

### Self-employment as a stepping stone to better labor market matching: a comparison between immigrants and natives

Ulceluse, Magdalena

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

#### Empfohlene Zitierung / Suggested Citation:

Ulceluse, M. (2020). Self-employment as a stepping stone to better labor market matching: a comparison between immigrants and natives. *Journal of Demographic Economics*, 86(4), 479-501. <https://doi.org/10.1017/dem.2020.1>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:

<https://creativecommons.org/licenses/by/4.0>

RESEARCH PAPER

# Self-employment as a stepping stone to better labor market matching: a comparison between immigrants and natives

Magdalena Ulceluse\* 

Faculty of Spatial Sciences Department of Cultural Geography, University of Groningen,  
Landleven 1 9747 AD Groningen, The Netherlands

\*Corresponding author. E-mail: [m.m.ulceluse@rug.nl](mailto:m.m.ulceluse@rug.nl)

(Received 6 August 2018; revised 7 January 2020; accepted 8 January 2020)

## Abstract

The paper investigates the relation between overeducation and self-employment, in a comparative analysis between immigrants and natives. Using the EU Labour Force Survey for the year 2012 and controlling for a list of demographic characteristics and general characteristics of 30 destination countries, it finds that the likelihood of being overeducated decreases for self-employed immigrants, with inconclusive results for self-employed natives. The results shed light on the extent to which immigrants adjust to labor market imperfections and barriers to employment and might help explain the higher incidence of self-employment that immigrants exhibit, when compared to natives. This is the first study to systematically study the nexus between overeducation and self-employment in a comparative framework. Moreover, the paper tests the robustness of the results by employing two different measures of overeducation, contributing to the literature of the measurement of overeducation.

**Keywords:** Immigrants; overeducation; self-employment; skills mismatch

**JEL classification:** J15; J24; J61

## 1. Introduction

Immigrants generally exhibit a higher incidence of overeducation *and* self-employment than the native population. This might not be a coincidence. When immigrants arrive in a new country, they often find it difficult to carry over their human capital to the new labor market. This can happen for many reasons<sup>1</sup>—like in the case of language abilities—or because the skills they have acquired in the country of origin are not perfectly transferable to the new context [Chiswick and Miller (2009, 1992)]. Moreover, since immigrants are usually positively self-selected, their average

<sup>1</sup>Reasons may include no recognition of qualifications, lack of language skills, delayed adjustment to the new context, etc.

educational level will likely be higher than that of the native population [Chiswick (1978)]. But overeducated individuals often endure wage penalties, experience less job satisfaction, and have a higher probability to quit than well-matched individuals.<sup>2</sup> It seems thus intuitive to assume that they would try to find or create opportunities that would match their level of education and skills.

One such opportunity is self-employment. If the existing paid employment opportunities do not adequately meet their educational level and experience, by starting a business, they can create a job for themselves that matches their level of skills and education. In this case, self-employment becomes a strategy through which they reduce the incidence of overeducation. On the other hand, however, there is the possibility that immigrants become self-employed because they cannot find *any* paid employment, not necessarily an ill-fitted one. This becomes a type of necessity self-employment,<sup>3</sup> in which case the incidence of overeducation may in fact increase. *Is self-employment, therefore, increasing or decreasing skills mismatch? Moreover, is this effect stronger for immigrants than for the native population?*

The present study intends to provide an answer to precisely these questions. It investigates how the probability of being overeducated increases or decreases for self-employed individuals, whether immigrants or natives. It does so in an attempt to enrich our understanding of three critical areas of policy interest: immigrant integration, skills mismatch and self-employment/entrepreneurship. I employ two different measures of overeducation, in order to test the robustness of the results. The analysis includes both an aggregate, cross-country analysis and an individual country analysis that considers country-specific institutional arrangements and how they can interact with and create incentives and opportunities for immigrants and natives. Moreover, the analysis compares immigrant and native self-employment, comparison motivated by the assumption that by virtue of being outsiders to the labor market, immigrants encounter more barriers to finding a job, which might increase their mismatch and by extension their propensity to become self-employed. This phenomenon might help explain the significantly higher incidence of both overeducation and self-employment that immigrants generally exhibit compared to the native population.

Given the high policy relevance of matching skills to jobs and promoting self-employment, we know surprisingly little about the way these two phenomena interact. To date, there are only two studies that analyze directly the relationship between mismatch and self-employment, and they present contradictory results. In a cross-sectional study, and using a sample of workers in the science and engineering fields, Bender and Roche (2013) investigate whether mismatch differs across different types of employment—salary and self-employment jobs—and what are the effects of mismatch on wages and job satisfaction. They focus on the US and utilize the 2003 National Survey of College Graduates, from the US National Science Foundation. The dataset comprises workers who have obtained at least a Bachelors' degree in hard or social science, technology, engineering, or

<sup>2</sup>For more details, see: Mavromaras and McGuinness (2012); Verhaest and Omeij (2010); Bennett and McGuinness (2009); Battu and Sloane (2004); Chevalier (2003); Allen and Van der Velden (2001); Hartog (2000); Tsang and Levin (1985); Duncan and Hoffman (1981).

<sup>3</sup>See Reynolds et al. (2005) and Ulceluse and Kahanec (2018).

mathematics field and/or are currently working in that field. The study employs a subjective measure of mismatch,<sup>4</sup> and the analysis is conducted using three models: a probit, a linear model with instrumental variables<sup>5</sup> and a recursive bivariate probit model. They find that self-employed individuals are more likely to report being mismatched than employed individuals. Moreover, there seems to be a larger wage penalty for mismatched self-employees, although they find this does not affect job satisfaction.

In a longitudinal study this time, Sanchez et al. (2015) analyze the impact of the transition from salaried employment to self-employment on self-reported skill mismatches. They employ the European Community Household Panel (ECHP) for the period 1994–2001, for the EU-15<sup>6</sup> countries. They too use a subjective measure of mismatch, and estimate a random-effects probit model, complemented by a pooled bivariate probit model to account for endogeneity.<sup>7</sup> They find that self-employed individuals are less likely to declare being skill mismatched, and that individuals who transition from salaried employment to self-employment reduce their probability of being mismatched after the transition.

The two studies present a rather inconsistent picture of the relationship between skills mismatch and self-employment, which might be explained by significant differences in their respective research designs. While Bender and Roche (2013) focus on the US, analyze a specific dataset of college graduates, and employ a cross-sectional analysis, Sanchez et al. (2015) analyze the EU-15 member states, utilize a representative sample of these countries' populations, and conduct a longitudinal analysis. Nevertheless, the contradictory results of these two studies reflect our lack of clear understanding of the self-employment–overeducation relationship. In this context, the current study intends to improve our current knowledge of the dynamic between the two processes, and to further it, by systematically comparing natives and immigrants. Since the latter generally exhibit a higher incidence of overeducation than the native population, I expect to observe significant differences between the two groups.

The paper is structured as follows. Section 2 provides a theoretical incursion into the existing knowledge of immigrant overeducation and the potential mechanism behind the overeducation–self-employment relationship. Section 3 presents the data sources with descriptive statistics of the main variables, and the methodology employed. Section 4 presents the results of the analysis, while section 5 discusses the implications and relevance of these results and identifies new research directions.

---

<sup>4</sup>Defined by the question “Thinking about the relationship between your work and your education, to what extent is your work related to your highest degree? Closely related, somewhat related, or not at all related.”

<sup>5</sup>They use as instruments: (1) the number of published articles (grouped at zero, 1–10, 11–20, and 21 plus), assuming that research is less likely to be necessary in self-employment, and (2) the month that the highest degree was awarded, assuming that firms will hire entry-level jobs cyclically and so wage and salary jobs will not be as available in nonstandard graduation months (such as May, June, or December) [Bender and Roche (2013), p. 90].

<sup>6</sup>Denmark, the Netherlands, Belgium, France, Ireland, Italy, Greece, Spain, Portugal, Austria, Germany, Sweden, Luxembourg, UK, and Finland

<sup>7</sup>The authors use a variable that indicates whether the individual holds a permanent labor contract, as an exclusion restriction. A Wald test leads them to conclude there is no endogeneity.

