

ORIGINAL ARTICLE

Open Access

Overeducation among immigrants in Sweden: incidence, wage effects and state dependence

Pernilla Andersson Joona^{1*}, Nabanita Datta Gupta² and Eskil Wadensjö¹

* Correspondence:

Pernilla.Andersson.Joona@sofi.su.se

¹Swedish Institute for Social Research, Stockholm University, SE-106 91 Stockholm, Sweden

Full list of author information is available at the end of the article

Abstract

The utilization and reward of the human capital of immigrants in the labor market of the host country has been studied extensively. Using Swedish register data from 2001–2008, we extend the immigrant educational mismatch literature by analyzing incidence, wage effects and state dependence in overeducation among natives and immigrants. In line with previous research we find a higher incidence and a lower return to overeducation among immigrants indicating that immigrants lose more from being overeducated. We find a high degree of state dependence in overeducation both among natives and immigrants, but considerably higher among immigrants.

JEL codes: J61, I21, J24, J31, F22

Keywords: Educational mismatch; Immigrants; State dependence; Wages

1 Introduction

Sweden has been an immigration country since WW2. Up to the early 1970s mainly labor migrants arrived, but since the 1980s the majority of entrants has been refugees and tied movers. At the end of 2012, 1.5 million individuals born outside Sweden lived in the country, corresponding to 15 percent of the population (Statistics Sweden 2012). Although employment has traditionally been higher among labor migrants than among refugees or tied movers, since the 1970s both groups have had employment rates below that of natives. In 2009 the employment rate for foreign born men was about 68 percent compared to about 78 percent among native men. Among foreign born men and women from countries outside Europe the employment rate was 57 percent (Eriksson 2011). Differences in human capital characteristics such as education, work experience and Swedish language proficiency but also discrimination are likely explanations for this gap.

There is not only a problem of low employment rates among non-Western immigrants but also a problem of underutilization of their skills once employed. Previous studies, both for Sweden and for other countries, find that the incidence of overeducation, i.e. the degree to which workers have an education that exceeds the one that is required for the job, is larger among immigrants than among natives (Dahlstedt 2011; Nielsen 2011; Wald and Fang 2008). It is often assumed that due to limited transferability of human capital an initial mismatch, or overeducation, among immigrants can be expected. However, if there is a higher degree of persistence and state dependence in overeducation among immigrants this would be a sign of a more severe problem. By

state dependence we mean that overeducation in a previous period will have an independent effect on current overeducation.

Why is there a problem with being overeducated? First of all this means that the skills of workers are not fully used. But it is also a problem because the returns to education for those overeducated are lower than the returns to the same level of education for workers that are correctly matched. This is one of the most persistent findings in the literature on overeducation; overeducated workers earn more than correctly matched workers in the same types of jobs but earn less than correctly matched workers with the same years of schooling see e.g. (Chiswick and Miller 2008; Chiswick and Miller 2010a; Duncan and Hoffman 1981; Hartog 2000; Korpi and Tåhlin 2009). Studies focusing on the wage effects of overeducation for immigrants tend to find that the return to overeducation is lower for immigrants than for natives meaning that immigrants lose more from being overeducated than natives do (Nielsen 2011; Wald and Fang 2008).¹

In this paper we address the questions of the incidence and wage effects of overeducation, thus bringing evidence on these issues in the Swedish case, and we also extend the analysis to include estimations of state dependence. In particular, we study if state dependence is a more severe problem for immigrants than for natives, which is a novel question in the literature on overeducation of immigrants. If this is the case, this is an indication that overeducation is not only a passing problem for immigrants but that it may have scarring effects affecting the integration of immigrants also in the long run. This is the major contribution of the paper. The use of rich Swedish register data gives us the possibility of estimating panel models that control for unobserved heterogeneity. Furthermore, the Swedish case is interesting from an international standpoint because even though it has had a fairly long history of receiving immigrants, its high-wage, high-labor cost economic regime may present barriers to immigrants to utilizing their skills.

