I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17027

Immigrant Overeducation across Generations: The Role of Gender and Part-Time Work

Kevin Pineda-Hernández
François Rycx
Mélanie Volral

MAY 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17027

Immigrant Overeducation across Generations: The Role of Gender and Part-Time Work

Kevin Pineda-Hernández
ULB, UMONS and FRS-FNRS

François Rycx
ULB, UMONS, UCLouvain, GLO and IZA

Mélanie Volral
UMONS and ULB

MAY 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Immigrant Overeducation across Generations: The Role of Gender and Part-Time Work*

A large body of literature shows that first-generation immigrants born in developing countries experience a higher likelihood of being overeducated than natives (i.e. immigrant overeducation). However, evidence is remarkably scarce when it comes to the overeducation of second-generation immigrants. Using a matched employer-employee database for Belgium over the period 1999-2016 and generalized ordered logit regressions, we contribute to the literature with one of the first studies on the intergenerational nexus between overeducation and origin among tertiary-educated workers. We show that immigrant overeducation disappears across two generations when workers work full-time. However, immigrant overeducation is a persistent intergenerational phenomenon when workers work part-time. Our gender-interacted estimates endorse these findings for female and male immigrants.

JEL Classification: 21, I22, J15, J24, J61, J62, J71

Keywords: immigrants, intergenerational studies, labour market integration, overeducation, generalized ordered logit, moderating factors

Corresponding author:
Kevin Pineda-Hernández
Solvay Brussels School of Economics and Management (SBS-EM)
Université libre de Bruxelles
Avenue F.D. Roosevelt 50 - CP 140
1050 Brussels
Belgium
E-mail: kevin.pineda.hernandez@ulb.be

* The authors gratefully acknowledge funding from the National Fund for Scientific Research (FRS-FNRS, 2022 FRESH grantee), the Belgian Science Policy Office (BELSPO, Brain-BE 2.0: LAMIE project) and the Fondation Bernheim. They also thank Statistics Belgium for giving access to the data.

1. Introduction

According to Eurostat (2023), the employment rate of foreign-born people in the European Union (EU) was 65.9% in 2021, around 5% points lower than that of native-born people (70.7%).[1] Difficulties in accessing the labour market also extend to the descendants of immigrants, whose employment rate was 65.3%. Zooming in on the EU employed population, several studies show that the wages of immigrants are lower than those of natives, albeit the extent of this wage gap significantly depends on immigrants' origin (e.g. Athari et al., 2019 for France; Hammarstedt and Palme, 2012 for Sweden; Pineda-Hernández et al., 2022 for Belgium). More precisely, while immigrants from developed countries earn similar wages to natives across two generations, immigrants from developing countries face persistent intergenerational wage inequalities. However, regarding the employment conditions of immigrants, intergenerational studies are scarce (e.g. Belfi et al., 2022; Belzil and Poinas, 2010), especially in the context of overeducation (i.e. the condition of having a higher level of education than that required to perform a specific job).

Moreover, it is of general interest to investigate whether origin contributes to the likelihood of being overeducated for a worker (i.e. immigrant overeducation) and whether this phenomenon holds across generations, as overeducation can have negative micro and macroeconomic consequences. In essence, overeducation can lead to: i) wage penalties and lower job satisfaction for overeducated workers because they are not paid according to their level of education, ii) lower productivity for firms due to the underutilisation of overeducated workers' skills, higher labour turnover and an increase in absenteeism, iii) higher inequality and poverty for societies because overeducated workers may replace less-qualified workers, pushing them into low-paying occupations or unemployment; and iv) lower economic growth for countries as a result of funding non-productive education and misallocation of human capital (Brunello and Wruuck, 2019; Davia et al., 2017; McGuiness, 2006; Nielsen, 2011; Nugent, 2022).

[1] Unless mentioned otherwise, this paper henceforth uses the words i) 'first-generation immigrants' and 'foreign-born people' for people born abroad; ii) 'second-generation immigrants', 'children of immigrants' and 'descendants of immigrants' for people born in the host country with at least one foreign-born parent; iii) 'immigrants' for first- and second-generation immigrants; and iv) 'natives' for people born in the host country with both parents born in the host country.

The nexus between origin and overeducation can be established through different underlying mechanisms. First, human capital theory specifies that first-generation (F-G) immigrants face a disadvantageous position in the host country's labour market, as their foreign education and experience are unlikely to be perfectly transferred across borders (e.g. educational qualifications acquired in developing countries are often non-recognized in developed countries) (Basilio et al., 2017). Second, screening theory underlines the poor signal that a foreign diploma may send to employers (i.e. employers tend to undervalue schooling and language capabilities acquired in developing countries) (Chiswick and Miller, 2009; Zwysen and Demireva, 2018). Third, job search theory states that F-G immigrants may remain clustered in low-paying occupations (i.e. jobs that require a low level of education) due to their insufficient knowledge of the functioning of the host country's labour market (Akkaymak, 2017). Several empirical studies accord with these theories. In sum, they show that: i) F-G immigrants are more likely to be overeducated than natives (i.e. evidence of immigrant overeducation) (e.g. Jacobs et al., 2021; Lindley, 2009; Nielsen, 2011; Wen and Maani, 2018); ii) F-G immigrants experience a higher probability of state dependence in overeducation than natives[2] (e.g. Joona et al., 2014; Kalfa and Piracha, 2017); and iii) the negative effect of overeducation on earnings is more substantial for F-G immigrants than for natives (e.g. Jacobs et al., 2022a; Maani and Wen, 2021; Nielsen, 2011).

However, these underlying mechanisms are unlikely to explain the overeducation of second-generation (S-G) immigrants. Indeed, classical assimilation theory stipulates that since S-G immigrants are born, educated and socialised from childhood to adulthood in the host country, their likelihood of being overeducated should be lower than that of their F-G peers and similar to that of natives (Alba et al., 2011; Park and Myers, 2010). In other words, immigrant overeducation should disappear across two generations, as S-G immigrants possess human and social capital linked to the host country's labour market. However, segmented-assimilation theory sees this as a very optimistic assumption and instead suggests that S-G immigrants may still have to deal with marginalisation and discrimination due to the parental transmission of ethnic traits (e.g. religiosity, skin colour and patronymic) and social characteristics (e.g. low-income families and concentration in immigrant-dense neighbourhoods) (Blau et al., 2013; OECD, 2017b; Phalet and Heath, 2010). OECD (2020c)

[2] By state dependence in overeducation, we mean that being overeducated for a worker in the previous period increases his/her likelihood of being overeducated in the current period.

further highlights that S-G immigrants may not wholly benefit from the host country's education due to school segregation and less parental support in their learning process.[3] Therefore, differences in employment conditions between natives and S-G immigrants cannot be completely ruled out.

Furthermore, immigrant overeducation could be a persistent intergenerational phenomenon because: i) employers can make recruitment decisions based on imperfect information and ethnic stereotypes (i.e. statistical discrimination), ii) employers can have ethnic preferences for certain occupations, regardless of observed abilities and qualifications (i.e. taste-based discrimination), and iii) employers can take advantage of the barriers that immigrants face in accessing the primary labour market to hire them for jobs that do not match their level of education (e.g. monopsonistic discrimination) (Becker, 1957; Zschirnt and Ruedin, 2016).
As far as we know, only a few empirical papers have explicitly considered S-G immigrants in the relationship between overeducation and origin. This is understandable because databases that examine this research topic rarely contain information on the country of birth of workers' parents. Using population data for Sweden in 2007 and logit regressions, Dahlstedt (2015) shows that the likelihood of being overeducated for S-G female immigrants is similar to that of female natives. In contrast, there is still an overeducation gap between S-G male immigrants and male natives. Using register data for Norway in 2014 and OLS regressions, Larsen et al. (2018) find that S-G female and male immigrants are just as likely to be overeducated as same-gender natives, thus performing better than their F-G same-gender counterparts. Using the Labour Force Survey for Spain in the years 2008 and 2014 and logit regressions, Fernández-Reino et al. (2018) find that while immigrant overeducation significantly reduces across two generations among male workers, S-G female immigrants reverse the overeducation gap that their F-G same-gender peers experience. Using survey data for the Netherlands over the period 2006-2014 and multinomial logit regressions, Falcke et al. (2020) show that S-G female and male immigrants from non-Western countries (i.e. emerging and developing economies) are more likely to be overeducated than same-gender natives.

Although these studies are the first to bring light to the overeducation of S-G immigrants, we believe there is room for significant improvement in the empirical strategy and data used for

[3] F-G immigrants tend to be less educated, less proficient in the host country's language and less informed about how the school system works, which reduces the degree of support in their children's learning.

this research topic. First, most evidence is based on small samples or unbalanced data. Second, most studies use econometric regressions that may lead to biased estimations. For instance, binary non-linear models (e.g. logit and probit regressions) are unsuited for analysing educational mismatches, as a worker can be undereducated[4], adequately educated or overeducated. Similarly, results based on OLS regressions may face under- or over-estimation issues, as these regressions impose a linear relationship between origin and overeducation and do not guarantee that every predicted probability stands between 0 and 1. Therefore, multiple-choice non-linear models (e.g. ordered or multinomial logit regressions) must be used to avoid misspecification issues. Third, none of the existing studies uses interactions between origin and gender in the regression analysis to explore a potential double penalty for female immigrants across generations (i.e. being penalised due to gender and foreign background).

Before delving into the details of our research, it is worth noting that we focus exclusively on tertiary-educated workers because they are, *per se,* individuals who can be reasonably at risk of being overeducated. Indeed, considering all individuals in a study on overeducation could lead to misleading conclusions, as workers holding at most a secondary diploma are much less likely to be overeducated. Several papers follow a similar strategy (e.g. Nielsen, 2011; Nugent, 2022; Shi et al., 2022). In addition, we also devote particular attention to immigrants from developing countries[5], as a large body of literature shows that F-G immigrants born in developed countries and their S-G counterparts appear not to represent a significant integration concern for Western societies (e.g. Abramitzky et al., 2021; Algan et al., 2010; Fays et al., 2021).

Using a granular, matched employer-employee database for Belgium between 1999 and 2016, containing around 400,000 tertiary-educated workers, we aim to contribute to the literature with one of the first empirical studies on the intergenerational interplay between overeducation and origin. The novelty of our database is that it contains information on workers' country of birth and that of their parents, which allows us to identify F-G immigrants and their S-G peers correctly. The rich detail of our database also allows us to

[4] Workers are considered undereducated if their educational attainment is lower than that required for their job. Undereducation may notably result from periods of labour shortages (i.e. bottleneck vacancies) and technologically induced changes in job content and complexity.

[5] By 'developing countries', we mean either transition and developing countries listed in the United Nations' (2020) classification and/or emerging market and developing economies listed in the IMF's (2020) classification. See Appendix 1.

identify overeducation cases across more than 13,500 occupation-age-sector cells using a realized matches approach. Moreover, we also have access to the 2021 Labour Force Survey for Belgium and its ad-hoc module on 'Migration and labour market'. This complementary data allows us to improve the quality of our paper using more detailed statistics. From an econometric viewpoint, as far as we know, our paper is the first to employ generalized ordered logit (GOLOGIT) regressions in the context of educational mismatches for the immigrant population. The main advantage of this econometric technique is that it relaxes the parallel regression assumption (i.e. the primary condition of an ordered logit regression) in the coefficients that do not meet it. The estimates of GOLOGIT regressions are also more parsimonious and interpretable than those of multinomial logit regressions (Williams, 2016; Williams and Quiroz, 2020).

