IZA Journal of Migration
a SpringerOpen Journal

ORIGINAL ARTICLE

Open Access

CrossMark

Skill mismatch among migrant workers: evidence from a large multi-country dataset

Stefano Visintin[*], Kea Tijdens and Maarten van Klaveren

* Correspondence:
stefano.visintin@gmail.com
Amsterdam Institute for Advanced
Labour Studies (AIAS), University of
Amsterdam, Amsterdam, the
Netherlands

Abstract

This article unravels the migrants' incidence of skill mismatch taking into consideration different migration flows. Mismatch is the situation in which workers have jobs for which lower skill levels are required compared to their education. We use a dataset (from a large multi-country web survey) particularly suited to investigate differences in skill mismatch between native and migrant workers. The main advantages are its ample size and the large variety of country of origin and destination combinations, which allows for detailed analysis of different migration flows. This provides an innovative multi-country perspective, including nations and migrants from all continents. We also identify the relation between overeducation and some of the most widely accepted theoretical explanations for the phenomenon among native workers and test whether it holds for migrants. These results are achieved by fulfilling three research objectives, which are to investigate (1) the factors affecting overeducation and whether migrants are more often overqualified, (2) the relation between overeducation and different country of origin and destination combinations, and (3) whether a range of theoretically based assumptions affect the incidence of overeducation and the extent to which they are relevant in the case of migrant workers. Skill mismatch is found to be more common among migrants compared to native workers, although the incidence differs across migrants depending on the country of residence. Differences in the incidence of overeducation between native and migrant workers are not only related to the country of residence but also to the combination of country of origin and destination. When theoretically based assumptions are used to explain overeducation, the relation found for the total population does not always hold in the case of migrants. All these findings are confirmed by both an explorative and a in-depth analysis.

JEL codes: J24; J61; J15

Keywords: Overeducation; Mismatch; Migration; Workers; Job-skill match; Job-mobility

1. Introduction

Is overeducation more common among migrants compared to native workers? If so, is the overeducation incidence alike across migrants from various home countries and across various host countries? Are the arguments behind overeducation the same for native and migrant workers? This article unravels the migrants' incidence of skill mismatch, defined as the situation in which workers have jobs for which lower skill levels are required compared to their current educational level. The focus is on the skill mismatch of almost 700,000 native and migrant workers in 86 countries[1] over the period

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 2 of 34

2008–2013. Migrants are defined as workers not born in the country where they are currently living. In the sample, they originate from 79 countries and therefore represent a highly heterogeneous group, ranging from refugees and those who have migrated for economic reasons, to expats, intercultural married couples and others.

The academic discourse on mismatch in the labour market covers such issues as residential mismatch and hours mismatch, but this article focuses on what is generally considered the most relevant sort of mismatch: skill mismatch. Furthermore, the literature in the field often considers skill mismatch a unique phenomenon, treating overeducation and undereducation as one. Since they are two separate circumstances, in this paper we solely focus on overeducation and its explanations, leaving undereducation to further studies. The literature on skill mismatch can be classified into three categories. A number of studies have investigated the incidence of over- and undereducation, some of which provide breakdowns for specific groups in the labour market, such as gender and firm size. Many studies have addressed the impact of over- and undereducation, mostly on wages. Finally, an important body of knowledge relates to the dynamics of overeducation, that is, how educational requirements and the educational composition of the workforce have changed over time. As Leuven and Oosterbeek (2011) point out in their overview study, few studies have addressed the incidence of overeducation among migrants, although the literature has grown since they performed their study.

In our study, we use a dataset that is particularly suited to investigate differences in skill mismatch between native and migrant workers and that allows us to distinguish between different country of origin and destination combinations. This article extends the body of knowledge on migrants' overeducation in two ways. Firstly, we provide a multi-country perspective including nations and migrants from all continents, whereas most of the existing literature relies on single-country studies. Secondly, we test the specific relation between migrants' overeducation and some of the most widely accepted theoretical explanations for the phenomenon. We achieve these results by fulfilling three research objectives, which are to investigate (1) the factors affecting overeducation and whether migrants are more often overqualified compared to native workers, (2) the relation between overeducation and different country of origin and destination combinations, and (3) whether a range of theoretically based assumptions affect the incidence of overeducation and the extent to which they are relevant in the case of migrant workers.

The outline of this article is as follows. Section 2 concerns the theoretical and empirical literature with regard to the skill mismatch of migrant and native workers. In section 3, we describe the data and methods used. We present our results in section 4, and in section 5 we discuss our findings and conclusions.

2. Review of the literature on migrants' skill mismatch

2.1 What is skill mismatch?

Skill mismatch refers to the mismatch between a worker's educational attainment and the requirements of his or her job, whereby several types of skill mismatch are distinguished (for example, McGuinness and Sloane 2011). A vertical mismatch refers to a worker whose level of education is either above or below the educational level required for his or her job. Here, the terms 'overeducation' (also referred to as 'overschooling') and 'undereducation' are used. Educational level is a crude measure for indicating an individual's educational attainment or

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 3 of 34

job requirements. For jobs, the skill-based approach seems more adequate, as are the terms 'overskilling' and 'underskilling'. However, skills are more difficult to measure than educational attainment. The most common method is to measure an individual's generic skills, for example in cognitive tests or in the OECD's IALS and PIAAC literacy surveys, whereas job-specific skill requirements are hardly used because they are far more difficult to measure. A horizontal mismatch refers to a worker who is educated in a field other than the one that his or her job requires. Particularly in Germany, the concept of occupational mismatch is clearly distinguished from that of educational mismatch because of the country's widespread vocational training system, which provides the majority of the labour force with a generally accepted qualification for a wide range of occupations (Burkert and Seibert 2007). This article focuses solely on vertical skill mismatch, and particularly on overeducation.

Studying skill mismatch requires information about the educational attainment of individuals as well as insight into the educational level required for jobs. The former is less subject to dispute than the latter. In country-specific surveys the educational attainment of individuals is measured mainly in terms of national educational categories. For cross-country comparisons, the ISCED classification—which distinguishes seven educational attainment levels—is most often used (OECD 1999). In order to collect information about the educational requirements of jobs, the most frequently applied method is asking individual workers to indicate the educational attainment required for their job or whether they have sufficient skills to perform their job. This is called the subjective method because it is based on surveys entailing workers' self-assessment (Van der Velden and van Smoorenburg 1997; Groot and Maassen van den Brink 2000; Jensen et al. 2007; Leuven and Oosterbeek 2011; Piracha et al. 2012).

A second method is called the objective method because it is based on expert classification of the education and skills required to perform particular jobs. Here, a wide range of approaches can be noted. One approach is to classify jobs according to broad job levels, for example, the four skill levels ranging from unskilled to highly skilled, distinguished by the International Labour Organisation (ILO) in the first digit of its ISCO-08 occupational classification (ILO 2007). In many countries, national statistical agencies have adopted ISCO in their labour market surveys, either by classifying occupations directly in terms of ISCO or by using crossover tables from a national occupational classification. Statistics Netherlands has attempted to classify the 1,200 occupations in its SBC classification in terms of seven job levels (CBS 1993). O*net, the occupations database in the United States that is based on desk research and company visits, indicates skill requirements for a large range of occupations (O*net 2002).[2]

A third method is called the empirical method, whereby the mean (alternatively the median or the mode) number of years of schooling of all workers in a given occupation or group of occupations is compared to the schooling of an individual in the occupation. Individuals are defined as overeducated if their schooling level is more than one standard deviation above the mean (or median or mode) of all individuals in that occupation (Clogg and Shockey 1984; Verdugo and Verdugo 1989; Van der Velden and van Smoorenburg 1997). Chevalier (2003) applies a mixed method based on subjective and objective overeducation to distinguish between apparently overeducated and truly overeducated workers.

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 4 of 34

Objections have been raised to all three methods. The first method is criticised because workers may be inclined to over- or understate the educational requirements of their job or simply to equate these requirements to their own level of education (Hartog and Jonker 1997). Furthermore, respondents may not always have a good insight into the level of education required to perform a job (Cohn and Khan 1995; Halaby 1994). The second method, the objective one, is criticised because skill requirements within a given occupation cannot vary (Halaby 1994). Based on a survey of school leavers, Van der Velden and van Smoorenburg (1997) conclude that job analysts systematically overestimate the level of required education, probably because they do not use the 'real' situation as the basis of their rating, but descriptions of the tasks and the nature and required level of knowledge and skills. The third method also ignores the variation in terms of educational requirements within an occupation. Additionally, the choice of the reference measure (mean, median or mode) and the choice of one standard deviation seem rather arbitrary (Halaby, 1994). Therefore, Hartog and Jonker (1997) and Verhaest and Omey (2006) conclude that this should be the least preferred method for determining overschooling.

2.2 The incidence of skill mismatch

All studies on skill mismatch confirm the existence of some rate of overeducation among workers. Leuven and Oosterbeek (2011) carried out a meta-analysis of more than 180 studies covering countries in Asia, Europe (predominantly the EU15), the Americas and Australia over a period of five decades and concluded that, on average, 30% of the workforce is overeducated and 26% is undereducated. Overeducation is found less often in Latin America and most often in the USA/Canada. From the 1970s to the 1990s, overeducation declined, but then increased in the 2000s, although the authors note that this might be due to a single 2008 study. In an earlier meta-analysis, Groot and Maassen van den Brink (2000) concluded that the overall incidence of overeducation in the labour market appears to be about 26%.

The incidence of overeducation is likely to be affected by the measurement method. According to Leuven and Oosterbeek (2011), studies based on self-assessment and job analysis methods do not point to large differences in this respect, but the method based on the mean reveals lower levels of overeducation. Groot and Maassen van den Brink (2000) find that overeducation is more frequent when self-reported rather than when objective measures are used. Leuven and Oosterbeek (2011) found that although many studies find statistically significant relations between overeducation and individual characteristics, the specifications of these characteristics vary widely. More or less consistent findings across studies are that young people, women and migrants are more likely to be overeducated. Remarkably few findings refer to the incidence of overeducation for specific educational categories. Mavromaras et al. (2009) analysed the Australian HILDA Survey 2001–2006 and found that overeducation occurs more often in the top half of education brackets than in the lower half, pointing to a relative lack of high-skilled jobs.[3]

According to Leuven and Oosterbeek (2011), only a few studies have addressed the incidence of over- and undereducation among migrants. The available evidence indicates that migrants are more likely to be overeducated. Different arguments are presented as the source of divergence, for example, the imperfect international transferability of human capital (Chiswick and Miller, 2009a), low destination-country

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 5 of 34

language skills (Green et al., 2007) or previous work experience in the home country (Piracha et al. 2012). Most of these contributions are at a national level. This is the case for the labour markets of Canada (Wald and Fang, 2008), Sweden (Dahlstedt, 2011; Joona et al., 2014) and Denmark (Nielsen, 2011).

