either born in a developed country (excluding Belgium) or having at least one parent born in a developed country (excluding Belgium)); and iii) immigrants from developing countries¹⁶ (*i.e.* workers born in a developing country or having at least one parent born in a developing country). Next, we computed the variables 'extent of over-education' and 'extent of under-education', which appear in equations (1) and (2), separately for each of these three groups. Put differently, we calculated mean years

¹⁵ More precisely, to determine the region of origin of second-generation immigrants, the order of priority is based on the father's region of birth (see *e.g.* Corluy *et al.*, 2015; FPS Employment and Unia, 2017; Piton and Rycx, 2021). Put differently, for second-generation immigrants, the father's region of birth is used, unless the father was born in Belgium and the mother born abroad, in which case the mother's region of birth is retained. This choice stems from the fact that: i) children born in Belgium before June 1st, 2014, were named after their father (the legislation has become more flexible since then), and ii) correspondence studies have shown that call-back rates depend upon the origin of job seekers' names (see *e.g.* Baert, 2017).

¹⁶ By 'developing' countries, we refer to both transition and developing countries, as listed in the UNCTAD (2020) classification.

of over- and under-education within firms respectively for: i) natives $\left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,N}}{m_{j,t}}; \sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,N}}{m_{j,t}}\right)$, ii) immigrants from developed countries $\left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,IMd}}{m_{j,t}}; \sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,IMd}}{m_{j,t}}\right)$, and iii) immigrants from developing countries $\left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,IMdev}}{m_{j,t}}; \sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,IMdev}}{m_{j,t}}\right)$. Equations (1) and (2) can thus be reformulated as follows:

$$\ln(W_{j,t}) = \delta_0 + \delta_1 \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,N}}{m_{j,t}}\right) + \delta_2 \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,IMd}}{m_{j,t}}\right) + \delta_3 \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,IMdev}}{m_{j,t}}\right) + \delta_4 \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^R}{m_{j,t}}\right) + \delta_5 \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,N}}{m_{j,t}}\right) + \delta_6 \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,IMd}}{m_{j,t}}\right) + \delta_7 \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,IMdev}}{m_{j,t}}\right) + \delta_8 X_{j,t} + \delta_9 Z_{j,t} + \gamma_t + u_{j,t} \quad (1')$$

$$\ln(VA_{j,t}) = \delta_0^* + \delta_1^* \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,N}}{m_{j,t}}\right) + \delta_2^* \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,IMd}}{m_{j,t}}\right) + \delta_3^* \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{O,IMdev}}{m_{j,t}}\right) + \delta_4^* \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^R}{m_{j,t}}\right) + \delta_5^* \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,N}}{m_{j,t}}\right) + \delta_6^* \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,IMd}}{m_{j,t}}\right) + \delta_7^* \left(\sum_{i=1}^{m_{j,t}} \frac{E_{i,j,t}^{U,IMdev}}{m_{j,t}}\right) + \delta_8^* X_{j,t} + \delta_9^* Z_{j,t} + \gamma_t^* + u_{j,t}^* \quad (2')$$

Then, by estimating equations (1') and (2'), we are able to assess whether the impact of over- (and under-) education on wages and productivity is different when focusing on immigrants originating from developed and developing countries, respectively.

Our second sensitivity test analyses the role of immigrants' generation. More precisely, we aim to investigate whether results vary for first- and second-generation immigrants. For this purpose, we computed mean years of over- (and under-) education separately for first- and second-generation immigrants, that is, we re-estimated equations (1) and (2) by explicitly taking into account the extent of over- (and under-) education among first- and second-generation immigrants, respectively.

Next, to assess the role of immigrants' age on arrival in the host country, we further split first-generation immigrants into two subgroups according to whether they arrived in Belgium before or after their 18th birthday. We re-estimated equations (1) and (2) including information on the extent of over- (and under-) education respectively among: i) natives; ii) second-generation immigrants; iii) first-generation immigrants who arrived in Belgium before (or at) the age of 18; and iv) first-generation immigrants who arrived in Belgium after the age of 18.

Finally, to examine the role of tenure (*i.e.* the number of years an employee has been working for her/his current employer), we calculated the extent of over- (and under-) education among natives and immigrants with less than or at least 10 years of tenure, respectively. More precisely, we re-estimated equations (1) and (2) by including specific explanatory variables for

the average years of over- (and under-) education among natives and immigrants, respectively, depending on whether or not they were below the 10-year tenure threshold.

3 Data and descriptive statistics

Our empirical analysis is based on a combination of three large datasets covering the period 1999-2016. The first is the “Structure of Earnings Survey” (SES), carried out by Statistics Belgium. It covers all firms that are operating in Belgium, that employ at least 10 workers, and whose economic activities fall within sections B to N of the NACE Rev. 2 nomenclature.¹⁷ The SES contains a wealth of information on both the characteristics of firms (*e.g.* sector of activity, number of employees, region, level of collective agreement) and the individuals working in these firms (*e.g.* gross hourly wage, education, gender, age, occupation, tenure, working time), as provided by the firms’ HR departments. The SES has 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.* firm-level hourly value added and gross operating surplus).¹⁸ The third dataset is the Belgian Population Register, which contains, among other variables, information on the country of birth of workers and their parents and on workers’ date of registration in the national register. The three datasets have been merged by Statistics Belgium using workers’ national register and firms’ social security numbers.

To ensure that the required level of education was calculated on the basis of a sufficient volume of data, we first checked that the number of observations in each occupation (at the

¹⁷ The SES is a stratified sample. The stratification criteria refer to the region (NUTS-groups), the main economic activity (NACE-groups), and the size of the firm. The sample size in each stratum depends on the size of the firm. The sampling percentages of firms are equal to 10, 50, and 100% when the number of workers is between 10 and 50, between 50 and 99, and above 100, respectively. Within a firm, the sampling percentages of employees also depend on the size of the firm. The sampling percentages of employees are equal to 100, 50, 25, 14.3, and 10% 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, respectively. Firms employing at least 300 workers are required to report information for an absolute number of employees, which ranges from 30 (for firms with 300 to 349 workers) to 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 gives 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 the workers’ names in their list). If they reach the letter Z and still have to provide information on some of their employees, then 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 list in alphabetical order for each of these categories and to report information on a number of workers in these categories 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, we should point out that there is no threshold at the upper limit for wages in the SES. In other words, wages are not censored. For an extended discussion, see Demunter (2000).

¹⁸ Note that the coverage of the SBS is not the same as that of the SES, as the former does not cover the entire financial sector.

ISCO 3-digit level) was equal to or greater than 10.¹⁹ Then, we dropped a small number of workers and/or firms for which data were missing or inaccurate.²⁰ Finally, our use of the ‘perpetual inventory method’ (OECD, 2009), in order to estimate capital stock at the firm level, requires the observation of firms (and particularly of their investments) over a minimum of two consecutive periods. Given that the sampling percentages of firms in our dataset increase with the size of the firms (see footnote 17), medium-sized and large firms are over-represented in our econometric investigation. Our final sample, covering the period 1999-2016, consists of an unbalanced panel of 21,086 firm-year-observations from 5,704 firms. It is representative of medium-sized and large firms in the Belgian private sector, with the exception of large parts of the financial sector (NACE K) and the electricity, gas, and water supply industry (NACE D & E).

[Insert Table 1 about here]

Table 1 presents descriptive statistics of selected variables at the firm (see column (1)) and worker (see columns (2) to (4)) levels. At the firm level, we observe that the mean hourly value added, which is our direct measurement of productivity, stands at 67 EUR, whereas the average hourly wage reaches 18.7 EUR.²¹ The mean number of required years of education at the firm level equals 12.5, and the proportions of over- and under-educated workers are around 20 and 27%, respectively. The average years of over- and under-education within firms are equal to 0.58 and 0.89, respectively. Moreover, we find that around 26% of workers within firms are (first- and second-generation) immigrants, 75% of workers are aged between 25 and 49, 37% have at least 10 years of tenure, 29% are women, 45% are working part-time, 95% have an open-ended contract, and 52% are blue-collarers. The firms in our sample have, on average, 332 employees and a capital stock of 220.4 thousands EUR, and they are concentrated essentially in the following sectors: manufacturing (48%); wholesale and retail trade, repair of motor vehicles and motorcycles (15%); construction (11%); administrative and support service activities (7%); transport and storage (5%); and information and communication (4%).

To examine potential differences between natives and immigrants, we present their respective means, measured at the worker level, in columns (2) and (3) of Table 1. Among the

¹⁹ This threshold of 10 observations was almost always met, so this constraint left our sample virtually unaffected.

²⁰ For instance, we dropped some workers with no information on their level of education and a very small number of firms for which the recorded value added was negative.

²¹ All variables measured in monetary terms have been deflated to constant prices of 2013 by the consumer index taken from Statistics Belgium.

26% of immigrants in our sample, 55% were born abroad (first-generation immigrants) and 45% were born in Belgium but have at least one parent who was born abroad (second-generation immigrants). Moreover, we find that 42% of (first- and second generation) immigrants originate from developing countries, and 39% of first-generation immigrants arrived in Belgium before the age of 18. Natives are working in firms with higher average hourly productivity than those employing immigrant workers (around 69 vs. 64 EUR). Natives also earn higher gross hourly wages than immigrants (around 19 vs. 17 EUR). In addition, immigrants have fewer years of attained and required education. The share of over-educated workers is larger among natives than among immigrants (21 vs. 18%)²², but the opposite result is observed for the share of under-educated workers (26 vs. 30%). On average, immigrants have fewer years of tenure, are more likely to work part-time, in a blue-collar job, with a temporary contract, and in administrative and support service activities (19 vs. 10%) as well as in the accommodation and food service sector (5 vs. 2%). Conversely, immigrants are under-represented in manufacturing (31 vs. 38%) and in wholesale and retail trade (17 vs. 20%).