Our empirical strategy starts with a GOLOGIT regression where, conditioning on a wide range of covariates (e.g. worker, employment and firm characteristics), we estimate the likelihood of being overeducated for immigrants from developing countries. Using a more fine-grained classification, we further explore how geographical origins[6] (e.g. the Maghreb, Sub-Saharan Africa and the Near and Middle East) shape the likelihood of being overeducated for immigrants from developing countries. We follow this approach as source-country characteristics at the individual level (e.g. patronymic, physical appearance, religion and cultural manners) and the macro level (e.g. economic and political stability, the quality of the education system and reasons for migration in the source country) can shape the integration of immigrants (Fleischmann and Dronkers, 2010; Levels and Dronkers, 2008). Several studies on immigrants' employment and earnings outcomes accord with this premise (e.g. Athari et al., 2019; Lindley, 2009; Pineda-Hernández et al., 2022a; Piton and Rycx, 2021).

The contributions of our paper also extend to the analysis of two moderating variables (i.e. gender and part-time work) in the intergenerational relationship between overeducation and origin. There are theoretical and empirical justifications for this additional analysis. Regarding the role of gender, when female workers become mothers, they may revisit their labour market trajectories or employment conditions to deal with motherhood and childcare

[6] For the sake of accuracy in correctly classifying immigrants by geographical origin and economic development level, we constructed our geographical classification of countries based on the United Nations' (2020) classification and the IMF's (2020) classification. See Appendix 2.

(i.e. accepting jobs that do not often match their level of education but allow them to spend more time with their children or leaving the labour market) (Kifle et al., 2014; Petrongolo, 2019). The European Commission (2019) further reveals that in the EU, tertiary-educated women are still underrepresented in professional or managerial occupations (i.e. potential overeducation issues). Recent papers underpin these statements as they find evidence of a gender overeducation gap in Europe (e.g. Nugent, 2022; Santiago-Vela and Mergener, 2022).

Moreover, traditional values and ethnic behaviour (e.g. family hierarchy and fertility and partner choices) can affect the working time decisions of F-G female immigrants, as they tend to be more involved in informal childcare and household production than natives (Baudin and Kondi, 2023; Ferrer and Mascella, 2022). Put differently, F-G female immigrants may find more difficulties than female natives in balancing their professional careers and ethnic identity (Blau et al., 2013; Fernández and Fogli, 2009), leading to a higher likelihood of being overeducated. There is also evidence that the influence of traditional values and ethnic norms extends to S-G female immigrants (Biegel et al., 2016; Maes et al., 2023; Noghanibehambari et al., 2022). Indeed, Jacobs et al. (2022c) point out that S-G female immigrants still face discrimination and prejudices in the workplace due to the construction of professional identities that match their migration background. However, these statements are not fully supported by the few papers that analyse the moderating role of gender in the relationship between origin and overeducation (e.g. Falckle et al., 2020; Jacobs et al., 2021; McGuinness and Byrne, 2015).

Regarding the role of part-time work, Davia et al. (2017) and Wen and Maani (2018) state that the level of education required for part-time jobs rarely corresponds to that of workers, as they may have chosen them for family reasons or personal preferences rather than career aspirations. Moreover, it could be argued that F-G immigrants may struggle to find full-time jobs due to poor human capital, imperfect transferability of qualifications or discrimination (Zschirnt and Ruedin, 2016), leading them to accept part-time jobs in a disadvantageous position (i.e. involuntary part-time work). Similarly, if employers present ethnic stereotypes or preferences for natives when recruiting full-time workers, an employment selection could also be applied to S-G immigrants. The findings of Fernández-Reino et al. (2018) and Green et al. (2007) go in this direction, suggesting that people with a migration background are more likely to be in a part-time job than natives. Therefore, since immigrants are more likely to be

in part-time jobs, it is interesting to investigate whether immigrants working part-time also experience more educational mismatches than natives working part-time.

Last but not least, several papers show that women are more likely to work part-time than men (e.g. de Quinto et al., 2020; McIntosh et al., 2012; Piton, 2022). According to Fernández-Kranz et al. (2013) and Kifle et al. (2014), this overrepresentation of women in part-time jobs is mainly driven by mothers. Furthermore, Piton (2022) highlights that the size effect of motherhood on the likelihood of working part-time significantly depends on women's foreign background. Based on these premises, we also investigate how part-time work affects the likelihood of being overeducated for female and male immigrants from developing countries across two generations. To achieve this goal, we re-estimate our GOLOGIT regression using a three-way interaction (gender, part-time work and origin).

Our paper shows that, unlike their F-G peers, S-G immigrants from developing countries experience the same likelihood of being overeducated as natives (i.e. immigrant overeducation disappears across two generations). This finding applies to all geographical groups, with the exception of S-G immigrants from the Maghreb, who are somewhat more likely to be overeducated than natives. Our gender-interacted estimates suggest that F-G female immigrants born in developing countries face a double penalty in their likelihood of being overeducated due to their gender and migration background. In contrast, S-G female immigrants from developing countries perform on par with female natives (i.e. only evidence of a gender penalty in their likelihood of being overeducated). Moreover, part-time work is positively associated with immigrant overeducation, and this relationship largely persists across two generations. Finally, we find that immigrant overeducation is a persistent intergenerational issue when workers work part-time, irrespective of gender.

2. Belgian context

Belgium is one of the developed countries with the largest immigrant population. According to the 2021 Labour Force Survey, F-G immigrants accounted for 21.4% of the total population aged 20-64 in Belgium, while S-G immigrants represented 13.3%. Among this immigrant population, 44.7% had an EU background, while 55.3% came from non-EU countries (primarily immigrants from developing countries). This demographic context makes

Belgium an interesting case study to assess the labour market integration of immigrants across generations.

In the developed world, there is a large consensus regarding the positive role of education in boosting employment and wages (OECD, 2022). However, holding a tertiary diploma in Belgium seems more profitable for natives than for people with a migration background. Indeed, tertiary-educated people born in developing countries experienced a sizeable immigrant-native employment gap in 2018, ranging between 17% and 30% points according to their geographical origin (FPS Employment and Unia, 2022). Moreover, several papers show that this integration issue extends to S-G immigrants (e.g. Corluy et al. 2015; De Cuyper et al. 2018). Piton and Rycx (2021) also find that F-G female immigrants born in developing countries and their same-gender descendants experience a double penalty in accessing the Belgian labour market.

Once in employment, F-G immigrants born in developing countries earn less than natives and face wage discrimination (e.g. Fays et al. 2021; Kampelmann and Rycx 2016; Grinza et al. 2020). From an intergenerational perspective, Pineda-Hernández et al. (2022) show that the wages of S-G immigrants from developing countries are similar to those of their F-G peers, being substantially lower than those of natives. The authors also find evidence of a significant double penalty in the wages of female immigrants from developing countries across two generations. However, controlling for observables at the firm level, Pineda-Hernández et al. (2022) show that the adjusted wage gap (e.g. wage discrimination) for immigrants from developing countries disappears across two generations (from 2.7% to nil). In terms of employment conditions, Jacobs et al. (2021) show that F-G immigrants born in developing countries are much more likely to be overeducated than natives. Their estimates also suggest that F-G female immigrants born in developing countries face a similar likelihood of being overeducated as their F-G male peers (i.e. no evidence of a double penalty). Nevertheless, as far as we know, intergenerational migration issues related to overeducation remain unaddressed in Belgium.

3. Methodology

3.1. Realized matches approach

Given the structure of our data (i.e. a matched employer-employee database), we follow a realized matches (RM) approach (also known as a statistical approach) to measure overeducation.[7] The RM approach uses workers' distribution of levels of education (ISCED: 7 categories) to calculate the modal value within each occupational group (i.e. workers' educational attainment that repeats the most in an occupation) (Kiker et al., 1997; Sellami et al., 2018; Verdugo and Verdugo, 1989).[8] Then, the modal value is used as a reference to identify educational mismatches (i.e. undereducation or overeducation).[9] For instance, a taxi driver is overeducated if she holds a bachelor's degree, whereas most taxi drivers only hold a secondary diploma.

However, it should be noted that the level of education required in an occupation may vary over workers' careers (i.e. age cohort effects) (Lindley, 2009). Put differently, as young workers' skills increase over time, the probability of being overeducated for old workers mechanically increases. In this regard, Voets (2022) shows that in the EU, while overeducation decreases over time for the youngest age groups, an opposite path is observed for the oldest ones. Moreover, the possibility of occupational skill upgrading or downgrading across sectors should also be considered (i.e. sector cohort effects). For instance, while the required level of education for a consultant in the real estate sector is a bachelor's degree, a master's degree is needed for the same occupation in the banking sector. Therefore, in order to minimize a potential bias related to these cohort effects, we sort workers by occupation (ISCO classification at a three-digit level), age group (five categories)[10] and sector (NACE classification at a two-digit level). Then, we identify educational mismatches using the modal

[7] Two other approaches are also available in the literature to measure overeducation: i) a job analysis approach that gauges overeducation by occupation based on analysts' criteria and ii) a worker self-assessment approach that uses surveys to ask workers if they consider or not themselves in a situation of overeducation given their current educational attainment. However, our data do not contain the necessary information to implement them.
[8] Since firms' human resources departments have provided information on workers' educational attainment to the construction of our dataset, this might be somewhat underestimated for F-G immigrants who obtained their educational qualifications abroad. Therefore, the likelihood of being overeducated for F-G immigrants should be considered as a lower bound.
[9] In an RM approach, the mean by occupation can also be used as a reference. However, the mean is very likely to be influenced by outliers.
[10] We consider the following age groups: 20-24, 25-29, 30-39, 40-49 and > 50.

value of the level of education in each occupation-age-sector cell.[11] Our granular employer-employee database allows us to identify 13,628 occupation-age-sector cells.[12] Therefore, the realized matches approach of our paper is notably more precise than those implemented in previous studies on immigrant overeducation.

3.2. Generalized ordered logit regressions

Ordered logit (OLOGIT) regressions and multinomial logit (MLOGIT) regressions are the two main econometric methods used in the literature to estimate educational mismatches (e.g. the probability of being under- or over-educated for a job). The choice between these two models depends on the possibility of establishing a ranking for the ordinal dependent variable. Initially, assuming an unequivocal order for educational mismatches (1 – undereducated, 2 – adequately educated, and 3 – overeducated), OLOGIT regressions are the right econometric choice. However, OLOGIT regressions depend on the parallel regression assumption (i.e. the effect of any explanatory variable is consistent or proportional across the different categories), which is often violated (Williams and Quiroz, 2020).[13] An alternative solution is to leverage MLOGIT regressions, which are well-suited to estimate likelihoods without the need to satisfy the parallel regression assumption. However, MLOGIT regressions depend on the assumption of independence of irrelevant alternatives (IIA) (i.e. the characteristics of one particular category do not impact the relative probabilities of choosing other categories), which is unlikely to hold in the analysis of educational mismatches.[14]

Against this background, Williams (2016) states that "*generalized ordered logit (GOLOGIT) regressions can fit estimates that are less restrictive than OLOGIT regressions, whose parallel regression assumption is often violated, but more parsimonious and interpretable*

[11] When several modes are identified within an occupation-age-sector cell, the minimum mode is used as the unique reference.
[12] The distribution of occupation-age-sector cells by size is as follows: 7.3% have less than ten observations, 38.7% have between 10 and 100 observations, 45.2% have between 101 and 1000 observations, and 8.8% have more than 1000 observations.
[13] The null hypothesis of the Brant test for parallel regression assumption specifies that the coefficients should be the same in each cumulative logistic regression. If the null hypothesis of this test is rejected, OLOGIT estimates can be highly misleading. Using our database, we reject the null hypothesis of the Brant test. The results of the test can be obtained on request.
[14] The null hypothesis of the Hausman test for the IIA assumption states that there is no systematic change in the coefficients if a category of the dependent variable is excluded from the model. If the null hypothesis of this test is rejected, the disturbances of the categories are not independent. Using our database, we reject the null hypothesis of the IIA assumption. The results of the test can be obtained on request.