In a study based on the Labour Force Survey in the United Kingdom, Lindley and Lenton (2006) suggest that immigrants initially experience higher overeducation but that this difference is eroded with time spent in the UK. The result of an analysis run on immigrants in Sweden (Joona et al., 2014), however, suggests that the persistence over time of overeducation among migrant workers is, at least, higher than among native workers. In a study based on the Longitudinal Survey of Immigrant Australians (LSIA), Green et al. (2007) conclude that migrants are more likely to be overeducated than the native population, even if the migrants entered the country in question on skill-based visas. They were better educated than the native-born population but were relatively less likely to be found in managerial and professional occupations and were overrepresented in unskilled work. The authors find that overeducation is greater for migrants from non-English speaking backgrounds (Korpi, 2012). Further details on home countries are provided by Battu and Sloane (2002), using a survey of ethnic minorities in the UK. They conclude that different ethnic groups have varying levels of overeducation, with the highest incidence of overeducation among Indian and African–Asian groups.

However, the results of a study of the US high-skilled labour market by Chiswick and Miller (2009b) show that overeducation is widespread among both migrants and native-born. In the USA, the extent of overeducation declines with job-tenure as high-skilled migrants obtain jobs commensurate with their educational level. Using the Longitudinal Survey of Immigrants to Australia, Piracha et al. (2012) reveal that a significant part of the variation in the migrants' probability of being over- or undereducated in the Australian labour market can be explained by having been over- or undereducated in their last job in the home country. Home country mismatch was notably large in the case of undereducation. We could not find in the literature arguments relating migrant mismatch to the country of origin/destination combination. Nevertheless, we believe that this distinction could be the cause of different mismatch incidence since, firstly, the extent of the international transferability of human capital might vary according to this factor; secondly, it could reflect different levels of dissimilarities among labour markets; and thirdly, language and ethnic differences might vary consistently among different combinations.

Turning to the dynamics of over- and undereducation over time and their methodological implications, there is a massive literature on upgrading and downgrading with regard to occupations. In the past 15 years, much of this literature has been devoted to so-called skill-biased technological change, assuming—and largely confirming—that in developed countries the educational requirements for a similar job within industries have increased over time, mainly due to technological developments (Berman et al. 1998; Machin 2001; Autor et al. 2003). Upgrading entails that, with tenure, the incidence of undereducation increases, whereas downgrading works the other way round. A second dynamic process refers to the inflation of qualifications, implying that new entrants are more likely to be overeducated. Third, dynamics over time may also be caused by fluctuations in labour market conditions, with alternating periods of scarce and excess labour supply: in periods of scarce supply, new entrants are more likely to be undereducated, whereas the reverse holds for entrants in

periods of excess supply. No studies have yet revealed the impact of the economic crisis on the skill structure of the labour market, that is, whether more high-skilled than low-skilled jobs have been lost, or vice versa. Finally, in a study of skill mismatch among migrants, the dynamics over time caused by national migration policies should be taken into account. Policies that stimulate access for high-skilled migrants may affect the educational composition of relevant cohorts of migrants, but this also applies to more restrictive policies towards migration (Korpi 2012). Our study does not consider these dynamic processes.

Few empirical attempts have been made to investigate the longitudinal impact of over- and undereducation, while a legitimate question is whether job allocation frictions diminish over an individual's lifecycle. Korpi and Tahlin (2009) did not find support for the assumption that mismatch dissolves with the time individuals spent in the labour market. Using cross-sectional and panel data from the Swedish Standard of Living Surveys 1974–2000, the authors conclude that the overeducated are penalised early on by an inferior rate of return to schooling, from which this group does not recover.

A final caveat must be made here. Following Piracha et al. (2012), a match or mismatch is observed only for employed individuals. Skill mismatches may be larger for the unemployed labour force, for example, if the educational level of the unemployed does not match the educational requirements of relevant job vacancies. When assuming a higher incidence of mismatch for migrants, the fact that they may constitute a self-selected sub-sample might be overlooked.

2.3 Explanations of skill mismatch and their relation to the migrant condition

In this section, we explore some theoretical explanations of overeducation and the implications of such explanations for the higher incidence of overeducation among migrants. Most of the literature points to explanations related to job allocation frictions. Here, we present five explanations for overeducation that appear to be widely accepted.

A first explanation refers to the assumption that, to begin with, entry-level workers might have jobs for which they are overeducated and later on move to jobs that better match their educational attainment. In their overview studies, Leuven and Oosterbeek (2011) and Cedefop (2010) conclude that, according to many studies, younger workers are more likely to be overeducated than older workers. This supports the assumption that overeducation is part of an adaptation process in the early stages of a working career, in which it compensates for the lack of other human capital endowments, such as ability, experience or on-the-job training. Following this explanation, we investigated job allocation frictions in our empirical study by testing the assumption that the incidence of overeducation is higher among workers who have recently entered the labour market.

A second explanation details the assumption of job allocation frictions. This explanation refers to specific groups of workers when entering the labour market. It is assumed that workers with low bargaining power—for example, students with a job on the side, re-entering housewives for whom a job–education match has a low priority, or workers who have had unemployment spells and involuntary quits—will have jobs for which they are overeducated. This assumption is supported by a range of research results. According to Groot and Maassen van den Brink (2000), workers who have experienced a career break are more likely to be found in jobs for which they are overeducated. Sloane et al. (1999) found that overeducated workers had more unemployment spells and

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 7 of 34

involuntary quits than others. The evidence of Sicherman (1991) showed that overeducated workers changed jobs more frequently and that they had less experience, tenure, and on-the-job training than correctly matched workers. In our empirical analysis, we investigated this type of job allocation friction by testing the assumption that the incidence of overeducation is higher among females and workers who have experienced unemployment spells and quits.

A third theoretical explanation refers to job allocation frictions due to labour market discrimination: employers have a preference for workers from the 'same group'. Field experiments show pervasive ethnic discrimination in many countries (OECD 2007). The condition of ethnic disparity is often embedded in the condition of migrant. Nevertheless, to better isolate this effect in our empirical study, we also investigated whether second-generation migrants (nationals whose parents were born in a foreign country) are more likely to be overeducated compared to native workers.

A fourth theoretical explanation of overeducation refers to job allocation frictions that are related to career mobility. This explanation assumes that individuals accept a lower-level job if the probability of promotion is higher (Sicherman and Galor, 1990). In our empirical study, we tested whether the incidence of overeducation is higher for jobs with good promotion prospects compared to jobs with average or poor promotion prospects.

A fifth theoretical explanation concentrates on job allocation frictions due to the poor abilities of individual workers. This assumption goes beyond the crude measurement of educational attainment and details a worker's ability as well as the skill requirements of a job. As regards the migrant population, a single ability has been investigated, namely the worker's mastery of the native language or lingua franca of the host country. Thus, in this approach the language ability of the worker is critical. According to a study carried out in Australia, workers from non-native language speaking backgrounds showed a higher and more persistent incidence of overeducation than those from native-language speaking backgrounds (Kler, 2005). In our empirical study, we tested whether migrants from home countries where the native language or lingua franca does not match that of the home country are more likely to experience overeducation. The theories exposed and the proposed tests are summarised in Table 1.

Table 1 Overeducation theoretical explanations and proposed tests

Theories of overeducation	Proposed test
1 Overeducation is part of an adaptation process in the early stages of a working career.	Overeducation is higher in workers who have recently entered the labour market.
2 Given job allocation frictions, workers with low bargaining power are more likely to be overeducated.	Overeducation is higher among females and workers who have experienced unemployment spells.
3 Labour market discrimination turns to overeducation.	Overeducation is higher among second generation migrants.
4 Individuals accept overeducation if the probability of promotion is higher.	Overeducation is higher for jobs with good promotion prospects.
5 Overeducation is related with personal abilities (not measured by education attainment).	Overeducation is higher among migrants whose native language or lingua franca does not match that of the home country.

3. Data and methodology

3.1 Data and definitions

This article is based on statistical analyses of the WageIndicator dataset (Tijdens et al. 2010). The WageIndicator project is currently running in 86 countries on five continents. It consists of national websites, each of which receives large numbers of visitors, primarily because the websites post a 'salary check' that provides free information on occupation-specific wages. Worldwide, the national WageIndicator websites attract large numbers of web visitors—more than 20 million in 2013. The websites are consulted by workers when making job mobility decisions or before annual performance talks or wage negotiations. The sites are also consulted by school pupils, students and re-entrant women facing occupational choices, and by employers in small and medium-sized companies when recruiting staff or negotiating wages with their employees.[4]

The WageIndicator dataset is derived from a web survey on work and wages that is posted on all national WageIndicator websites and is comparable across all countries. In return for the free provision of information, visitors are asked to complete the survey. Thus, the survey is voluntary, continuous, and multi-country.[5] It contains detailed questions about, for example, education, occupation, skill mismatch, industry, country of birth, country of birth of parents, and, in some countries, ethnic group. Respondents are asked if they were born in the country of survey; if not, they can select a country from a list. In this article we use 'native workers' and 'migrant workers' to identify the two groups. The web survey does not allow the identification of return migration. The large sample size allows, for each country, a breakdown of migrant groups according to country of birth in order to better capture the heterogeneity of migrants. In order to improve the intelligibility of our results, countries of residence and countries of birth were grouped into continent classes (see section 3.2 for further details)[6].

This source has been used to describe and study a wide variety of labour related issues. Often it has supplied information in circumstances where official sources (a) face technical difficulties, e.g., wage studies in African countries; (b) do not provide adequate sector breakdowns, e.g., De Vries and Tijdens (2010; Tijdens et al. (2013a), and Steinmetz et al. (2014), who approach several labour related issues in the health sector; or (c) do not provide specific information, e.g., Guzi and de Pedraza (2015) research workers well-being. One of its main advantages consists in providing a solid base for international comparisons. Fabo and Tijdens (2014), for example, measure the demand for specific skills in occupations across countries, while Tijdens et al. (2013b) internationally evaluate the tasks implemented within different occupations. The role of the Wageindicator is also discussed within the larger debate on the growth of internet-based information in the social sciences. Kureková et al. (2014) bring it into play when presenting advantages and disadvantages in using online data for labour market analysis. Askitas and Zimmermann (2015) mention this web survey among the internet sources of human resources data. The data source has also been subject of several methodological discussions (Andreadis, 2013, on response time analysis in web-surveys; Tijdens, 2014, on dropout rates; Steinmetz et al. 2009a, and Steinmetz et al. 2013a on potential biases and correction methods).

We used the pooled annual data on 86 countries for the years 2008–2013. Note, however, that some countries joined the survey later than 2008 and that in some of the

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 9 of 34

countries the question about skill mismatch has been asked only since 2009. We excluded respondents aged under 15 or over 70, unemployed people, school pupils, students and those who have never had a job. Altogether, 673,898 observations were included in the analysis. However, the response rate for some of the variables presented in the following section was less than 20% of the total, which led to a consistent reduction in the number of complete cases. In order to consider the possible bias arising from this, we present two analyses, one including and one excluding those variables.