We estimate models of state dependence in overeducation by regressing overeducation in period t on overeducation in period $t-1$. To handle the initial conditions problem we estimate random effects dynamic probit models by including controls for initial overeducation as has been suggested by Wooldridge (2005). To analyze these questions we use Swedish register data on all employed workers covering the period 2001 to 2008.

The remainder of the paper is structured as follows. In section 2 the literature in the area is reviewed, in section 3 the data, variables and methods are described, in section 4 the incidence of overeducation is discussed, in section 5 the returns to actual, required, over- and undereducation are analyzed, and in section 6 state dependence in overeducation is analyzed. Section 7 summarizes our findings and we draw some conclusions.

2 Previous research

The literature on educational mismatch among immigrants is small but increasing, linking the immigrant wage assimilation literature founded by Chiswick and Borjas to the literature on overeducation (Duncan and Hoffman 1981; Hartog 2000; Verdugo and Verdugo 1989). Many studies find that immigrants have higher rates of overeducation than natives, though the size of the gap varies by immigrant ethnicity and the destination labor market. For instance, while the rates of immigrant overeducation tend to be

higher than those for natives in many countries (see Battu and Sloane 2002 for Britain, Fernández and Ortega 2008 for Spain, Lianos 2007 for Greece, Green et al. 2007 for Australia, and Nielsen 2011 for Denmark), substantially higher rates of immigrant *undereducation* compared to natives are found for the US and Canada (Chiswick and Miller 2008, Chiswick and Miller 2010a).

In the search for reasons underlying the immigrant educational mismatch, one starting point is to decompose the lower payoff to schooling for foreign born from non-Western countries compared to natives. Chiswick and Miller (2008) present a new decomposition technique that links overeducation to a less than perfect transferability of immigrants' human capital and undereducation to favorable selection in migration. Applying data from the 2000 U.S. Census, they find that natives and immigrants receive about the same return to the level of required education in the occupation. The lower payoff to schooling for immigrants is largely accounted for by the higher proportion of undereducated immigrant men relative to native men and their relatively strong wage performance. Thus, the evidence is consistent with the notion of self-selection of immigrants with superior ability or motivation to the US.

The same decomposition technique applied to Canadian data shows that the lower return to schooling of immigrants is not as affected by undereducation as in the US, since the effect is only twice as large as the effect of overeducation, whereas it is about ten times as large in the US (Chiswick and Miller 2010b). Evidence from Australia reveals, somewhat paradoxically, that over and undereducation are equally important in accounting for the lower payoff to schooling for immigrants, indicating that the Australian strict screening policy does not necessarily improve skill transfers across countries (Chiswick and Miller 2010d).

Our paper makes two contributions to this literature. First, evidence on immigrant educational mismatch is sparse in Sweden. Korpi and Tåhlin (2009) track education, wages and wage growth in Sweden over the period 1974–2000 and find significant differences in returns to education across matched categories even after that variation in ability is taken into account. Furthermore, they find that wage growth among overeducated workers does not exceed that of other groups. This means that overeducated workers in Sweden are penalized early on in their careers by a lower rate of return to schooling, and that this effect persists over time. They estimate both cross-sectional and fixed effects models correcting for unobserved ability effects, but they do not perform separate analyses for natives and immigrants. A recent paper by Dahlstedt (2011) on occupational match based on logistic match regressions run on the LISA database from 2003 confirms that immigrants have lower rates of match and higher rates of overeducation than the native population. He also shows that it is important to differentiate between immigrants according to country of origin – immigrants from Iraq in particular display low levels of match and a high level of overeducation.

Second, we extend the immigrant educational mismatch research area by estimating the extent to which there exists state dependence in overeducation and if this appears to be a more severe problem among non-Western immigrants than among natives. According to the career mobility hypothesis (Sicherman and Galor 1990) overeducation is a temporary phenomenon and should be seen as an investment in work experience that could lead to better employment opportunities in the future. A higher incidence of overeducation among recently arrived immigrants is expected. Mavromaras and McGuinness

(2012) analyze state dependence in overskilling in the Australian labor market. Overskilling differs slightly from the concept of overeducation; overskilling is defined as a situation when wage-earners report that their skills are not fully utilized in their job. They find evidence of state dependence in particular among workers with higher education. Their study is based on working age employees but does not distinguish between natives and immigrants. A high degree of state dependence in overskilling contradicts the idea of overskilling being a temporary phenomenon.