that those fitted by MLOGIT regressions". Moreover, GOLOGIT regressions present an additional advantage compared to other non-linear models. They allow the parallel regression assumption to be relaxed for estimates that do not satisfy it while imposing it on those that do (i.e. estimating partial proportional odds) (Williams, 2006). Therefore, we use GOLOGIT regressions to assess the intergenerational nexus between overeducation and origin among tertiary-educated workers in Belgium. Our benchmark GOLOGIT regression is written as follows:

$$P(Y_{it} > j) = g(\beta_j X) = \frac{\exp\left(\alpha_j + \sum_{k=1}^{4} \beta_{jk} Origin_{itk} + \underline{z}_{it}^{T} \underline{\vartheta}_j + \underline{g}_{it}^{T} \underline{\lambda}_j + \underline{f}_{it}^{T} \underline{\xi}_j + \delta_{jt}\right)}{1+\left\{\exp\left(\alpha_j + \sum_{k=1}^{4} \beta_{jk} Origin_{itk} + \underline{z}_{it}^{T} \underline{\vartheta}_j + \underline{g}_{it}^{T} \underline{\lambda}_j + \underline{f}_{it}^{T} \underline{\xi}_j + \delta_{jt}\right)\right\}} \quad (1)$$

From equation (1), it can be established that the coefficients and fixed effects (summarised with the letter ϕ) associated with our regressors (X) vary for each ordered category j.[15] Therefore, the probability that a worker i at time t within an occupation-age-sector cell will be under-, adequately or over-educated is respectively equal to:

$$P(Y_{it} = 1 = undereducated) = 1 - g(\phi_1 X)$$

$$P(Y_{it} = 2 = adequately\ educated) = g(\phi_1 X) - g(\phi_2 X)$$

$$P(Y_{it} = 3 = overeducated) = g(\phi_2 X)$$

As we are interested in the overeducation of workers, GOLOGIT regressions estimate the probability of being in the category 3 (i.e. overeducated) compared to being in a lower category (i.e. under- or adequately educated). It should also be noted that we use the average marginal effects of GOLOGIT regressions to facilitate the interpretation of non-linear estimations.

In equation (1), our main explanatory variable is '$Origin_{itk}$', which categorizes workers into five groups k: i) natives, i.e. workers born in Belgium with both parents born in Belgium (the

[15] In equation (1), α_j represents the threshold parameter for each of the j ordered categories.

reference group), ii) F-G immigrants born in developing countries[16], iii) S-G immigrants from developing countries (i.e. workers born in Belgium with at least one parent born in a developing country)[17], iv) immigrants from developed countries, excluding Belgian natives[18], and v) others (i.e. workers born abroad with both parents born in Belgium)[19] (see Appendix 1 for a chart of developed and developing countries).

We also introduce an extensive range of covariates in our benchmark GOLOGIT regression to reduce a potential omitted variable bias. In order to facilitate the presentation of our covariates and their corresponding coefficients in equation (1), they are written as vectors in the following manner:

$$\underline{\mathbf{z}_{it}} = (z_{it1}, \dots, z_{itM})^{T} \qquad \underline{\boldsymbol{\vartheta}_j} = (\vartheta_{j1}, \dots, \vartheta_{jM})^{T}$$

$$\underline{\mathbf{g}_{it}} = (g_{it1}, \dots, g_{itL})^{T} \qquad \underline{\boldsymbol{\lambda}_j} = (\lambda_{j1}, \dots, \lambda_{jL})^{T}$$

$$\underline{\mathbf{f}_{it}} = (f_{it1}, \dots, f_{itQ})^{T} \qquad \underline{\boldsymbol{\xi}_j} = (\xi_{j1}, \dots, \xi_{jQ})^{T}$$

where $\underline{\mathbf{z}_{it}}$ is a M x 1 vector of observations on worker characteristics (i.e. gender, tenure, squared tenure, level of tertiary education and type of household); $\underline{\mathbf{g}_{it}}$ is a L x 1 vector of observations on employment characteristics (i.e. type of contract and dummies for part-time and overtime work); $\underline{\mathbf{f}_{it}}$ is a Q x 1 vector that contains observations on the firm where the worker is employed (i.e. size of the firm, region where the firm is located and dummies for

[16] See footnote 6.

[17] Regarding S-G immigrants from developing countries, it should be noted that their origin has been firstly determined by the father's country of birth, except if the father was born in Belgium and the mother in a developing country. In that case, the mother's country of birth has been retained.

[18] F-G immigrants born in developed countries and their descendants present similar labour market outcomes to those of Belgian natives (e.g. Piton and Rycx, 2021; Pineda-Hernandez et al., 2022). Therefore, as they do not represent an intergenerational integration issue for Belgium, we merge F-G immigrants born in developed countries and their descendants into a unique group.

[19] The category 'others' was created because workers born in developing countries with both parents born in developed countries earn more than any other group of origin (see Pineda-Hernández et al. (2022)) and have better worker and employment characteristics than immigrants from developing countries (see Appendix 3). One might expect these workers to be the children of expatriates, who are generally highly educated and have a high socio-economic level. Consequently, considering these workers as F-G immigrants born in developing countries could lead to misleading conclusions.

the existence of a firm-level collective agreement and the type of economic and financial control); and δ_{jt} represents year fixed effects.

In a subsequent specification, we split immigrants from developing countries into six geographical groups: i) Sub-Saharan African countries, ii) the Maghreb countries, iii) the Near and Middle Eastern countries, iv) non-EU Eastern European countries, v) emerging and developing Asian countries, and vi) Latin American and Caribbean countries (see Appendix 2 for a list of countries by geographical region). We aim to assess the effects of more fine-grained characteristics associated with workers' region of birth or that of their parents (e.g. patronymic, physical appearance, type of religion, political stability of the region and quality of higher education systems in the destination countries) on the likelihood of being overeducated for immigrants from developing countries.

Moreover, we investigate the moderating role of gender and part-time work in the intergenerational relationship between overeducation and origin in order to reveal potential immigrant-native differences that our benchmark specification may mask. To achieve this goal, we re-estimate equation (1) with an explanatory variable that varies according to origin and each moderator. Unlike gender-segregated regressions, this approach allows us to estimate potential double penalties resulting from being a woman with a foreign background or an immigrant in a part-time job. Finally, we simultaneously test the role of gender and part-time in the likelihood of being overeducated for immigrants from developing countries by re-estimating equation (1) with an explanatory variable that varies according to origin, gender and part-time work (i.e. a three-way interaction).

4. Data and descriptive statistics

Our empirical strategy uses a matched employer-employee database for the Belgian labour market between 1999 and 2016, provided by Statistics Belgium (STATBEL). This database was obtained by merging two datasets: the Structure of Earnings Survey (SES) and administrative data from the Belgian National Register (BNR). The SES covers all firms operating in Belgium with economic activities defined by the NACE nomenclature. Based on a sophisticated stratified sampling design, the SES provides a nationally representative sample of workers in the Belgian labour market.

The SES also contains granular information on the structural characteristics of the firms (e.g. sector of activity, number of employees and type of collective agreement) and the demographic and employment characteristics of workers (e.g. age, gender, educational attainment, tenure, occupation and type of contract). Regarding the administrative data from the BNR, they provide reliable information on workers' country of birth, that of their parents and the type of household they live. The original sample contains information on 1,604,835 workers employed in 20,375 firms.

After calculating educational mismatches by occupation-age-sector cell (see Section 3.1 for more details), two filters were applied to the original sample. First, we restricted our sample to tertiary-educated workers to focus exclusively on the population for which overeducation is more likely to be an issue (1,147,473 observations deleted).[20] Second, to avoid misclassification issues in the design of groups by origin and generation, we dropped workers born in Belgium but with missing information on the country of birth of at least one of their parents (33,826 observations deleted).[21] Therefore, the final sample consists of 396,462 workers employed in 15,628 firms.

Moreover, STATBEL has also given us access to the 2021 Labour Force Survey for Belgium and its ad-hoc module on '*Migration and labour market'*. This dataset contains around 22,000 observations and is representative of the working-age population in Belgium. We use descriptive statistics from this dataset as a complementary information source to better understand our results.

Appendix 3 shows the statistical profiles of tertiary-educated workers by origin and generation.[22] 78.9% of workers in our sample are Belgian natives, while F-G immigrants born in developing countries and their descendants represent 3.7% and 3%, respectively.

[20] Using our database, we find that within the cohort of workers with at most upper secondary education, F-G immigrants born in developing countries and their S-G peers are just as likely to be overeducated as natives. Those estimates can be obtained on request.

[21] Around 8% of tertiary-educated workers born in Belgium were dropped because of missing information on their parental country of birth. Our estimates show that these workers face a 2% points higher likelihood of being overeducated than natives. Therefore, if we assume that a substantial part of these workers born in Belgium have at least a foreign-born parent, our findings on the overeducation gap between natives and S-G immigrants must be interpreted as a lower bound.

[22] To shorten the term that describes tertiary-educated workers, tertiary-educated natives and tertiary-educated immigrants in the remainder of this paper, we refer to them simply as workers, natives and immigrants, respectively.

Moreover, the groups 'immigrants from developed countries' and 'others' constitute 13.3% and 1.1% of our sample, respectively. Within the cohort of immigrants from developing countries, we observe that around 3 in 4 workers are geographically originating from Sub-Saharan Africa and the Maghreb, which accords with the total distribution of the working-age population in Belgium by geographical origin between 2008 and 2016 (FPS Employment and Unia, 2019).

5. Results

5.1 Benchmark scenario

To investigate the intergenerational relationship between origin and overeducation, Table 1 shows the marginal effects of our benchmark GOLOGIT regression (see equation (1)). We find that the likelihood of being overeducated for a worker increases by 10.5% points if the worker was born in a developing country. In other words, F-G immigrants born in developing countries are much more likely to be overeducated than natives, whose overeducation incidence is 43.5% (see the incidences of overeducation by origin and generation in Appendix 4).[23] We cannot ultimately assert whether this labour market disadvantage for F-G immigrants born in developing countries is completely related to their origin or unobserved individual heterogeneity (e.g. motivation, language proficiency and organization skills). However, given the granular realized matches approach used in the identification of overeducation and the large number of covariates included in our GOLOGIT regression, we feel confident to attribute this finding at least partially to a penalty related to immigrants' origin.

[Insert Table 1 about here]

Moreover, it should be noted that the estimates of standard econometric models go in the same direction as those of our GOLOGIT regressions, although their magnitude differs (Appendix 5). For instance, OLS estimates suggest that F-G immigrants born in developing countries face a 12.1% points higher likelihood of being overeducated than natives, whereas OLOGIT estimates quantify this penalty in 9.4%. However, these standard econometric

[23] Our results in Table 3 also show that immigrants from developed countries are 1.3% less likely to be overeducated than natives, and the group 'others' is just as likely to be overeducated as natives.

models face significant constraints (e.g. linearity and parallel regression assumptions), which may lead to biased estimates. Therefore, our GOLOGIT regressions provide more precise estimates by dealing with these constraints.

Turning to S-G immigrants from developing countries, we find no significant relationship between their origin and their likelihood of being overeducated. Put differently, S-G immigrants from developing countries are just as likely to be overeducated as natives. Therefore, our GOLOGIT estimates align, on average, with the classical assimilation theory, which suggests that S-G immigrants from developing countries are treated as natives in the recruitment or promotion process.

The following assumptions can explain the disappearance of immigrant overeducation across two generations: i) S-G immigrants are likely to have good proficiency in at least one of the languages used in the Belgian labour market (i.e. French, Dutch or English), as they completed their whole education in Belgium; ii) employers may consider S-G immigrants as natives since they have accumulated human capital specific to Belgium (e.g. tertiary diplomas from Belgian universities or internships in firms located in Belgium); and iii) S-G immigrants have probably built better social networks than their parents due to a process of socialization from childhood to adulthood, which can be helpful in their job search and progress along the job ladder.

As workers' origin does not exclusively explain their likelihood of being overeducated, we also find that the coefficients associated with our covariates are statistically significant and have the expected signs. These complementary results further support our choice of moderating variables. Specifically, being a woman increases the likelihood of being overeducated for a worker by 5% points. Similarly, part-time workers are 15.6% points more likely to be overeducated than full-time workers.