The web survey is voluntary, and therefore using this data may have important drawbacks. By definition, a web survey is completed only by individuals with sufficient language and computer skills to read and answer the survey questions. This might be particularly off-putting for migrants and could lead to biased data due to the low representativeness of some specific demographic groups.[7] We did not employ within-country weights since previous studies that used this dataset have shown how weighting to correct for these groups scarcely affects the means of some of the variables under study and that, in general, weighting volunteer surveys to control for socio-demographic composition does not solve the small bias in some specific variables (such as wages); see Steinmetz et al. (2009b) for further details. The problem, however, is not as bad as it seems because it can be assumed that literacy skills are higher among employed migrants compared to unemployed migrants, and the size of the group of employed migrants with insufficient literacy skills is relatively small compared to the labour force as a whole.

On the other hand, because of the characteristics discussed above, our dataset can represent the internet population, although not the entire world population. When focusing attention on migrants, the worldwide internet users form a population from which interesting information can be extracted.

In addition, the dataset presents other clear advantages. The main advantage is the large size of the sample and the large variety of countries of birth and destination, which allowed analyses to be performed on large groups even when considering specific continent of provenance and destination combinations. In addition, when our dataset is compared to the means of demographic variables known from other sources, the sample variable means do not deviate to a large extent. For example, two meta-analysis studies (Groot and Maassen van den Brink 2000; and Leuven and Oosterbeek, 2011) found an average of overeducation ranging from 26% to 30%. Our dataset reveals 25% of overeducation in the overall sample. The EU Labour Force Survey shows an incidence of migrants among the total population aged between 15 and 64 years of 8.7% in the EU15 and 1.6% in the EU12. These figures compare with the 7.4% and 3.5%, respectively, registered for the employed population in our dataset. All things considered, we acknowledge the limitation of our dataset and therefore consider our findings to be exploratory rather than conclusive.

3.2 Methodology

We set up an analytical model in order to observe the relation between self-assessed skill mismatch and individual characteristics. The WageIndicator survey includes the question "Do your qualifications match your job?" The three response options are "Yes", "No, I am over-qualified for my job", and "No, I am under-qualified for my job". We only considered the difference between overqualification and the other options (proper

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 10 of 34

match and under-qualification) and obtained a binary variable: the dependent variable in our model. We then assessed the relation between the possibility of being overeducated and some idiosyncratic characteristics. These features can be classified into two groups. The first is composed of demographic characteristics (e.g., gender or country of residence), the second of proxy variables to test the five theoretical explanations of overeducation. The model specification is as follows:

$$y = X_d\boldsymbol{\beta_d} + X_t\boldsymbol{\beta_t} + \beta_m\boldsymbol{m} + \boldsymbol{e} \qquad (1)$$

and the coefficients of this logistic regression model were estimated using the method of maximum likelihood.

In equation 1, y is the logit of being overeducated or not, in other words, the natural logarithm of the odds ratio of being overeducated and the other options. X_d is the matrix of the demographic characteristics, and X_t is the matrix of the proxies of the theoretical explanations of overeducation presented in section 2.3; m is a categorical variable distinguishing between native workers, native workers whose mother or father were born in a foreign country (second-generation migrants) and migrant workers. Accordingly, β_d and β_t are the vectors of the corresponding coefficients, and β_m is the coefficient of the migrant condition.

The variables labelled as demographic characteristics are a mix of personal and professional characteristics. We took into consideration features such as gender (male being the reference category) or country of residence. We grouped the country of residence into seven continental categories: the EU15[8] (the reference category), the EU12,[9] Africa, Latin America, Asia, North America & Oceania, and non-EU European countries (predominantly Russia and CIS countries such as Belarus and Ukraine). We did this because for some countries the number of respondents is limited and to keep the number of categories manageable. Because mismatch is considered to be dependent on educational attainment, we included a variable describing the educational attainment of each individual. For the sake of international comparison, we recoded the national educational categories into the worldwide International Standard Classification of Education classification 1997, as designed by UNESCO.[10] The variable ranges from 1 (primary level of education and our reference value) to 6 (second stage of tertiary education, leading to an advanced research qualification). The analysis controlled for job-specific attributes. We considered the difficulties related to measuring job levels in section 2. In our study, we used a job level indicator, derived from the occupation variable. It is called 'corporate hierarchy' and is based on a mapping of the 1,700 occupations distinguished in the survey into six corporate hierarchical levels, ranging from 1 = helper to 6 = CEO, developed by the second author[11]. Furthermore, a firm-size categorical variable and an industry variable, where jobs are classified under three main categories (agriculture and manufacturing; distribution related services; and other public, commercial and personal services) were included.

Altogether, six variables were included to test the theoretical explanations of overeducation. The age of individuals was used to proxy the entry-level workers effect (theoretical explanation number 1). 'Breaks' is a variable denoting the number of career breaks an individual has experienced. It was used to test the assumption that particular groups of workers with low bargaining power are more likely to experience overeducation (theoretical explanation number 2). Female workers as opposed to male workers, measured

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 11 of 34

by the gender variable, also illustrate the same effect on overeducation. Theoretical explanation number 3 involves the role of ethnic discrimination. The categorical variable distinguishing between native, migrant, and second-generation migrant workers helps to figure out this effect on overeducation. Survey respondents are asked whether they perceive to have career opportunities in the organisation they are presently working for. The response to this question was included to test theoretical explanation number 4. Although job allocation frictions due to the poor abilities of individual workers for the whole sample of workers (theoretical explanation 5) were not tested on the native population, this aspect was taken into consideration for the case of migrant workers. Workers with lower language abilities can proxy workers with poor individual abilities. Here, they are defined as migrant workers born in a country with a native language or a lingua franca that does not match that of the host country.

The relation between overeducation and the migrant condition, which was the main focus of the present study, was studied from different points of view. First, we tested the effect of being a migrant controlling for all demographic and theory related characteristics (estimating equation 1). We then tested whether overeducation is similar across migrants from various home countries and across different host countries. To do so, we estimated equation 2, where the effect of country of residence, country of birth and their interaction is taken into consideration. We made use of the estimated coefficients to compute the predicted probabilities of all groups of migrant workers and compare them with the native workers' predicted probabilities on each continent.

$$y = \ X_d\boldsymbol{\beta_d} + X_t\boldsymbol{\beta_t} + \beta_{cr}CONT_{RES} + \beta_{cb}CONT_{BIRTH} + \ CONT_{BIRTH}CONT'_{RES}\boldsymbol{\beta_c} + \boldsymbol{e} \qquad (2)$$

Since equation 2 is estimated through logistic regression, predicted probabilities are computed as follows:

- for natives $\frac{e^{(intercept+cont.of\ residence\ coefficient)}}{1+e^{(intercept+cont.of\ residence\ coefficient)}}$
- for migrants in the EU15 $\frac{e^{(intercept+cont.oforigin\ coefficient)}}{1+e^{(intercept+cont.of\ origin\ coefficient)}}$
- for migrants in other continents $\frac{e^{(intercept+c.ofor.\ coefficient+c.of\ re.coefficient+interaction\ coefficient)}}{1+e^{(intercept+c.ofor.\ coefficient+c.of\ re.coefficient+interaction\ coefficient)}}$

Finally, we modified equation 1, introducing interaction terms between the migrant condition and all other covariates (matrix X_i in equation 3). Estimating equation 3 allowed us to observe whether the theoretical explanations for overeducation affect migrant and native workers in a different manner.

$$y = \ X_d\boldsymbol{\beta_d} + X_t\boldsymbol{\beta_t} + \beta_m\boldsymbol{m} + X_i\boldsymbol{\beta_i} + \ \boldsymbol{e} \qquad (3)$$

4. Empirical findings on skill mismatch

4.1 Descriptive analysis of skill mismatch and migration

Table 2 provides basic information about the age and gender distribution of the workers in the dataset. We made use of 673,898 observations. Individuals between 15 and 45 years represent around 80% of the total; within this age group, males comprise the largest share (around 56%).

Table 3 shows the respondents' continent of residence and continent of birth, and the matrix of continent of residence/birth. It is helpful to understand the incidence of the migrant population and its provenance on each continent. Our dataset is clearly

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 12 of 34

Table 2 Distribution of workers by age and gender

Broad age groups in years	Total	Per cent	Males	Females	% females	<NA>
Total	673,898		356,612	280,166	41.57	37,120
15-29	237,313	39.09	124,201	109,991	46.35	3,121
30-44	256,385	42.24	145,445	107,427	41.90	3,513
45-59	103,220	17.00	57,317	44,184	42.81	1,719
60+	10,113	1.67	7,351	2,439	24.12	323
<NA>	66,867		22,298	16,125		28,444

European focused, since around 60% of the respondents work in a European country. Within Europe, we distinguish between EU15, EU12 and non-EU European countries. Central & South American workers and Asian workers account for 19% and 14% of the surveyed population, respectively. Africa and North America & Oceania are clearly underrepresented, with a share of 6% and 1%, respectively. The incidence of migration varies widely across continents. In Central & South America and in the EU12, the presence of migrants is the lowest, accounting for around 3% of the surveyed population. Migrants represent 5% of the population in Asia and around 10% in the EU15, Africa and the non-EU European countries. Finally, migrants make up 25% of the (low number of) survey respondent from North America & Oceania. Not surprisingly, within-continent migration accounts for most of the flows observed. Only in the cases of Asia and North America & Oceania they do not represent the largest share of the phenomenon.[12] In the former case, the relevance of migrants from non-EU European countries is evident. In the EU15, our largest group—more than 40% of all migrants (2.40% of the total population)—come from another EU country; Central & South America is the second largest region of provenance (20% of migrants, 1.59% of the total population).

The incidence of overeducation in the various countries is shown in Figs. 1, 2 and 3, which help to disentangle the relation between overeducation and migration status. At first, it is worth noticing how overeducation is a phenomenon that affects countries, and therefore continents, differently. Figure 1 shows how the incidence of self-assessed overeducated

Table 3 Continent of residence and continent of birth. Percentage

	Total (n)	Per cent	Country natives	EU15	EU12	Africa	Central and South America	Asia	North America and Oceania	Non-EU European Countries
Continent of residence										
EU15	253,197	37.58	92.57	2.40	0.67	0.79	1.59	1.22	0.29	0.46
EU12	29,415	4.37	96.50	0.11	2.82	0.06	0.02	0.09	0.02	0.35
Africa	41,332	6.13	89.86	3.23	0.11	4.86	0.64	0.88	0.29	0.09
Central and South America	126,852	18.83	97.15	0.44	0.02	0.02	2.13	0.07	0.12	0.05
Asia	91,089	13.52	95.91	0.05	0.04	0.03	0.02	1.61	0.03	2.31
North America and Oceania	8,548	1.27	75.47	6.56	1.11	1.38	6.59	5.72	2.18	0.92
Non-EU European countries	123,375	18.31	90.93	0.31	0.54	0.02	0.03	1.80	0.03	6.30

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 13 of 34
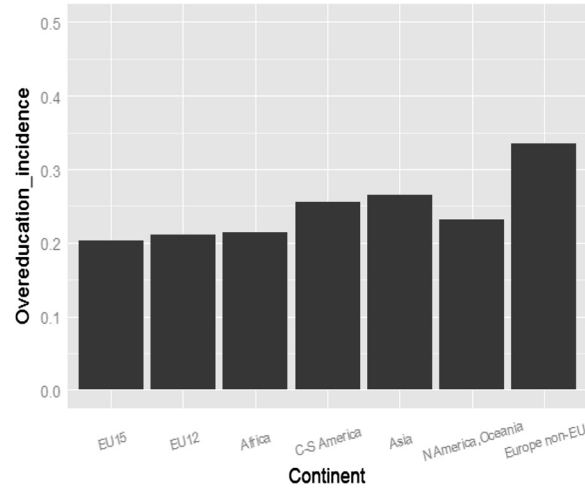
Fig. 1 Incidence of overeducation per continent

workers in the total survey population is around 20% in EU countries. Around the same figure is observed in African countries, while its incidence is slightly higher in North America & Oceania (23%), Central & South America (26%) and Asia (27%). More than one third of respondents from non-EU European countries stated that they were overeducated.