[Insert Table 2 about here]

Table 2 further reports firm-level descriptive statistics on the extent of over-education among natives and immigrants according to our moderating variables. The results show that the average share of over-educated workers within firms stands at 20.4%, of which 15.6% points are natives and 4.7% points are immigrants.²³ We also observe that most over-educated workers have less than 10 years of tenure, both among natives (10.6% points) and among immigrants (3.6% points). Breaking down the share of over-educated immigrants (*i.e.* 4.7% points), we further find that those from developing countries account for almost 40% of the total (1.8% points) and that about half of over-educated immigrants were born abroad (2.5% points). Finally, our results indicate that 36% of the foreign-born among over-educated immigrants arrived in Belgium before (or at) the age of 18 (0.9% point).

²² At first glance, this result may seem somewhat surprising, but it follows from the fact that the level of attained education is lower among immigrants than among natives. Indeed, workers with lower levels of attained education are less likely to be over-educated. Restricting our sample to workers with tertiary education, we find that the proportion of over-educated workers is, as expected, higher among immigrants than among natives.

²³ The lower contribution of immigrants to this statistic is of course due to the fact that immigrants represent only 26% of workers in our sample.

4 Results

4.1 Benchmark specification

We first estimated equations (1) and (2), *i.e.* the impact of ORU variables on wages and productivity at the firm level, using a pooled OLS estimator with clustered standard errors at the firm level. This approach allows us to focus on variation within and between firms while controlling for heteroscedasticity and serial correlation of the error term. We also added a large set of control variables for worker and firm characteristics (described in Section 2) as well as year-fixed effects. The results, reported in column (1) of Table 3, first indicate that firms employing more workers in jobs with higher educational requirements should expect to pay higher hourly wages: a one-year increase in the average required level of education leads to an 8.4% increase in hourly wages. Over-education is also estimated to have a positive and significant impact on wages. However, the wage premium associated with over-education appears to be significantly smaller among immigrants than among natives. More precisely, an additional year of over-education is found to increase the hourly wage by 8.1 and 5.1%, on average, among native and immigrant workers, respectively.²⁴ Although these wage premia are positive, they are smaller than the wage premium associated with the level of education required to perform a given job. Overall, this means that over-educated workers enjoy a wage premium compared to their adequately-educated colleagues but suffer a wage penalty compared to their former classmates employed in jobs matching their education level.

[Insert Table 3 about here]

The results reported in column (2) of Table 3 enable us to analyse whether the difference between the over-education wage premium for natives and that for immigrants can be explained by a corresponding difference in productivity. They suggest that years of required and over-education both exert a significantly positive influence on productivity. More precisely, we find that productivity increases on average by 11.3% following a one-year increase in the average firm-level required level of education. This outcome is compatible with the economic literature showing that education fosters output per worker both at the country and firm level (Krueger

²⁴ It should be reminded that native-born people are those born in Belgium with both parents born in Belgium, whereas immigrants include foreign-born people (*i.e.* first-generation immigrants) and those born in Belgium with at least one foreign-born parent (*i.e.* second-generation immigrants).

and Lindahl, 2011; Lebedinski and Vandenberghe, 2014; Kampelmann *et al.*, 2018a). As regards over-education, the results show that a one-unit increase among natives and immigrants improves firm productivity on average by 13.3 and 7.7%, respectively. From a firm perspective, years of over-education among immigrants are thus found to bring less value added than among natives. This result could be explained, at least in part, by the limited transferability of immigrants' human capital, especially when the host and home country contexts differ significantly. Indeed, the difference in returns to over-education according to workers' origin might notably result from human capital discrepancies attributed to immigrants' language abilities (*e.g.* Chiswick, 1991; Chiswick and Miller, 1995; Borjas, 1999; Carnevale *et al.*, 2001; Dustmann and van Soest, 2002), literacy skills (Ferrer *et al.*, 2006; Himmler and Jäckle, 2018), or schooling quality (Sweetman, 2004).

Qualitatively, our results for productivity are consistent with those obtained for wages. However, regardless of the workers' origin, we find that over-education has a greater impact on firms' productivity than on wages. Moreover, our estimates show that the productivity-wage gap associated with over-education is larger for natives than for immigrants. From a business perspective, this means that the positive economic rents (*i.e.* productivity-wage gaps) associated with over-education are greater when over-education concerns native workers than when it concerns immigrants. From a worker perspective, this implies that over-educated workers are more productive than their well-matched colleagues and that this productivity differential is only partially reflected in the wage premium earned by over-educated workers. Put differently, it indicates that, considering their productivity, over-educated workers are underpaid compared to their well-matched colleagues. Moreover, as the magnitude of the productivity-wage gap associated with over-education is greater for natives than for immigrants, our estimates also suggest that the former are relatively more underpaid than the latter.

To sum up, in line with the existing literature (based on worker-level regressions), our estimates show that the over-education wage premium is higher for natives than for immigrants. However, since the differential in productivity gains associated with over-education between natives and immigrants outweighs the corresponding wage premium differential, we conclude that, among over-educated workers, natives are actually more underpaid than immigrants.

4.2 Robustness tests

4.2.1 Sectoral and occupational segregation

The results we have presented so far appear to be quite robust, because we have controlled for many covariates, including the share of blue-collar workers in firms and the sectoral affiliation at the NACE 1-digit level (12 dummy variables). However, it could be argued that our approach may not be fine enough to properly account for sectoral and occupational segregation. Studies indeed show that native and immigrant workers are unequally distributed across industries and occupations, and that immigrants are over-represented in less productive jobs (Aydemir and Skuterud, 2008; Elliott and Lindley, 2008; Peri and Sparber, 2009; Grinza *et al.*, 2021). In order to address this point more accurately, as a robustness test, we re-estimated equations (1) and (2) controlling for the distribution of employment within firms in 8 occupational categories (instead of the white-collar/blue-collar dichotomy in our benchmark specifications) and for the sectoral affiliation at the NACE 3-digit level (instead of 1-digit), resulting in the inclusion of 201 sectoral dummies instead of 12.

[Insert Table 4 about here]

The results are presented in Table 4. Although they are smaller than in our benchmark specification, they confirm our earlier findings. More precisely, irrespective of workers' origin, they show that over-education has a stronger impact on productivity than on wages. They also indicate that the productivity-wage gap associated with over-education is larger among natives than among immigrants. Accordingly, they again suggest that over-educated workers, in view of their productivity, are underpaid compared to their well-matched colleagues and that this underpayment is more pronounced for natives than for immigrants. Controlling for more detailed occupations and sectors is thus not found to affect our main conclusions, despite the reduced size of the effects.

4.2.2. Unobserved heterogeneity and endogeneity at the firm level

Another potential weakness of our results is that they do not take unobserved firm-level heterogeneity (*i.e.* firm-level fixed effects) or endogeneity problems into account. One way to remove unobserved characteristics that remain unchanged during the observation period (*e.g.* an advantageous location, management quality, or the ownership of a patent) is to estimate a fixed effects (FE) model. However, neither pooled OLS nor the FE estimator address the

potential endogeneity of our main explanatory variables. Yet, there might be some cyclical ‘crowding out’, namely a process by which highly educated workers take 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 over-education (under-education) within firms may increase (decrease) because of lower firm productivity. To control for this potential endogeneity issue, in addition to the state dependence of firms’ productivity/wages²⁵ and the presence of firm-level fixed effects, we re-estimated equations (1) and (2) using the dynamic system generalized method of moments (GMM-SYS) proposed by Arellano and Bover (1995) and Blundell and Bond (1998). This approach is standard in the literature on firm-level determinants of productivity and wages (*e.g.* van Ours and Stoeldraijer, 2011; Buhai *et al.*, 2017; Mahy *et al.*, 2021). It boils down to simultaneously estimating a system of two equations (one in levels and one in first differences) and relying on internal instruments to control for endogeneity. More precisely, educational variables are instrumented by their lagged levels in the differenced equation and by their lagged differences in the level equation.²⁶ The implicit assumption is that differences in a firm’s productivity and wages in one period, although possibly correlated with contemporaneous differences in educational variables, are uncorrelated with lagged levels of the latter. In the same line, levels of productivity and wages in one period, although possibly correlated with contemporaneous levels of educational variables, are assumed to be uncorrelated with lagged differences of the latter. Moreover, differences in educational variables are assumed to be reasonably correlated with their past levels, and levels of educational variables are assumed to be reasonably correlated with their past differences.

[Insert Table 5 about here]

The GMM-SYS estimates of equations (1) and (2) are reported in Table 5.²⁷ First, as shown in column (2), we find that firms’ productivity grows on average by 3.8% following a one-year increase in the level of required education. Regarding over-education, the results vary

²⁵ The assumption of persistent productivity, both at the industry and firm level, finds some support 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 wages is also highlighted in the literature (see *e.g.* Le Bihan *et al.*, 2012). Wage stickiness is notably the outcome of labour market institutions, adjustment costs, and efficiency wages’ motives.

²⁶ Bond and Söderbom (2005) provide a review of the literature on the identification of production functions. The authors notably highlight that the adjustment costs of labour and capital can justify the use of lagged values (of the endogenous variables) as instruments.