## 2. A theoretical perspective on immigrant overeducation

Four main theories have been put forward to explain the existence of overeducation in the labor market, and their hypotheses can be extrapolated to explain immigrant overeducation too: search and match theory, human capital theory, signaling theory, and technological change theory.

According to *search and match theory*, immigrant overeducation is the result of imperfect (and asymmetric) information in the labor market. When immigrants arrive into a country, as outsiders, they have limited knowledge of the available jobs and of the functioning of the local labor market. To get their foot in the door, they may take up jobs for which they are overqualified, with the intention of advancing up the occupational ladder once they get acquainted with the new labor market structure and gain local job experience [see Groot and Maassen van den Brink (2000)]. The adjustment process is especially pronounced among immigrants originating from countries with significantly different labor markets and institutions [Chiswick and Miller (2009)]. According to the *search and match theory*, thus, overeducation appears as a necessary adjustment to new employment environments. The *searching* ability is impaired as compared to locals. Once immigrants familiarize with the local job market and have removed the necessary hurdles in adjusting to the new environment, they should, the theory concludes, be able to match employment to their education level. Overeducation in this case is viewed as a temporary phenomenon, as immigrants are expected to eventually find jobs that match their level of education.

The *human capital theory*,<sup>8</sup> too, considers overeducation to be a temporary phenomenon. When they arrive in the destination country, immigrants often find it difficult to transfer (or have recognized) the skills they have acquired in the country of origin [Chiswick and Miller (2009)]. Overeducation becomes then an adjustment mechanism, a strategy they employ to enter the new labor market, with the purpose of gaining experience that smooths out the path for a matching job in the future. Thus, in time, with residence length and the accumulation of locally recognized human capital, the incidence of overeducation is likely to decline [Piracha and Vadean (2013)].

In a similar vein, the *screening theory* [Arrow (1973); Spence (1974)] considers education to be a signal individuals send concerning their labor productivity and abilities. The theory presupposes that hiring someone represents an investment involving risk and uncertainty and that formal education reduces uncertainty by sending a signal about a person's abilities and skills. The theory is rooted in the asymmetric nature of information about employees' skills and the fact that employers face lots of uncertainties in assessing job applicant, thus they rely on their educational degrees, in which case they assume that individuals with a higher educational level (an observable signal) also have higher skills (initially difficult to observe for employers) [Ghaffarzadegan et al. (2017)]. Formal education becomes particularly relevant for immigrants, as they need to signal employers—who might be apprehensive about the quality and content of foreign education—a measure of their ability. Therefore, recent immigrants would experience a higher incidence of

---

<sup>8</sup>The premise of the human capital theory is similar to that of the career mobility theory of Sicherman and Galor (1990), according to which workers accept jobs for which they are overqualified in order to acquire work experience and enhance the chances of finding a better job match.

overeducation, which should however decrease over time once their skills are recognized.

The above theories of overeducation can be extrapolated to motivate the decision to become self-employed too. Self-employment itself can be a transitional process toward finding paid employment. Particularly for immigrants, who as outsiders often lack information about the local labor market, and whose hiring constitutes an investment implying greater risks, self-employment can represent a period of transition, in which they get accustomed to the new labor market and build up the necessary human capital to acquire paid employment in the new destination.

A latest explanation for the overeducation phenomenon focuses on the effects of *technological change* [see Kiker et al. (1997); Mendes de Oliveira et al. (2000)]. This theory argues that the rapid pace of technological development generates the need for more school-acquired skills than those possessed by other employees in the same position. If the requirements for the same positions are higher today than they used to be in the past, then those people employed today may seem overeducated in comparison to their older colleagues who were employed at a time at which the required skill level was lower. But, in this case, overeducation is more of a perceived phenomenon rather than an actual one; individuals in fact have the required level of education to keep up with technological advancements, they just seem overqualified when compared to previous employee cohorts. This implies that the incidence of overeducation is not expected to decrease with time, as there was none to begin with. According to this theory, the perceived incidence of overeducation is expected to be higher the larger the discrepancy in terms of technological advances between the immigrants' origin and destination countries.

The above theories and their predictions are not mutually exclusive, but rather different facets of the same process of immigrant labor market integration. When first arriving in the destination country, immigrants do have a limited knowledge of the local labor market (search and match theory), for which they need a strategy (human capital theory), while employers have limited knowledge of their abilities for which they need a signal (screening theory). Overeducation becomes thus an adjustment mechanism to overcome existing labor market inefficiencies, which should disappear over time.

These papers, however, do not explicitly study the self-employed. Yet, the nature of self-employment could have important spillover effects on the incidence of overeducation. If self-employment is necessity-based because there are no opportunities in paid employment (or if the gains associated with self-employment surpass those associated with a well-matched job), then the incidence of mismatch might increase. Conversely, if self-employment is taken up as an alternative to a mismatched job, then the incidence of overeducation might decrease. The next sections attempt to provide more information with regards to the direction of the self-employment–overeducation relationship.

### 3. Measuring overeducation

The concept of overeducation, as employed in this paper, refers to the instance in which workers have more years of education than required for the job they are performing. Relatively unambiguous and with an intuitive interpretation, the concept has been employed extensively in the studies of mismatch over the past decades. Yet measuring overeducation is not straightforward and previous studies have shown that

the incidence of overeducation is sensitive to the method of measurement [see Groot and Maassen van den Brink (2000)]. Currently, four main approaches to mismatch measurement have been identified in the existing literature.

The *job analysis* (or normative) approach is an objective method that derives information concerning the required level of education for an occupation from occupational classification databases, like the O\*NET or ISCO [e.g., Chevalier (2003); Piracha and Vadean (2012)]. The *realized matches* (statistical) approach derives the level of education necessary for a particular occupation by taking the mean (or mode) of years of schooling of all individuals employed in that occupation. Individuals with a standard deviation above the mean (mode) are considered overeducated [e.g., Chiswick and Miller (2007, 2009)]. The *income-ratio* approach equates overeducation with income inefficiency and computes overeducation as the ratio between potential and actual income [e.g., Guironnet and Peypoch (2007); Jensen et al. (2010)]. Proponents of this measure argue that income maximization is an important reason why individuals invest in education, and that this measure “allows the inclusion of income and efficiency aspects of overeducation ignored by the well-established objective or subjective measures focusing on some (ordinal) matching aspects” [Jensen et al. (2010, p. 34)]. The *self-assessment* approach consists in asking individuals whether they have more or less education than required for the job (direct assessment) or the minimum level of education required for the job they perform (indirect self-assessment).

In order to determine which immigrants and natives are overeducated, I employ both the normative and statistical approaches. Each method presents a number of benefits and drawbacks [Hartog (2000), Verhaest and Omeij (2010)], therefore the comparison enables me to test the robustness of the results. For the normative (job analysis) measure, I compare the required level of education for an occupation against the current level of education of the individual. For this purpose, I use the International Standard Classification of Occupations (henceforth ISCO-08) and the International Standard Classification of Education (henceforth ISCED-97) and their correspondence as developed by the ILO [ILO (2012, 2014)]. The nine Major Occupational Groups in ISCO correspond to four skill levels, which in turn correspond to the six educational classifications (see Annex A for correspondence). Individuals who exhibit an educational level above the corresponding one are considered overeducated. The approach has been successfully employed elsewhere, to measure skills mismatch and its determinants [see, for instance, Chevalier (2003); Sutherland (2012); Tarvid (2012)]. It presents a number of advantages, including a relative ease to measure mismatch and consistency over time. In addition, unlike the self-assessed and the income-ratio approaches, for instance, it is a rather objective measure. However, the approach has a number of limitations too. Firstly, it assumes constant mapping over all jobs of a given occupation, not taking into account that in some countries with a high share of educational attainment, the average educational level for a job would be higher [ILO (2014)]. Moreover, the approach clusters together groups of occupations for which the educational level required varies significantly (for instance, there is substantial variation between ISCO groups 4–8, which results in an underestimation of the number of overeducated individuals in this case).