When it comes to establishing whether overeducation has a different incidence among native and migrant workers (Fig. 2), it appears this is the case on every continent but not in the non-EU European countries. Nevertheless, the relation between migration and overeducation can be of two kinds. In Europe and Asia, overeducation affects migrants to a greater extent than it does native workers. Its incidence appears to be around 5% higher: 25% versus 20% in the EU15, 26% versus 21% in the EU12, and 33% versus 28% in Asian countries. Overeducation affects migrants to a lesser extent than native workers in African, Central & South American, and North American & Oceanian countries. On these continents the phenomenon is around 3% lower among immigrants. The migration phenomenon is different across countries in terms of magnitude as well as country of provenance, and these differences reflect differences in the

Fig. 2 Incidence of overeducation per continent and migrant status

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 14 of 34

Fig. 3 Incidence of overeducation per continent and continent of birth

reason behind the migration decision. Differences in the incidence of overeducation with respect to the native population might be a consequence of the different reasons for which people migrate.
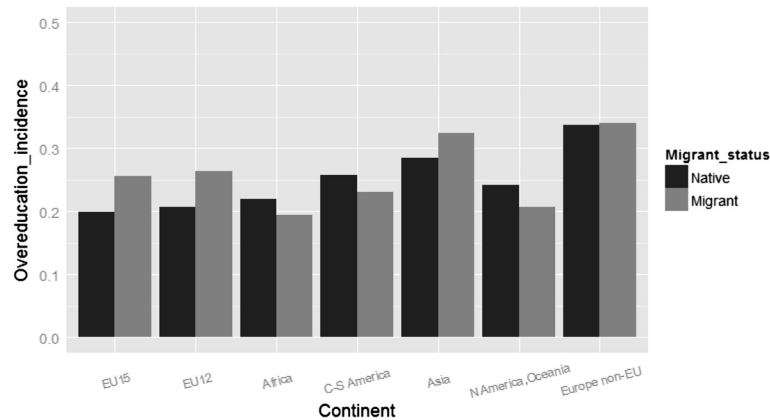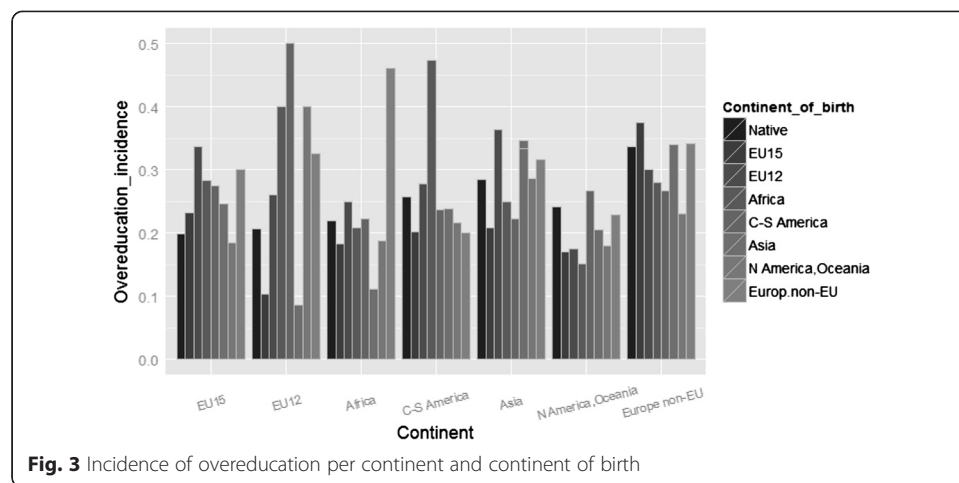
Figure 3 presents a comparison of the incidence of overeducation per continent of residence and distinguishes between native and migrant workers according to their continent of birth. There are some clear patterns. EU15 migrants, for example, show just a slightly higher overeducation incidence with respect to native workers in other EU15 countries and European non-EU countries, and they are clearly less affected by overeducation when they migrate to the rest of the world. North American & Oceanian migrant workers are in a similar situation, with a lower incidence of overeducation than native workers in all countries except the EU12 countries. Quite the opposite case is shown by EU12 migrants, who are always (except in North America & Oceania and non-EU Europe) more overeducated than native workers. Non-EU European migrants show a similar pattern. Differences in the incidence of overeducation with respect to native workers among African, Central & South American and Asian migrants depend on the continent of destination. It is clear that the differences in the incidence of overeducation depend not only on the country of residence but, as can be expected, also on the country of origin and country of destination bilateral relation.

The findings of our data exploration show the need for a deeper investigation into the relation between overeducation and migration, focusing on workers' idiosyncratic characteristics.

4.2 Skills mismatch in detail

4.2.1 Overeducation and idiosyncratic characteristics

Our first research objective was to assess whether personal demographic and job-related characteristics are related to the self-assessed overeducation. This analysis was aimed at obtaining evidence as to whether overeducation is related to personal characteristics (which, in turn, reflect national labour market characteristics), assuming that skill mismatch varies among workers' characteristics, such as educational attainment and job level. In the first column of Table 4 are the results of the logistic regression analysis including what we defined in section 3.2 as demographic characteristics run on

Table 4 Chance of overeducation compared to a correct skill match or undereducation, odds ratios, standard errors in brackets

	(Demographic variables)			(Demo and theoretical)			(Demo and theoretical)			(Demo, theoretical and MIGRATION)		
MIGRANT										1.109054	***	(0.014)
Age				0.9882	***	(0.001)	0.9941	***	(0.001)	0.9939876	***	(0.000)
Breaks				1.0934	***	(0.009)	1.0746	***	(0.002)	1.0744215	***	(0.002)
Second generation				1.1583	*	(0.070)	1.1143	***	(0.026)	1.1247619	***	(0.026)
Job prospects				0.4013	***	(0.027)						
Gender (female = 1)	1.3497	***	(0.013)	1.2388	***	(0.027)	1.1614	***	(0.008)	1.1615692	***	(0.008)
Continent (ref. EU15)												
EU12	1.0088		(0.028)	0.8973	**	(0.042)	1.0344	.	(0.018)	1.0400934	*	(0.018)
Africa	0.9394	**	(0.029)	0.9463		(0.066)	0.9466	**	(0.018)	0.9448265	**	(0.018)
Latin America	1.2374	***	(0.018)	1.0673	*	(0.033)	1.2084	***	(0.011)	1.2164861	***	(0.011)
Asia	1.2717	***	(0.021)	1.5446	***	(0.101)	1.2976	***	(0.014)	1.3027477	***	(0.014)
North America and Oceania	1.0676		(0.053)	1.3034	.	(0.150)	1.0099		(0.033)	0.9953363		(0.033)
Europe non-EU	1.9529	***	(0.018)	1.4598	***	(0.044)	1.6604	***	(0.011)	1.6590651	***	(0.010)
Education (ref. ISCED10)												
ISCED 20	1.3761	***	(0.061)	1.5837	***	(0.137)	1.3741	***	(0.035)	1.3743994	***	(0.035)
ISCED 30	1.3765	***	(0.059)	1.3747	*	(0.136)	1.3997	***	(0.034)	1.3981505	***	(0.034)
ISCED 40	1.6345	***	(0.063)	1.9162	***	(0.141)	1.6286	***	(0.036)	1.6268751	***	(0.036)
ISCED 50	1.7580	***	(0.059)	2.1475	***	(0.134)	1.8736	***	(0.034)	1.8685825	***	(0.034)
ISCED 60	1.8940	***	(0.064)	2.3162	***	(0.144)	1.9215	***	(0.037)	1.9129916	***	(0.037)
Corporate Hierarchy												
CH2	0.4914	***	(0.060)	0.5024	***	(0.122)	0.6957	***	(0.037)	0.6966276	***	(0.037)
CH3	0.5743	***	(0.072)	0.5946	***	(0.146)	0.7564	***	(0.044)	0.7577056	***	(0.044)
CH4	0.3733	***	(0.064)	0.3494	***	(0.133)	0.4760	***	(0.041)	0.4765683	***	(0.041)

Table 4 Chance of overeducation compared to a correct skill match or undereducation, odds ratios, standard errors in brackets *(Continued)*

CH5	0.3392	***	(0.064)	0.3714	***	(0.130)	0.4551	***	(0.042)	0.4553189	***	(0.043)
CH6	0.3225	***	(0.079)	0.4210	***	(0.152)	0.3892	***	(0.055)	0.3894885	***	(0.055)
Firm size (ref. 1–10)												
size 11-50	0.9187	***	(0.018)	0.9349	.	(0.036)	0.9470	***	(0.011)	0.947486	***	(0.011)
size 51-100	0.8855	***	(0.023)	0.9541		(0.045)	0.9037	***	(0.014)	0.904243	***	(0.014)
size 101-500	0.8694	***	(0.019)	0.9553		(0.039)	0.9038	***	(0.012)	0.9043934	***	(0.012)
Size 500+	0.8035	***	(0.019)	0.8641	***	(0.041)	0.9015	***	(0.012)	0.9018726	***	(0.012)
Industry (ref. agr., man. and constr.)												
Trade, transp. and hotels	1.6906	***	(0.023)	1.5635	***	(0.050)						
Commercial services	1.4966	***	(0.018)	1.4566	***	(0.040)						
No. of observations	145 684			34 447			368 564			368 564		
Nagelkerke presudo R2	0.045			0.097			0.034			0.034		
LR chi2	4 475.61			2369.28			8673.66			8726.23		
Pr(> chi2)	<0.0001			<0.0001			<0.0001			<0.0001		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

the total survey population. The covariates' coefficients perform as expected, and the corresponding odds ratios are presented. According to the results, female workers are over 30% more likely to be overeducated than male workers. It is also interesting to observe the (expected) significant relation between overeducation and education level: the higher the education level attained by a worker, the greater the possibility of being overeducated. When considering job-specific characteristics, it turns out that the corporate hierarchy has a significant negative relation to overeducation. This relation appears to be close to linear, since high-level workers (CEOs) have 77% less probability of being overeducated than low-level workers (helpers). The firm size also seems to influence the probability of being overeducated: the probability is lower in large firms than in small firms. Workers in trade and transport sectors are the most likely to experience overeducation (around 60% more likely than agriculture and manufacturing workers). Finally, as regards geographical differences, the results confirm our previous findings: non-EU European workers are the most likely to experience overeducation, followed by Asian and Latin American workers.