[Insert Table 2 about here]

Using a more fine-grained geographical classification in equation (1), Table 2 reports the marginal effects of our GOLOGIT estimates regarding the intergenerational relationship between geographical origin and overeducation. We find that F-G immigrants born in the Maghreb are the geographical group that experiences the highest likelihood of being overeducated (18.0% points higher than natives). To a lesser extent, F-G immigrants born in Sub-Saharan Africa are also more likely to be overeducated than natives by 12.1% points. Regarding F-G immigrants born in the Near and Middle East and non-EU Eastern Europe, their likelihood of being overeducated is around 9.0% points higher than that of natives. In contrast, F-G immigrants born in emerging and developing Asia, Latin America and the Caribbean are just as likely to be overeducated as natives.

Turning to S-G immigrants from developing countries, we find three intergenerational patterns. First, immigrant overeducation vanishes across two generations for workers from the Near and Middle East and non-EU Eastern Europe. In addition, in line with their F-G peers, S-G immigrants from emerging and developing Asia, Latin America and the Caribbean perform on a par with natives (i.e. no immigrant overeducation across two generations for these geographical groups). Second, although S-G immigrants from the Maghreb fare better than their F-G peers, their likelihood of being overeducated remains 3.2% points higher than that of natives. Third, S-G immigrants from Sub-Saharan Africa experience a 2.0% points lower probability of being overeducated than natives, thus reversing the initial penalty faced by their F-G peers.

5.2 Gender and immigrant overeducation

In this Subsection, we investigate the moderating effect of gender on the intergenerational nexus between origin and overeducation. Results are displayed in Table 3. Our gender-interacted estimates suggest that while F-G male immigrants born in developing countries experience a 14.5% points higher likelihood of being overeducated than male natives, S-G male immigrants from developing countries are just as likely to be overeducated as male natives. These estimates go in the same direction as those of our benchmark GOLOGIT regression in Table 1, where immigrant overeducation also disappears across two generations.

[Insert Table 3 about here]

However, when it comes to female workers, the findings are more nuanced. More precisely, our gender-interacted estimates suggest that female natives experience a 5.6% points higher likelihood of being overeducated than male natives (i.e. evidence of a gender overeducation gap). We also find that F-G female immigrants born in developing countries are 5.5% points more likely to be overeducated than female natives (i.e. F-G female immigrants born in developing countries experience an 11.1% points higher likelihood of being overeducated than male natives). In addition, the coefficient associated with F-G female immigrants born in developing countries is statistically different from that associated with female natives (see the test for equality of coefficients at the bottom of Table 3). Thus, there are solid grounds for suggesting that F-G female immigrants born in developing countries undergo a double penalty in their likelihood of being overeducated due to their gender and foreign background.

Regarding S-G female immigrants from developing countries, our gender-interacted estimates suggest that they no longer face a penalty in their likelihood of being overeducated due to their origin. More precisely, S-G female immigrants from developing countries are 1.8% points less likely to be overeducated than female natives (i.e. S-G female immigrants from developing countries undergo a 3.8% points higher likelihood of being overeducated than male natives). However, although S-G female immigrants from developing countries seem to perform better than female natives, it is worth noting that their coefficient associated with their origin is not statistically different from that for female natives (see the test for equality of coefficients at the bottom of Table 3). Therefore, S-G female immigrants from developing countries are just as likely to be overeducated as female natives (i.e. only evidence of a gender penalty in their likelihood of being overeducated).

5.3 Part-time work and immigrant overeducation

Column (5) in Appendix 5 shows that in our sample, about 60% of natives in part-time jobs are overeducated. In addition, this incidence is higher if the worker has a foreign background (80.1% for F-G immigrants born in developing countries and 71.7% for their S-G peers). Thus, we study the moderating role of part-time work in the intergenerational nexus between origin and overeducation. Results are shown in Table 4. We find that within the cohort of full-time workers, F-G immigrants born in developing countries are 9.4% points more likely to be overeducated than natives. In contrast, S-G immigrants from developing countries working

full-time perform similarly to natives working full-time in terms of overeducation. These findings largely mirror those of our benchmark GOLOGIT regression in Table 1.

[Insert Table 4 about here]

Moreover, our estimates suggest that natives working part-time face a 14.1% points higher likelihood of being overeducated than natives working full-time. However, it should be noted that working part-time affects more immigrants than natives in terms of overeducation. More precisely, F-G immigrants born in developing countries working part-time are 25.7% points more likely to be overeducated than natives working part-time (i.e. F-G immigrants born in developing countries working part-time face a 39.8% higher likelihood of being overeducated than natives working full-time). Although this overeducation gap between natives and immigrants reduces across two generations, it remains sizeable. In other words, S-G immigrants from developing countries working part-time are 12.4% points more likely to be overeducated than natives working part-time (i.e. S-G immigrants from developing countries working part-time experience a 26.5% higher likelihood of being overeducated than natives working full-time).

However, caution must be exercised when interpreting these results. Indeed, the relationship between part-time work and overeducation might suffer from endogeneity (e.g. reverse causality). Although 2SLS regressions could tackle this econometric issue, it remains challenging to find valid instruments, namely variables that are both relevant (i.e. goods predictors of part-time work) and exogenous (i.e. uncorrelated with the error of being overeducated). In addition, our potential endogenous variable is represented by nine categories in the regression (i.e. the interactions between origin and part-time work), which implies the search of at least nine instruments to guarantee identification in 2SLS regressions (i.e. having as many instruments as endogenous variables to satisfy one of 2SLS regressions' conditions). Accordingly, although our estimates provide solid evidence of a relationship between part-time work and immigrant overeducation, they should not be interpreted as causal.

5.4 Three-way interaction: origin, gender and part-time work

The literature suggests that part-time work is mainly an issue for women because they are more likely to work part-time due to motherhood and home production (e.g. Kifle et al., 2014; Piton, 2022). Our LOGIT estimates accord with this statement, as they suggest that the likelihood of working part-time for female workers significantly increases when they have children at home: by 7.1% points for female natives, by 7.2% points for F-G female immigrants born in developing countries and by 3.9% points for S-G female immigrants from developing countries (see Appendix 6). In contrast, we find no significant effect of parenthood on the likelihood of working part-time for male workers, irrespective of origin and generation. These results remain largely stable after controlling for a large set of covariates. Accordingly, we implement a three-way interaction (origin, part-time work and gender) in our GOLOGIT regression. Results are presented in Table 5.

[Insert Table 5 about here]

Within the cohort of full-time female and male workers, our findings largely align with the gender-interacted estimates in Table 3. However, focusing on part-time workers, our three-way interaction reveals interesting new outcomes. First, we find that female natives working part-time are 3.3% points more likely to be overeducated than their male peers working part-time, who already face a 15.7% points higher likelihood of being overeducated than male natives working full-time.

Second, being born in a developing country skyrockets the likelihood of being overeducated for a part-time worker, regardless of gender. More precisely, F-G male immigrants born in developing countries working part-time are 29.2% points more likely to be overeducated than male natives working part-time (i.e. F-G male immigrants born in developing countries working part-time face a 44.9% higher likelihood of being overeducated than male natives working full-time). Similarly, F-G female immigrants born in developing countries working part-time are 22.3% points more likely to be overeducated than female natives working part-time (i.e. F-G female immigrants born in developing countries working part-time experience a 35.9% higher likelihood of being overeducated than female natives working full-time). In this regard, it should be noted that there is no significant difference between the likelihood of

being overeducated for F-G male immigrants born in developing countries and that of their F-G female peers (see the test for equality of coefficients at the bottom of Table 5).

Third, immigrant overeducation is a persistent intergenerational phenomenon within the cohort of part-time female and male workers. More precisely, S-G male immigrants from developing countries working part-time are 21.6% points more likely to be overeducated than male natives working part-time (i.e. S-G male immigrants from developing countries working part-time undergo a 37.3% higher likelihood of being overeducated than male natives working full-time). Similarly, S-G female immigrants from developing countries working part-time are 8% points more likely to be overeducated than female natives working part-time (i.e. S-G female immigrants from developing countries working part-time face a 27% higher likelihood of being overeducated than female natives working full-time). In addition, the test for equality of coefficients at the bottom of Table 5 shows that S-G male immigrants from developing countries working part-time fare worse than their S-G female peers working part-time in terms of overeducation.

The findings associated with our three-way interaction are in line with the literature, which suggests that family duties (e.g. informal childcare and home production) and ethnic norms (e.g. fertility rate and family hierarchy) affect the labour market expectations of female immigrants more than those of female natives (Blau et al., 2013; Fernández and Fogli, 2009; Jacobs et al. 2022b). However, it seems less clear that male immigrants revisit the labour market trajectories due to parenthood or ethnic identities (Nadim and Midtbøen, 2023). Therefore, other underlying mechanisms linking part-time work, origin and overeducation among male workers need to be considered. For instance, an excess supply of labour from male immigrants, which could foster employers' incentive to offer them jobs below their level of education or generate additional barriers preventing them from finding full-time jobs (i.e. jobs where the likelihood of being overeducated is less substantial).

Incidentally, it should be noted that descriptive statistics of the 2021 ad-hoc module of the Labour Force Survey (LFS) concerning migration in Belgium seem to be supportive of this interpretation. Appendix 7 shows indeed that in 2021, the labour market participation rate of F-G male immigrants born in developing countries was more than 20% points higher than that of their F-G female peers. Moreover, this gender gap appears to be intergenerational persistent

among immigrants (i.e. the gender gap in participation rates of S-G immigrants from developing countries is still as high as 14% points).

Next, Appendix 8 reveals that in 2021, the involuntary part-time employment rate of F-G male immigrants born in developing countries was substantially higher than that of natives (57.4% vs. 15.4%). This employment issue extends to S-G male immigrants from developing countries, whose involuntary part-time employment rate was 45.6%. However, when it comes to female immigrants, involuntary part-time work seems less problematic. Indeed, 20% of F-G female immigrants born in developing countries and 21.7% of their S-G female peers declared themselves in an involuntary part-time job. Thus, the additional difficulties immigrants face to find full-time jobs are likely to explain their extremely high likelihood of being overeducated when working part-time.

Finally, we cannot exclude an employment selection according to female and male workers' origin (i.e. taste-based, statistical or monopsonistic discrimination) as a potential channel to explain the intergenerational connection between part-time work and immigrant overeducation. For instance, it might be possible that when hiring tertiary-educated workers, some employers have preferences for natives or negative stereotypes against immigrants.

6. Conclusion

In the OECD area, the population with tertiary education increased from 26.1% in 2000 to 47.1% in 2020 (OECD, 2022). However, although educational expansion accords with the growing demand for tertiary-educated people in developed countries, it could also lead to many overeducation cases (e.g. an economist employed as a cashier in a supermarket) (Green and Henseke, 2016). Indeed, overeducation has become a persistent social and economic phenomenon in the developed world (McGuinness et al., 2018; Nugent, 2022). Moreover, the literature shows that overeducation rates are higher in developed countries with more immigrant labour because F-G immigrants are more likely to be overeducated than natives (e.g. Davia et al., 2017; Jacobs et al., 2020; Wen and Maani, 2018). However, the evolution of immigrant overeducation across generations remains largely unexplored. Put differently, very little is known about the overeducation of S-G immigrants (e.g. Falcke et al., 2020; Fernández-Reino et al., 2018). Therefore, we leverage rich employer-employee data, covering almost two decades (1999-2016), a granular realized matches approach and GOLOGIT

regressions to investigate the intergenerational interplay between origin and overeducation among tertiary-educated workers in Belgium.