For the realized matches approach (statistical measure), I compute the mode<sup>9</sup> of educational level for each particular occupation and consider those individuals that

<sup>9</sup>To reflect the most common level of education for an occupation.

present an educational attainment level one standard deviation above this mode, to be overeducated. The approach has been successfully employed elsewhere [e.g., Kiker et al. (1997); Chiswick and Miller (2009)] and presents the advantage of considering the actual educational level of workers within a particular occupation, at any given time.

## 4. Data and methodology

### 4.1 The data

The analysis in this paper relies on the European Union Labour Force Survey's (EU LFS) for the year 2012. The EU LFS is the largest European household sample survey, providing annual data on labor participation of people aged 15 and over and on persons outside the labor force [Eurostat (2007)]. The data provide information on individual socio-economic characteristics, occupation, education, as well as on individual's country of birth, which enables the distinction between natives and immigrants, and length of residence in the country. Further, the study only considers immigrants from outside the EU and EFTA,<sup>10</sup> as the latter technically share the same labor market rights as the native population. There are 22 countries covered in the sample.<sup>11</sup> The sample includes 74,727 non-EU immigrants, 12% of which are self-employed.

### 4.2 The dependent variable

The dependent variable is overeducation, a dummy variable equal to 1 if the individual is overeducated and 0 otherwise. The variable is derived using information on occupations, educational levels, and country of origin from the EULFS. Tables 1–3 compare the incidence of overeducation between immigrants and natives across a number of demographic characteristics, using both the normative and the statistical approach to computing overeducation.

Table 1 presents the incidence of overeducation disaggregated by major regions of origin. By far the highest incidence of overeducation seems to be experienced, surprisingly, by immigrants from the EU10 Member States,<sup>12</sup> followed by immigrants from South East Asia, South America, and the Near Middle East. Perhaps not surprisingly, the native population exhibits the lowest incidence of mismatch. There are substantial differences in the incidence of overeducation for each region, when we compare the two different measurements, yet no clear pattern emerges. If we compare the statistical measure of overeducation against the baseline normative measure, some origin regions or group of countries experience an increase in the incidence of overeducation (e.g., Australia and Oceania, EU10, and EU15), while others experience a decrease (e.g., East Asia, EU3, or South East Asia).

In terms of occupations (Table 2), immigrants register a significantly higher level of overeducation in all but one major occupational group. Notably, individuals employed in elementary occupations present a disproportionate level of overeducation compared with the other major groups, and in this case only, more natives seem to be mismatched than immigrants. The incidence of overeducation for both groups increases

<sup>10</sup>European Free Trade Association.

<sup>11</sup>AT, BE, BG, CZ, DE, DK, EE, ES, FI, FR, GR, HR, HU, IE, IT, LU, NO, PL, PT, RO, SE, UK.

<sup>12</sup>Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, and Slovenia.



**Table 1.** The incidence of overeducation by major region of origin, %

| Major geographical region | Normative measure of overeducation % | Statistical measure of overeducation % | Total obs. |
|---------------------------|--------------------------------------|--|------------|
| Native                    | 16                                   | 19                                     | 1,511,594  |
| EU15                      | 21                                   | 22                                     | 30,079     |
| EU10                      | 45                                   | 48                                     | 10,546     |
| EU3                       | 34                                   | 28                                     | 9,896      |
| EFTA                      | 19                                   | 19                                     | 2,641      |
| North Africa              | 25                                   | 20                                     | 10,931     |
| Near Middle East          | 34                                   | 33                                     | 2,143      |
| East Asia                 | 25                                   | 22                                     | 2,148      |
| South East Asia           | 38                                   | 29                                     | 8,329      |
| North America             | 29                                   | 32                                     | 1,547      |
| Central America           | 33                                   | 27                                     | 855        |
| South America             | 34                                   | 26                                     | 4,648      |
| Australia and Oceania     | 22                                   | 28                                     | 542        |
| Total                     | 17                                   | 19                                     | 1,642,096  |

Note: The figures cover 22 destination countries.

Source: Own calculations using EU LFS 2012.

**Table 2.** The incidence of overeducation by major occupational ISCO-08 group

| ISCO-08   | Immigrants |      |        | Natives |      |         |
|---|------------|------|--------|---------|------|---------|
|   | Norm       | Stat | Obs.   | Norm    | Stat | Obs.    |
| 1. Managers   | 0          | 1    | 2,883  | 0       | 1    | 81,254  |
| 2. Professionals                                      | 9          | 9    | 10,184 | 5       | 5    | 266,457 |
| 3. Technicians and support workers                    | 31         | 46   | 6,565  | 20      | 36   | 241,464 |
| 4. Clerical support workers                           | 37         | 38   | 3,836  | 21      | 23   | 143,813 |
| 5. Services and sales workers                         | 20         | 21   | 17,127 | 12      | 14   | 251,448 |
| 6. Skilled agricultural, forestry and fishery workers | 11         | 33   | 1,081  | 5       | 29   | 93,099  |
| 7. Craft and related trades workers                   | 9          | 12   | 9,512  | 5       | 9    | 184,632 |
| 8. Plant and machine operators and assemblers         | 13         | 16   | 6,719  | 4       | 6    | 118,622 |
| 9. Elementary occupations                             | 79         | 43   | 16,449 | 88      | 50   | 117,201 |

Norm, normative measure of overeducation; Stat, statistical measure of overeducation.

Note: The figures cover 22 destination countries.

Source: Own calculations using EU LFS 2012.

**Table 3.** The incidence of overeducation by gender and occupations status

|                         | Normative | Statistical |
|-------------------------|-----------|-------------|
| Female native           | 53.9      | 51.0        |
| Female immigrant        | 53.8      | 51.8        |
| Self-employed native    | 10.7      | 18.3        |
| Self-employed immigrant | 20.5      | 20.1        |
| Recent immigrant        | 40.4      | 36.6        |

Note: The figures cover 22 destination countries. Source: Own calculations using EU LFS 2012.

substantially when the statistical measure of overeducation is employed, sometimes significantly so, as in the case of native skilled agricultural, forestry, and fishery workers. The disparity is to be expected if we bear in mind that the normative measure groups a number of occupations into the same skill level (see [Annex A](#) for reference), which means less variability and by extension, a tendency to underestimate the level of mismatch.

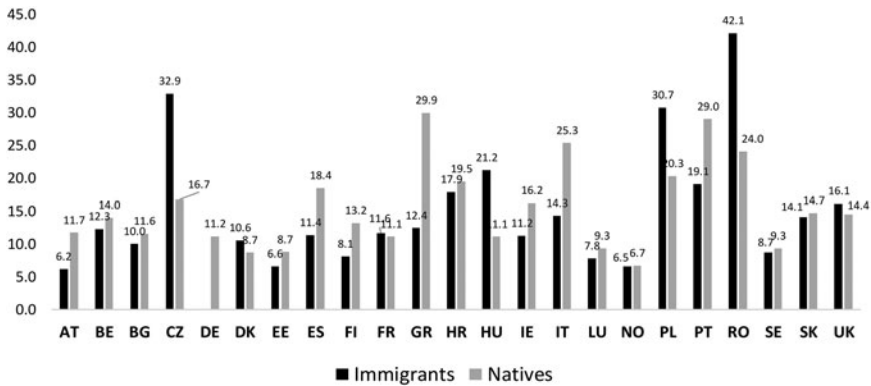
In terms of gender, women experience more overeducation than men, although interestingly, there does not seem to be much of a difference between native and immigrant women ([Table 3](#)). The incidence of overeducation among self-employed immigrants is higher than that of the corresponding native population, and almost half of all recent immigrants (with <5 years residence in the destination country) are mismatched. There are interesting differences to be noted between the two measures of overeducation, especially the sudden increase in the incidence of overeducation for self-employed natives.