Piracha et al. (2012) analyze the correlation between educational mismatch in home and host countries for immigrants arriving to Australia. They show that part of the higher incidence of overeducation among immigrants can be explained by that they have been overeducated already in their home country.

Dolton and Vignoles (2000) study persistence in overeducation and find that 38 percent of U.K. graduates were overeducated in their first job and that 30 percent were still overeducated six years later. Frenette (2004) investigates overqualification among Canadian workers and finds that graduates who enter jobs for which they are overqualified shortly after graduation often remain overqualified in the near future. Analyzing transitions from overeducation using data for two consecutive years, Rubb (2003) finds that three out of four overeducated workers in year t are still overeducated in year $t + 1$. Cuesta and Budría (2012) use the German Socio-Economic Panel to analyze overeducation dynamics and personality. Their results are in line with previous research suggesting that overeducation mainly reflects unobserved differences in personal characteristics such as ability or motivation. Regarding state dependence in overeducation, their calculations suggest that almost 18 percent of the overeducation risk is due to individual state dependence.

Our overview of the existing research on persistence and state dependence in overeducation indicates that there is evidence of a high degree of persistence. None of the above mentioned studies have however focused on heterogeneous effects across different groups of workers, such as natives and immigrants. This paper adds to the literature in analyzing whether state dependence is a more severe problem for non-Western immigrants than for natives in Sweden.

3 Data and method

3.1 Definition of overeducation

To determine whether someone is overeducated the norm within the occupation has to be determined. Each individual's educational attainment is then compared to the norm in that occupation. Individuals with higher educational attainment than the norm are defined as overeducated, while individuals with lower educational attainment than the norm are defined as undereducated. Individuals whose education is the same as the norm in the occupation are defined as correctly matched, or as having the required level of education.

There are different methods to determine the occupational norm. One is the so-called realized matches approach with the norm defined as the number of years of schooling within a one standard deviation range around the mean; individuals are defined as being undereducated, overeducated or having the required education in relation to this norm (Verdugo and Verdugo, 1989). A second method is to use the most frequently occurring number of years of schooling, i.e. the modal value, within

occupations to define the norm instead of the mean. These two approaches are used in the current study.

A third method is to define the norm by using job analysis. Professional job analysts determine the educational requirements for a job and the individual's educational attainment is compared to this. A fourth method is worker self-assessment where workers are asked in surveys about the educational requirements of their job.

All methods have their weaknesses and strengths (see Hartog 2000 for a discussion) but in many cases the choice of method is driven by data availability. In our case, we do not have access to survey data so we cannot use self-assessed educational requirements for an individual's job as a way of measuring overeducation. On the other hand we have detailed register data on all employees and workers have been divided into more than 110 distinct occupations. Results from ORU earnings equations have been found to be robust to whether the reference level of education is measured according to realized matches or worker self-assessment (Chiswick and Miller 2010c).

3.2 Data and sample restrictions

We use Swedish register data for the period 2001–2008. The data are collected primarily for administrative purposes and maintained by Statistics Sweden. It includes all individuals 16–64 years who were registered residing in Sweden by the end of December each year. In the main analysis, regressions are estimated separately by gender and separately for natives (those born in Sweden independently of where their parents are born), Western immigrants (those born in a Nordic country, within EU15 and in North America), and non-Western immigrants (those born elsewhere). The population used in our analysis is restricted to those aged 25–57, who were employed in November each year and for whom we have information on both occupation and education. The year 2001 has been chosen as the starting year since this is the first year for which information on occupations exists in the registers. Occupations are classified using the SSYK-code in the Swedish registers. We define occupations at the three-digit level which leaves us with 113 occupations. Occupations with fewer than 100 workers are excluded and so are military personnel. Following previous literature we also exclude the self-employed.

When defining the norm we include workers who are between 25 and 57 years of age, who have not been enrolled in education during the year and who have been in Sweden for three years or more. The most recently arrived immigrants are excluded when we calculate the norm but are included in the analysis of overeducation.