²⁷ The time dummies have been considered as exogenous, and we used the first and second lags of other explanatory variables as instruments.

according to workers' origin. Among natives, an increase in mean years of over-education is found to foster firms' productivity significantly. Conversely, among immigrants, the results show no significant relationship between over-education and productivity. The reliability of these estimates is assessed through the Arellano and Bond (1991) test for second-order autocorrelation in first-differenced errors and the Hansen (1982) test for overidentifying restrictions. As shown at the bottom of column (2), we cannot reject either the null hypothesis of valid instruments or the null hypothesis of no autocorrelation. The estimates for wages are presented in column (1) of Table 5. They show that both required and over-education exert a significantly positive effect on mean workers' wages. Again, they also indicate that the return to over-education is bigger among natives than among immigrants. However, these results should be taken with caution because the results of the Arellano and Bond (1991) test suggest the presence of some auto-correlation issue and the results of the Hansen (1982) test show that instruments are not valid. As a robustness test, we re-estimated our benchmark equation (1) using firm-level average hourly wage costs²⁸ instead of hourly wages as the dependent variable. By making this adjustment, as shown in column (1'), the GMM-SYS estimates become more reliable as they now pass both the Arellano and Bond (1991) and Hansen (1982) tests. The results using wage costs are quite similar to those obtained with wages. Indeed, they show that wage costs positively depend on mean years of required education within firms and that the effect of over-education on wage costs is bigger among natives than among immigrants. However, as with the results for productivity (see column (2)), we now find that the effect of over-education on wage costs is only significant for native workers. For immigrants, the over-education wage differential is indeed not statistically different from zero.

Turning to the impact of over-education on productivity-wage cost gaps, (*i.e.* comparing results in columns (2) and (1')), we observe that the regression coefficients associated with over-education among natives are equal to 0.041 and 0.045 in the productivity and wage cost regressions, respectively. Given that the mean sample values of productivity and wage costs reach 67 and 43.4 EUR, respectively, our GMM-SYS estimates suggest that, for firms, a one-year increase in mean years of over-education among native workers results, on average, in rises of 2.7 EUR (*i.e.* $0.041 * 67$ EUR) for hourly productivity and of 1.9 EUR ($0.045 * 43.4$ EUR) for hourly wage costs. Put differently, we find that firm rents (*i.e.* the productivity-wage

²⁸ It is the firm-level ratio of the total labour costs (*i.e.* the total expenditure borne by employers for employing staff) to the total number of hours worked. Total labour costs consist of: employee compensation (including wages, salaries in cash and in kind, employers' social security contributions); vocational training costs; other expenditure such as recruitment costs, spending on working clothes and employment taxes regarded as labour costs; minus any subsidies received. Information on this variable has been taken from the SBS.

cost gap) positively depend on the incidence of over-education among native workers. Again, this means that, in view of their productivity, over-educated native workers are underpaid compared to their well-matched colleagues. Conversely, the estimates for years of over-education among immigrants are significant neither in the productivity nor in the wage cost regressions. They therefore suggest that over-educated immigrants are just as productive and expensive as their well-matched colleagues. Overall, these results confirm our earlier findings that, despite a higher over-education wage premium for natives than for immigrants, the underpayment is in fact more pronounced for natives than for immigrants when productivity differentials are taken into account. Although the results based on wages (rather than on wage costs) should be interpreted with more caution (see our discussion above), it is worth noting that they lead to a similar conclusion.

4.3 Moderating factors

In this section, we examine the sensitivity of our benchmark results to immigrants' background. Specifically, we investigate the moderating role of the following factors: a) region of origin, b) immigrants' generation, c) age at arrival in the host country, and d) years of tenure.²⁹

4.3.1. *Immigrants' region of origin*

First, we investigate whether the effects of over-education on wages and productivity depend on immigrants' region of origin. More precisely, we test whether the consequences of over-education among immigrants differ according to whether they originate from developed or developing countries. To this end, as outlined in Section 2.2, we re-estimated equations (1) and (2) by including firm-level average years of over- (and under-) education among three distinct groups of workers: i) natives (workers born in Belgium with both parents born in Belgium), ii) immigrants from developed countries (workers either born in a developed country (excluding Belgium) or having at least one parent born in a developed country (excluding Belgium)), and

²⁹ The use of moderating variables, *i.e.* interaction effects between our educational variables and immigrants' background, mechanically reduces the number of data points in each category and for each period. In addition, for some categories of immigrants, the average number of years of over-education within firms becomes much more limited. The year-to-year variations in the value of these categories are often also smaller. This reduces the precision of the GMM-SYS estimator. In particular, micro-numerosity increases the standard errors of the corresponding estimates and makes statistical inference more difficult. Therefore, to be conservative, in this section, we will rely on the pooled OLS estimator. This implies that our results should not be interpreted in a causal way, but rather as conditional expectations aiming to complement and deepen our understanding of our benchmark findings.

iii) immigrants from developing countries (*i.e.* workers born in a developing country or having at least one parent born in a developing country).

[Insert Table 6 about here]

The results, presented in columns (1) to (3) of Table 6, show that over-education increases both productivity and wages among all three groups of workers. However, we also observe that the effects are the strongest among natives, followed by immigrants from developed countries, and finally immigrants from developing countries. In addition, we find that over-education is systematically associated with significantly positive productivity-wage gaps. The results therefore again suggest that, irrespective of their origin, over-educated workers are underpaid (in view of their productivity) compared to their well-matched colleagues. However, the estimates also indicate that the extent of this underpayment is not constant across the three groups. The underpayment of over-educated workers is the highest among natives, somewhat lower among immigrants from developed countries, and the lowest among immigrants from developing countries. Because of the estimated productivity differentials, the hierarchy of underpayment for these three groups thus appears to be identical to that suggested by the corresponding over-education wage premia. Overall, these results corroborate and refine our benchmark estimates.

4.3.2. *Immigrants' generation*

Our second sensitivity test focuses on the moderating role of immigrants' generation. More precisely, we aim to assess whether the effects of over-education on productivity and wages are different for first- and second-generation immigrants. Therefore, as discussed in Section 2.2, we re-estimated equations (1) and (2) including firm-level information on mean years of over- (and under-) education among the following groups of workers: i) natives (workers born in Belgium with both parents born in Belgium), ii) second-generation immigrants (workers born in Belgium with at least one parent born abroad), and iii) first-generation immigrants (workers born abroad).

The results are presented in columns (4) to (6) of Table 6. For all three groups of workers, we find that over-education results in increased productivity, wages, and productivity-wage gaps. Over-educated workers, irrespective of the group to which they belong, are therefore again found to be underpaid. However, the magnitude of this underpayment varies significantly

from one group to another: i) it is the greatest among natives, ii) somewhat lower among second-generation immigrants, and iii) the lowest among first-generation immigrants. Again, by explicitly taking productivity differentials into account, we can conclude that the hierarchy of the underpayment for these three groups is the same as that associated with their over-education wage premia (see the estimates in column (4) of Table 6).

Overall, this sensitivity analysis not only confirms our benchmark results but, more importantly, attests to the role of immigrants' country of birth. As expected, we indeed find that, among the over-educated, second-generation immigrants are significantly closer to natives in terms of productivity and wage premia associated with over-education. However, we also observe that: i) the higher productivity of over-educated second-generation immigrants compared to their first-generation counterparts is only partially reflected in their wage gains, and ii) over-educated second-generation immigrants close less than half the gap, both in terms of wages and productivity, between over-educated native and first-generation immigrants.

4.3.3. *Age at arrival in the host country*

We now question whether the productivity and wage effects of over-education among first-generation immigrants depend on their age of arrival in the host country. More precisely, we test whether the results differ between those who arrived in Belgium before (or at) the age of 18 and those who arrived after their majority. For this purpose, we re-estimated equations (1) and (2) including information on the extent of over- (and under-) education respectively among: i) natives; ii) second-generation immigrants; iii) first-generation immigrants who arrived in Belgium before (or at) the age of 18; and iv) first-generation immigrants who arrived in Belgium after the age of 18.

The results, presented in column (7) of Table 6, first show that a year of over-education is better rewarded for first-generation immigrants who arrived in Belgium before (or at) the age of 18. The regression coefficients are indeed equal to 0.053 and 0.043 for immigrants who arrived in Belgium before and after reaching their majority, respectively.³⁰ As regards the impact of over-education on productivity (see column (8) of Table 6), whereas it is significantly positive for those who arrived in Belgium before (or at) the age of 18, it appears to be non-significant for those who arrived at a later age. Furthermore, we find that the productivity differential of over-education for first-generation immigrants who arrived in Belgium as minors

³⁰ A *t*-test for equality of regression coefficients confirms that the return to over-education is significantly bigger for first-generation immigrants who arrived in Belgium before (or at) the age of 18 ($t = 13.2$).

is close to but still significantly lower than that estimated for second-generation immigrants (8.3 vs. 9.3%).

This sensitivity test reinforces some of our previous findings and qualifies others. Indeed, it confirms that over-education is associated with positive productivity-wage gaps for natives and second-generation immigrants. However, it also shows that the picture is more nuanced for first-generation immigrants. For those who arrived in Belgium before (or at) the age of 18, we find that the productivity gain associated with an additional year of over-education outweighs the corresponding wage differential, which means that the underpayment hypothesis can still be validated. For those who arrived after their majority, the result is different: since a year of over-education generates a significant wage premium while leaving productivity unchanged, we conclude that they are, in fact, overpaid. Indeed, our results suggest that they are paid significantly more than their well-matched colleagues despite being equally productive.

In the end, we see once again that the hierarchy of productivity-wage gaps associated with over-education by workers' origin is very similar to that observed for over-education wage premia. At one end of the spectrum, we find over-educated natives, with the highest underpayment (*i.e.* productivity-wage gap), and at the other end, we find over-educated first-generation immigrants arrived in Belgium after the age of 18, who turn out to be overpaid considering their relative productivity.

4.3.4. *Years of tenure*

With our last sensitivity test, we investigate the moderating role of tenure (*i.e.* the number of years an employee has been working for her/his current employer). In particular, we seek to assess whether the returns to over-education among natives and immigrants vary according to whether the latter have less than or at least 10 years of tenure. Therefore, we re-estimated equations (1) and (2) by explicitly taking into account the average years of over- (and under-) education among natives and immigrants respectively, depending on whether or not they were below the 10-year tenure threshold.