### 4.3 The independent variable

The main independent variable is self-employment, a dummy variable equal to 1 if the individual is self-employed and 0 otherwise. The variable is derived from the “Professional status” variable in the EULFS, which includes three options: (i) self-employed with or without employees, (ii) employee, and (iii) family worker. The variable is based on the International Standard Classification of Status in Employment (ISCE)<sup>13</sup> developed by the ILO to measure the professional status of employed persons.

[Figure 1](#) presents the self-employment rates of both immigrants and natives by country of destination. Some interesting patterns seem to emerge. To begin with, some countries seem to generally exhibit higher rates of self-employment, regardless of the group analyzed. Consider, for instance, the case of Croatia, Romania, Czechia, Greece as opposed to Norway, Estonia, or Sweden. This would seem to point to specific institutional context and labor market policies which would shape the entrepreneurial decisions of individuals. A second pattern that emerges is that of countries with a significantly higher share of self-employment for immigrants than

<sup>13</sup>See [www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms\\_087562.pdf](http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087562.pdf)



**Figure 1.** Self-employment rates for non-EU immigrants and natives by country, 2012.

Source: Own calculations using EU-LFS 2012.

Notes: No observations on immigrants for Germany.

natives, where we can include Czechia, Poland, Hungary, Romania,<sup>14</sup> as well as to a lower extent, the UK, France, and Denmark. Although the first four countries also exhibit high self-employment rates for natives as well, the significantly higher share for immigrants seems to indicate the presence of labor market mechanisms that either push or pull immigrants into self-employment.

Figure 2 presents the incidence of overeducation (the normative measure) among self-employed individuals in each country. Strikingly, all countries exhibit a (sometimes very) large share of overeducated self-employed immigrants, larger than the share of overeducated self-employed natives. This, irrespective of whether the country presents a higher share of immigrant or native self-employment as seen in the previous graph.

#### 4.4 Control variables

The existing theories of immigrant overeducation already point to a number of relevant explanatory factors. The incidence of overeducation should decrease the longer the individual has been residing in the country, which enables the accumulation of local work experience and human capital. Previous literature has also found significant differences in mismatch by gender [see Groot and Maassen van den Brink (2000)]. General characteristics of the destination country economies, such as gross domestic product per capita and the unemployment rate of the native population, are also considered, factors found relevant by the existing literature. High levels of unemployment have direct implications for the assignment of workers to available jobs [Sattinger (1993)]. Competition for jobs is more intense generally and educated workers may compete with the less educated for any job available, irrespectively of occupation. Hence, we expect a higher overall incidence of overeducation in an economy with higher levels of unemployment [Aleksynska and Tritah (2013)].

<sup>14</sup>Romania presents a small overall sample of immigrants, which might drive the high rate of self-employment in this case.

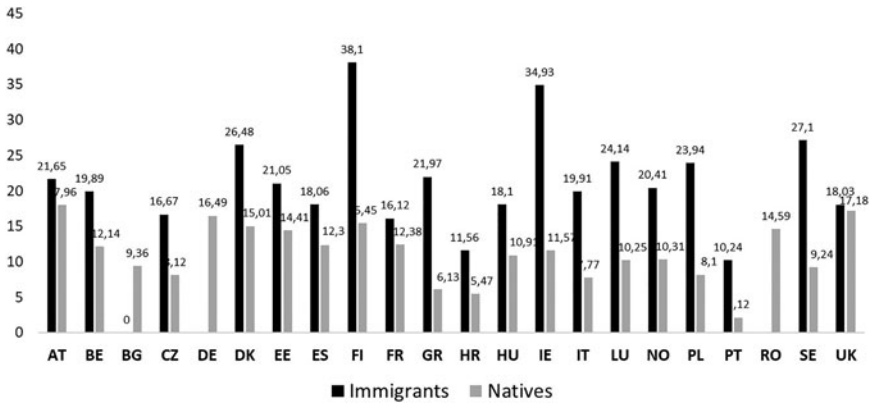


Figure 2. Incidence of overeducation for self-employed individuals, normative measure, 2012. Source: Own calculations using EU-LFS 2012.

4.5 The empirical model

In order to disentangle the effects of various factors on individual’s propensity to be overeducated, I use a three-step approach, specifically, a probit baseline model of all aggregate countries, a probit model of individual country effects, and a biprobit model of aggregate country effects. The baseline model is estimated as follows:

$$Y_i = I(Y_i > 0) = I(\beta_1 X_i + \beta_2 Z_i + \varepsilon_i > 0), \tag{1}$$

where  $Y_i$  is the main outcome variable, a dummy equal to 1 if the individual is overeducated and zero otherwise;  $I(.)$  is a binary indicator function taking the value 1 if the argument is true and 0 otherwise;  $X_i$  represents the explanatory variable self-employment, a dummy variable equal to 1 if the individual is self-employed,  $\beta_1$  its slope and the main parameter of interest,  $i$  refers to the cross-national units, while  $\varepsilon$  is the error term.  $Z_i$  represents a vector of the control variables previously mentioned, which include both individual- and country-level characteristics. Since the dependent variable has a discrete outcome, a probability model is more suitable than a linear regression model. Using the latter would result in biased and inconsistent estimates, because the fitted probabilities can be  $<0$  or  $>1$  (as they are not constrained to the unit interval), the model imposes heteroscedasticity and the partial effect of the explanatory variables (appearing in level form) is constant [Wooldridge (2013)].

An adjusted version of model (1) is used for individual country analyses, in the second step.<sup>15</sup> The difference this time is that the vector of control variables  $Z$  only includes individual-level characteristics—gender, years of residence in the country, and marital status.

$$Y_{AT} = I(Y_{AT} > 0) = I(\beta_1 X_{AT} + \beta_2 Z_{AT} + \varepsilon_{AT} > 0). \tag{2}$$

<sup>15</sup>AT stands for Austria, for illustrative purposes.

Equations (1) and (2) do not account for a potential endogeneity issue, which might stem from the fact that several unobserved factors could affect both the probability of being self-employed and the probability of being overeducated. If left unaccounted for, endogeneity will lead to inconsistent and biased estimates of equations (1) and (2). Given that both the dependent and the independent variables have discrete outcomes, thus both the first stage and the second stage equations are probit models, a maximum likelihood bivariate probit [Heckman (1978)] is the optimal choice. Any other two-stage model which would mimic 2SLS would produce inconsistent estimators<sup>16</sup> [Wooldridge (2002); Greene (2012)].

To account for endogeneity bias, I estimate the following empirical model, which simultaneously estimates equation (1) and the stage defined below:

$$X_i = I(X_i > 0) = I(\beta_3 Z_i + \mu_i > 0), \quad (3)$$

where  $X_i$  is a dummy variable equal to 1 if the individual is self-employed and 0 otherwise,  $Z_i$  is a vector of the same explanatory variables as used in equation (1), and  $\mu_i$  is the error term.

While the data source does not contain suitable candidates for a strong instrument<sup>17</sup> that would satisfy the exclusion restriction, two potential variables, derived from external sources, are included: (1) the number of patents per million population, and (2) expenditure on research and development as share of GDP, both variables at the regional level.<sup>18</sup> Since both instruments are regional-level variables, the biprobit model can only be applied to model (1), and not model (2) on individual country analyses.

There is an extensive literature that positively links the number of patents to increased entrepreneurship and self-employment [see Lee et al. (2004); Allred and Park (2007); Acs et al. (2009); Acs and Sanders (2012)]. The underlying mechanism behind this relationship has been formalized in innovation-driven models which argue that intellectual property rights, and thus patents, are key institutions that allow investors to market their inventions and thereby recover their costs [Acs and Sanders (2012)]. Patent creation should thus provide incentives for business formation to collect the benefits of this initial investment [Kitch (1977)]. The second instrument is derived from previous studies which have found that spillover effects from research and development contribute to business creation [see Acs and Varga (2005); Kirchoff et al. (2007)]. Research and development produce knowledge and ideas, which contribute to the creation of new services or goods, and thus new entrepreneurial opportunities. Expenditure on research and development as a share of GDP is employed in this context as a proxy for these entrepreneurial opportunities.

As mentioned, I obtain unbiased and asymptotically efficient estimates of the simultaneous equation model consisting of equations (1) and (3), by employing a maximum likelihood estimation of a bivariate probit model.