In section 5 we analyze the wage-effects of over, under and required education. Information on wages exists in the Swedish registers for all employees in the public sector and for a sample of about 45 percent of those employed in the private sector.

The probability of being overeducated is analyzed using the whole sample while the ORU-regressions used for analyzing wage effects of overeducation are based on the sample of employees for whom information on monthly wages exists. The analysis of state dependence in overeducation is based on a balanced panel for the period 2001–2008.

Information on both occupation and education has to be present for an individual to be included in the sample. Information on education is missing for more immigrants than natives, especially newly arrived immigrants. Information is missing on education for less than 0.1 percent of natives but for around 3 percent of non-Western immigrants.

Information on education is collected in different ways for different segments of the population. For those being educated in Sweden, either natives or foreign-born, information stems from reports continuously received by Statistics Sweden from the educational institutions. This information is generally of high quality. For those with education dating back to before 1990, the 1990 census (the latest census in Sweden) has been used.²

One challenge is that the quality of the education variable may be more variable for the foreign-born who have immigrated after completing their education in their home country or in another country than Sweden. Those registered as new immigrants in Sweden are asked by Statistics Sweden to fill out a questionnaire with questions regarding their education, but many who receive the questionnaire do not answer it, which means that information is lacking for many newly arrived immigrants. However, the information received through the questionnaire is gradually complemented by other data sources; from the Public Employment Service for those who have been searching for work through an employment office, from the National Health Board for those who apply for a permit to work as medical doctors, dentists, nurses etc. Still, individuals for whom we have complete information on education may differ from those where information on education is lacking, meaning a potential selection problem. In part this selection problem is mitigated by omitting the most recently arrived immigrants when constructing the educational norms.

Note that measurement errors in education exist also for those who have grown up in Sweden and been educated in Sweden if they also have studied abroad. They are not covered by the questionnaire sent out by Statistics Sweden. For example, a person with a BA from a university in Sweden and a PhD from a university in the US will have a BA recorded as their highest degree according to the statistics, as the Swedish degree is the only recorded one. There could be a difference, however, between the foreign-born growing up in Sweden and native Swedes regarding how often they study abroad and receive their highest degree from a country other than Sweden.

A final issue with comparing an education from Sweden with those received in other countries is that the quality could differ. The quality of the education could be higher or lower if achieved in another country even if it is labeled as the same. Even if educations are at the same level, education acquired in Sweden may be preferred by the employers. It may also be such that some educations acquired in another country are not possible to use in Sweden directly but have to be validated by an authority. It can be complicated and take time, especially for those coming from countries outside EU/EEA. It may also be the case that immigrants when arriving in Sweden have qualifications in occupations with an excess supply in the Swedish labor market. Finally, for some occupations it may be necessary to have good knowledge of the Swedish language, making it impossible to get a job before this requirement is fulfilled. We will address these issues via the empirical specification.

3.3 Econometric analysis

In the empirical part of the paper, we first present the incidence of overeducation among natives and immigrants in Sweden. Second, we analyze wage-effects of over,

under and required education by estimating the ORU-model first developed by Duncan and Hoffman (1981).

$$\ln w_{it} = \beta_0 + \beta_1 UE_{it} + \beta_2 RE_{it} + \beta_3 OE_{it} + \delta X_{it} + u_{it} \quad (1)$$

Undereducation is measured as years of deficit education in relation to the “norm” in the occupation which we either derive using the mean plus/minus one standard deviation or the modal years of schooling. Overeducation is in turn measured as the number of excess years of schooling an individual has. Years of undereducation is set to zero for all except for those who are defined as undereducated and years of overeducation is set to zero for all except for those who are overeducated. Required education corresponds to the number of years that is the norm within the occupation.