[Insert Table 7 about here]

The results, reported in columns (1) and (2) of Table 7, first show that the productivity and wage effects of over-education are: i) significantly positive for all four categories of workers; ii) conditional on tenure, always higher for natives than for immigrants; and iii) found

to increase with years of tenure among both natives and immigrants. Moreover, we observe that the productivity gains associated with over-education are systematically greater than the corresponding wage premia, so that all categories of over-educated workers are found to be underpaid. However, the extent of this underpayment for over-educated natives and immigrants varies with the number of years of tenure. Below the 10-year tenure threshold, the underpayment (*i.e.* productivity-wage gap) is more pronounced among natives. Conversely, at the 10-year tenure threshold and beyond, the underpayment is greater among immigrants. This reversal in the outcome can be explained by the combination of two elements. First, we find that the productivity gains associated with over-education increase much faster with tenure for immigrants than for natives. Indeed, the estimates indicate that, when crossing the 10-year tenure threshold, the productivity return to over-education doubles for immigrants (from 6.6 to 12.3%) and increases by only slightly more than a quarter for natives (from 12.3 to 15.5%). Second, we observe that the impact of tenure on the wage return to over-education is much smaller for immigrants than for natives. When crossing the 10-year tenure threshold, the over-education wage premium increases by less than 25% for immigrants (from 4.9 to 6.1%) and by more than 40% for natives (from 7.2 to 10.3%).

Overall, this sensitivity test highlights that our benchmark results, according to which over-educated native workers are more underpaid than their immigrant counterparts, hold true only for low to moderate levels of tenure. Beyond 10 years of tenure, the results are reversed because, despite the much higher productivity gains associated with tenure for over-educated immigrants, it is the wages of over-educated natives that increase more rapidly with years of tenure.

5 Discussion and conclusion

This paper is the first to estimate the direct impact of over-education among natives and immigrants, respectively, on productivity and wages at the firm level. It fills an important gap in the literature on educational mismatch as existing studies have so far not been able to directly address the question of whether the wage premia associated with over-education according to workers' origin are offset by corresponding changes in productivity. Moreover, we provide first evidence of the role played by a large range of moderators related to immigrants' background (*i.e.* region of origin, immigrant generation, age at arrival in the host country, and years of tenure) in the relationship between over-education, productivity, and wages. To this end, we used detailed Belgian linked employer-employee panel data, covering all years from 1999 to

2016, and relied on the methodological framework pioneered by Hellerstein *et al.* (1999) to estimate ORU (over-, required, and under-education) equations aggregated at the firm level.

Our benchmark results first show that a higher incidence of over-education, among both natives and immigrants, is associated with improved firm-level productivity and wages. However, they also indicate that these productivity gains and wage premia are significantly higher for natives than for immigrants. Second, irrespective of workers' origin, we find that the impact of over-education on firm's productivity is greater than that on wages. Over-education is thus found to generate a positive gap between productivity and wages, which implies that, in view of their productivity, over-educated workers appear to be underpaid compared to their well-matched colleagues. Third, our estimates show that the productivity-wage gap (*i.e.* underpayment) associated with over-education is larger for natives than for immigrants. To sum up, in line with the existing literature based on worker-level regressions (*e.g.* Battu and Sloane, 2004; Chiswick and Miller, 2008, 2010; Nielsen, 2011; Jacobs *et al.*, 2021b; Maani and Wen, 2021), we find that the over-education wage premium is higher for natives than for immigrants. However, since the differential in the productivity gains associated with over-education between natives and immigrants outweighs the corresponding wage premium differential, we conclude – based on OLS and dynamic GMM-SYS estimates – that over-educated native workers are, in fact, underpaid to a greater extent than their immigrant counterparts.

Our sensitivity analyses reinforce and refine this conclusion. Indeed, they first highlight the moderating role of immigrants' region of origin and generation, showing that the underpayment of over-educated workers is respectively: i) the highest among natives, somewhat lower among immigrants from developed countries, and the lowest among immigrants from developing countries; and ii) higher among second-generation immigrants than among first-generation immigrants. Through a sensitivity analysis, we also assess the role of the age of first-generation immigrants at their arrival in the host country. For those who arrived in Belgium before (or at) the age of 18, we find that the productivity gain associated with an additional year of over-education outweighs the corresponding wage differential, indicating that the underpayment hypothesis can still be validated. For those who arrived after their majority, the result is different: our estimates suggest that they are paid significantly more than their well-matched colleagues despite being equally productive. Accordingly, we conclude that they are relatively overpaid.

Our last sensitivity test, focusing on the role of tenure (*i.e.* the number of years an employee has been working for her/his current employer), is also of particular interest. It highlights that our benchmark results, according to which over-educated native workers are

more underpaid than their immigrant counterparts, hold true only for low to moderate levels of tenure. Among over-educated workers with at least 10 years of tenure, the results are reversed: the underpayment (*i.e.* productivity-wage gap) is found to be larger for immigrants than for natives. The reason is that, although tenure-related productivity gains are much higher for over-educated immigrants, it is the over-education wage premium of natives that increases more rapidly with years of tenure.

How can these results be interpreted? First, the positive effect of over-education on productivity, irrespective of workers' origin, is in line with the predictions of human capital theory (Becker, 1964). Indeed, it suggests that over-educated workers are more productive than their adequately educated colleagues in similar jobs due to their additional skills and abilities acquired through schooling. On the basis of our estimates, we cannot rule out that, for a given job, over-education may lead to lower job satisfaction and worse correlated attitudes and behaviours (*e.g.* absenteeism, shirking, or turnover). However, our results indicate that the *net* effect of over-education is significantly positive, *i.e.* the productivity gains associated with over-education are on average greater than the potential losses.

Second, our results showing that the productivity effects of over-education are higher for natives than for immigrants support the assumption that education levels obtained in different countries cannot be treated as perfect substitutes. More precisely, the results suggest that, on average, the returns to domestic education are higher than those to foreign education. This could be explained by the lower quality of education in immigrants' countries of origin and/or by the imperfect transferability of these skills (Chiswick, 1978; Borjas, 1985). The imperfect transferability of human capital refers not only to the difficulties immigrants face in having their diplomas and certificates recognized in the host country, but also to the fact that the skills valued in their home country are not always as productive in the host country (Schmidt, 1997).

Moreover, our estimates indicating that the productivity returns to over-education are higher for immigrants from developed countries than for those from developing countries support the idea that the transferability of human capital depends on the proximity between the home and host countries, particularly in terms of economic conditions, industrial structures, educational systems, languages, and culture. They might also suggest that, because developing countries have fewer resources to invest in their education systems, the degrees they award are, on average, of lower quality (Basilio and Bauer, 2010). According to this reasoning, the human capital of natives would therefore be less substitutable with that of immigrants from developing countries than with that of immigrants from developed countries. Our results by immigrants'

generation and age of arrival in the host country can be interpreted in a similar way. Given that second-generation immigrants are born, educated, and socialized in the host country, it is not surprising to find that the return to their years of over-education is closer to that of natives. The fact that second-generation immigrants are still far from closing the gap with natives is, however, striking and can be understood notably in the light of the study by Danhier and Jacobs (2017), who point out, based on the PISA survey, that the level of equity in terms of origin in the Belgian school system is one of the lowest among industrialized countries. First-generation immigrants who arrived in Belgium before (or at) the age of 18 spent at least the last years of their schooling (and sometimes all their schooling) in the host country. Therefore, in line with our results, we expected the returns to their years of over-education to be quite similar to those of second-generation immigrants. As regards the smaller and insignificant productivity effect of over-education for immigrants who entered Belgium at a later age, this effect again supports the hypothesis that returns to foreign education are generally lower than those to domestic education (Bratsberg and Ragan, 2000).

Our results concerning the moderating role of tenure can be interpreted as follows. On the one hand, the estimates showing that the productivity effect of over-education is smaller for immigrants than for natives at relatively low levels of tenure support the idea that over-educated immigrants have an initial knowledge deficit compared to their native counterparts at the time of hiring, which might be due to the lower quality of education in their country of origin and/or the imperfect transferability of their skills (Chiswick, 1978; Borjas, 1985).³¹ On the other hand, the fact that the productivity effect of over-education increases much faster with tenure for immigrants than for natives suggests that the former have steeper learning curves than the latter, *i.e.* that over-educated immigrants are able to make up (to a significant extent) their initial knowledge deficit relative to over-educated natives as years of service increase.³²

The empirical results presented in this paper also highlight that the wage premia associated with over-education vary widely according to workers' origin. As in previous studies (*e.g.* Chiswick and Miller, 2008, 2010; Nielsen, 2011; Jacobs *et al.*, 2021b; Maani and Wen, 2021), we find that the over-education wage premium is the greatest for natives, intermediate for immigrants from developed countries, and the smallest among immigrants from developing

³¹ At relatively low levels of tenure, over-educated immigrants might be less proficient in the languages used at work and/or have less business experience in the home country than their native counterparts. Put differently, they might have less "host-country-specific" human capital.