<sup>16</sup>Sometimes called the “forbidden regression” [Wooldridge (2002)].

<sup>17</sup>Since the survey (EU LFS) concerns labor market conditions and experience, most variables are related to both overeducation and self-employment.

<sup>18</sup>Data sourced from Eurostat’s regional statistics.

**Table 4.** Correlation matrix

| Variable           | Normative measure of overeducation | Statistical measure if overeducation |
|--------------------|------------------------------------|--------------------------------------|
| Self-employment    | -0.0677                            | -0.0108                              |
| Years of residence | 0.0606                             | 0.0260                               |
| Female             | 0.0678                             | 0.0439                               |
| Married            | -0.0353                            | -0.0210                              |
| GDP per capita     | 0.0233                             | 0.0445                               |
| Unemployment       | 0.0279                             | 0.0213                               |

**Table 5.** Determinants of overeducation, univariate models

|               | Normative measure of overeducation |                    | Statistical measure of overeducation |                 |
|---------------|------------------------------------|--------------------|--------------------------------------|-----------------|
|               | Probit immigrants                  | Probit natives     | Probit immigrants                    | Probit natives  |
| Self-employed | -0.11***<br>(0.12)                 | -0.07***<br>(0.01) | -0.06***<br>(0.1)                    | -0.01<br>(0.01) |
| Observations  | 74,727                             | 1,511,594          | 74,727                               | 1,511,594       |

*Data Source:* Data are from the 2012 EU LFS and cover 22 destination countries.

*Note:* Robust standard errors in parentheses, clustered at regional level. All coefficients have been transformed in average marginal effects.

\*Statistical significance at the 10% level. \*\*Statistical significance at the 5% level. \*\*\*Statistical significance at the 1% level.

## 5. Results

The paper investigates the effect of self-employment on immigrants' and natives' probability to be overeducated. This section presents the results of the empirical analysis.

I begin by exploring the correlation between overeducation and the variables used in the empirical specifications (Table 4). Self-employment appears to be negatively correlated with both measures of overeducation, albeit rather weakly. Overeducation also seems to be higher among women and to decrease with for married individuals.

The probability of a self-employed immigrant or native to be overeducated is summarized in a parsimonious model in Table 5, where both measures of overeducation are presented, for comparison purposes (the table presents average marginal effects). To begin with, if we consider the normative measure, the probability of being overeducated decreases for the self-employed, by 11 percentage points for immigrants and 7 percentage points for natives. For immigrants, the effect seems to be slightly larger, although a *t*-test indicates the difference is not statistically significant. The analysis using the statistical measure of overeducation seems to confirm the results for immigrants, albeit with a slightly lower magnitude, but not for natives.

Table 6 presents a multivariate model, in which a number of control variables are added. The same pattern emerges, although overall, the magnitude of the effects is lowered by the introduction of covariates. Being female increases the likelihood of

**Table 6.** Determinants of overeducation, multivariate models

|                    | Normative measure of overeducation |                     | Statistical measure of overeducation |                     |
|--------------------|------------------------------------|---------------------|--------------------------------------|---------------------|
|                    | Probit immigrants                  | Probit natives      | Probit immigrants                    | Probit natives      |
| Self-employed      | -0.07***<br>(0.01)                 | -0.06***<br>(0.01)  | -0.03***<br>(0.01)                   | 0.004<br>(0.01)     |
| Female             | 0.09***<br>(0.02)                  | 0.04***<br>(0.003)  | 0.06***<br>(0.02)                    | 0.03***<br>(0.004)  |
| Years of residence | -0.06***<br>(0.004)                | /                   | -0.04***<br>(0.004)                  | /                   |
| Married            | 0.01*<br>(0.01)                    | -0.03***<br>(0.003) | 0.09<br>(0.01)                       | -0.02***<br>(0.004) |
| Unemployment       | -0.001<br>(0.001)                  | 0.003***<br>(0.001) | 0.004*<br>(0.002)                    | 0.003<br>(0.002)    |
| GDP per cap        | 1.60<br>(4.97)                     | 1.47<br>(3.04)      | 1.77**<br>(6.92)                     | 1.09**<br>(5.09)    |
| Observations       | 69,657                             | 1,501,432           | 69,657                               | 1,501,433           |

*Data Source:* Data are from the 2012 EU LFS and cover 22 destination countries.

*Note:* Robust standard errors in parentheses, clustered at regional level. All coefficients have been transformed in average marginal effects. \*Statistical significance at the 10% level. \*\*Statistical significance at the 5% level. \*\*\*Statistical significance at the 1% level.

being overeducated for both immigrants and natives, the likelihood that seems to slightly decrease with age for natives. As hypothesized, the years of residence in the country seem to decrease overeducation. GDP per capita, a proxy for the level of economic development of a country, seems to positively contribute to mismatch.

Table 7 presents the probit regressions for individual countries. A number of interesting observations can be made based on these results. To begin with, and as expected, there is substantial variation in the effect for individual countries as opposed to the overall sample. While the probability of immigrants being overeducated decreases with self-employment for countries such as Denmark, Estonia, or Spain, it seems to be increasing for countries such as Finland or Luxembourg, although these results are not statistically significant. The same applies in the case of natives, with Austria experiencing a positive effect, while Belgium, for instance, experiencing a negative one.

Secondly, in some countries, including Denmark, Belgium, Spain, Italy, Luxembourg, and Portugal, we observe comparable effects for natives and immigrants. This observation would imply the existence of a broader institutional context shaping the interaction between self-employment and overeducation for both groups, in a similar way. Conversely, in countries such as Estonia or the UK, the probability of being overeducated decreases more for self-employed immigrants than for self-employed natives (for which the results are inconsistent across the two overeducation measures), pointing to the existence of group-specific differences.

**Table 7.** Determinants of overeducation, country analysis, probit regression

| Country        | Normative measure of overeducation  |                                  | Statistical measure of overeducation |                                  |
|----------------|-------------------------------------|----------------------------------|--------------------------------------|----------------------------------|
|                | Natives                             | Immigrants                       | Natives                              | Immigrants                       |
| Overall sample | -0.06***<br>(0.01)<br>N = 1,501,432 | -0.07***<br>(0.01)<br>N = 69,657 | 0.004<br>(0.01)<br>N = 1,501,432     | -0.03***<br>(0.01)<br>N = 69,657 |
| AT             | 0.05***<br>(0.004)<br>N = 71,120    | -0.08***<br>(0.02)<br>N = 6,458  | 0.22***<br>(0.01)<br>N = 71,120      | 0.004<br>(0.02)<br>N = 6,458     |
| BE             | -0.06***<br>(0.01)<br>N = 34,128    | -0.09***<br>(0.03)<br>N = 2,747  | -0.01**<br>(0.01)<br>N = 34,128      | -0.05**<br>(0.03)<br>N = 2,747   |
| BG             | -0.09***<br>(0.01)<br>N = 11,169    | Insufficient obs                 | 0.05***<br>(0.01)<br>N = 11,169      | Insufficient obs                 |
| CZ             | -0.04***<br>(0.06)<br>N = 18,476    | 0.00001<br>(0.06)<br>N = 213     | -0.01*<br>(0.01)<br>N = 18,476       | 0.01<br>(0.06)<br>N = 213        |
| DE             | 0.03***<br>(0.003)<br>N = 198,843   | No obs                           | 0.08***<br>(0.003)<br>N = 198,843    | No obs                           |
| DK             | -0.14***<br>(0.01)<br>N = 63,690    | -0.15***<br>(0.03)<br>N = 2,129  | -0.05***<br>(0.01)<br>N = 63,690     | -0.08**<br>(0.03)<br>N = 2,129   |
| EE             | -0.05***<br>(0.01)<br>N = 10,161    | -0.11**<br>(0.06)<br>N = 1,046   | 0.05***<br>(0.02)<br>N = 10,161      | -0.17***<br>(0.06)<br>N = 1,046  |
| ES             | -0.16***<br>(0.01)<br>N = 32,921    | -0.16***<br>(0.03)<br>N = 1,846  | -0.09***<br>(0.01)<br>N = 32,921     | -0.10***<br>(0.03)<br>N = 1,846  |
| FI             | -0.02**<br>(0.01)<br>N = 12,461     | 0.12<br>(0.12)<br>N = 250        | 0.07***<br>(0.01)<br>N = 12,461      | 0.09<br>(0.11)<br>N = 250        |
| FR             | -0.04***<br>(0.003)<br>N = 165,063  | -0.06***<br>(0.01)<br>N = 13,608 | -0.003<br>(0.003)<br>N = 165,063     | -0.02<br>(0.01)<br>N = 13,608    |