After accounting for differences in worker, employment and firm characteristics (e.g. gender, tenure, level of tertiary education, type of contract, overtime, firm size and firm-level collective agreements), our GOLOGIT estimates suggest that F-G immigrants born in developing countries are substantially more likely to be overeducated than natives, who face an important likelihood of being overeducated (43.5%). However, when it comes to S-G immigrants from developing countries, their likelihood of being overeducated is not statistically different from that of natives. In other words, immigrant overeducation disappears across two generations, thus illustrating the positive effect of being born, studying and socialising from childhood to adulthood in the host country (i.e. the classical assimilation theory). Our findings also hold for all geographical groups of immigrants, with the exception of workers from the Maghreb. More precisely, although S-G immigrants from the Maghreb perform much better than their F-G peers, their likelihood of being overeducated remains somewhat higher than that of natives.

Moreover, our gender-interacted estimates suggest that F-G female immigrants born in developing countries experience a double penalty in their likelihood of being overeducated due to their gender and migration background. In contrast, S-G female immigrants from developing countries perform similarly to female natives, who experience a higher likelihood of being overeducated than male natives (i.e. a gender overeducation gap). Put differently, our gender-interacted estimates show that S-G female immigrants from developing countries only face a gender penalty in their likelihood of being overeducated.

However, it should be noted that all these findings only apply to the cohort of full-time female and male workers. Indeed, our GOLOGIT estimates show that part-time work makes immigrant overeducation intergenerationally persistent, in line with the segmented assimilation theory. More precisely, although S-G female and male immigrants from developing countries working part-time are less likely to be overeducated than their F-G same-gender peers, their likelihood of being overeducated is much higher than that of same-gender natives working part-time. The following underlying channels are likely to explain the nexus between part-time work and immigrant overeducation: i) the intergenerational transmission of ethnic and gender norms among female immigrants (e.g. fertility choices,

informal childcare and family hierarchy); ii) the high incidence of involuntary part-time work among male immigrants from developing countries; and iii) an employment selection based on workers' origin (i.e. taste-based, statistical or monopsonistic discrimination).

Overall, our paper provides solid evidence on the disappearance of immigrant overeducation across two generations in Belgium. However, the main finding of our paper highly depends on the type of employment (i.e. working full-time or part-time), which emphasizes the importance of accounting for employment conditions when designing integration policies aimed at promoting the labour market integration of immigrants. Last but not least, our paper also opens promising avenues for further intergenerational research on other barriers that workers of foreign origin may face at the workplace (e.g. mismatches between the job and the field of education and overeducation persistence).

References

Abramitzky, R., Boustan, L., Jácome, E., and Pérez, S. (2021). Intergenerational mobility of immigrants in the United States over two centuries. *American Economic Review*, *111*(2), 580-608.

Akkaymak, G. (2017). A Bourdieuian analysis of job search experiences of immigrants in Canada. *Journal of International Migration and Integration*, *18*(2), 657-674.

Alba, R., Kasinitz, P., and Waters, M. C. (2011). The kids are (mostly) alright: Second-generation assimilation: Comments on Haller, Portes and Lynch. *Social Forces*, *89*(3), 763-773.

Algan, Y., Dustmann, C., Glitz, A., and Manning, A. (2010). The Economic Situation of First and Second-Generation Immigrants in France, Germany and the United Kingdom. *The Economic Journal*, *542*(120), F4-F30.

Athari, E., Brinbaum, Y., and Lê, J. (2019). Le rôle des origines dans la persistance des inégalités d'emploi et de salaire. In *Emploi, chômage, revenus du travail – Edition 2019,* pp. 29–46.

Basilio, L., Bauer, T. K., and Kramer, A. (2017). Transferability of human capital and immigrant assimilation: An analysis for Germany. *Labour*, *31*(3), 245-264.

Baudin, T., and Kondi, K. (2023). Integration Vs Cultural Persistence: Fertility and Working Time among Second-Generation Migrants in France (No. 2023002). Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).

Becker, G. S. (1957). The economics of discrimination. University of Chicago Press, Chicago.

Belfi, B., Allen, J., Jacobs, B., Levels, M., and van Oosterhout, K. (2022). Early career trajectories of first-and second-generation migrant graduates of professional university. *Journal of Ethnic and Migration Studies*, *48*(10), 2415-2435.

Belzil, C., and Poinas, F. (2010). Education and early career outcomes of second-generation immigrants in France. *Labour Economics*, *17*(1), 101-110.

Biegel, N., Wood, J., and Neels, K. (2021). Migrant-native differentials in the uptake of (in) formal childcare in Belgium: The role of mothers' employment opportunities and care availability. *Journal of Family Research*, *33*(2), 467-508.

Blau, F. D., Kahn, L. M., Liu, A. Y. H., and Papps, K. L. (2013). The transmission of women's fertility, human capital, and work orientation across immigrant generations. *Journal of Population Economics*, *26*(2), 405-435.

Brunello, G. and Wruuck, P. (2019) Skill Shortages and Skill Mismatch in Europe: A Review of the literature. IZA Discussion Paper Series No. 12346, Bohn.

Chiswick, B. R., and Miller, P. W. (2009). The international transferability of immigrants' human capital. *Economics of Education Review*, *28*(2), 162-169.

Corluy, V., Haemels, J., Marx, I., and Verbist, G. (2015). The labour market position of second-generation immigrants in Belgium. NBB working paper, No. 285, National Bank of Belgium, Brussels.

Dahlstedt, I. (2015). Over-education amongst the children of immigrants in Sweden. *Nordic Journal of Migration Research*, *5*(1), 36-46.

Davia, M. A., McGuinness, S., and O'Connell, P. J. (2017). Determinants of regional differences in rates of overeducation in Europe. *Social science research*, *63*, 67-80.

De Cuyper, P., Havermans, N., and Vandermeerschen, H. (2018). Labour market outcomes and activation of second generation jobseekers in Belgium (Working paper 2018-10). Leuven: HIVA-KU Leuven.

de Quinto, A., Hospido, L., and Sanz, C. (2020). The child penalty in Spain. Occasional Paper, N° 2017. Banco de Espana, Madrid.

European Commission (2019). 2019 report on equality between women and men in the EU. Directorate-General for Justice and Consumers of the European Commission, Publications Office. https://data.europa.eu/doi/10.2838/395144

Eurostat (2023). Migrant integration statistics. Labour market indicators, Luxembourg.

Falcke, S., Meng, C., and Nollen, R. (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., Mahy, B., Rycx, F., and Volral, M. (2021). Wage discrimination based on the country of birth: do tenure and product market competition matter?. *Applied Economics*, *53*(13), 1551-1571.

Fernández-Kranz, D., Lacuesta, A., and Rodríguez-Planas, N. (2013). The motherhood earnings dip: Evidence from administrative records. *Journal of Human Resources*, *48*(1), 169-197.

Fernández, R., and Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American economic journal: Macroeconomics*, 1(1), 146-177.

Fernández-Reino, M., Radl, J., and Ramos, M. (2018). Employment outcomes of ethnic minorities in Spain: towards increasing economic incorporation among immigrants and the second generation?. *Social Inclusion*, *6*(3), 48-63.

Ferrer, A. M., and Mascella, A. (2022). Immigrant gaps in parental time investments into children's human capital activities (No. 48). Canadian Labour Economics Forum (CLEF), University of Waterloo.

Fleischmann, F., and Dronkers, J. (2010). Unemployment among immigrants in European labour markets: an analysis of origin and destination effects. *Work, employment and society*, 24(2), 337-354.

FPS Employment and Unia (2019), Socio-economic Monitoring 2019: Labour Market and Origin, FPS Employment and Unia Publishing, Brussels.

FPS Employment and Unia (2022), Socio-economic Monitoring 2022: Labour Market and Origin, FPS Employment and Unia Publishing, Brussels.

Green, C., Kler, P., and Leeves, G. (2007). Immigrant overeducation: Evidence from recent arrivals to Australia. *Economics of Education Review*, *26*(4), 420-432.

Green, F., and Henseke, G. (2016). Should governments of OECD countries worry about graduate underemployment?. *Oxford Review of Economic Policy*, *32*(4), 514-537.

Grinza, E., Kampelmann, S. and Rycx, F. (2020). "L'union fait la force? Evidence for wage discrimination in firms with high diversity". *Journal of Economic Inequality,* 18 (2), 181–211.

Hammarstedt, M., and Palme, M. (2012). Human capital transmission and the earnings of second-generation immigrants in Sweden. *IZA Journal of Migration*, *1*(1), 1-23.

IMF (2020), World Economic Outlook: A Long and Difficult Ascent, Washington, DC, p. 142-147.

Jacobs, V., Mahy, B., Rycx, F., and Volral, M. (2021). Over-education among immigrants: the role of demographics, time, and firm characteristics. *Applied Economics*, 53(1), 61-78.

Jacobs, V., Rycx, F., and Volral, M. (2022a). Wage effects of educational mismatch according to workers' origin: the role of demographics and firm characteristics. *De Economist*, 1-43.

Jacobs, C., Van De Mieroop, D., and Van Laar, C. (2022b). A reversed gender bias? Exploring intersectional identity work by Belgian women with a Turkish or Moroccan migration background. In *Globalisation, Geopolitics, and Gender in Professional Communication* (pp. 106-126). Routledge.

Joona, P. A., Gupta, N. D., and Wadensjö, E. (2014). Overeducation among immigrants in Sweden: incidence, wage effects and state dependence. *IZA Journal of Migration*, *3*(1), 1-23.

Kalfa, E., and Piracha, M. (2017). Immigrants' educational mismatch and the penalty of over-education. *Education Economics*, *25*(5), 462-481.

Kampelmann, S., and Rycx, F. (2016). Wage discrimination against immigrants: measurement with firm-level productivity data. *IZA Journal of Migration, 5*(1), 1-24.

Kifle, T., Kler, P., and Shankar, S. (2014). The power of the pram: do young children determine female job satisfaction?. *Applied Economics Letters*, *21*(4), 289-292.

Kiker, B. F., Santos, M. C., and De Oliveira, M. M. (1997). Overeducation and undereducation: evidence for Portugal. *Economics of Education Review*, *16*(2), 111-125.

Larsen, E. N., Rogne, A. F., and Birkelund, G. E. (2018). Perfect for the job? Overqualification of immigrants and their descendants in the Norwegian labor market. *Social Inclusion*, *6*(3), 78-103.

Levels, M., and Dronkers, J. (2008). Educational performance of native and immigrant children from various countries of origin. *Ethnic and Racial Studies*, *31*(8), 1404-1425.

Lindley, J. (2009). The over-education of UK immigrants and minority ethnic groups: Evidence from the Labour Force Survey. *Economics of Education Review*, *28*(1), 80-89.

Maani, S. A., and Wen, L. (2021). Over-education and immigrant earnings: a penalized quantile panel regression analysis. *Applied Economics*, *53*(24), 2771-2790.

Maes, J., Neels, K., Biegel, N., and Wood, J. (2023). Uptake of formal childcare among second generation and native mothers in Belgium: can increasing local childcare availability narrow migrant-native gaps?. *Genus*, *79*(1), 7.

McGuinness, S. (2006). Overeducation in the labour market. *Journal of economic surveys*, *20*(3), 387-418.

McGuinness, S., Bergin, A., and Whelan, A. (2018). Overeducation in Europe: trends, convergence, and drivers. *Oxford Economic Papers*, *70*(4), 994-1015.

McGuinness, S., and Byrne, D. (2015). Born abroad and educated here: examining the impacts of education and skill mismatch among immigrant graduates in Europe. *IZA Journal of Migration*, *4*, 1-30.

McIntosh, B., McQuaid, R., Munro, A., and Dabir-Alai, P. (2012). Motherhood and its impact on career progression. *Gender in Management: An International Journal*.

Nadim, M., and Midtbøen, A. H. (2023). Gendered Consequences of Social Mobility: Second-Generation Immigrants' Work–Care Considerations in High-Status Occupations in Norway. *Sociology*, 0(0). https://doi.org/10.1177/00380385221139682

Nielsen, C. P. (2011). Immigrant over-education: evidence from Denmark. *Journal of Population Economics*, *24*(2), 499-520.