There are several problems related to the ORU specification of the wage equation, in particular problems with omitted variables and measurement error. In the literature on the return to acquired education the endogeneity problem has since long been acknowledged and specifications trying to correct for this is almost always estimated, primarily using instrumental variable analysis. In the ORU specification, the problem arises if sorting into years of under, over or required education is correlated with the error term, i.e. correlated with some unobservable variable that also is correlated with wages. If this is not taken into account in the empirical specification, we cannot claim to have estimated a causal effect of overeducation on wages. Although the focus of this paper is not primarily on estimating a causal wage effect, we estimate wage regressions controlling for individual fixed effects. This will not, however, correct for the fact that some unobservables change over time; the individual fixed effects will only take care of the problem with time invariant unobservables that are correlated both with wages and years of overeducation.

A second problem in the ORU-specification is measurement error. We have already discussed the problems surrounding the educational variable in the Swedish registers which is used to determine both an individual’s acquired number of years of schooling and number of years of schooling that is required for a job. Leuven and Oosterbeek (2011) points out that the measurement error in the key variables in the ORU analysis is likely to contain an even larger measurement error since both over and undereducation are defined as the difference between acquired and required years of schooling and that this leads to the measurement error becoming more severe. We have also discussed the possibility that measurement error is even more severe for immigrants than for natives, in particular if their education was received in the country of origin.

In spite of these problems, many researchers have estimated the ORU-model and the results are remarkably consistent both over time and space (see Hartog 2000): (1) The returns to actual years of schooling are lower than the returns to required years of schooling; (2) The returns to overeducation are positive, but smaller than the returns to required education, i.e. $\beta_3 > 0$ but $\beta_3 < \beta_2$. This means that overeducated workers earn more than correctly matched workers in the same types of jobs but less than correctly matched workers with the same years of schooling; (3) The returns to under-education are negative, but the estimate is smaller than the estimate for the returns to required education, i.e. $\beta_1 < 0$ but $|\beta_1| < \beta_2$.

One concern that has been raised in previous studies is if unobserved heterogeneity can influence the results (e.g. Chevalier 2003; Bauer 2002; Korpi and Tählin 2009; Nielsen 2011). Bauer (2002) argues that controlling for unobserved heterogeneity might be important if individuals with lower ability need more education to acquire a job for

which they are formally overeducated. He further argues that if there is a negative correlation between the probability of being overeducated and ability, then one would expect that we underestimate the returns to overeducation and overestimate the returns to undereducation when not controlling for unobserved heterogeneity.

In the case of immigrants, it can also be argued that some employers might require a stronger signal, i.e. more formal education for the same job from an immigrant applicant than from a native one. In the hiring process, a high level of education is an indication of high ability and conscientiousness, but this may be offset by a general skepticism towards people with a foreign background. Thus, it is not a priori clear how the results are expected to change by controlling for unobserved heterogeneity, in particular for immigrants, given that many studies point to a tendency of immigrants being discriminated against in the hiring process in the Swedish labor market (Carlsson and Rooth 2007; Bursell 2007; Arai et al. 2010).

Leuven and Oosterbeek (2011) are critical of the attempts that have been made to control for unobserved heterogeneity, both using fixed-effect models and instrumental variables. As a result, they argue that it is very difficult to get a credible estimate of the causal wage-effect of being over or undereducated.

The discussion above has mostly been about selection into over- and undereducation and how it may be correlated with ability, given employment. Another type of selection stems from the fact that we observe the occupation only for those who are employed. In Sweden, one of the main issues in the debate about integration of immigrants is that employment rates are substantially lower. A general tendency on the Swedish labor market is that employment increases with educational attainment (Eriksson 2011). Among highly educated individuals education (and thereby overeducation) may be positively correlated with the probability of being employed.

In our sample, almost 90 percent of native men were employed in November 2008. Among native women, employment is slightly lower except for those with higher education of three years or more and those with post-graduate education where employment rates for women and men are about the same. Employment is about 25 percentage point higher among native men and women compared to immigrants. A number of factors affect the immigrants' probability of getting a job given their education; where they live (Zenou et al. 2010), which type of job that they apply for (Carlsson and Rooth 2007)³, and the period of arrival to Sweden (Åslund and Rooth 2007).

3.4 Estimating state dependence in overeducation

In the introduction it was argued that state dependence in overeducation might be a more severe problem than the incidence of overeducation. If it exists and is higher among immigrants than natives, this indicates that a high incidence of overeducation among newly arrived immigrants is not only an initial problem but can have long-lasting negative effects on their labor market integration. Therefore, it is important to estimate the effect of earlier overeducation on future overeducation.