(Continued)



Table 7. (Continued.)

| Country | Normative measure of overeducation |            | Statistical measure of overeducation |            |
|---------|------------------------------------|------------|--------------------------------------|------------|
|         | Natives                            | Immigrants | Natives                              | Immigrants |
| GR      | -0.08***                           | -0.03      | -0.09***                             | 0.09***    |
|         | (0.004)                            | (0.03)     | (0.01)                               | (0.03)     |
|         | N = 24,595                         | N = 2,433  | N = 24,595                           | N = 2,433  |
| HR      | -0.05***                           | -0.06*     | -0.01                                | 0.003      |
|         | (0.01)                             | (0.03)     | (0.01)                               | (0.03)     |
|         | N = 11,464                         | N = 946    | N = 11,464                           | N = 946    |
| HU      | -0.06***                           | -0.02      | 0.04***                              | 0.02       |
|         | (0.004)                            | (0.05)     | (0.004)                              | (0.05)     |
|         | N = 82,031                         | N = 477    | N = 82,031                           | N = 477    |
| IE      | -0.12***                           | -0.02      | -0.02***                             | 0.02       |
|         | (0.004)                            | (0.03)     | (0.005)                              | (0.03)     |
|         | N = 70,067                         | N = 3,502  | N = 70,067                           | N = 3,502  |
| IT      | -0.08***                           | -0.11***   | -0.03***                             | -0.07***   |
|         | (0.002)                            | (0.01)     | (0.002)                              | (0.01)     |
|         | N = 181,487                        | N = 15,460 | N = 181,487                          | N = 15,460 |
| LU      | 0.03*                              | 0.03       | 0.11***                              | 0.09       |
|         | (0.01)                             | (0.07)     | (0.02)                               | (0.07)     |
|         | N = 5,142                          | N = 718    | N = 5,142                            | N = 718    |
| NO      | -0.02                              | -0.14**    | -0.01                                | -0.09      |
|         | (0.01)                             | (0.06)     | (0.01)                               | (0.06)     |
|         | N = 11,897                         | N = 721    | N = 11,897                           | N = 721    |
| PL      | -0.09***                           | 0.03       | 0.04***                              | 0.03       |
|         | (0.002)                            | (0.07)     | (0.003)                              | (0.07)     |
|         | N = 159,639                        | N = 209    | N = 159,639                          | N = 209    |
| PT      | -0.09***                           | -0.09***   | -0.04**                              | -0.04**    |
|         | (0.002)                            | (0.02)     | (0.002)                              | (0.02)     |
|         | N = 39,845                         | N = 1,937  | N = 39,845                           | N = 1,937  |
| RO      | 0.03***                            | No obs     | 0.04***                              | No obs     |
|         | (0.003)                            |            | (0.003)                              |            |
|         | N = 88,367                         |            | N = 88,367                           |            |
| SE      | -0.03***                           | -0.001     | -0.01***                             | 0.003      |
|         | (0.003)                            | (0.01)     | (0.003)                              | (0.01)     |
|         | N = 151,012                        | N = 13,041 | N = 151,012                          | N = 13,041 |

(Continued)

Table 7. (Continued.)

| Country | Normative measure of overeducation |            | Statistical measure of overeducation |            |
|---------|------------------------------------|------------|--------------------------------------|------------|
|         | Natives                            | Immigrants | Natives                              | Immigrants |
| UK      | -0.04***                           | -0.11***   | -0.02***                             | -0.08***   |
|         | (0.07)                             | (0.02)     | (0.01)                               | (0.02)     |
|         | N = 29,493                         | N = 2,859  | N = 29,493                           | N = 2,859  |

Data Source: EU-LFS, 2012.

Note: Robust standard errors in parentheses. All coefficients have been transformed in average marginal effects. All regressions include controls for gender, years of residence in the country and marital status. \*Statistical significance at the 10% level. \*\*Statistical significance at the 5% level. \*\*\*Statistical significance at the 1% level.

Lastly, as previously noted, differences can be observed between the direction and magnitude of the effect between the two measures of overeducation for most of the countries.

### 5.1 Endogeneity

As previously mentioned, immigrants might become self-employed precisely because they are overeducated for the job they perform, in which case overeducation has an influence on the decision to become self-employed. Thus, the dependent and main explanatory variables might be endogenous. To account for a potential endogeneity bias, I employ a maximum likelihood bivariate probit model. Table 8 presents a parsimonious model which includes only the main independent variable. Although it maintains the same direction, the effect of being self-employment on the probability of being an overeducated immigrant decreases and loses its significance for the normative measure of overeducation, while it increases for the statistical measure, when compared to the baseline probit model. The same effect happens in the case of natives, in the statistical model, while we observe a quite abnormal result in the normative model.

These results change significantly, when I introduce covariates (see Table 9). In the case of immigrants, the covariates drive the effect down, with loss of significance, whereas in the case of natives, the covariates seem to push the effect up, with strong significance levels across the board.

Overall, in the case of immigrants, both the probit and the biprobit analyses seem to indicate that being self-employed decreases the probability of being overeducated. This effect is seen across all specifications and with both measures of overeducation, albeit with various magnitudes and significance levels. The magnitude of the effect decreases with the introduction of covariates, while the significance level disappears in the biprobit models, which present the strictest specifications. The results are slightly more heterogeneous in the case of natives, with no clear pattern emerging.

Importantly, because the instruments used present regional-level values, I cannot conduct the same biprobit analysis for individual countries. This methodology would restrict the sample to, in some cases, two or even one cluster, not enough for an analysis. Table 7, therefore, presents a descriptive (and quite interesting) perspective on whether being self-employed decreases or increases the probability of being overeducated for immigrants and natives, across individual countries.

**Table 8.** Determinants of overeducation, accounting for endogeneity

|               | Normative measure of overeducation |                  | Statistical measure of overeducation |                  |
|---------------|------------------------------------|------------------|--------------------------------------|------------------|
|               | Biprobit immigrants                | Biprobit natives | Biprobit immigrants                  | Biprobit natives |
| Self-employed | -0.04                              | 0.002            | -0.09***                             | -0.10***         |
|               | (0.08)                             | (0.02)           | (0.004)                              | (0.03)           |
| Observations  | 74,727                             | 1,511,594        | 74,727                               | 1,511,594        |

Data Source: Data are from the 2012 EU LFS and covers 22 destination countries.

Note: Robust standard errors in parentheses, clustered at regional level. All coefficients have been transformed in average marginal effects. \*Statistical significance at the 10% level. \*\*Statistical significance at the 5% level. \*\*\*Statistical significance at the 1% level.