Noghanibehambari, H., Tavassoli, N., and Noghani, F. (2022). Intergenerational Transmission of Culture Among Second-and-Higher Generation Immigrants: the Case of Age at First Birth and Nonmarital Childbirth. *Journal of Economics, Race, and Policy*, 1-18.

Nugent, C. (2022). Overqualified third-level graduates in the Irish labour market. NERI Report Series. No. 19, Dublin.

OECD (2017), Catching Up? Intergenerational Mobility and Children of Immigrants, OECD Publishing, Paris.

OECD (2020). What is the impact of the COVID-19 pandemic on immigrants and their children?. OECD Policy Responses to Coronavirus (COVID-19). OECD Publishing, Paris.

OECD (2022). Education at a Glance 2022: OECD Indicators. OECD Publishing, Paris.

Park, J., and Myers, D. (2010). Intergenerational mobility in the post-1965 immigration era: Estimates by an immigrant generation cohort method. *Demography*, 47(2), 369-392.

Phalet, K., and Heath, A. (2010). From ethnic boundaries to ethnic penalties: Urban economies and the Turkish second generation. *American Behavioral Scientist*, *53*(12), 1824-1850.

Petrongolo, B. (2019). The gender gap in employment and wages. *Nature Human Behaviour*, *3*(4), 316-318.

Pineda-Hernández, K., Rycx, F., and Volral, M. (2022). Moving up the Social Ladder? Wages of First-and Second-Generation Immigrants from Developing Countries. IZA Discussion Paper No. 15770, Bohn.

Piton, C. (2022). The labour market performance of vulnerable groups: towards a better understanding of the main driving forces (No. 2013/352519). Doctoral thesis. ULB-Universite Libre de Bruxelles.

Piton, C., and Rycx, F. (2021). A Broken Social Elevator? Employment Outcomes of First- and Second-generation Immigrants in Belgium. *De Economist*, 1-47.

Santiago-Vela, A., and Mergener, A. (2022). Gender overeducation gap in the digital age: Can spatial flexibility through working from home close the gap?. *Social Science Research*, *106*, 102727.

Sellami, S., Verhaest, D., and Van Trier, W. (2018). How to Measure Field-of-Study Mismatch? A Comparative Analysis of the Different Methods. *Labour*, *32*(4), 141-173.

Shi, Y., Yu, H., Huang, Y., Shen, X., and Guo, W. (2022). Over-education and Job Satisfaction among New Graduates in China: A Gender Perspective. *Social Indicators Research*, 1-24.

United Nations (2020). World Economic Situation Prospects. New-York: United Nations, p. 163-166.

Verdugo, R. R., and Verdugo, N. T. (1989). The impact of surplus schooling on earnings: Some additional findings. *Journal of Human Resources*, 629-643.

Voets, I. (2022). *Youth Underemployment Across the EU* (Doctoral dissertation, Universitat Pompeu Fabra).

Wen, L., and Maani, S. A. (2018). A panel study of immigrants' overeducation and earnings in Australia. *International Migration*, *56*(2), 177-200.

Williams, R. (2006). Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *The Stata Journal*, *6*(1), 58-82.

Williams, R. (2016). Understanding and interpreting generalized ordered logit models. *The Journal of Mathematical Sociology*, *40*(1), 7-20.

Williams, R. A., and Quiroz, C., (2020). Ordinal Regression Models, In P. Atkinson, S. Delamont, A. Cernat, J.W. Sakshaug, and R.A. Williams (Eds.), *SAGE Research Methods Foundations*.

Zschirnt, E., and Ruedin, D. (2016). Ethnic discrimination in hiring decisions: a meta-analysis of correspondence tests 1990–2015. *Journal of Ethnic and Migration Studies*, *42*(7), 1115-1134.

Zwysen, W., and Demireva, N. (2018). An examination of ethnic hierarchies and returns to human capital in the UK. *Social Inclusion*, *6*(3), 6-33.

Table 1. Benchmark: average marginal effects - GOLOGIT regression

	Probability of being overeducated
Tertiary-educated workers born in/from:	(1)
Belgium (n = 311,222)	Reference
Developing countries[a]	
First generation (n = 14,459)	0.105***
	(0.010)
Second generation (n = 11,748)	0.008
	(0.007)
Developed countries[b] (n = 54,524)	-0.013***
	(0.005)
Others[c] (n = 4,509)	0.008
	(0.010)
Control variables	
Women	0.050***
	(0.005)
Tenure	-0.002***
	(0.001)
Squared tenure	0.000***
	(0.000)
Level of education (ref. Bachelor)	
Master	-0.060***
	(0.010)
Advanced Master or PhD	0.543***
	(0.009)
Type of household (ref. without children living at home)	
With children living at home	0.009***
	(0.003)
Other households	-0.014*
	(0.008)
Type of contract (ref. permanent)	
Fixed term	0.104***
	(0.029)
Internship	0.067**
	(0.028)
Part-time work	0.156***
	(0.011)
Overtime work	0.145***
	(0.015)
Size of the firm (FTE number of employees in log)	-0.021***
	(0.003)
Firm-level collective agreement (Yes)	0.018**
	(0.008)
More than 50% privately owned	0.008
	(0.024)
Region (ref. Brussels)	
Flanders	0.076***
	(0.010)
Wallonia	0.084***
	(0.013)
Year fixed effects[d]	Yes
Occupation-age-sector cells[e]	13,628
Observations	396,462

Notes: *** p<0.01, ** p<0.05, * p<0.1. Worker and firm weights are used in the GOLOGIT regression. Robust standard errors are in parentheses, which are clustered at the firm level. [a] Second-generation immigrants' origin is defined according to the father's country of birth. However, if the father was born in Belgium and the mother was born in a developing country, the mother's country of birth is retained. [b] The group 'immigrants from developed countries' includes F-G immigrants born in developed countries and their S-G

peers. [c] The group ‘others’ refers to workers born in developing countries with both parents born in Belgium (e.g. children of expatriates). [d] 17 year dummies. [e] Using a realized matches approach, overeducation is identified within each occupation-age-sector cell (see Section 3.1 for more details). Source: STATBEL, 1999-2016.

Table 2. Geographical origin: average marginal effects - GOLOGIT regression

		Probability of being overeducated
Tertiary-educated workers born in/from:		(1)
Belgium (n = 311,222)		Reference
Developing countries[a]		
Sub-Saharan African countries	First generation (n = 3,913)	0.121***
		(0.015)
	Second generation (n = 5,403)	-0.020**
		(0.009)
Maghreb countries	First generation (n = 3,634)	0.180***
		(0.016)
	Second generation (n = 3,517)	0.032**
		(0.015)
Near and Middle Eastern countries	First generation (n = 1,589)	0.093***
		(0.023)
	Second generation (n = 1,138)	0.032
		(0.021)
Emerging and developing Asian countries	First generation (n = 2,175)	0.020
		(0.016)
	Second generation (n = 634)	-0.023
		(0.029)
Non-EU Eastern European countries	First generation (n = 1,733)	0.092***
		(0.019)
	Second generation (n = 742)	0.026
		(0.024)
Latin American and Caribbean countries	First generation (n = 1,415)	0.018
		(0.021)
	Second generation (n = 314)	0.019
		(0.035)
Control variables		
Year fixed effect[b]		Yes
Worker characteristics[c]		Yes
Employment characteristics[d]		Yes
Firm characteristics[e]		Yes
Occupation-age-sector cells[f]		13,628
Observations		396,462

Notes: *** p<0.01, ** p<0.05, * p<0.1. Worker and firm weights are used in the GOLOGIT regression. Robust standard errors are in parentheses, which are clustered at the firm level. The groups 'immigrants from developed countries' and 'others' are also included in the regression, but their estimates are not portrayed in this table. [a] Second-generation immigrants' origin is defined according to the father's country of birth. However, if the father was born in Belgium and the mother was born in a developing country, the mother's country of birth is retained. [b] 17 year dummies. [c] Gender, level of tertiary education, tenure, squared tenure and type of household. [d] Type of contract and dummies for part-time and overtime work. [e] Size of the firm (FTE number of workers in log), dummy for more than 50% privately owned, dummy for a firm-level collective agreement and region where the firm is located (Brussels, Flanders or Wallonia). [f] Using a realized matches approach, overeducation is identified within each occupation-age-sector cell (see Section 3.1 for more details). Source: STATBEL, 1999-2016.

Table 3. Gender and origin: average marginal effects - GOLOGIT regression

		Probability of being overeducated
Tertiary-educated workers born in/from:		(1)
Belgium		
Men (n = 191,913)		Reference
Women (n = 119,309) [1]		0.056***
		(0.005)
Developing countries[a]		
Men	First generation (n = 9,382)	0.145***
		(0.012)
	Second generation (n = 6,541)	0.015
		(0.010)
Women	First generation (n = 5,077) [2]	0.111***
		(0.012)
	Second generation (n = 5,207) [3]	0.038***
		(0.013)
Control variables		
Year fixed effects[b]		Yes
Worker characteristics[c]		Yes
Employment characteristics[d]		Yes
Firm characteristics[e]		Yes
Test for equality of coefficients (p-value)[f]		
[1] = [2]		0.00
[1] = [3]		0.12
[2] = [3]		0.00
Occupation-sector-age cells[g]		13,628
Observations		396,462

Notes: *** p<0.01, ** p<0.05, * p<0.1. Worker and firm weights are used in the GOLOGIT regression. Robust standard errors are in parentheses, which are clustered at the firm level. The groups 'immigrants from developed countries' and 'others' are also included in the regression, but their gender-interacted estimates are not portrayed in this table (available on request). [a] Second-generation immigrants' origin is defined according to the father's country of birth. However, if the father was born in Belgium and the mother was born in a developing country, the mother's country of birth is retained. [b,c,d,e] All control variable included in Table 2, with the exception of gender. [f] The null hypothesis of the test specifies that the estimates are not statistically different from each other if the p-value is higher than 0.10. [g] Using a realized matches approach, overeducation is identified within each occupation-age-sector cell (see Section 3.1 for more details). Source: STATBEL, 1999-2016.

Table 4. Part-time work and origin: average marginal effects - GOLOGIT regression

		Probability of being overeducated
Tertiary-educated workers born in/from:		(1)
Belgium		
Full-time work (n = 297,169)		Reference
Part-time work (n = 14,053)		0.141***
		(0.011)
Developing countries[a]		
Full-time work	First generation (n = 13,463)	0.094***
		(0.010)
	Second generation (n = 11,247)	-0.001
		(0.008)
Part-time work	First generation (n = 996)	0.398***
		(0.022)
	Second generation (n = 501)	0.265***
		(0.033)
Control variables		
Year fixed effects[b]		Yes
Worker characteristics[c]		Yes
Employment characteristics[d]		Yes
Firm characteristics[e]		Yes
Occupation-sector-age cells[f]		13,628
Observations		396,462