Following Mavromaras and McGuinness (2012), the model to be estimated is

$$OE_{it} = X'_{it}\beta + \gamma OE_{it-1} + \varepsilon_i + u_{it} \quad (2)$$

where ε_i is the unobserved heterogeneity which together with u_{it} , which is assumed to be *iid*, are components of the error term. The dependent variable is a dummy variable

taking the value one if the individual is overeducated in period t and zero if not. Since the left-hand side variable is a dummy variable we would like to estimate a probit model and since we use panel data we could choose between a fixed and a random effect model. However, since we are not only interested in the effect of time-varying covariates on the outcome but the variable of interest is in fact a lagged dependent variable a random effects model is the one we should estimate. But estimating a simple random effect probit model would lead to biased estimates of the effect of previous overeducation on present overeducation. To be able to establish if there is a direct effect of lagged overeducation on present overeducation net of all factors that affect the probability of being overeducated in the first place, we need to address two problems. The first is the so called initial conditions problem and occurs since the lagged dependent variable is likely to be correlated with the individual effect, ε_i . Unobservables that are correlated with the outcome will in almost all cases be correlated with the lagged dependent variable. Three different methods have been suggested to correct for this developed by Heckman (1981), Orme (2001) and Wooldridge (2005). A comparison of these three estimators has shown that none of them outperforms the other two, and all three estimators display in most cases satisfactory results. However, the Heckman estimator for which Stewart (2006) has developed a STATA code, is more time consuming than the other two (Arulampalam and Stewart 2009). We have therefore chosen to follow Wooldridge (2005), where the relationship between the individual effect and the lagged dependent variable is modeled conditional on the initial value of overeducation and exogenous explanatory variables.

The second problem arises because of the assumption of independence between the covariates and the error term. This is resolved by applying the Mundlak correction which in practice means that we include individual means of each of the time varying variables that are assumed to be correlated with the unobserved heterogeneity (Mundlak 1978). In our case individual means over age, number of children, years of schooling, and years in Sweden (for immigrants) are included. The model to be estimated then becomes:

$$OE_{it} = X'_{it}\beta + \gamma OE_{it-1} + \delta OE_{it=0} + \overline{X'_i}\alpha + \varepsilon_i + u_{it} \quad (3)$$

There are basically two ways in which the incidence of overeducation can change: (i) if the individual's years of schooling change or (ii) if the norm within the occupation changes. The norm within the individuals occupation can in turn change for two reasons: (i) if the individual stays in the same occupation and the norm within that occupation changes and (ii) if the individual him or herself change occupation. The primary source for changes in overeducation is likely to be job changes. However, changing job is no guarantee for improving the match. Even if one changes job one could end up in another job that one is overeducated for. This makes us wanting to investigate the effect of a job change between period $t-1$ and period t on overeducation in period t . The full model to be estimated then becomes

$$OE_{it} = X'_{it}\beta + \gamma OE_{it-1} + \delta OE_{it=0} + \rho \Delta job_{it} + \overline{X'_i}\alpha + \varepsilon_i + u_{it} \quad (4)$$

Where Δjob_{it} is defined as $job_{it} - job_{it-1}$. This is a dummy variable that takes the value one if the worker has changed job and zero otherwise.

4 The incidence of educational mismatch in Sweden

In Table 1 we present the incidence of under and overeducation by gender and birth region for the entire Swedish population. We present the shares calculated using the realized matches method both using the mode and a one standard deviation range around the mean as the norm.

First, we find that the incidence of overeducation differs depending on the method used for defining the occupational norm. Using the modal value as the norm has also in previous studies produced estimates of overeducation that are higher compared to the other methods, while the mean method and job analysis give similar results (see for example Bauer 2002 and Nielsen 2011). However, all methods show that the incidence of overeducation is higher among immigrants than among natives. Using the mode method, the incidence of overeducation among males is about 33 percent among natives and between 42 and 50 percent among non-Western immigrants, depending on birth region. Overeducation levels are slightly lower for female natives and immigrants. These numbers are in line with what has been found for Denmark (Nielsen 2011), the US (Chiswick and Miller 2008) and Germany (Bauer 2002). Dahlstedt (2011) reports a similar incidence of overeducation using job analysis and Swedish register data for 2003.