**Table 9.** Determinants of overeducation, accounting for endogeneity

|                    | Normative measure of overeducation |                  | Statistical measure of overeducation |                  |
|--------------------|------------------------------------|------------------|--------------------------------------|------------------|
|                    | Biprobit immigrants                | Biprobit natives | Biprobit immigrants                  | Biprobit natives |
| Self-employed      | -0.01                              | -0.14***         | -0.02                                | -0.13***         |
|                    | (0.01)                             | (0.004)          | (0.02)                               | (0.004)          |
| Female             | -0.01**                            | -0.08***         | -0.02*                               | -0.08***         |
|                    | (0.01)                             | (0.004)          | (0.01)                               | (0.004)          |
| Years of residence | -0.0004                            | /                | 0.001                                | /                |
|                    | (0.001)                            | /                | (0.002)                              | /                |
| Married            | 0.01**                             | 0.05***          | 0.01**                               | 0.05***          |
|                    | (0.003)                            | (0.003)          | (0.003)                              | (0.003)          |
| Unemployment       | 0.0003                             | 0.004***         | 0.001                                | 0.004***         |
|                    | (0.0004)                           | (0.001)          | (0.001)                              | (0.001)          |
| GDP per cap        | -3.33**                            | -1.42***         | -1.56                                | -1.32***         |
|                    | (1.64)                             | (2.63)           | (1.93)                               | (2.93)           |
| Observations       | 69,657                             | 1,501,433        | 69,657                               | 1,501,433        |
| Wald statistic     | $P = 0.9702$                       | $P = 0.0049$     | $P = 0.4356$                         | $P = 0.0000$     |

Data Source: Data are from the 2012 EU LFS and covers 22 destination countries.

Note: Robust standard errors in parentheses, clustered at regional level. All coefficients have been transformed in average marginal effects.

\*Statistical significance at the 10% level. \*\*Statistical significance at the 5% level. \*\*\*Statistical significance at the 1% level.

## 6. Discussion

The paper explores the effect of being self-employed on the probability of being overeducated, in a comparative analysis between immigrants and natives. Controlling for a list of demographic characteristics and general characteristics of the destination

country, the results seem to suggest that in the case of immigrants, self-employed individuals are generally less likely to be overeducated. This probability seems to decrease with the number of years of residence in the country of destination for immigrants, and to be higher for females in both groups. This would confirm the findings of Sanchez et al. (2015), who conduct a similar analysis in a longitudinal study. If correct, the results would imply that self-employment represents a strategy to minimize overeducation, at least for immigrants. By virtue of being outsiders to the labor market, immigrants encounter more barriers to employment, which make them more likely to be overeducated. In order to minimize or avoid overeducation altogether, immigrants can become self-employed. This hypothesis could help explain the higher incidence of self-employment that immigrants exhibit, when compared to natives. To confirm it, however, a longitudinal study, in a similar fashion to Sanchez et al. (2015), following immigrants in and out of self-employment and investigating how overeducation fluctuates, would be necessary and desirable. Nevertheless, the results are important and provide insight into a phenomenon which has been long hypothesized, but little researched.

Importantly, there are significant cross-national differences in this effect. While some countries exhibit a negative relation between self-employment and overeducation for both immigrants and natives, others present different effect for each group. These differences point to the existence of different labor market regimes and institutional settings with which immigrants interact and which create incentives and opportunities for self-employment and labor market matching. The differences might be also caused by the different ways in which countries measure education and economic activities. Although the EU-LFS provides a high degree of cross-country comparability by using international standards of classification, each country presents its own idiosyncratic system of qualifications and requirements. For instance, a vocational training and education college can be perceived as tertiary education in one country, or as post-secondary, non-tertiary education in another. This would lead to a different counting of overeducation in the two different countries, at least in the case of the normative measure.

The findings have also broader research and policy implications and contribute to scholarship in a number of ways. To start with, they confirm overeducation's sensitivity to definition and measurement. The normative measure of overeducation seems to generally underestimate the incidence of overeducation, with some exceptions. Further, while self-employment seems to decrease the probability of an individual being overeducated when we employ the normative measure, the results are not as clear-cut when the statistical measure is used instead. This sensitivity has been remarked in previous studies [see Groot and Maassen van den Brink (2000); CEDEFOP (2010)], and should be accounted for when translating these studies into policy-making.

Further research, however, should look into the nature of self-employment, as it is unclear at the moment whether this type would be productive, or more akin to necessity self-employment. The difference has important implications for policy-making. The latter has been associated with low productivity, job creation, and job satisfaction, which in the long term would represent an underutilization of human resources and a failure to tap into the potential that immigration represents. The former is the type of self-employment that policy-makers would want to incentivize, that brings about innovation and job creation.

Another implication of these results is that, by implementing measures to promote opportunity self-employment, policy-makers could achieve two objectives with one instrument—increase entrepreneurship *and* decrease mismatch. There is no doubt that if countries intend to make themselves attractive destinations for “the best and the brightest”, they need to tackle these labor market inefficiencies and promote a business creation-friendly environment. This in turn would help smooth out the socio-economic integration of immigrants, who could more easily become productive members of society. More research, however, is needed to understand the exact dynamic between these two labor market processes and how it changes over time and space.

No study is bound to be without limitations, and the present one is no exception. One significant issue right from the start is the potential endogeneity bias, addressed in the methodological section with maximum likelihood bivariate probit estimation. This is the most fitting model for analyses including both a binary-dependent and a binary-independent variable, as it is this case. The model includes two additional variables used as regressors of self-employment, which fulfil the exclusion restriction of not being correlated with the error term. Lastly, the results of the study and their implications are bound to be dependent on the context and the time of the analysis.

**Acknowledgements.** The author acknowledges the financial support of the EDUWORKS Marie Curie Initial Training Network Project (PITN-GA-2013-608311) of the European Commission’s 7th Framework Program. The author is particularly grateful to Michael Dorsch, Martin Kahanec, and two anonymous referees for useful discussions that helped improve this manuscript significantly.

## References

- Acs, Zoltan J., Pontus Braunerhjelm, David B. Audretsch, and Bo Carlsson (2009) The knowledge spillover theory of entrepreneurship. *Small Business Economics* 32(1), 15–30. <https://doi.org/10.1007/s11187-008-9157-3>.
- Acs, Zoltan and Attila Varga (2005) Entrepreneurship, agglomeration and technological change. *Small Business Economics* 24(3), 323–334. <https://doi.org/10.1007/s11187-005-1998-4>.
- Acs, Zoltan J. and Mark Sanders (2012) Patents, knowledge spillovers, and entrepreneurship. *Small Business Economics* 39(4), 801–817. <https://doi.org/10.1007/s11187-011-9322-y>.
- Aleksynska, Mariya and Ahmed Tritah (2013) Occupation–education mismatch of immigrant workers in Europe: context and policies. *Economics of Education Review* 36(October), 229–244. <https://doi.org/10.1016/j.econedurev.2013.06.001>.
- Allen, Jim and Rolf Van der Velden (2001) Educational mismatches versus skill mismatches: effects in wages, job satisfaction, and on-the-job search. *Oxford Economic Papers* 3, 434–452.
- Allred, Brent B. and Walter G. Park (2007) The influence of patent protection on firm innovation investment in manufacturing industries. *Journal of International Management* 13(2), 91–109. <https://doi.org/10.1016/j.intman.2007.02.001>.
- Arrow, Kenneth (1973) Higher education as a filter. *Journal of Public Economics* 2, 193–216.
- Battu, H. and P. J. Sloane (2004) Over-education and ethnic minorities in Britain. *The Manchester School* 72(4), 535–559. <https://doi.org/10.1111/j.1467-9957.2004.00407.x>.
- Bender, Keith and Kristen Roche (2013) Educational mismatch and self-employment. *Economics of Education Review* 34, 85–95.
- Bennett, Jessica and Seamus McGuinness (2009) Assessing the impact of skill shortages on the productivity performance of high-tech firms in Northern Ireland. *Applied Economics* 41(6), 727–737.
- CEDEFOP (2010) The Skill Matching Challenge. Analysing Skill Mismatch and Policy Implications. Luxembourg: Publications Office of the European Union.
- Chevalier, Arnaud (2003) Measuring over-education. *Economica* 70(279), 509–531.
- Chiswick, Barry R (1978) The effect of Americanization on the earnings of foreign-born men. *Journal of Political Economy* 86(5), 897–921. <https://doi.org/10.1086/260717>.