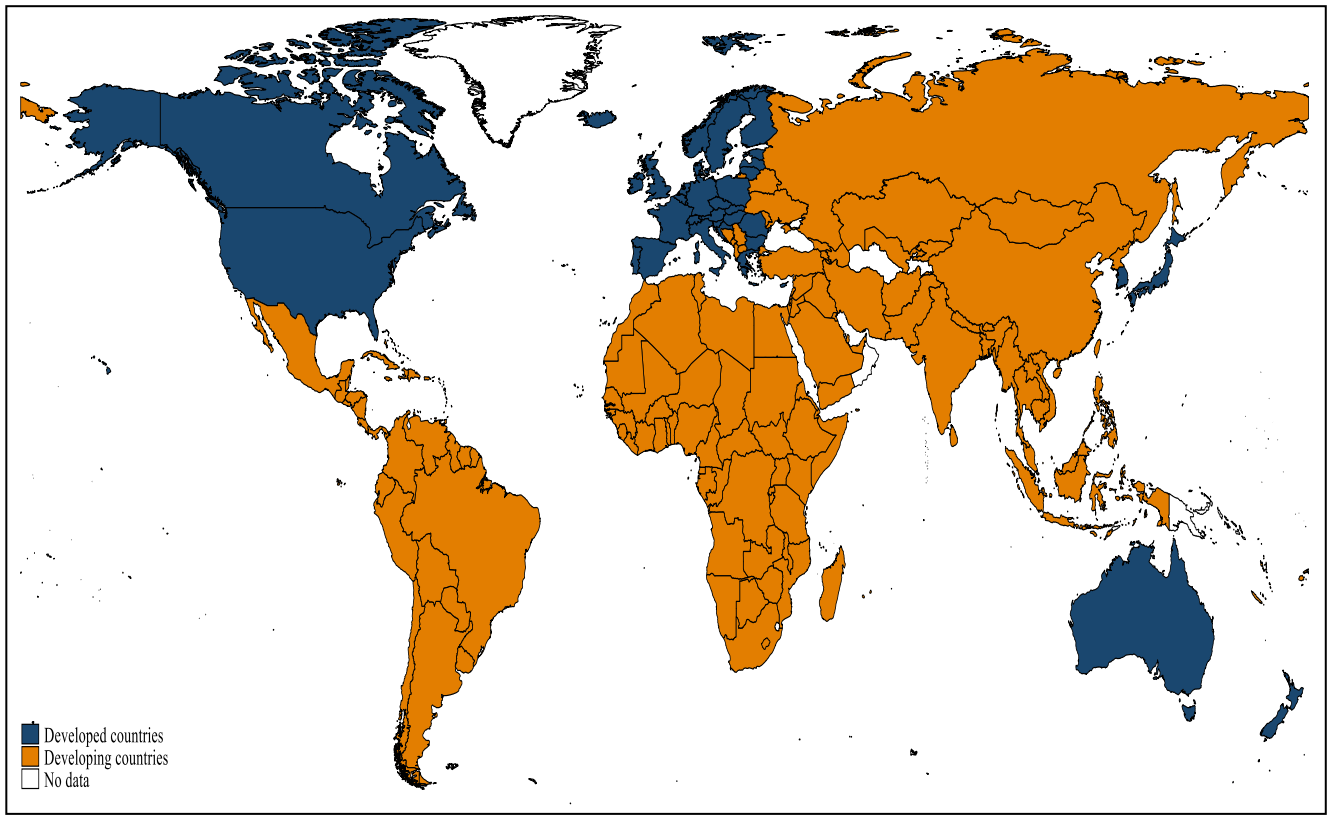
Notes: *** p<0.01, ** p<0.05, * p<0.1. Worker and firm weights are used in the GOLOGIT regression. Robust standard errors are in parentheses, which are clustered at the firm level. The groups 'immigrants from developed countries' and 'others' are also included in the regression, but their estimates associated with full- or part-time work are not portrayed in this table (available on request). [a] Second-generation immigrants' origin is defined according to the father's country of birth. However, if the father was born in Belgium and the mother was born in a developing country, the mother's country of birth is retained. [b,c,d,e] All control variable included in Table 2, with the exception of part-time work. [f] Using a realized matches approach, overeducation is identified within each occupation-age-sector cell (see Section 3.1 for more details). Source: STATBEL, 1999-2016.

Table 5. Gender, part-time work and origin: average marginal effects - GOLOGIT regression

Tertiary-educated workers born in/from:			**Probability of being overeducated**
			(1)
Belgium			
Men	Full-time jobs (n = 188,859)		Reference
	Part-time jobs (n = 3,054)		0.157*** (0.011)
Women	Full-time jobs (n = 108,310)		0.054*** (0.002)
	Part-time jobs (n = 10,999)		0.190*** (0.006)
Developing countries[a]			
Men	Full-time jobs	First generation (n = 8,995)	0.128*** (0.007)
		Second generation (n = 6,393)	0.018** (0.007)
	Part-time jobs	First generation (n = 387) [1]	0.449*** (0.024)
		Second generation (n = 148) [2]	0.373*** (0.048)
Women	Full-time jobs	First generation (n = 4,468)	0.080*** (0.010)
		Second generation (n = 4,854)	0.028*** (0.010)
	Part-time jobs	First generation (n = 609) [3]	0.413*** (0.018)
		Second generation (n = 353) [4]	0.270*** (0.035)
Control variables			
Year fixed effects[b]			Yes
Worker characteristics[c]			Yes
Employment characteristics[d]			Yes
Firm characteristicse			Yes
Test for equality of coefficients (p-value)[f]			
[1] = [3]			0.23
[2] = [4]			0.08
Occupation-sector-age cells[g]			13,628
Observations			396,462

Notes: *** p<0.01, ** p<0.05, * p<0.1. Worker and firm weights are used in the GOLOGIT regression. Robust standard errors are in parentheses, which are clustered at the firm level. The groups 'immigrants from developed countries' and 'others' are also included in the regression, but their gender-interacted estimates associated with full- or part-time work are not portrayed in this table (available on request). [a] Second-generation immigrants' origin is defined according to the father's country of birth. However, if the father was born in Belgium and the mother was born in a developing country, the mother's country of birth is retained. [b,c,d,e] All control variable included in Table 2, with the exception of gender and part-time work [f] The null hypothesis of the test specifies that the estimates are not statistically different from each other if the p-value is higher than 0.10. [g] Using a realized matches approach, overeducation is identified within each occupation-age-sector cell (see Section 3.1 for more details). Source: STATBEL, 1999-2016.

Appendix 1: Chart of developed and developing countries



Notes: Overseas territories are classified depending on their neighbouring countries. No data stipulates that no observation for workers born in or from these countries (Greenland (Denmark), Oman, Papua New Guinea, Tajikistan and Turkmenistan) is presented in our database.

Appendix 2: List of countries by geographical region in our database

Developed countries

Belgium

EU-14 countries[a]: Austria, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden and United Kingdom.

Other EU countries: Bulgaria, Croatia, Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovak Republic and Slovenia.

Other developed countries: Andorra, Australia, Canada, Iceland, Japan, Liechtenstein, Monaco, New Zealand, Norway, Saint-Marin, Singapore, South Korea, Switzerland, Taiwan and United States.

Developing countries

The Maghreb countries: Algeria, Libya, Mauritania, Morocco, and Tunisia.

Sub-Saharan African countries: Angola, Benin, Botswana. Burkina Faso, Burundi, Cabo Verde, Cameroon, Central Africa Republic, Chad, Comoros, Congo, Congo DRC, Côte d'Ivoire, Djibouti, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Liberia, Madagascar, Malawi, Mali, Mauritius, Mozambique, Namibia, Niger, Nigeria, Reunion (French Department), Rwanda, Sao Tome and Principe, Sierra Leone, Somalia, South Africa, Sudan, Tanzania, Togo, Uganda, Zambia and Zimbabwe.

Near and Middle Eastern countries: Afghanistan, Bahrein, Egypt, United Arab Emirates, Iran, Iraq, Israel, Jordan, Kuwait, Pakistan, Palestine, Qatar, Saudi Arabia, Syria, Turkey and Yemen.

Non-EU Eastern European countries: Albania, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Georgia, Kosovo, North Macedonia, Moldova, Montenegro, Russia, Serbia and Ukraine.

Emerging and developing Asian countries: Bangladesh, Bhutan, Brunei Darussalam, Cambodia, China, Fiji, French Polynesia (French Department), India, Indonesia, Kazakhstan, Kyrgyz Republic, Laos, Malaysia, Mongolia, Myanmar, Nauru, Nepal, North Korea, Nauru, New Caledonia, Philippines, Sri Lanka, Thailand, Uzbekistan, Vanuatu, Vietnam and Wallis and Futana (French Department).

Latin American and Caribbean countries: Argentina, Bahamas, Barbados, Belize, Bermuda, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Dutch Antilles, Ecuador, Grenada, Guadeloupe (French Department), Guatemala, Guyana, French Guyana (French Department), Haiti, Honduras, Jamaica, Martinique (French Department), Mexico, Nicaragua, Panama, Paraguay, Peru, Saint Lucia, Suriname, Trinidad and Tobago, Uruguay and Venezuela.

[a] EU countries are defined as during the time span of the database (1999-2016). Therefore, the United Kingdom is still considered an EU country.

Appendix 3. Descriptive statistics by origin and generation – means and percentages

	Sample of tertiary-educated workers born in or from:		
	Belgium	**Developing countries[a]**	
		First generation	**Second generation[b]**
Share of the sample (%)[c]	78.9	3.7	3.0
Observations	311,222	14,459	11,748
Region of birth (%)			
Belgium	100.0		
Developing countries[d]			
Sub-Saharan African countries		27.1	46.0
Maghreb countries		25.1	29.9
Near and Middle Eastern countries		11.0	9.7
Emerging and developing Asian countries		15.1	5.4
Other Eastern European countries		12.0	6.3
Latin American and Caribbean countries		9.8	2.7
Worker characteristics			
Women (%):	37.4	34.5	43.4
Age categories (%):			
20-24	6.2	2.7	10.3
25-29	18.0	15.5	31.8
30-39	36.8	40.6	42.9
40-49	26.7	27.1	12.1
50+	12.4	14.0	2.9
Tenure in years	7.8	5.2	4.3
Education (%):			
Bachelor	57.6	57.5	57.3
Master	39.7	38.1	39.9
Advanced Master or PhD	2.7	4.4	2.8
Household (%):			
Without children living at home	30.3	33.3	32.7
With children living at home	68.1	59.4	64.3
Other households[e]	1.6	7.3	3.1
Employment characteristics			
Type of contract (%):			
Permanent	95.9	90.6	91.9
Fixed-term	3.6	8.5	7.3
Internship or apprenticeship	0.5	1.0	0.8
Part-time work (%)[f]	4.7	7.7	4.4
Overtime work (%)[g]	1.9	2.7	2.0
Occupational categories - ISCO1 (%):			
Managers	12.5	9.0	8.9
Professionals	38.1	35.2	39.0
Technicians and associate professionals	17.9	14.0	19.1
Clerical support	23.1	19.4	23.4
Service and sales workers	4.0	6.2	5.0
Craft and related trades workers	2.0	4.2	1.9
Plant and machine operators and assemblers	1.8	4.6	1.8
Elementary Occupations	0.6	7.4	0.9

Table 1. (Continued)

Appendix 3. Continued

	Sample of tertiary-educated workers born in or from:		
	Belgium	**Developing countries[a]**	
		First generation	**Second generation[b]**
Firm characteristics			
Sector of activity - NACE1 (%):			
B - Mining and Quarrying	0.1	0.1	0.1
C - Manufacturing	29.8	24.5	20.0
D - Electricity, gas, steam, and air conditioning supply	2.1	0.8	1.9
E - Water supply, sewerage, waste management and remediation activities	0.8	0.6	0.5
F - Construction	4.2	2.5	3.5
G - Wholesale and retail trade, repair of motor vehicles and motorcycles	18.5	15.7	16.7
H - Transportation and storage	5.2	6.2	5.0
I - Accommodation and food service activities	0.9	3.4	1.6
J - Information and communication	11.7	12.1	15.9
K - Financial and insurance activities	2.6	4.3	4.3
L - Real Estate activities	0.3	0.4	0.4
M - Professional, scientific, and technical activities	12.4	14.5	14.2
N - Administrative and support service activities	8.3	11.6	13.3
P - Education	0.3	0.3	0.3
Q - Human Health and social work activities	2.4	2.0	1.8
R - Arts, entertainment, and recreation	0.1	0.3	0.1
U - Other service activities	0.3	0.6	0.5
Size of the firm (FTE number of employees)	459.0	505.3	481.2
Firm-level collective agreement (Yes) (%)	28.2	27.0	26.0
More than 50% privately owned (%)	95.1	96.1	94.9
Region where the firm is located (%):			
Brussels	19.2	34.6	37.3
Flanders	64.8	45.3	38.8
Wallonia	16.1	20.1	24.0

Notes: Worker and firm weights are used in the calculation of means and percentages. [a] By 'developing countries', we mean either transition and developing countries listed in the United Nations' (2020) classification and/or emerging market and developing economies listed in the IMF's (2020) classification. [b] Second-generation immigrants' origin is defined according to the father's country of birth. However, if the father was born in Belgium and the mother was born in a developing country, the mother's country of birth is retained. [c] The groups 'immigrants from developed countries' and 'others' are also considered in our empirical strategy. They represent 13.3% and 1.1% of our sample, respectively. Descriptive statistics for these groups are shown in Appendix 3. [d] Appendix 2 shows the list of developing countries by region of birth. For the sake of accuracy in correctly classifying immigrants by geographical origin and economic development level, we construct our geographical classification of countries based on both the United Nations' (2020) classification and the IMF's (2020) classification. [e] 'Other households' refer to brothers/sisters living together, friends living together, students or workers' homes, etc. [f] A worker is recognized as a part-time employee if he/she works less than 30 hours per week. [g] Overtime is when an employee works more than his/her contractual working hours. Source: STATBEL, 1999-2016.