4.2.2 Theoretical explanations

Our first research objective also comprised testing whether a range of theoretically based assumptions affect the incidence of overeducation. We tested whether the theoretical explanations for overeducation presented in section 2.3 hold in our dataset. For this, we made use of the proxies presented in section 3.2. The second column of Table 4 shows the regression results when these covariates are included. As expected, overeducation is a phenomenon that affects younger workers to a greater extent: as each year passes, the probability of the typical worker being overeducated decreases by more than 1%. Poor bargaining power—measured here by the number of career breaks workers experience—has a significant effect on the probability of having jobs for which they are overeducated. For each additional career break, this probability increases by almost 10%, confirming how overeducated workers changed jobs more frequently. Ethnic differences also matter, according to our findings: second-generation migrants who are nationals are 15% more likely to be overeducated than other nationals.

The fourth theoretical explanation refers to job allocation frictions related to career mobility, assuming a direct relation between overeducation and probability of promotion. Our results do not seem to support this idea. There is no trade-off between good job prospects and accepting being overeducated. Indeed, the antithesis seems to be confirmed: workers who believe they have good job prospects in their present organisation are 60% less likely to be overeducated. This outcome suggests that there could be a clear distinction concerning the quality of jobs, with good skill match and proper career development opportunities opposed to low quality positions with poor match and career opportunities. It has to be highlighted, however, that the job prospects variable has a non-negligible number of missing values, with a response rate of only 18%, which might bias the whole sample. A similar response rate is found in the industry variable (25%). Results omitting these variables are therefore reported in the third column of Table 4. In order to avoid these biases, these variables were excluded from the subsequent analysis.

4.2.3 Overeducation and the migrant condition

The main aim of this study was to assess whether overeducation is more common among migrants compared to native workers. The last column of Table 4 shows the regression outcome when considering the full equation 1 including a discrete variable describing the migrant condition. It confirms that overeducation occurs more often among migrant workers. When considering the whole dataset, it is clear that being a migrant has a significant relation to the probability of experiencing overeducation: migrants are 10% more likely to be overeducated than native workers.

We clearly observed in section 4.1 that the migrant condition has a different effect on the overeducation incidence depending on the continent of residence and that of birth. We therefore estimated equation 2 taking into consideration the personal characteristics presented above as well as the interaction between the continent of birth and that of residence. We then computed the overeducation predicted probability for each country of birth/residence combination and compared it with the probability of native workers. The differences between migrant and native workers are reported in Table 5.[13]

The analysis so far had relied on the strong assumption that the covariates' effects have the same intensity across several countries, in which case the covariates' coefficients should be the same (or at least similar) across different countries. Furthermore, we are treating countries with high and low survey penetration similarly. In order to test the extent to which these assumptions held, we ran the regressions based on equation 1 for 10 country-specific datasets[14] and a dataset composed by the information collected in the 46 countries with the lowest survey penetration. We than compared the odds obtained. The control countries were Germany, the Netherlands, Belgium, Belarus, Ukraine, Kazakhstan, Indonesia, Brazil, Argentina, and South Africa. The results are presented in Appendix 1. The mean and standard deviations presented suggest that the assumptions essentially hold. Demographic variables' coefficients show a very

Table 5 Estimated difference of the incidence of overeducation for migrants (per provenance) with respect to natives, after controlling for demographic and theoretical aspects

Continent of residence	Continent of provenance						
	EU15	EU12	Africa	C-S America	Asia	NAO	Eur not-EU
EU15	0.03	0.11	0.08	0.07	0.05	−0.05	0.09
	**	***	***	***	***	*	***
EU12	−0.11	0.05	0.07	0.32	−0.16	0.42	0.10
		**		.	.		
Africa	0.01	0.00	−0.05	−0.09	−0.14	0.04	0.18
	**		***		***		.
C-S America	−0.02	−0.03	0.02	0.00	−0.06	−0.03	−0.09
	**			***			.
Asia	−0.03	−0.02	0.01	−0.33	0.01	0.28	0.01
							*
NAO	−0.13	−0.11	−0.22	−0.01	−0.06	0.06	−0.15
	***	**	***		**		*
Eur not-EU	0.12	−0.14	−0.19	−0.01	0.01	0.04	0.04
		***			**		***

Note: Results obtained on the basis of equation [2] estimation. Predicted probabilities difference between different groups.
Significance of the interaction term coefficients are reported; codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 19 of 34

low variation across countries. The age effect shows a standard deviation of 0.07; the gender effect is negative for female workers in all but non-EU European countries. Job-related variables, such as the role of breaks and position in the corporate hierarchy, are very similar across the world. The coefficients concerning the role of educational attainment vary the most; nevertheless, almost all countries share an increasing effect as education levels grow. In conclusion, we acknowledge that the strong assumption of the same intensity for the covariates in different countries can be accepted.

It is well known that migration leads different types of individuals to move to certain destination places according to their personal (observable and unobservable) characteristics (Borjas 1987, 1989). This selection process can, in turn, affect the observed probabilities of migrants to be overeducated. According to the literature, individuals make the migration decision upon consideration of the utility they can get from the different options, which in turn are driven by both host and source countries' socio-economic conditions. Studying the migration assessment process goes well beyond the purpose of this paper. Nevertheless, we need to acknowledge that migrants across countries are non-randomly selected individuals and that the characteristics that drive the migration decision might also be related to the probabilities of these individuals to be overeducated. It is possible that potential migrants compare the professional development perspectives in the various countries and choose to move to the country with the best professional fit, of which overeducation is a measure. In equation 1 we can control for those observable characteristics influencing the fit; however, there could be some unobserved characteristics that affect simultaneously the migration decision and the overeducation outcome. Not controlling for these characteristics might lead to biased results. Since Table 3 shows how observed characteristics explain only limitedly overeducation (given the low pseudo-R2), unobserved idiosyncrasy differences among individuals could therefore be relevant on our study.

In order to take into account this self-selection mechanism and observe to which extent it affects our results, we estimated a system of equations where the probability of overeducation is jointly estimated with the probability of being a migrant. This two-stage method (first developed by Heckman, 1976) corrects for selection bias. Furthermore it produces an estimation of the extent to which the individuals' unobservable characteristics are simultaneously related with both the probability of being a migrant and the probability of being overeducated. Indeed, this is measured by the *rho* coefficient, which can vary in absolute values from 0 to 1.

Results of our estimation are presented in Appendix 2. The significant and negative *rho* coefficient observed when the system is estimated on the whole dataset confirms the presence of a self-selection bias. There are personal characteristics among migrants responsible for the migration decision, as well as for the fact that the person consider herself overeducated in the present job even though they affect those probabilities antithetically. However, the value of the coefficient is limited, from 0.23 to 0.36. When the system is estimated for two specific countries (the Netherlands and Germany), the *rho* coefficient is not significantly different from 0. We therefore recognised the relevance of the self-selection process in the overeducation of migrants workers; nevertheless, we consider that assuming its triviality does not change the fundamental results of the analysis that follows.

The second research objective was to observe whether the probability that the migrant status affects overeducation varies consistently by continent of residence and of origin.

Despite the large variety among different combinations, observing the situation by migrant groups (by columns in Table 5), some patterns emerge. For example, when EU15 citizens migrate they seem to show a lower overeducation incidence than native workers, with the exception of other EU15 countries, where they are slightly more affected. African migrants, on the other hand, are more overeducated, especially when moving to European countries. Central & South American migrants are more likely than natives to experience overeducation when migrating to EU15 and EU12 countries but less likely when migrating to Asia, Central & South America, and Africa. Given the limited number of respondent residents in North America & Oceania, we can hardly draw conclusions on the figures exposed.

If personal characteristics and the migrant condition are related to overeducation, it is worth studying how these features interact. Our third research objective was to examine the relation between the theoretical explanations of overeducation presented in section 2.3 and the migrant condition. The effect of the interaction between migrant condition and the proxy variables standing for the theoretical explanations are presented in Table 6.[15] Note how young migrant workers are no more overeducated than older migrants, ceteris paribus. It seems that if a native worker's probability of being overeducated decreases over his or her lifecycle, this is not the case for migrants. The effect is indeed the opposite: the odds of being overeducated increase slightly as the years go by. Time does not 'cure' migrants' overeducation as it does with native workers. This finding, which is in line with previous studies, indicates that overeducation is a problem that affects migrants not only during their integration process, but also in the long run. Although the low bargaining power assumption still holds for migrants, its effect is significantly less. A native worker is 10% more likely to be overeducated per each career break experience, while this value is reduced to half in the case of

Table 6 Estimated coefficients and odd rations considering the migrant status and its interaction with other covariates

	(Interaction)				(Interaction)			
	Estimate coeff.		Std.	Odds	Estimate coeff.		Std.	Odds
Age	−0.0133	***	0.00147	0.9868	−0.0068	***	0.0004	0.9932
Breaks	0.0946	***	0.00936	1.0992	0.0726	***	0.0025	1.0752
Gender (female = 1)	0.1981	***	0.0276	1.2191	0.1489	***	0.008	1.1605
Job prospects	−0.9032	***	0.02802	0.4053				
[…]								
MIGRANT	−1.8293	*	0.04494	0.1605	−0.0914		0.0716	0.9126
Age * Migrant	0.0193	***	0.00074	1.0060	0.0094	***	0.0015	1.0025
Breaks * Migrant	−0.0529	*	0.03251	1.0426	−0.0116		0.0086	1.0628
Gender * Migrant	−0.2068		0.18685	0.9914	−0.0234		0.03	1.1337
Language mismatch	0.0894		0.68452	1.0936	0.0894	*	0.0375	1.0936
Job prosp. * Migrant	0.1456		0.06303	0.4688				
No. of observations	34 391				365 545			
Nagelkerke presudo R2	0.099				0.035			
LR chi2	2 435.13				8 843.44			
Pr(> chi2)	<0.0001				<0.0001			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
Note: Results concerning the interaction between the migrant condition and demographical covariates are omitted, available from authors upon request

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 21 of 34

migrants. When observing the poor bargaining power associated with gender, it seems that there is no significant difference between native and migrant workers. In section 2.3 it was also hypothesised that workers facing discrimination by employers are more likely to report overeducation. This third theoretical explanation (ethnic discrimination) was not tested for migrants, given the absence of a specific proxy variable that allowed for ethnic distinctions. Career mobility has the same (unexpected) effect for both native and migrant workers: the better the career prospects, the lower the propensity towards overeducation. We then tested the job allocation frictions due to the poor abilities of migrant workers. Not sharing the same mother language with native workers is assumed to be a proxy for the poor abilities. The language mismatch increases the probability of being overeducated by approximately 10%. In conclusion, given the exposed positive relation to age, career breaks, and poor job prospects, overeducation appears to be a persisting problem that affects immigrants in the long run.