³² Besides, the results showing that the effect of over-education on productivity increases with tenure for both natives and immigrants is consistent with human capital theory, especially with the concepts of firm-specific human capital (Becker, 1964) and tacit knowledge (Polanyi, 1966).

countries. Moreover, we add to the existing literature by showing that the over-education wage premium is: i) the highest among natives, intermediate among second-generation immigrants, and the smallest among first-generation immigrants; and ii) higher among first-generation immigrants who arrived in Belgium before (or at) the age of 18 than among those who arrived at a later age. These results seem to be fairly consistent with human capital theory insofar as the ranking of over-educated workers in terms of wage premia is almost identical to that obtained for their productivity gaps. Put differently, they corroborate the idea that over-educated workers (of different origins) have additional skills that make them more productive and that the wages reflect these differences in productivity. However, our results do not support the hypothesis that the productivity gains driven by over-education are well aligned with wage differentials. Indeed, our estimates almost systematically highlight that over-education is associated with positive productivity-wage gaps.³³ This outcome is compatible with a series of theoretical arguments supporting the existence of a “wage compression effect” (Cardoso, 2010; Kampelmann and Rycx, 2012b), *i.e.* a situation in which the distribution of wages by level of (required and over-) education is more compressed than the corresponding education-productivity profile. More precisely, our results are in line with the literature on social norms and the hysteresis of the wage structure³⁴, fairness theories³⁵, and especially arguments that labour market regulations, such as minimum wages and collective bargaining, reduce wage inequalities by pushing the earnings of low-educated workers upward and capping wage increases for more highly (over-) educated workers (Cahuc and Zylberberg, 2014; Kampelmann *et al.*, 2020).

Considering Belgium’s strong labour market institutions³⁶ and recent empirical evidence (Kampelmann *et al.*, 2018a, 2018b, 2020), this “wage compression effect” is an

³³ The only exception is for first-generation immigrants who arrived in Belgium after the age of 18. For this group, the effect of over-education is more significant on wages than on productivity.

³⁴ Given that the technological changes of the past decades appear to be skill/task biased and that low-educated workers are typically less skilled and often doing more routine tasks (Goos *et al.*, 2014), the hysteresis in social norms discussed by Doeringer and Piore (1985) and Skott (2005) could lead to the overpayment of low-educated workers whose productivity might have been negatively affected by technological change, and the underpayment of highly (over-) educated workers whose productivity might have increased.

³⁵ Following Hamermesh (1975) and Akerlof and Yellen (1988), there is an efficiency argument for firms to pay high-productivity jobs below and low-productivity jobs above their marginal products, with a view to compress the overall wage structure.

³⁶ The collective wage bargaining system in Belgium is characterized by a high degree of centralization across firms and wage co-ordination across sectors. Wage bargaining occurs at three levels: the national (interprofessional) level, the sectoral level, and the firm level. Negotiations generally occur every two years on a pyramidal basis. In principle, the bargaining rounds are inaugurated by a national collective agreement defining the national minimum wage and, since 1996, a maximum margin for wage cost growth that may be bargained at lower levels (the “wage norm”). The objective of this “wage norm” is to make sure that all parties in the negotiations take into consideration the need for wage restraint in an open economy. After the national agreement, sector-level agreements are negotiated and concluded within Joint Committees that bring together employer and

expected outcome.³⁷ However, it has important consequences for the interpretation of over-education wage premia. Indeed, given that the productivity gains associated with over-education are almost systematically higher than the corresponding wage premia, our results suggest that, in view of their productivity, almost all over-educated workers are underpaid compared to their adequately-educated colleagues employed in similar jobs. Yet, not all groups of over-educated workers are found to be underpaid to the same extent. Indeed, our results show that the ranking of over-educated workers according to the size of their underpayment (*i.e.* productivity-wage gap) is very similar to that obtained for their wage premia. Accordingly, at one end of the spectrum, we find the over-educated natives whose underpayment is the largest and, at the other end, over-educated first-generation immigrants who arrived in Belgium after the age of 18, who turn out to be overpaid in relation to their relative productivity.

Finally, our results show that tenure plays a crucial role in determining which of the two groups of over-educated workers – natives or immigrants – is the most underpaid. During the first years of their employment in a company, over-educated natives turn out to be more underpaid than their immigrant counterparts. However, the result is gradually reversed as tenure increases: i) the productivity gains are greater for over-educated immigrants, and ii) the wage increases are more pronounced for over-educated natives. Among workers having at least 10 years of tenure, over-educated immigrants are indeed found to be more underpaid than their native counterparts. Overall, this highlights that the evolution of wage and productivity over the years of over-educated native and immigrant workers' careers should not be neglected when assessing their relative underpayment. Moreover, given that over-educated immigrants are found to catch up (to a large extent) with over-educated natives in terms of productivity as their

union representatives. Sector-level agreements set industry-wide standards, including very detailed pay scales in terms of wage levels and progress (*e.g.* promotion, seniority), for all workers covered by the Joint Committee. Sector-level agreements apply compulsorily to all companies in the sector and to their workers, whether or not they are members of the signatory organizations (employers' organizations or unions). As a result, practically the entire workforce in Belgium is covered by a sector-level agreement. Finally, firm-level agreements can complement sector-level agreements and fix wages and working time, as well as work organization and other aspects of the working life when a union delegation is present. Due to the so-called "favorability principle", firm-level bargaining can only improve (or confirm) the conditions set in the sectoral agreement. The Belgian bargaining system has remained practically unchanged since 1968 and, despite some debates, continues to enjoy a broad support. However, a concern raised by the OECD (2018) is that the maximum wage norm and the detailed pay scales in Belgium restrain the extent to which high-performing firms can offer higher wages, attract skilled workers, innovate, and grow. At the other end, legally-binding wage indexation (in combination with regulation of job protection) makes it difficult for firms in economic difficulty to reduce the wage bill. Moreover, the extension of sectoral agreements to non-signatory firms, although it reduces differences in working conditions among workers in the same sector and ensures a level-playing field for companies, may also have an adverse effect on employment and firms' performance (Hijzen and Martins, 2016). For a more detailed discussion, see *e.g.* Kampelmann and Rycx (2013), OECD (2018), and Garnero *et al.* (2020).

³⁷ It could also be expected in other EU countries whose collective bargaining translates into more rigid wage setting, such as France, Germany, Italy, or Spain, for example (Cardoso, 2010).

tenure increases, statistical discrimination theory (Phelps, 1972) predicts that the gap in over-education wage premia between natives and immigrants should also decrease with tenure. However, our results suggest the opposite. Therefore, other models, such as taste-based discrimination (Becker, 1964), might be more relevant to explain why the productivity-wage gap (*i.e.* the extent of underpayment) becomes larger for over-educated immigrants than for their native counterparts as the length of their employment relationship within a firm increases.

Reference

- Akerlof, G., and J. Yellen (1988), "Fairness and unemployment", *American Economic Review*, 78: 44-49.
- Alba, R., and V. Nee (2003), *Remaking the American mainstream: assimilation and contemporary immigration*, Harvard University Press, Cambridge (Ma.).
- Aleksynska, M., and A. Tritah (2013), "Occupation-education mismatch of immigrant workers in Europe: Context and policies", *Economics of Education Review*, 36: 229-244.
- Arellano, M., and S. Bond (1991), "Some tests of specification for panel data: Monte Carlo evidence and application to employment equations", *Review of Economic Studies*, 58 (2): 277-297.
- Arellano, M., and O. Bover (1995), "Another look at the instrumental variable estimation of error-component models", *Journal of Econometrics*, 68 (1): 28-51.
- Aydemir, A., and M. Skuterud (2008), "The immigrant wage differential within and across establishments", *ILR Review*, 61 (3): 334-352.
- Baert, S. (2017), "Hiring discrimination: an overview of (almost) all correspondence experiments since 2005", IZA Discussion paper, No. 10738, Bonn.
- Bartelsman, E. and M. Doms (2000), "Understanding productivity: lessons from longitudinal micro- data", *Journal of Economic Literature*, 38 (3): 569-594.
- Bartolucci, C. (2014), "Understanding the native-immigrant wage gap using matched employer-employee data: evidence from Germany", *ILR Review*, 67 (4): 1166-1202.
- Basilio, L., and T. Bauer (2010), "Transferability of human capital and immigrant assimilation: an analysis for Germany", IZA Discussion Paper, No. 4716, Bonn.
- Basilio, L., T. Bauer, and A. Kramer (2017), "Transferability of human capital and immigrant assimilation: An analysis for Germany", *Labour*, 31 (3): 245-264.
- Battu, H., and P. Sloane (2004), "Overeducation and ethnic minorities in Britain", *The Manchester School*, 72 (4): 535-559.
- Becker, G. (1964), *Human Capital*, National Bureau of Economic Research, New-York.
- Blinder, A. (1973), "Wage discrimination: reduced and structural estimates", *Journal of Human Resources*, 8 (4): 436-455.
- Blundell, R. and S. Bond (1998), "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, 87 (1), 115-143.
- Bond, S. and M. Söderbom (2005), "Adjustment costs and the identification of Cobb Douglas production functions", IFS Working Paper, No. 05/04, London.