Looking at the incidence of undereducation, we find a slightly different pattern. Immigrants from the Nordic countries experience the highest degree of undereducation while immigrants from North America have the lowest incidence of undereducation. The difference in the incidence between natives and non-Western immigrants is not

Table 1 Incidence of over and undereducation (%) by birth region and gender among those aged 25–57 years in 2008

Region	Mean			Mode		
	Correctly matched	Over-educated	Under-educated	Correctly matched	Over-educated	Under-educated
Men						
Sweden	71.9	11.9	16.3	36.0	33.2	30.8
Other Nordic countries	64.5	11.2	24.3	36.5	25.3	38.2
EU15	56.7	26.6	16.7	23.9	43.7	32.4
Rest of Europe	65.9	20.4	13.7	22.7	47.7	29.6
Africa	55.1	26.2	18.6	18.3	48.6	33.1
North America	54.9	32.6	12.5	22.0	50.0	28.1
South America	62.9	20.2	16.9	24.9	44.3	30.8
Asia	53.6	23.9	22.5	20.9	42.2	36.9
Women						
Sweden	76.4	12.4	11.2	39.3	31.1	29.6
Other Nordic countries	70.6	12.8	16.6	35.1	27.7	37.2
EU15	62.1	23.6	14.3	29.9	40.5	29.5
Rest of Europe	62.4	23.6	14.0	25.4	48.6	26.1
Africa	60.1	17.9	22.0	23.1	39.9	37.0
North America	59.2	30.3	10.6	27.9	48.1	24.0
South America	64.3	21.6	14.1	27.5	44.5	27.9
Asia	55.8	22.2	22.0	24.4	41.3	34.3

Note: Overeducation among immigrants from Oceania and Soviet Union is not shown due to small sample sizes.

that large which is different from the incidence of overeducation, which we found to be much larger among non-Western immigrants.

To see how different variables jointly affect the probability of being overeducated as well as the probability of being undereducated compared to being correctly matched, we estimate a multinomial regression for 2001–2008. In Table 2 the dependent variable

Table 2 Probability of being over- and undereducated estimated with a multinomial logit model, relative risk ratios

Base outcome: correctly matched	Men		Women	
	Over-educated	Under-educated	Over-educated	Under-educated
Sweden	Ref.	Ref.	Ref.	Ref.
Other Nordic countries	2.273*** (0.021)	2.819*** (0.029)	2.342*** (0.019)	2.509*** (0.024)
EU15	2.409*** (0.024)	3.469*** (0.039)	2.288*** (0.025)	3.595*** (0.046)
Rest of Europe	3.837*** (0.026)	1.765*** (0.014)	3.785*** (0.024)	1.825*** (0.014)
Africa	4.597*** (0.050)	1.859*** (0.023)	3.528*** (0.043)	2.232*** (0.029)
North America	3.086*** (0.052)	3.537*** (0.071)	2.761*** (0.047)	3.359*** (0.070)
South America	3.459*** (0.039)	1.716*** (0.022)	3.355*** (0.036)	1.999*** (0.024)
Asia	3.454*** (0.025)	2.324*** (0.019)	3.493*** (0.025)	2.602*** (0.021)
Oceania	3.262*** (0.135)	4.862*** (0.231)	2.497*** (0.125)	4.000*** (0.230)
Soviet union	2.189*** (0.098)	3.964*** (0.229)	3.149*** (0.086)	2.409*** (0.092)
Years since migration/10 ^a	0.758*** (0.002)	0.702*** (0.002)	0.777*** (0.002)	0.759*** (0.002)
Age	0.641*** (0.001)	0.866*** (0.001)	0.785*** (0.001)	0.972*** (0.001)
Age squared/100	1.649*** (0.002)	1.197*** (0.001)	1.277*** (0.001)	1.062*** (0.001)
Number of children	0.982*** (0.001)	0.982*** (0.001)	1.000 (0.001)	0.930*** (0.001)
Married	0.974*** (0.002)	1.262*** (0.003)	0.895*** (0.002)	1.035*** (0.002)
Years of schooling	1.533*** (0.001)	0.480*** (0.000)	1.357*** (0.001)	0.536*** (0.000)
Firm size/1000	0.954*** (0.001)	1.048*** (0.001)	0.958*** (0.001)	1.069*** (0.001)
N	9,921,204	9,921,204	10,058,041	10,058,041