5. Conclusion

Skill mismatch is more common among migrants compared to native workers, although the incidence differs across migrants depending on the country of residence and the country of origin. Furthermore, idiosyncratic workers' characteristics affect native and migrant workers' overeducation differently. To achieve these exploratory results, this study used data from a survey in which workers themselves assess whether they are qualified or overqualified for their job.

The data stem from a continuous and voluntary multi-country web survey, extracting 673,898 observations for the years 2008–2013 from 86 countries. The dataset is European focused: almost 60% of the respondents work in the EU or in a non-EU eastern European country. The share of migrants varies, depending on the continent, from 3% to 10%.[16] Within-continent migration represents the greatest share of migrants. The main disadvantage of the dataset is embodied in the low representativeness of some specific demographic groups; nevertheless, when our dataset is compared to the means of demographic variables known from other sources, the sample's demographic means do not deviate to a large extent. The main advantages are its ample size and the large variety of country of origin and destination combinations, which allows for detailed analysis of different migration flows. Nevertheless, given the limitations of our dataset, we consider our findings exploratory rather than conclusive.

Overeducation affects migrants to a different extent according to the country of residence. Migrants working in Europe and Asia are more likely than native workers to be overeducated, whereas migrants in Africa and Latin America are less likely to be overeducated compared to native workers. These findings might be the sign of two kinds of migration. Nevertheless, differences in the incidence of overeducation between native and migrant workers are related not only to the country of residence, but also to the combination of country of origin and destination. Immigrants experience a higher incidence of overeducation than native workers, although the disparity differs by migrant origin and destination. For example, the incidence of overeducation is higher compared to native workers for EU12 migrants living in a EU15 country, and it is lower for EU15 migrants working in the EU12. This finding is confirmed by both an explorative and a in-depth analysis.

As regards the relation between personal (demographic and job-related) characteristics and overeducation, our dataset confirms the results most frequently highlighted by the literature. Our analysis shows, not surprisingly, that female workers are more likely to be over-educated. The higher the individual's level of education, the more overeducation can be expected; and the higher the individual's job level, the less overeducation can be expected. Controls for firm size and industry reveal that overeducation occurs more often in small firms compared to large firms and more often in trade, transport, and hospitality compared to the other commercial services or to primary and secondary economic activities. Thus, the characteristics of both workers and national labour markets influence the incidence of overeducation.

A few theoretically based assumptions, grounded in the literature on the subject, are used to explain overeducation. When tested, they reveal that recent labour market entrants are more likely to be overqualified. This inclination decreases over time. Workers with poor bargaining power (e.g., workers with several career breaks) are also inclined towards a higher overeducation incidence. Employer discrimination is assumed to increase the incidence of overeducation as well. Indeed, second-generation migrants are prone to labour market discrimination, and this in turn seems to relate to the likelihood of overqualification. Finally, a trade-off between job prospects and overeducation is hypothesised, with workers accepting temporarily a job for which they are overqualified. Our analysis refutes this hypothesis.

When the migrant condition was introduced to test whether it is related to the probability of being overeducated, our study confirmed the higher incidence of overeducation among the migrant population. In the search for personal characteristics determining the migrant educational mismatch, we tested the same theoretically based assumptions specifically for the migrant population. The relation between overeducation and these personal features found for the total population does not always hold in the case of migrants. Age is not negatively related to the skill mismatch as it is in the case of native workers, while being part of a poor bargaining power group (for gender or career history reasons) seems to have less effect on the overeducation incidence among migrant than native workers. Good job prospects are related to lower overeducation, as in the case of native workers. Furthermore, other migrant-specific conditions—such as sharing the same mother language with native workers—matter, reducing the chances of overeducation. We conclude that overeducation is a persisting problem that affects immigrants over the long run to a greater extent than it does native workers.

This study confirms the existing literature on the effect of personal characteristics on overeducation and extends it to the case of migrant workers. In addition, our analysis classifies migrants into groups according to their country of origin and destination combination—something that had not been done before—and shows that this sort of classification needs to be controlled for when researching overeducation among migrants. However, this article calls for further investigations into the theoretical underpinnings behind the country of origin and destination combination as a factor affecting overeducation. It also underlines the need for further empirical research on comparative cross-national differences.

Endnotes

[1]Of whom, 42% live in an EU country, 18% in a non-EU European country, 18% in Latin America, 13% in Asia, 6% in Africa, and 1% in North America & Oceania.

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 23 of 34

[2]For the purpose of matching jobseekers to vacancies, skill requirements need to be far more detailed. This is usually done by professional job analysts, who analyse skill requirements in job advertisements, study realised job matches, or undertake company studies of required skills. However, this method typically addresses a selected set of occupations and does not cover all occupations in a national labour market, as the latter is a huge undertaking.

[3]The latter finding could also be justified by the fact that a certain level of education is a necessary condition to observe overeducation.

[4]The project's website is www.wageindicator.org.

[5]Note that also web surveys that are based on email invitations from a large database (panel) of respondents are volunteer surveys. Only a very few web surveys, such as the LISS panel from Tilburg University, are randomly sampled using non-internet sampling frames. Note further that random sampled surveys may also be biased in the case of substantial non-response, which nowadays drops below 50% in many surveys.

[6]More than 80% of the Wageindicator websites visitors are first-time visitors. Given that respondents can win a prize to complete the survey, multiple entries with the same email are excluded. Therefore, although we cannot assure that a visitors enter the survey multiple times, we believe that chances are little.

[7]The most underrepresented groups are small groups, for example, workers with a part-time job of less than 10h per week.

[8]Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden, and the United Kingdom.

[9]Bulgaria, Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovak Republic, and Slovenia.

[10]For details about ISCED, see: www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm

[11]The six levels are: six levels within the corporate hierarchy, ranging from 1 = Helper, 2 = Occupation, 3 = First line supervisor, 4 = Departmental manager, 5 = Manager, company director, or chief executive, 6 = CEO. Respondents self-identify their occupation from a database of 1,700 occupational titles, which is sufficiently large for the vast majority of respondents, through text string matching and through a search tree. The second author has classified these 1,700 occupations according to the corporate hierarchy.

[12]This is not surprising in the case of the North America & Oceania group, which is composed of only four countries.

[13]Predicted probabilities for natives and migrant groups are computed and compared. Table 5 presents the difference of the predicted probability of each migrant group with respect to the corresponding native population on each continent. The regression results and the predicted probabilities are presented in table Table 10 in Appendix 3.

[14]Countries were chosen taking into consideration continent representativeness and data availabilities criteria.

[15]The regression results are reported in table Table 11 in Appendix 3.

[16]With the exeption of North America & Oceania, where we registered an incidence of 25% of migrant population. Nevertheless, this value is probably biased by the low number of respondents from that geographical area.

Appendix 1

Table 7 Country specific regressions/1

	Netherlands		Belarus		Germany		Brazil		Ukraine		Argentina	
MIGRANT	1.521	***	1.064	.	1.158	*	0.935		1.051		1.001	
Age	0.987	***	1.001		1.005	**	0.988	***	0.991	***	0.998	
Breaks	1.188	***	1.070	***	1.142	***	1.077	***	1.042	***	1.027	**
Second generation	1.187	***	1.164	**	1.908				1.141			
Gender (female = 1)	1.411	***	0.959	*	1.284	***	1.327	***	0.858	***	1.222	***
Education (ref. ISCED10)												
ISCED 20	1.936	***	0.491	.	1.182	**	1.387	***	0.486	.	1.516	***
ISCED 30	3.387	***	1.002		1.929	***			0.819		1.704	***
ISCED 40			0.943		1.663	***			1.195			
ISCED 50	4.801	***	1.200		1.925	***	1.245	*	1.641		1.757	***
ISCED 60	2.779	***	1.194		2.232	***	0.705	**	1.978	*	1.603	***
Corporate hierarchy												
CH2	0.871	*	0.516	***	0.453	***	0.585	***	0.422	***	1.054	
CH3	0.901		0.572	***	0.692		0.490	***	0.422	***	0.808	
CH4	0.286	***	0.453	***	0.303	*	0.315	***	0.396	***	0.576	*
CH5	0.308	***	0.316	***	0.377	***	0.375	***	0.299	***	0.667	.
CH6	0.149	***	0.289	***	0.363	**	0.292	***	0.235	***	0.434	**
Firm size (ref. 1–10)												
size 11-50	0.989		1.016		1.091		0.871	***	0.930	.	1.006	
size 51-100	1.034		1.012		1.054		0.740	***	0.942		0.835	**
size 101-500	0.959		0.991		1.079		0.737	***	0.989		0.891	*
size 500+	0.818	***	1.157	***	0.920		0.767	***	1.075	.	0.983	
No. of observations	53 564		43 007		22 848		26 267		29 327		20 852	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

Table 8 Country specific regression/2

	Indonesia		Kazakhstan		South Africa		Belgium		Low data-density countries		Mean	st.dev.
MIGRANT	1.658		1.139	*	0.910		1.542	***	1.450		1.221	0.270
Age	1.005		0.994	***	1.004		0.997		0.996		0.997	0.006
Breaks	1.025	*	1.051	***	1.045	***	1.126	***	0.994		1.072	0.058
Second generation	1.110		1.169	.			0.940			*	1.231	0.310
Gender (female = 1)	1.083	.	0.964		1.468	***	1.296	***	1.193		1.188	0.198
Education (ref. ISCED10)												
ISCED 20	3.488		2.647		0.329	**	1.299		2.559		1.575	1.005
ISCED 30	2.577		4.288	*	0.769		1.816	*	3.870	.	2.216	1.272
ISCED 40			5.111	*	1.427		1.894	.	3.462		2.242	1.507
ISCED 50	3.428	.	5.589	*	1.660	*	2.439	**	4.660	*	2.759	1.588
ISCED 60	3.541		5.969	*	2.087	***	2.495	**	5.817	*	2.764	1.723
Chorporate Hierarchy												
CH2	1.342		0.549	*	0.998		0.849		1.452		0.826	0.357
CH3	1.681	.	0.641		1.512		1.126		1.321		0.924	0.428
CH4	1.092		0.552	*	0.639		0.412	***	0.758		0.526	0.240
CH5	1.200		0.394	***	0.899		0.376	***	0.762		0.543	0.299
CH6	1.153		0.246	***	0.541	.	0.062	***	1.571		0.485	0.462
Firmsize (ref. 1–10)												
size 11-50	1.079		0.854	**	0.989		0.978		0.828		0.966	0.087
size 51-100	1.191	*	0.762	***	0.865	*	0.830	*	1.056		0.938	0.143
size 101-500	1.096		0.888	*	0.853	*	0.774	***	0.902		0.924	0.114
size500+	1.151	.	0.912	.	0.818	***	0.798	**	0.801		0.927	0.145
n. of observations	12 654		19 126		16 232		12 585		2 140			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 26 of 34