- Borjas, G. (1985), "Assimilation, changes in cohort quality, and the earnings of immigrants", *Journal of Labor Economics*, 3 (4): 463-489.
- Borjas, G. (1999), "Chapter 28 - The economic analysis of immigration", in Ashenfelter, O. and Card, D. (eds.) *Handbook of Labor Economics*, North-Holland, 3 (A): 1697-1760.
- Bratsberg, B., and J. Ragan (2000), "The impact of host-country schooling on earnings: a study of male immigrants in the United States", *Journal of Human Resources*, 37 (1), 63-105.
- Brinbaum, Y. (2018), "Incorporation of immigrants and second generations into the French labour market: changes between generations and the role of human capital and origins", *Social Inclusion*, 6 (3): 104-118.
- Buhai, S., E. Cottini, E. and N. Westergaard-Nielsen (2017), "How productive is workplace health and safety?", *Scandinavian Journal of Economics*, 119 (4), 1086-1104.
- Cahuc, P., and A. Zylberberg (2014), *Labor Economics*, MIT Press, Cambridge.
- Card, D. (2005), "Is the new immigration really bad ?", *Economic Journal*, 115 (507): F300-F323.
- 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*, Palgrave Macmillan: 202-218.
- Carnevale, A., R. Fry, and B. Lowell (2001), "Understanding, speaking, reading, writing, and earnings in the immigrant labor market", *American Economic Review*, 91 (2): 159-163.
- Cataldi, A., S. Kampelmann, and F. Rycx (2011), "Productivity-wage gaps among age groups: does the ICT environment matter ?", *De Economist*, 159 (2): 193-222.
- Chiswick, B. (1978), "The effect of Americanization on the earnings of foreign-born men", *Journal of Political Economy*, 86 (5): 897-921.
- Chiswick, B. (1991), "Speaking, reading and earnings among low-skilled immigrants", *Journal of Labor Economics*, 9 (2): 149-170.
- Chiswick, B., and P. Miller (1995), "The endogeneity between language and earnings: international analyses", *Journal of Labor Economics*, 13 (2): 246-288.
- Chiswick, B., and P. Miller (2008), "Why is the payoff to schooling smaller for immigrants ?", *Labour Economics*, 15 (6): 1317-1340.
- Chiswick, B., and P. Miller (2009), "The international transferability of immigrants' human capital", *Economics of Education Review*, 28 (2): 162-169.
- Chiswick, B., and P. Miller (2010), "The effects of educational-occupational mismatch on immigrant earnings in Australia, with international comparisons", *International Migration Review*, 44 (4): 869-898.

- Corluy, V., J. Haemels, I. Marx, and G. Verbist (2015), “The labour market position of second-generation immigrants in Belgium.” National Bank of Belgium Working Paper, No. 285, Brussels.
- Danhier, J., and D. Jacobs (2017), “Aller au-delà de la ségrégation scolaire. Analyse des résultats à l’enquête PISA 2015 en Flandre et en Fédération Wallonie-Bruxelles”, Fondation Roi Baudoin, Brussels.
- Demunter, C. (2000), Structure and distribution of earnings survey: Analysis 1995. Statistics Belgium Working Paper.
- Doeringer, P., and M. Piore (1985), *Internal Labor Markets and Manpower Analysis*, ME Sharpe, Armonk.
- Dolado, J., F. Felgueroso, and J. Jimeno (2000), “Explaining youth labor market problems in Spain: crowding-out, institutions, or technology shifts ?”, *European Economic Review*, 44 (4-6): 943-956.
- Duncan, G., and S. Hoffman (1981), “The incidence and wage effects of overeducation”, *Economics of Education Review*, 1 (1): 75-86.
- Dustmann, C., and A. van Soest (2002), “Language and the earnings of immigrants”, *ILR Review*, 55 (3): 473-492.
- Elliott, R., and J. Lindley (2008), “Immigrant wage differentials, ethnicity and occupational segregation”, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171 (3): 645-671.
- Eurostat (2021), *Over-qualification Rate by Economic Activity: Whole Economy*, Luxembourg: Eurostat Database.
- Falcke, S., C. Meng, and R. Nollen (2020), “Educational mismatches for second generation migrants. An analysis of applied science graduates in the Netherlands”, *Journal of Ethnic and Migration Studies*, 46 (15): 3235-3251.
- Fays, V., B. Mahy, F. Rycx, and M. Volral (2021), “Wage discrimination based on the country of birth: do tenure and product market competition matter ?”, *Applied Economics*, 53 (13): 1551-1571.
- Ferrer, A., W. Riddell, and D. Green (2006), “The effect of literacy on immigrant earnings”, *Journal of Human Resources*, 42 (2): 380-410.
- Flabbi, L., and A. Ichino (2001), “Productivity, seniority and wages: new evidence from personnel data”, *Labour Economics*, 8 (3), 359-387.
- FPS Employment and Unia. (2017), *Monitoring socio-économique: marché du travail et origine*. FPS Employment and Unia Publishing, Brussels.

- Gagliardi, N., B. Mahy, and F. Rycx (2021), “Upstreamness, Wages and Gender: Equal Benefits for All?.” *British journal of industrial relations*, 59 (1): 52-83.
- Garnero, A., S. Kampelmann, and F. Rycx (2014), “Part-time work, wages and productivity: evidence from Belgian matched panel data”, *ILR Review*, 67 (3): 115-130.
- Garnero, A., F. Rycx, and I. Terraz (2020), “Productivity and wage effects of firm-level collective agreements: evidence from Belgian linked panel data”, *British Journal of Industrial Relations*, 58 (4): 936-972.
- Giuliano, R., S. Kampelmann, B. Mahy, and F. Rycx (2017), “Short notice, big difference? The effect of temporary employment on firm competitiveness across sectors”, *British Journal of Industrial Relations*, 55 (2): 421-449.
- Goos, M., A. Manning, A., and A. Salomons (2014), “Explaining job polarization: routine-biased technological change and offshoring”, *American Economic Review*, 104, 2509-2526.
- Grinza, E., S. Kampelmann, and F. Rycx (2020), “L’union fait la force? Evidence for wage discrimination in firms with high diversity”, *The Journal of Economic Inequality* 18 (2): 181-211.
- Grunau, P. (2016), “The impact of overeducated and undereducated workers on firm-level productivity: first evidence for Germany”, *International Journal of Manpower*, 37 (2): 258-283.
- Hamermesh, D. (1975), “Interdependence in the labour market”, *Economica*, 42 (168) : 420-429.
- Hansen, L. (1982), “Large sample properties of generalized method of moments estimators.” *Econometrica*, 50: 1029-1054.
- Hartog, J. (2000), “Over-education and earnings: where are we, where should we go ?”, *Economics of Education Review*, 19 (2): 131-147.
- Hellerstein, J., D. Neumark, and K. Troske (1999), “Wages, productivity and worker characteristics: evidence from plant-level production functions and wage equations”, *Journal of Labor Economics*, 17 (3): 409-446.
- Hellerstein, J., and D. Neumark (2004), “Production function and wage equation estimation with heterogeneous labor: evidence from a new matched employer-employee data set”, NBER Working Paper, No. 10365, Cambridge (Ma.).
- Hijzen, A., and P. Martins (2016), “No extension without representation? Evidence from a natural experiment in collective bargaining”, IMF Working Paper, No. 16/143, International Monetary Fund, Washington, DC.

- Himmler, O., and R. Jäckle (2018), “Literacy and the migrant- native wage gap”, *Review of Income and Wealth*, 64 (3): 592-625.
- Jacobs, V., B. Mahy, F. Rycx, and M. Volral (2021a), “Over-education among immigrants: the role of demographics, time and firm characteristics”, *Applied Economics*, 53 (1): 61-78.
- Jacobs, V., F. Rycx, and M. Volral (2021b), “Wage effects of educational mismatch according to workers’ origin: the role of demographics and firm characteristics”, IZA Discussion Paper, No. 14813, Bonn.
- Joona, P.A., N.D. Gupta, and E. Wadensjö (2014), “Overeducation among immigrants in Sweden: incidence, wage effects and state dependence”, *IZA Journal of Migration*, 3 (1): 9.
- Kampelmann, S., B. Mahy, F. Rycx, and G. Vermeulen (2020), “Over-, required-, and undereducation: consequences on the bottom lines of firms”, *Labour*, 34 (1): 80-112.
- Kampelmann, S., and F. Rycx (2012a), “The impact of educational mismatch on firm productivity: Evidence from linked panel data”, *Economics of Education Review*, 31 (6): 918-931.
- Kampelmann, S., and F. Rycx (2012b), “Are occupations paid what they are worth ? An econometric study of occupational wage inequality and productivity”, *De Economist*, 160: 257-287.
- Kampelmann, S., and F. Rycx (2013), “Does institutional diversity account for pay rules in Germany and Belgium ?”, *Socio-Economic Review*, 11 (1): 131-157.
- Kampelmann, S., and F. Rycx (2016), “Wage discrimination against immigrants: measurement with firm-level productivity data”, *IZA Journal of Migration*, 5 (1): 1-24.
- Kampelmann, S., F. Rycx, Y. Saks, and I. Tojerow (2018a), “Does education raise productivity and wages equally? The moderating role of age and gender”, *IZA Journal of Labor Economics*, 7 (1): 1-37.
- Kampelmann, S., F. Rycx, Y. Saks, and I. Tojerow (2018b), “Misalignment of productivity and wages across regions: evidence from Belgium”, *Regional Studies*, 52 (12): 1697-1707.
- Kiker, B., M. Santos, and M. Mendes de Oliveira (1997), “Overeducation and undereducation: evidence for Portugal”, *Economics of Education Review*, 16 (2): 111-125.
- Konings, J., and S. Vanormelingen (2015), “The impact of training on productivity and wages: firm level evidence”, *Review of Economics and Statistics*, 97 (2): 485-497.
- Krueger, A. and M. Lindahl (2011), “Education for growth: why and for whom ?”, *Journal of Economic Literature*, 39 (4): 1101-1136.
- Lebedinski, L., and V. Vandenberghe (2014), “Assessing education’s contribution to productivity using firm-level evidence”, *International of Manpower*, 35 (8): 1116-1139.