Appendix 4. Incidence of overeducation (%) by origin and generation across moderating variables

	Total	Gender		Employment		Gender and employment			
						Full-time work		Part-time work	
		Men	Women	Full-time work	Part-time work	Men	Women	Men	Women
Tertiary-educated workers born in/from:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Belgium	43.5	40.5	48.2	42.7	59.9	40.3	46.9	55.4	61.1
	(135,295 obs.)	*(77,760 obs.)*	*(57,535 obs.)*	*(126,884 obs.)*	*(8,411 obs.)*	*(76,067 obs.)*	*(50,817 obs.)*	*(1,693 obs.)*	*(6,718 obs.)*
Developing countries[a]									
First generation	53.7	52.7	51.0	50.0	80.1	51.4	47.3	83.5	78.0
	(7,534 obs.)	*(4,944 obs.)*	*(2,590 obs.)*	*(6,736 obs.)*	*(798 obs.)*	*(4,621 obs.)*	*(2,115 obs.)*	*(323 obs.)*	*(475 obs.)*
Second generation	42.1	40.7	44.8	41.7	71.7	40.7	43.1	79.7	68.3
	(5,052 obs.)	*(2,718 obs.)*	*(2,334 obs.)*	*(4,693 obs.)*	*(359 obs.)*	*(2,600 obs.)*	*(2,093 obs.)*	*(118 obs.)*	*(241 obs.)*

Notes: Using a realized matches approach, overeducation is identified within each occupation-age-sector cell (see Section 3.1 for more details). [a] S-G immigrants' origin is defined based on the father's country of birth. However, if the father was born in Belgium and the mother was born in a developing country, the mother's country of birth is retained. Source: STATBEL, 1999-2016.

Appendix 5. Average marginal effects - OLS and OLOGIT regressions

	Probability of being overeducated	
Tertiary-educated workers born in/from:	(1)	(2)
Belgium (n = 311,222)	Reference	Reference
Developing countries[a]		
First-generation (n = 14,459)	0.121***	0.094***
	(0.014)	(0.011)
Second-generation (n = 11,748)	0.011	0.007
	(0.010)	(0.007)
Developed countries[b] (n = 54,524)	-0.039***	-0.025***
	(0.006)	(0.004)
Others[c] (n = 4,509)	0.008	0.007
	(0.014)	(0.010)
Estimator	**OLS**	**OLOGIT**
Control variables		
Year fixed effects[d]	Yes	Yes
Worker characteristics[e]	Yes	Yes
Employment characteristics[f]	Yes	Yes
Firm characteristics[g]	Yes	Yes
Occupation-age-sector cells[h]	13,628	13,628
Observations	396,462	396,462

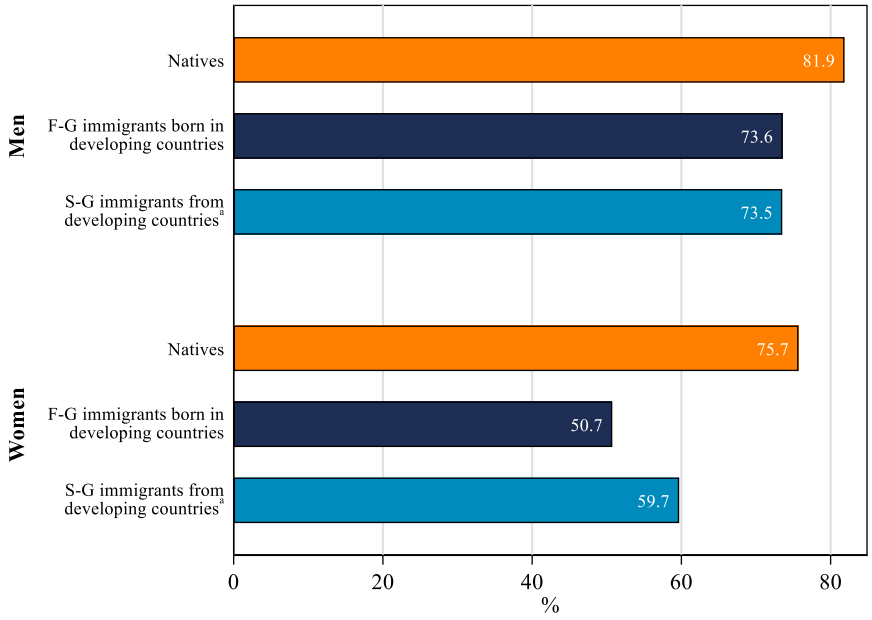
Notes: *** p<0.01, ** p<0.05, * p<0.1. Worker and firm weights are used. Robust standard errors are in parentheses, which are clustered at the firm level. [a] Second-generation immigrants' origin is defined according to the father's country of birth. However, if the father was born in Belgium and the mother was born in a developing country, the mother's country of birth is retained. [b] The group 'immigrants from developed countries' includes F-G immigrants born in developed countries and their S-G counterparts. [c] The group 'others' refers to workers born in developing countries with both parents born in Belgium. [d] 17 year dummies. [e] Gender, level of tertiary education, tenure, squared tenure and type of household. [f] Type of contract and dummies for part-time work and overtime work. [g] Size of the firm (FTE number of workers in log), dummy for more than 50% privately owned, dummy for a firm-level collective agreement and region where the firm is located (Brussels, Flanders or Wallonia). [h] Using a realized matches approach, overeducation is identified within each occupation-age-sector cell (see Section 3.1 for more details). Source: STATBEL, 1999-2016.

Appendix 6. Part-time work and parenthood: average marginal effects – LOGIT regression

			Probability of being in a part-time job	
Tertiary-educated workers born in/from:			(1)	(2)
Belgium				
Men	Without child(-ren) (n = 56,415)		Reference	Reference
	With child(-ren) (n = 132,102)		-0.002***	-0.002*
			(0.001)	(0.001)
Women	Without child(-ren) (n = 38,762)		0.031***	0.024***
			(0.002)	(0.002)
	With child(-ren) (n = 78,868)		0.102***	0.078***
			(0.002)	(0.002)
Developing countries[a]				
Men	Without child(-ren)	First generation (n = 2,969)	0.034***	0.011***
			(0.005)	(0.004)
		Second generation (n = 2,139)	-0.003	0.001
			(0.003)	(0.004)
	With child(-ren)	First generation (n = 5,676)	0.028***	0.008***
			(0.004)	(0.003)
		Second generation (n = 4,190)	0.008**	0.012***
			(0.003)	(0.004)
Women	Without child(-ren)	First generation (n = 1,840)	0.072***	0.042***
			(0.009)	(0.006)
		Second generation (n = 1,779)	0.029***	0.031***
			(0.007)	(0.006)
	With child(-ren)	First generation (n = 2,932)	0.144***	0.064***
			(0.009)	(0.005)
		Second generation (n = 3,294)	0.068***	0.061***
			(0.007)	(0.006)
Control variables				
Year fixed effects[b]			Yes	Yes
Worker characteristics[c]			No	Yes
Employment characteristics[d]			No	Yes
Firm characteristics[e]			No	Yes
Observations			386,555	386,555

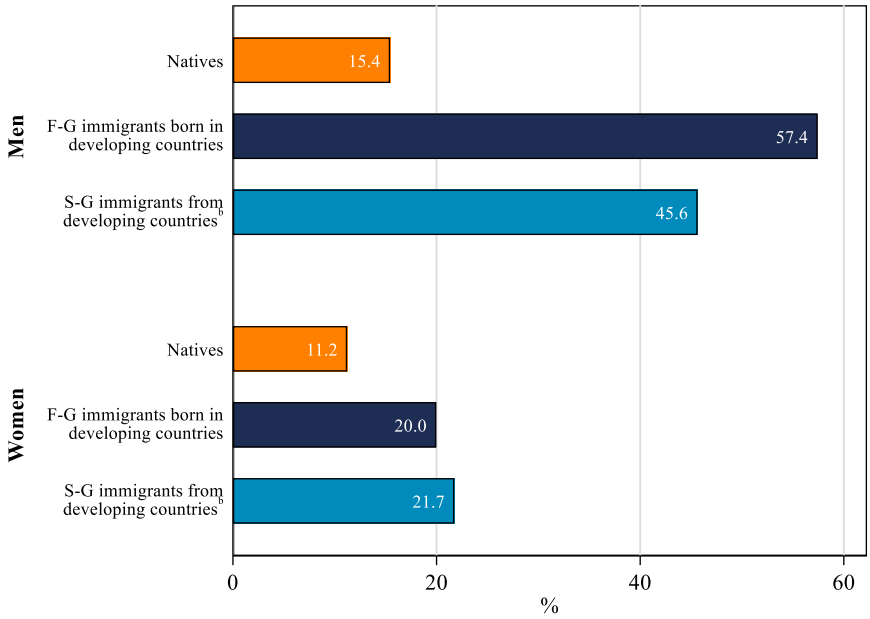
Notes: *** p<0.01, ** p<0.05, * p<0.1. Worker and firm weights are used in the LOGIT regression. Robust standard errors are in parentheses, which are clustered at the firm level. The groups 'immigrants from developed countries' and 'others' are also included in the regression but their gender-interacted estimates associated with or without having children are not portrayed in this table (available on request). Workers living in other types of households (e.g. brothers/sisters living together, friends living together, students or workers' homes) are not included in the regression [a] Second-generation immigrants' origin is defined according to the father's country of birth. However, if the father was born in Belgium and the mother was born in a developing country, the mother's country of birth is retained. [b] 17 year dummies. [c] age, level of tertiary education, tenure and squared tenure. [d] occupation, type of contract and dummy for overtime. [e] sector of activity, size of the firm (FTE number of workers in log), dummy for more than 50% privately owned, dummy for a firm-level collective agreement and region where the firm is located (Brussels, Flanders or Wallonia). Source: STATBEL, 1999-2016.

Appendix 7. Labour market participation rate in 2021 among people aged 20-64 in Belgium.



Notes: The labour market participation rate is calculated as the share of employed and unemployed persons aged 20-64 in the total population (active and inactive persons) of that same age. [a] Second-generation immigrants' origin is defined according to the father's country of birth. However, if the father was born in Belgium and the mother was born in a developing country, the mother's country of birth is retained. Source: 2021 Labour Force Survey, STATBEL.

Appendix 8. Involuntary part-time employment rate in 2021 among part-time workers aged 20-64 in Belgium.



Notes: The involuntary part-time employment rate is defined as the percentage of part-time workers (i.e. workers who work less than 30h per week) who want to work more hours and are available in the next two weeks at the time of the survey. [a] Second-generation immigrants' origin is defined according to the father's country of birth. However, if the father was born in Belgium and the mother was born in a developing country, the mother's country of birth is retained. Source: 2021 Labour Force Survey, STATBEL.