Appendix 2

Table 9 Two-stages regression correcting for selection bias

Probability of being a migrant	Whole sample						Netherlands				Germany			
	Estimate		Std.	Estimate		Std.	Estimate		Std.	Estimate		Std.		
Age	0.010	***	(0.000)	0.010	***	(0.000)	0.000		(0.001)	−0.013	***	(0.001)		
Gender (female = 1)	0.026	***	(0.006)	0.014	*	(0.006)	0.137	***	(0.017)	−0.007		(0.027)		
Education	YES			YES			YES			YES				
Continent of provenance (ref. EU15)														
EU12				0.376	***	(0.013)	0.701	***	(0.039)	0.842	***	(0.050)		
Africa				0.330	***	(0.013)	0.760	***	(0.033)	0.256	***	(0.073)		
C-S America				0.056	***	(0.010)	0.561	***	(0.023)	0.003		(0.044)		
Asia				0.457	***	(0.010)	0.742	***	(0.026)	0.438	***	(0.041)		
NAO				0.464	***	(0.022)	0.969	***	(0.059)	0.281	*	(0.113)		
Eur not-EU				0.428	***	(0.009)	−0.104	**	(0.031)	0.277	***	(0.034)		
Intercept	−1.937	***	(0.026)	−2.137	***	(0.027)	−1.317	***	(0.068)	−1.150	***	(0.059)		
Probability of overeducation														
Age	0.000		(0.001)	−0.001		(0.001)	−0.003		(0.002)	0.009	*	(0.004)		
Gender (female = 1)	0.096	***	(0.016)	0.088	***	(0.016)	0.180	***	(0.044)	0.258	**	(0.075)		
Breaks	0.038	***	(0.004)	0.036	***	(0.004)	0.065	***	(0.012)	0.064	*	(0.020)		
Education (ref. ISCED10)	YES			YES			YES			YES				
Corporate hierarchy	YES			YES			YES			YES				
Firmsize	YES			YES			YES			YES				
Intercept	−0.413	***	(0.010)	−0.063		(0.134)	−0.956	***	(0.227)	−0.933	**	(0.287)		
ath(rho)	−0.232	***	(0.034)	−0.381	***	(0.049)	0.013		(0.076)	−0.148		(0.147)		
rho	−0.228			−0.364			0.013			−0.147				
N	368564			368564			53583			22868				

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 ' '

Appendix 3

Table 10 Regression results equation 2 and predicted probabilities with continent of origin and continent of residence interactions

Continent of residence	Continent of provenance	Predicted probability	Prob. difference with natives		Estimate	Odds	Std.	Error	z-value	Pr(>|z|)
EU15	native	0.269	0.000	(Intercept)	−0.9981869	0.36854705	0.0186913	−53.404	<2E-16	***
EU15	EU15	0.299	0.029		0.1439838	0.42562223	0.0439123	3.279	0.001042	**
EU15	EU12	0.384	0.114		0.524447	0.62266919	0.0704424	7.445	9.69E-14	***
EU15	Africa	0.347	0.078		0.3661908	0.53152975	0.0729783	5.018	5.23E-07	***
EU15	C-S America	0.335	0.066		0.3119829	0.50348367	0.0548767	5.685	1.31E-08	***
EU15	Asia	0.321	0.052		0.2510214	0.47370738	0.0623645	4.025	0.000057	***
EU15	NAO	0.219	−0.050		−0.2705588	0.28118409	0.1359705	−1.99	0.046609	*
EU15	Eur not-EU	0.357	0.088		0.4111457	0.55596986	0.0904394	4.546	0.00000547	***
EU12	native	0.279			0.0485574	0.38688434	0.018798	2.583	0.009791	**
Africa	native	0.266			−0.0171214	0.36229072	0.0193865	−0.883	0.37715	
C-S America	native	0.313			0.2125563	0.45583216	0.0117194	18.137	<2E-16	***
Asia	native	0.326			0.2700074	0.4827871	0.0143084	18.87	<2E-16	***
NAO	native	0.285			0.0794261	0.39901319	0.037688	2.107	0.035077	*
Eur not-EU	native	0.384			0.5275274	0.62459022	0.0114767	45.965	<2E-16	***
EU12	EU15	0.169	−0.110		−0.7838634	0.20402574	0.6239081	−1.256	0.20898	
EU12	EU12	0.327	0.048		−0.2980656	0.4851738	0.1138719	−2.618	0.008856	**
EU12	Africa	0.349	0.071		−0.0378377	0.53725824	0.7061025	−0.054	0.957265	
EU12	C-S America	0.595	0.316		1.0227046	1.46969956	1.0145244	1.008	0.313424	
EU12	Asia	0.124	−0.155		−1.2598065	0.14108191	0.751244	−1.677	0.09355	.
EU12	NAO	0.696	0.417		2.047265	2.28662447	1.233512	1.66	0.096974	.
EU12	Eur not-EU	0.344	0.105		−0.1065295	0.52465556	0.2704692	−0.394	0.693678	

Visintin et al. IZA Journal of Migration (2015) 4:14

Page 27 of 34

Table 10 Regression results equation 2 and predicted probabilities with continent of origin and continent of residence interactions *(Continued)*

Africa	EU15	0.279	0.013	−0.3022772	0.38620201	0.1022032	−2.958	0.0031	**
Africa	EU12	0.263	−0.002	−0.3247473	0.35769568	0.4474864	−0.726	0.468014	
Africa	Africa	0.221	−0.045	−0.4960295	0.28356429	0.110506	−4.489	0.00000717	***
Africa	C-S America	0.174	−0.092	−0.2747152	0.21001374	0.1996193	−1.376	0.168761	
Africa	Asia	0.126	−0.140	−1.3354641	0.1437576	0.2667113	−5.007	5.52E-07	***
Africa	NAO	0.309	0.043	0.1631174	0.44769929	0.335223	0.487	0.626546	
Africa	Eur not-EU	0.446	0.180	0.8147534	0.80438578	0.4397227	1.853	0.0639	.
C-S America	EU15	0.296	−0.017	−0.3935247	0.42013785	0.1438612	−2.735	0.00623	**
C-S America	EU12	0.284	−0.029	−0.3885103	0.3972778	0.5983563	−0.649	0.516147	
C-S America	Africa	0.334	0.021	0.3664767	0.50172019	0.6122682	0.599	0.549469	
C-S America	C-S America	0.317	0.003	−0.3951961	0.46316079	0.0806993	−4.897	9.72E-07	***
C-S America	Asia	0.253	−0.060	−0.3456583	0.33866919	0.3403193	−1.016	0.309778	
C-S America	NAO	0.283	−0.030	−0.1267557	0.394748	0.280802	−0.451	0.651697	
C-S America	Eur not-EU	0.220	−0.093	−0.6899687	0.28279218	0.4093418	−1.686	0.091881	.
Asia	EU15	0.295	−0.031	−0.3958794	0.41768091	0.5167439	−0.766	0.443615	
Asia	EU12	0.302	−0.024	0.160168	0.43232818	0.5122427	0.313	0.754524	
Asia	Africa	0.333	0.008	−0.3755587	0.50027742	0.6969014	−0.539	0.589957	
Asia	C-S America	0.000	−0.325	−7.927744	0.00018275	17.7138962	−0.448	0.654482	
Asia	Asia	0.339	0.013	0.0772516	0.51270774	0.0998388	0.774	0.439071	
Asia	NAO	0.607	0.281	0.9495242	1.54326608	0.6850835	1.386	0.165748	
Asia	Eur not-EU	0.337	0.011	−0.2198606	0.50761464	0.1107282	−1.986	0.047079	*
NAO	EU15	0.153	−0.133	−0.5252812	0.18003558	0.1547359	−3.395	0.000687	***
NAO	EU12	0.173	−0.112	−1.0551452	0.20955683	0.3423818	−3.082	0.002058	**
NAO	Africa	0.065	−0.220	−1.7958227	0.06952797	0.47199	−3.805	0.000142	***

Table 10 Regression results equation 2 and predicted probabilities with continent of origin and continent of residence interactions *(Continued)*

NAO	C-S America	0.272	−0.013		−0.0461287	0.37455714	0.1386469	−0.333	0.739356	
NAO	Asia	0.224	−0.061		−0.5347483	0.28910862	0.1635464	−3.27	0.001077	**
NAO	NAO	0.347	0.061		0.0147423	0.5304598	0.2818781	0.052	0.958289	
NAO	Eur not-EU	0.136	−0.150		−1.0135648	0.15678192	0.4223469	−2.4	0.016402	*
Eur not-EU	EU15	0.502	0.118		0.0686529	1.00918099	0.128235	0.535	0.592396	
Eur not-EU	EU12	0.247	−0.137		−0.6905012	0.32870214	0.1216985	−5.674	1.4E-08	***
Eur not-EU	Africa	0.198	−0.186		−0.9095634	0.24725272	0.5742092	−1.584	0.113187	
Eur not-EU	C-S America	0.375	−0.009		−0.2520283	0.60041662	0.4433546	−0.568	0.569724	
Eur not-EU	Asia	0.397	0.012		−0.218502	0.65760285	0.0822371	−2.657	0.007885	**
Eur not-EU	NAO	0.424	0.040		0.0852799	0.73642086	0.5073541	0.168	0.866514	
Eur not-EU	Eur not-EU	0.426	0.042		−0.3533918	0.74339786	0.0949924	−3.72	0.000199	***
				AGE	−0.0060567		0.0004046	−14.968	<2E-16	***
				BREAK	0.071511		0.0024265	29.471	<2E-16	***
				GENDER	0.1467668		0.0077171	19.018	<2E-16	***
				Education level	YES					
				Corporate hierarchy	YES					
				Firm size	YES					

Note: Continuous variables where set constant to their mean, categorical variables reference values were set to the most common group

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

Visintin et al. IZA Journal of Migration (2015) 4:14

Page 29 of 34

Table 11 Regression results equation 3, interaction between the migrant condition and theoretical explanations of overeducation

	(Interaction)			(Interaction)		
	Estimate coeff.		Std.	Estimate coeff.		Std.
Age	−0.0133	***	0.00147	−0.0068	***	0.0004
Breaks	0.0946	***	0.00936	0.0726	***	0.0025
Gender (female = 1)	0.1981	***	0.0276	0.1489	***	0.008
Job prospects	−0.9032	***	0.02802			
MIGRANT	−1.8293	*	0.04494	−0.0914		0.0716
Age * Migrant	0.0193	***	0.00074	0.0094	***	0.0015
Breaks * Migrant	−0.0529	*	0.03251	−0.0116		0.0086
Gender * Migrant	−0.2068		0.18685	−0.0234		0.03
Language mismatch	0.0894		0.68452	0.0894	*	0.0375
Job prosp. * Migrant	0.1456	.	0.06303			
Continent (ref. EU15)						
EU12	−0.0955	*	0.04325	0.0507	**	0.0188
Africa	−0.0565		0.06962	−0.0190		0.0194
Latin America	0.0860	*	0.03355	0.2090	***	0.0118
Asia	0.5055	***	0.10587	0.2676	***	0.0144
North America and Oceania	0.3725	*	0.17095	0.0819		0.0377
Europe non-EU	0.4115	***	0.04667	0.5254	***	0.0115
Continent (ref. EU15) * Migrant						
EU12	0.0324		0.19144	−0.0614		0.088
Africa	−0.1048		0.22151	−0.4024	***	0.0665
Latin America	−0.1872		0.19152	−0.3391	***	0.0611
Asia	−0.7032	.	0.3797	0.0042		0.0608