- Le Bihan, H., J. Montornes, and T. Heckel (2012), “Sticky wages: evidence from quarterly microeconomic data”, *American Economic Journal: Macroeconomics*, 4 (3): 1-32.
- Maani, S., and L. Wen (2021), “Over-education and immigrant earnings: a penalized quantile panel regression analysis”, *Applied Economics*, 53 (24): 2771-2790.
- Mahy, B., F. Rycx, and G. Vermeulen (2015), “Educational mismatch and firm productivity: do skills, technology and uncertainty matter?”, *De Economist*, 163 (2): 233-262.
- Mahy, B., F. Rycx, M. Volral, and G. Vermeulen (2021), “Productivity and wage effects of firm-level upstreamness: evidence from Belgian Linked panel data”, *The World Economy*, forthcoming.
- Matano, A., S. Nieto, and R. Ramos (2015), “Educational mismatches in the EU: immigrants vs natives.” *International Journal of Manpower*, 36 (4): 540-561.
- Mattoo, A., I. Neagu, and Ç. Özden (2008), “Brain waste? Educated immigrants in the U.S. labor market”, *Journal of Development Economics*, 87 (2): 255-269.
- McGowan, M., and D. Andrews (2015), “Labour market mismatch and labour productivity: evidence from PIAAC data”, OECD Economics Department Working Paper, No. 1209, OECD Publishing Paris.
- McGuinness, S. (2006), “Overeducation in the labour market”, *Journal of Economic Surveys*, 20 (3): 387-418.
- Meurs, D., A. Pailhé, and P. Simon (2006), “Persistence des inégalités entre générations liées à l’immigration: l’accès à l’emploi des immigrés et de leurs descendants en France”, *Population* 61 (5-6): 763-802.
- Mincer, J. (1974), *Schooling, Experience, and Earnings*, NBER Publishing.
- Nielsen, C. (2011), “Immigrant over-education: evidence from Denmark”, *Journal of Population Economics*, 24 (2): 499-520.
- Oaxaca, R. (1973), “Male-female wage differential in urban labor market”, *International Economic Review*, 14 (3): 693-709.
- OECD (2009), *Measuring Capital*, Paris.
- OECD (2011), *Employment outlook*, Paris.
- OECD (2018), “The role of collective bargaining systems for good labour market performance.” In *OECD Employment Outlook 2018*, OECD Publishing, Paris.
- van Ours, J., and L. Stoeldraijer (2011), “Age, wage and productivity in Dutch manufacturing”, *De Economist*, 159 (2): 113-137.

- Peri, G., and C. Sparber (2009), “Task specialization, immigration, and wages”, *American Economic Journal: Applied Economics*, 1 (3): 135-169.
- Phelps, E. (1972), “The statistical discrimination theory of racism and sexism”, *American Economic Review*, 62 (4): 659-661.
- Piton, C., and F. Rycx (2021), “A broken social elevator? Employment outcomes of first- and second-generation immigrants in Belgium”, *De Economist*, 169 (3): 319-365.
- Polanyi, M. (1966), “The logic of tacit inference”, *The Journal of the Royal Institute of Philosophy*, 41 (155): 1-18.
- Portes, A. and R. Rumbaut (2001), *Legacies: the story of immigrant second generation*, University of California Press, California.
- Quintini, G. (2011), “Over-qualified and under-skilled: a review of existing literature”, OECD Social, Employment and Migration Working Paper, No. 121, OECD Publishing, Paris.
- Sanromá, E., R. Ramos, and H. Simón (2015a), “How relevant is the origin of human capital for immigrant wages? Evidence from Spain”, *Journal of Applied Economics*, 18 (1): 149-172.
- Sanromá, E., R. Ramos, and H. Simón (2015b), “Portability of human capital and immigrant overeducation in Spain”, *Population Research and Policy Review*, 34 (2): 223-241.
- Schmidt, C. (1997), “Immigrant performance in Germany: labor earnings and ethnic German migrants and foreign guest-workers”, *Quarterly Review of Economics and Finance*, 37, 683-740.
- Schwientek, C. (2016), “Are immigrants overeducated in Germany ? Determinants and wage effects of educational mismatch”, FAU Discussion Paper in Economics, No. 07/2016, Nürnberg.
- Skott, P. (2005), “Fairness as a source of hysteresis in employment and relative wages”, *Journal of Economic Behavior and Organization*, 57, 305-331.
- Sweetman, A. (2004), “Immigrant source country educational quality and Canadian labour market outcomes”, *Research Paper*, No. 234, Statistics Canada, Ottawa.
- Syverson, C. (2011), “What determines productivity?”, *Journal of Economic Literature*, 49 (2), 326-365.
- UNCTAD (2020), *World Economic Situation and Prospects 2020*, United Nations, New York, p. 160-200.
- Vandenberghe, V. (2013), “Are firms willing to employ a greying and feminizing workforce?”, *Labour Economics*, 22: 30-46.
- Verdugo, R., and N. Verdugo (1989), “The impact of surplus schooling on earnings. Some additional findings”, *Journal of Human Resources*, 24 (4): 629-643.

- Wald, S., and T. Fang (2008), “Overeducated immigrants in the Canadian labour market: evidence from the workplace and employee survey”, *Canadian Public Policy*, 34 (4): 457-479.
- Zorlu, A., and J. Hartog (2012), “Employment assimilation of immigrants in the Netherlands: dip and catchup by source country”, *International Journal of Population Research*, 23p.

Table 1: Firm- and worker-level descriptive statistics for selected variables, 1999-2016

Variables	Firm level	Worker level		
	Total (1)	Natives (2)	Immigrants (3)	Total (4)
Productivity and wages				
Hourly value added (€, at 2013 constant prices)	67.0	68.8	64.2	67.0
Hourly wage (€, at 2013 constant prices)	18.7	18.7	17.4	18.7
ORU education				
Attained education (years)	12.2	12.2	11.5	12.0
Required education (years)	12.5	12.4	12.0	12.3
Over-education				
Percentage of workers	20.4	20.8	18.4	20.1
Years	0.58	0.60	0.54	0.58
Under-education				
Percentage of workers	27.1	26.2	29.5	27.0
Years	0.89	0.86	1.01	0.90
Worker and job characteristics				
Natives	73.8	100		74.0
Immigrants	26.2		100	26.0
<i>Immigrants originating from:</i>				
Developed countries	15.7		58.1	15.1
Developing countries	10.4		41.9	10.9
<i>Generations of immigrants:</i>				
Second-generation immigrants	11.9		45.5	14.2
First-generation immigrants	14.3		54.5	11.9
<i>Age at arrival in Belgium:</i>				
< 18 years	5.5		39.1	5.6
≥ 18 years	8.8		60.9	8.6
Workers younger than 25 years (%)	8.1	10.4	9.4	10.1
Workers between 25 and 49 years (%)	75.1	73.8	74.9	74.1
Workers older than 49 years (%)	16.9	15.8	15.7	15.8
Workers with at least 10 years of tenure (%)	37.1	37.2	26.1	34.3
Women (%)	29.0	31.1	31.6	31.2
Part-time workers (%)	44.5	34.9	49.7	36.4
Open-ended contracts (%)	95.1	92.9	88.7	91.8
Blue-collar workers (%)	52.0	46.4	55.2	48.7
Firm characteristics				
Firm size (number of FTE employees)	331.9	519.8	498.1	514.2
Capital stock (€, at 2013 constant prices)	220,412			
Firm-level collective agreement (%)	29.6	26.7	24.3	26.1
Sector of activity (%)				
Mining and quarrying (B)	1.0	0.8	0.8	0.8
Manufacturing (C)	48.0	38.2	31.2	36.3
Electricity, gas, steam and air conditioning supply (D)	0.1	0.6	0.3	0.5
Water supply, sewerage, waste management and remediation activities (E)	0.9	1.0	0.7	1.0
Construction (F)	10.9	9.6	9.5	9.6
Wholesale and retail trade, repair of motor vehicles and motorcycles (G)	15.2	19.8	16.7	19.0
Transportation and storage (H)	5.0	8.7	8.1	8.6
Accommodation and food service activities (I)	2.3	2.0	4.9	2.7
Information and communication (J)	4.1	4.1	3.6	4.0
Financial and insurance activities (K)	1.4	1.0	1.2	1.1
Real estate activities (L)	0.9	1.1	1.4	1.2
Professional, scientific and technical activities (M)	2.8	2.8	2.5	2.7
Administrative and support service activities (N)	7.4	10.2	19.0	12.5
Share of sample (%)	100	74.0	26.0	100
Number of observations	21,086	906,970	321,150	1,228,120

Table 2: Firm-level incidence of over-educated workers (shares and mean years), decomposed according to origin and moderators, 1999-2016

	Share of over-educated workers	Mean years of over-education
<i>Among:</i>		
All workers	20.4	0.58
Natives	15.6	0.44
<i>According to tenure:</i>		
< 10 years	10.6	0.30
≥ 10 years	5.1	0.14
Immigrants	4.7	0.14
<i>According to tenure:</i>		
< 10 years	3.6	0.10
≥ 10 years	1.1	0.03
<i>According to origin:</i>		
Developed countries	3.0	0.08
Developing countries	1.8	0.05
<i>According to generations:</i>		
Second-generation immigrants	2.3	0.06
First-generation immigrants	2.5	0.07
<i>According to age at arrival in Belgium:</i>		
< 18 years	0.9	0.03
≥ 18 years	1.5	0.04

Table 3: Over-education, wages and productivity, benchmark OLS estimates

<i>Dependent variables:</i>	Hourly wage (log)	Hourly value added (log)	<i>t</i> -test for equality of coefficients
	(1)	(2)	(3)
Required-education (in years)	0.084*** (0.001)	0.113*** (0.004)	<i>t</i> = 1021.3
Over-education among natives (in years)	0.081*** (0.003)	0.133*** (0.010)	<i>t</i> = 723.3
Over-education among immigrants (in years)	0.051*** (0.008)	0.077*** (0.024)	<i>t</i> = 149.2
Worker characteristics ^a	YES	YES	
Firm characteristics ^b	YES	YES	
Year dummies	YES	YES	
Adjusted R ²	0.704	0.392	
Sig. Model (<i>p</i> -value)	0.000	0.000	
Number of firm-year observations	21,086	21,086	
Number of firms	5,704	5,704	

Robust standard errors are reported in parentheses. ^a Years of under-education among natives and immigrants respective, and firm-level share of the workforce that: (i) is immigrant, (ii) is female, (iii) is younger than 25 and older than 49 years, respectively, (iv) occupies blue-collar jobs, (v) has at least 10 years of tenure, (vi) works part-time, and (vii) has a fixed term contract, an apprenticeship contract or works under contract with a temporary employment agency. ^b Sectorial affiliation (13 categories), size of the company (log of FTE workers), the level of wage bargaining (2 categories), age and location of the firm, and log of capital stock per worker.