- Chiswick, Barry R. and Paul W. Miller (1992) Language in the immigrant labour market. In B. R. Chiswick (ed.), *Immigration, Language and Ethnicity: Canada and the United States*, pp. 229–296. Washington, DC: AEI Press.
- Chiswick, Barry R. and Paul W. Miller (2007) The international transferability of immigrants' human capital skills. IZA Discussion Paper No. 2670, March.
- Chiswick, Barry R. and Paul W. Miller (2009) The international transferability of immigrants' human capital. *Economics of Education Review* 28(2), 162–169. <https://doi.org/10.1016/j.econedurev.2008.07.002>.
- Duncan, Greg and Saul Hoffman (1981) The incidence and wage effects of overeducation. *Economics of Education Review* 1(1), 75–86.
- Eurostat (2007) Labour Force Survey in the EU, Candidate and EFTA Countries. Luxembourg: Office for Official Publications of the European Communities.
- Ghaffarzadegan, Navid, Yi Xue, and Richard C. Larson (2017) Work-education mismatch: an endogenous theory of professionalization. *European Journal of Operational Research* 261(3), 1085–1097. <https://doi.org/10.1016/j.ejor.2017.02.041>.
- Greene, William H. (2012) *Econometric Analysis*. Upper Saddle River, NJ 07458: Pearson Education Inc. Seventh ed.
- Groot, Wim and Henriëtte Maassen van den Brink (2000) Overeducation in the labor market: a meta-analysis. *Economics of Education Review* 19(2), 149–158. [https://doi.org/10.1016/S0272-7757\(99\)00057-6](https://doi.org/10.1016/S0272-7757(99)00057-6).
- Guironnet, J.-P. and N. Peypoch (2007) Human capital allocation and overeducation: a measure of French productivity (1987, 1999). *Economic Modelling* 24(3), 398–410. <https://doi.org/10.1016/j.econmod.2006.09.003>.
- Hartog, Joop (2000) Over-education and earnings: where are we, where should we go? *Economics of Education Review* 19, 131–147.
- Heckman, James J (1978) Dummy endogenous variables in a simultaneous equation system. *Econometrica* 46(4), 931. <https://doi.org/10.2307/1909757>.
- ILO (2012) ISCO-08. Volume 1. International Standard Classification of Occupations. Structure, Group Definitions and Correspondence Tables. International Labour Office, Geneva.
- ILO (2014) Skills Mismatch in Europe. International Labour Office, Geneva. Statistics Brief.
- Jensen, Uwe, Hermann Gartner, and Susanne Rässler (2010) Estimating German overqualification with stochastic earnings frontiers. *ASta Advances in Statistical Analysis* 94(1), 33–51. <https://doi.org/10.1007/s10182-009-0121-6>.
- Kiker, B.F., Maria C. Santos, and M. Mendes de Oliveira (1997) Overeducation and undereducation: evidence for Portugal. *Economics of Education Review* 16(2), 111–125. [https://doi.org/10.1016/S0272-7757\(96\)00040-4](https://doi.org/10.1016/S0272-7757(96)00040-4).
- Kirchhoff, Bruce A., Scott L. Newbert, Iftekhar Hasan, and Catherine Armington (2007) The influence of university R & D expenditures on new business formations and employment growth. *Entrepreneurship Theory and Practice* 31(4), 543–559. <https://doi.org/10.1111/j.1540-6520.2007.00187.x>.
- Kitch, Edmund W. (1977) The nature and function of the patent system. *The Journal of Law and Economics* 20(2), 265–290. <https://doi.org/10.1086/466903>.
- Lee, Sam Youl, Richard Florida, and Zoltan Acs (2004) Creativity and entrepreneurship: a regional analysis of new firm formation. *Regional Studies* 38(8), 879–891. <https://doi.org/10.1080/0034340042000280910>.
- Mavromaras, Kostas and Seamus McGuinness (2012) Overskilling dynamics and education pathways. *Economics of Education Review* 31(5), 619–628. <https://doi.org/10.1016/j.econedurev.2012.02.006>.
- Mendes de Oliveira, M., M. C. Santos, and B. F. Kiker (2000) The role of human capital and technological change in overeducation. *Economics of Education Review* 19(2), 199–206. [https://doi.org/10.1016/S0272-7757\(99\)00020-5](https://doi.org/10.1016/S0272-7757(99)00020-5).
- Piracha, Matloob and Florin Vadean (2012) Migrant educational mismatch and the labour market. IZA Discussion Paper No. 6414.
- Piracha, Matloob and Florin Vadean (2013) Migrant educational mismatch and the labor market. In A. Constant & K. Zimmermann (eds), *International Handbook on the Economics of Migration*, pp. 176–192. Cheltenham, UK and Massachusetts, USA: Edward Elgar Publishing. <https://doi.org/10.4337/9781782546078>.

- Reynolds, Paul, Niels Bosma, Erkki Autio, Steve Hunt, Natalie De Bono, Isabel Servais, Paloma Lopez-Garcia, and Nancy Chin (2005) Global entrepreneurship monitor: data collection design and implementation 1998–2003. *Small Business Economics* 24(3), 205–231.
- Sanchez, Judit Albiol, Luis Diaz-Serrano, and Mercedes Teruel. (2015) Is self-employment a way to escape from skill mismatches? IZA Discussion Paper No. 9008.
- Sattinger, Michael (1993) Assignment Models of the Distribution of Earnings. *Journal of Economic Literature* 31(2), 125–154.
- Sicherman, Nachum and Oded Galor (1990) A theory of career mobility. *Journal of Political Economy* 98(1), 169–192.
- Spence, Michael (1974) *Market Signalling: Informational Transfer in Hiring and Related Screening Processes*. Cambridge, Mass: Harvard University Press.
- Sutherland, John (2012) Qualifications mismatch and skills mismatch. *Education & Training* 54(7), 619–632.
- Tarvid, Alexander. (2012) Effects of the field of higher education on the prevalence of over-education in European Countries. In J. E. Kesner (ed.), *Education: Evaluation, Reform and Policy*, pp. 167–184. Athens: Athens Institute for Education and Research.
- Tsang, Mun C. and Henry M. Levin (1985) The economics of overeducation. *Economics of Education Review* 4(2), 93–104.
- Ulceluse, Magdalena and Martin Kahanec (2018) Self-employment as a vehicle for labour market integration of immigrants and natives: the role of employment protection legislation. *International Journal of Manpower* 39(8), 1064–1079. <https://doi.org/10.1108/IJM-10-2018-0332>.
- Verhaest, Dieter and Eddy Omey (2010) The determinants of overeducation: different measures, different outcomes? *International Journal of Manpower* 31(6), 608–625. <https://doi.org/10.1108/01437721011073337>.
- Wooldridge, Jeffrey (2002) *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts, London, England: MIT Press.
- Wooldridge, Jeffrey (2013) *Introductory Econometrics. A Modern Approach*. Mason, Ohio: South-Western, Cengage Learning. Fifth ed.

**Annex A:** Correspondence between ISCO-08 and ISCED-11

| International Standard Classification of Occupations Major Groups (ISCO-08) | ISCO-08 Skill Level | International Standard Classification of Education (ISCED-97)                         |
|---|---------------------|---|
| 9. Elementary Occupations   | 1                   | 1. Primary level of education   |
| 8. Plant and Machine Operators, and Assemblers                              | 2                   | 2. Lower secondary level of education   |
| 7. Craft and Related Trades Workers   |                     | 3. Upper secondary level of education   |
| 6. Skilled Agricultural, Forestry and Fishery Workers                       |                     | 4. Post-secondary, non-tertiary education   |
| 5. Services and Sales Workers   |                     |   |
| 4. Clerical Support Workers   |                     |   |
| 3. Technicians and associate professional                                   | 3                   | 5b. First stage of tertiary education (short or medium duration)                      |
| 2. Professionals  | 4                   | 5a. First stage of secondary education, 1 <sup>st</sup> degree (medium duration)      |
| 1. Managers <sup>19</sup>   |                     | 6. Second stage of tertiary education (leading to an advanced research qualification) |

Source: ILO (2012).

<sup>19</sup>Managers fall under the 3+4 skill levels.

**Cite this article:** Ulceluse M (2020). Self-employment as a stepping stone to better labor market matching: a comparison between immigrants and natives. *Journal of Demographic Economics* 86, 479–501. <https://doi.org/10.1017/dem.2020.1>