Table 11 Regression results equation 3, interaction between the migrant condition and theoretical explanations of overeducation *(Continued)*

North America and Oceania	−0.8123	*	0.38313	−0.4641	***	0.0872
Europe non-EU	−0.2827	.	0.16117	−0.1813	***	0.0425
Education (ref. ISCED10) * Migrant						
ISCED 20	0.5112	*	0.6938	−0.1934		0.157
ISCED 30	0.2788	**	0.68213	−0.1789	*	0.0712
ISCED 40	0.3750	.	0.71559	0.0670		0.041
ISCED 50	0.3062	*	0.67516	−0.1529	*	0.0608
ISCED 60	0.7363	.	0.70176	−0.0431		0.0583
Corporate hierarchy * Migrant						
CH2	1.2925		0.64551	−0.0313		0.1381
CH3	2.0681		0.72598	0.2629	**	0.0948
CH4	1.3311		0.68186	0.0012		0.0682
CH5	1.3178		0.66506	0.1043		0.0754
CH6	1.3776		0.72548	−0.1036		0.1524
Firm size (ref. 1–10) * Migrant						
size 11-50	−0.1232		0.15772	0.0259		0.0448
size 51-100	0.3917	*	0.18919	−0.0176		0.0516
size 101-500	−0.0166		0.16223	0.0432		0.0432
size 500+	0.0019		0.16801	−0.0836	.	0.0436
Industry (ref. agr., man. and constr.) * Migrant						
Trade, transp. and hotels	0.0572		0.19539			
Commercial serv.	−0.1712		0.16287			
Education level	YES					
Corporate hierarchy	YES					

Visintin et al. IZA Journal of Migration (2015) 4:14

Page 31 of 34

Table 11 Regression results equation 3, interaction between the migrant condition and theoretical explanations of overeducation *(Continued)*

Firm size	YES	
Industry	YES	
Nagelkerke presudo R2	0.099	0.035
LR chi2	2,435.13	8,843.44
Pr(> chi2)	<0.0001	<0.0001

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

References
Andreadis I (2013) Can timestamp analyses show the bottlenecks in web surveys? *WageIndicator Conference August 26–28 2013*. The Netherlands, Amsterdam
Askitas N and Zimmermann K (2015). The internet as a data source for advancement in social sciences. Int J Manpow 36(1).
Autor DH, Levy F, Murnane RJ (2003) The Skill Content of Recent Technological Change: An Empirical Investigation. Quarterly Journal of Economics 118(4):1279–1333
Battu H, Sloane PJ (2002) To what extent are ethnic minorities in Britain overeducated? Int J Manpow 23(3):192–208
Berman E, Bound J, Machlin S (1998) Implications of skill-biased technological change: international evidence. Q J Econ 113(4):1245–1279
Borjas GJ (1987) Self-Selection and the Earnings of Immigrants. Am Econ Rev 77(4):531–553
Borjas GJ (1989) Economic theory and international migration. Int Migr Rev 23(3):457–485
Burkert C, Seibert H (2007) Labour Market Outcomes after Vocational Training in Germany: Equal Opportunities for Migrants and Natives? IAB, IAB Discussion Paper 200731, Nuremberg
CBS (Centraal Bureau voor de Statistiek) (1993) Standaard Beroepsclassificatie 1992. CBS (Statistics Netherlands), Voorburg/Heerlen
Cedefop (2010) The Skill Matching Challenge. Analysing Skill Mismatch and policy Implications. Publications Office of the European Union, Luxembourg
Chevalier A (2003) Measuring Over-education. Economica 70:509–531
Chiswick BR, Miller PW (2009a) The international transferability of immigrants' human capital skills. Econ Educ Rev 28(2):162–169
Chiswick BR, Miller PW (2009b) Educational Mismatch: Are High-Skilled Immigrants Really Working at High-Skilled Jobs and the Price They Pay If They Aren't? IZA, IZA Discussion Paper No. 4280, Bonn
Clogg CC, Shockey JW (1984) Mismatch between occupation and schooling: a prevalence measure, recent trends and demographic analysis. Demography 21:235–257
Cohn E, Khan S (1995) The wage effects of overschooling revisited. Labour Econ 2(1):67–76
Dahlstedt I (2011) Occupational Match. Over- and Undereducation among immigrants in the Swedish Labor Market. J Int Migr Integr 12(3):349–367
De Vries DH and Tijdens K (2010). Using the WageIndicator web-survey to monitor Human Resources for Health, Paper presented at the International Conference on Research in Human Resources for Health, Rio de Janeiro, June.
Fabo B and Tijdens K (2014). Using Web Data to Measure the Demand for Skills, CELSI Discussion Paper No. 21, Central European Labour Studies Institute.
Green C, Kler P, Leeves G (2007) Immigrant overeducation: evidence from recent arrivals to Australia. Econ Educ Rev 26:420–432
Groot W, Maassen van den Brink H (2000) Overeducation in the labor market: a meta-analysis. Econ Educ Rev 19(2):149–158
Guzi M, de Pedraza P (2015) A web survey analysis of the subjective well-being of spanish workers, IZA Discussion Paper No. 7618. IZA, Bonn, Available at SSRN: http://ssrn.com/abstract=2332633
Halaby C (1994) Overeducation and skill mismatch. Sociol Educ 67(1):47–59
Hartog J, Jonker N (1997) A job to match your education: does it matter? In: Heijke H, Borghans L (eds) Towards a Transparent Labour Market for Educational Decisions. Aldershot, Avebury
Heckman JJ (1976) The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. Ann Econ Soc Meas 5(4):475–492
ILO (2007) Annex 1 International Standard Classification of Occupations (ISCO-08) – Conceptual Framework. Draft for consultation through second questionnaire on updating ISCO-88. ILO, Geneva, Available at: http://www.ilo.org/public/english/bureau/stat/isco/index.htm
Jensen U, Gartner H, Rässler S (2007) Measuring Overeducation with Earnings Frontiers and Multiply Imputed Censored Income Data. IAB Discussion Paper No. 11/2006, Nuremberg
Joona PA, Gupta ND, Wadensjö E (2014) Overeducation among immigrants in Sweden: incidence, wage effects and state dependence. IZA J Migr 3:9
Kler P (2005) Graduate overeducation in Australia: a comparison of the mean and objective methods. Educ Econ 13(1):47–72
Korpi T (2012) Importing Skills: Migration Policy, Generic Skills and Earnings among Immigrants in Australasia, Europe and North America. In: Galgóczi B, Leschke J, Watt A (eds) Migration and Labour Markets in Troubled Times: Skills Mismatch, Return Migration and Policy Responses. Ashgate, Aldershot, pp 261–284
Korpi T, Tahlin M (2009) Educational mismatch, wages, and wage growth: overeducation in Sweden, 1974–2000. Labour Econ 16:183–193

Visintin *et al. IZA Journal of Migration* (2015) 4:14

Page 34 of 34

Kureková L, Beblavy M, Thum A (2014) Using internet data to analyse the labour market: a methodological enquiry, IZA Discussion Paper 8555. IZA, Bonn, Available at SSRN: http://ssrn.com/abstract=2514748

Leuven E, Oosterbeek H (2011) Overeducation and Mismatch in the Labor Market. Hand Econ Educ 4:283–326

Lindley J, Lenton P (2006) The Over-Education of UK Immigrants: Evidence from the Labour Force Survey. Sheffield Economic Research Paper Series SERP Number: 2006001, Sheffield

Machin S (2001) The changing nature of labour demand in the new economy and skill-biased technological change. Oxf Bull Econ Stat 63(S1):753–776

Mavromaras K, McGuinness S, Fok YK (2009) Overskilling Dynamics and Education Pathways. IZA Discussion Paper 4321, Bonn

McGuinness S, Sloane PJ (2011) Labour market mismatch among UK graduates: an analysis using REFLEX data. Econ Educ Rev 30(1):130–145

Nielsen CP (2011) Immigrant overeducation evidence from Denmark. J Popul Econ 24(2):499–520

O*net (2002). Appendix D – The Development of the Occupational Information (O*NETTM)Analyst Database.

OECD (1999) Classifying Educational Programmes Manual for ISCED-97 Implementation in OECD Countries. OECD, Paris

OECD (2007) On the move. International migration. DELSA Newsletter, Issue 5. OECD, Paris

Piracha M, Tani M, Vadean F (2012) Immigrant Over- and Under-education: The Role of Home Country Labour Market Experience. IZA J Migr 1(1):1–21

Sicherman N (1991) 'Overeducation' in the labor market. J Labor Econ 9(2):101

Sicherman N, Galor O (1990) A theory of career mobility. J Polit Econ 98(1):169–192

Sloane PJ, Battu H, Seaman PT (1999) Overeducation, undereducation and the British labour market. Appl Econ 31:1437–1453

Steinmetz S, Tijdens K and de Pedraza P (2009) Comparing Different Weighting Procedures for Volunteer Online Panel: Lessons to Be Learned from German and Dutch WageIndicator Data. University of Amsterdam, AIAS Working Paper 09–76, Amsterdam.

Steinmetz S, Raess D, Tijdens K and de Pedraza P (2013a). Measuring Wages Worldwide: Exploring the Potentials and Constrains or Volunteer Web Based Surveys, in Sappleton, N. Advancing Research Methods with New Technologies, IGI Global (100–119).

Steinmetz S, de Vries DH, Tijdens K (2014) Should I stay or should I go? The impact of working time and wages on retention in the health workforce. Hum Resour Health 12(1):23

Tijdens K (2014) Dropout rates and response times of an occupation search tree in a web survey. J Off Stat 30(1):23–43

Tijdens KG, van Zijl S, Hughie-Williams M, van Klaveren M, Steinmetz S (2010) Codebook and explanatory note on the WageIndicator dataset: A worldwide, continuous, multilingual web-survey on work and wages with paper supplements. AIAS Working Paper 10–102, Amsterdam, University of Amsterdam

Tijdens K, De Vries DH, Steinmetz S (2013a). Health workforce remuneration: comparing wage levels, ranking, and dispersion of 16 occupational groups in 20 countries. Hum Resour Health 11(11)

Tijdens K, De Ruijter E, De Ruijter J (2013b) Comparing tasks of 160 occupations across eight European countries. Empl Relat 36(2):110–127

Van der Velden RKW, van Smoorenburg MSM (1997) The Measurement of Overeducation and Undereducation: Self-Report vs. Job-Analyst Method. Maastricht University, Research Centre for Education and the Labour Market (ROA), Maastricht

Verdugo RR, Verdugo NT (1989) The impact of surplus schooling on earnings. J Hum Resour 24(4):629–643

Verhaest D, Omey E (2006) Discriminating between alternative measures of over-education. Appl Econ 38(18):2113–2120

Wald S, Fang T (2008) Overeducated immigrants in the Canadian labour market: Evidence from the workplace and employee survey. Can Public Policy 34(4):457–479