*** *p*<0.01, ** *p*<0.05, * *p*<0.1.

Table 4: Robustness tests for sectoral and occupational segregation, OLS estimates

<i>Dependent variables:</i>	Hourly wage (log) (1)	Hourly value added (log) (2)	<i>t</i> -test for equality of coefficients (3)
Required-education (in years)	0.033*** (0.002)	0.046*** (0.006)	<i>t</i> = 298.5
Over-education among natives (in years)	0.049*** (0.003)	0.077*** (0.010)	<i>t</i> = 389.4
Over-education among immigrants (in years)	0.029*** (0.007)	0.041* (0.023)	<i>t</i> = 72.5
Share of workforce by occupation	8 categories	8 categories	
Firm sectoral affiliation	201 dummies	201 dummies	
Worker characteristics ^a	YES	YES	
Firm characteristics ^b	YES	YES	
Year dummies	YES	YES	
Adjusted R ²	0.774	0.450	
Sig. Model (<i>p</i> -value)	0.000	0.000	
Number of firm-year observations	21,086	21,086	
Number of firms	5,704	5,704	

Robust standard errors are reported in parentheses. ^a Years of under-education among natives and immigrants respectively, and firm-level share of the workforce that: (i) is immigrant, (ii) is female, (iii) is younger than 25 and older than 49 years, respectively, (iv) has at least 10 years of tenure, (v) works part-time, and (vi) has a fixed term contract, an apprenticeship contract or works under contract with a temporary employment agency. ^b Size of the company (log of FTE workers), the level of wage bargaining (2 categories), age and location of the firm, and log of capital stock per worker.

*** *p*<0.01, ** *p*<0.05, * *p*<0.1.

Table 5: Robustness test for unobserved heterogeneity and endogeneity at the firm level, GMM-SYS estimates

<i>Dependent variables:</i>	Hourly wage (log) (1)	Hourly wage cost (log) (1')	Hourly value added (log) (2)
Lagged dependent variable (ln)	0.203*** (0.018)	0.384*** (0.118)	0.403*** (0.050)
Required-education (in years)	0.063*** (0.003)	0.045*** (0.011)	0.038*** (0.007)
Over-education among natives (in years)	0.046*** (0.004)	0.045*** (0.015)	0.041*** (0.011)
Over-education among immigrants (in years)	0.022** (0.010)	0.012 (0.024)	-0.010 (0.018)
Worker characteristics ^a	YES	YES	YES
Firm characteristics ^b	YES	YES	YES
Year dummies	YES	YES	YES
Sig. Model (p-value)	0.000	0.000	0.000
Hansen statistic	1423.5	990.83	1246.65
<i>p-value</i>	0.000	0.225	0.146
Arellano-Bond statistic (AR2)	2.98	1.63	1.64
<i>p-value</i>	0.003	0.102	0.102
Number of firm-year observations	21,086	21,086	21,086
Number of firms	5,704	5,704	5,704

Robust standard errors are reported in parentheses. ^a Years of under-education among natives and immigrants respectively, and firm-level share of the workforce that: (i) is immigrant, (ii) is female, (iii) is younger than 25 and older than 49 years, respectively, (iv) occupies blue-collar jobs, (v) has at least 10 years of tenure, (vi) works part-time, and (vii) has a fixed term contract, an apprenticeship contract or works under contract with a temporary employment agency. ^b Sectorial affiliation (13 categories), size of the company (log of FTE workers), the level of wage bargaining (2 categories), age and location of the firm, and log of capital stock per worker.

*** p<0.01, ** p<0.05, * p<0.1.

Table 6: The moderating role of immigrants' region of birth, generation and age at arrival in host country, OLS estimates

Immigrants' region of birth				Immigrants' generation				Age at arrival in host country			
<i>Dependent variables:</i>	Hourly wage (log)	Hourly value added (log)	<i>t</i> -test for equality of coefficients	<i>Dependent variables</i>	Hourly wage (log)	Hourly value added (log)	<i>t</i> -test for equality of coefficients	<i>Dependent variables</i>	Hourly wage (log)	Hourly value added (log)	<i>t</i> -test for equality of coefficients
	(1)	(2)	(3)		(4)	(5)	(6)		(7)	(8)	(9)
Required-education (in years)	0.084*** (0.001)	0.113*** (0.004)	<i>t</i> = 1021.3	Required-education (in years)	0.084*** (0.001)	0.113*** (0.004)	<i>t</i> = 1021.3	Required-education (in years)	0.084*** (0.001)	0.113*** (0.004)	<i>t</i> = 1021.3
Over-education among natives (in years)	0.081*** (0.003)	0.132*** (0.010)	<i>t</i> = 709.3	Over-education among natives (in years)	0.081*** (0.003)	0.133*** (0.010)	<i>t</i> = 723.3	Over-education among natives (in years)	0.081*** (0.003)	0.132*** (0.010)	<i>t</i> = 709.3
Over-education among immigrants from developed countries (in years)	0.059*** (0.014)	0.097*** (0.031)	<i>t</i> = 162.2	Over-education among 2nd generation immigrants (in years)	0.055*** (0.016)	0.094** (0.042)	<i>t</i> = 126.0	Over-education among 2nd generation immigrants (in years)	0.055*** (0.016)	0.093** (0.043)	<i>t</i> = 120.3
Over-education among immigrants from developing countries (in years)	0.039*** (0.009)	0.050* (0.028)	<i>t</i> = 54.3	Over-education among 1st generation immigrants (in years)	0.046*** (0.010)	0.064** (0.030)	<i>t</i> = 82.7	Over-education among 1st generation immigrants arrived < 18 years (in years)	0.053*** (0.015)	0.083* (0.048)	<i>t</i> = 78.8
								Over-education among 1st generation immigrants arrived ≥ 18 years (in years)	0.043*** (0.013)	0.056 (0.038)	<i>t</i> = 47.0
Worker characteristics ^a	YES	YES		Worker characteristics	YES	YES		Worker characteristics	YES	YES	
Firm characteristics ^b	YES	YES		Firm characteristics	YES	YES		Firm characteristics	YES	YES	
Year dummies	YES	YES		Year dummies	YES	YES		Year dummies	YES	YES	
Adjusted R ²	0.704	0.392		Adjusted R ²	0.705	0.393		Adjusted R ²	0.705	0.392	
Sig. Model (p-value)	0.000	0.000		Sig. Model (p-value)	0.000	0.000		Sig. Model (p-value)	0.000	0.000	
Number of firm-year observations	21,086	21,086		Number of firm-year observations	21,086	21,086		Number of firm-year observations	21,086	21,086	
Number of firms	5,704	5,704		Number of firms	5,704	5,704		Number of firms	5,704	5,704	

Robust standard errors are reported in parentheses. ^a Years of under-education among natives and immigrants respectively, and firm-level share of the workforce that: (i) is immigrant (first and second-generation immigrants are considered separately when testing the moderating role of immigrant generation and age at arrival in host country), (ii) is female, (iii) is younger than 25 and older than 49 years, respectively, (iv) occupies blue-collar jobs, (v) has at least 10 years of tenure, (vi) works part-time, and (vii) has a fixed term contract, an apprenticeship contract or works under contract with a temporary employment agency. ^b Sectorial affiliation (13 categories), size of the company (log of FTE workers), the level of wage bargaining (2 categories), age and location of the firm, and log of capital stock per worker.

*** p<0.01, ** p<0.05, * p<0.1.

Table 7: The moderating role of tenure, OLS estimates

<i>Dependent variables:</i>	Hourly wage (log) (1)	Hourly value added (log) (2)	<i>t</i> -test for equality of coefficients (3)
Required-education (in years)	0.084*** (0.001)	0.113*** (0.004)	<i>t</i> = 1021.3
Over-education among natives with tenure < 10 years (in years)	0.072*** (0.004)	0.123*** (0.014)	<i>t</i> = 508.6
Over-education among natives with tenure ≥ 10 years (in years)	0.103*** (0.006)	0.155*** (0.020)	<i>t</i> = 235.9
Over-education among immigrants with tenure < 10 years (in years)	0.049*** (0.008)	0.066** (0.028)	<i>t</i> = 84.8
Over-education among immigrants with tenure ≥ 10 years (in years)	0.061** (0.025)	0.123** (0.048)	<i>t</i> = 166.4
Worker characteristics ^a	YES	YES	
Firm characteristics ^b	YES	YES	
Year dummies	YES	YES	
Adjusted R ²	0.705	0.393	
Sig. Model (<i>p</i> -value)	0.000	0.000	
Number of firm-year observations	21,086	21,086	
Number of firms	5,704	5,704	

Robust standard errors are reported in parentheses. ^a Years of under-education among natives and immigrants respectively, and firm-level share of the workforce that: (i) is immigrant, (ii) is female, (iii) is younger than 25 and older than 49 years, respectively, (iv) occupies blue-collar jobs, (v) has at least 10 years of tenure, (vi) works part-time, and (vii) has a fixed term contract, an apprenticeship contract or works under contract with a temporary employment agency. ^b Sectorial affiliation (13 categories), size of the company (log of FTE workers), the level of wage bargaining (2 categories), age and location of the firm, and log of capital stock per worker.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.