Note: ^aYear since migration is set to zero for all workers born in Sweden. The models also include controls for sector (five categories), industry (11 categories), county (21) and year dummies. Robust standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

is labor market match defined using the mode method.⁴ Regressions are estimated separately for men and women and relative risk ratios are presented. Most variables affect overeducation among men and women in a similar way. All groups of immigrants are more likely than natives to be overeducated, however, the relative risk is higher for immigrants from regions where Sweden traditionally has received refugee migrants and tied movers, for example Africa, South America and Asia. The results also indicate that immigrants are more likely than natives to be undereducated compared to being correctly matched. The risk of being mismatched, i.e. both under- and overeducated decreases with years since migration for immigrants. A more stable family situation, as indicated by being married and larger family responsibilities measured by the number of children, is associated with a lower probability of being overeducated for men. Overeducation decreases with age in a non-linear way. Years of schooling is strongly and positively correlated with the risk of being overeducated, which is not surprising since overeducation does not exist among those with lower education. On the other hand, years of schooling reduces the risk of being undereducated. Firm size, here measured as the number of employees in 1000's, decreases the risk of being overeducated while it increases the risk of being undereducated.

In Table 3 educational attainment and the ORU-variables are described in more detail. Native women have most schooling in terms of years (12.9), and are the group with the second largest share with a long university education (30.0 percent). Only Western immigrant women have a larger share (31.1 percent). Native women have jobs with the highest required education and the lowest number of years of overeducation. Non-Western immigrant men have the highest number of years of overeducation while Western immigrant men have the highest number of years of under-education, at least when using the mode method of measuring years of under- and overeducation.

5 The returns to educational mismatch

With the descriptive evidence as a background, in this section we analyze the returns to over- and undereducation for natives and immigrants separately. Overall, the pattern for men and women are very similar.

In the first three columns in Tables 4 and 5 the results for the returns to actual schooling is presented separately for natives, Western immigrants and non-Western immigrants. In line with previous research we find that the returns to actual schooling are lower for non-Western immigrants than for natives and lower for women than for men. The returns to schooling for Western immigrants are between that of natives and non-Western immigrants.

The estimates of the returns to actual schooling may be affected by the degree of mismatches in the labor market as well as by the returns to over- and undereducation. For the US, Chiswick and Miller (2008) find that the payoff to schooling is lower for immigrants than for natives and a decomposition reveals that this is primarily a result of lower payoff to schooling among low-educated workers (upper secondary education for two years and below). The returns to over, under and required education in Sweden have previously been analyzed by Böhlmark (2003) and Korpi and Tählin (2009), but they do not compare natives and immigrants. A recent paper by Katz and Österberg (2013) analyze over and undereducation among childhood immigrants in Sweden. Their

Table 3 Sample means for natives and immigrants in 2008

	Men			Women		
	Natives	Western immigrants	Non-Western immigrants	Natives	Western immigrants	Non-Western immigrants
Years of schooling (std.dev)	12.3 (2.3)	12.5 (3.0)	12.2 (2.8)	12.9 (2.2)	12.8 (2.7)	12.4 (2.9)
Level of education						
Primary school less than 9 years	0.8	4.0	7.1	0.3	2.6	9.1
Primary school 9 (10) years	10.9	12.0	9.4	6.1	8.2	7.2
Upper secondary 2 years or less	30.2	28.6	24.5	25.1	26.7	21.4
Upper secondary more than 2 years	22.0	13.6	21.6	21.4	14.5	20.0
Higher education less than 3 years	15.0	13.0	14.3	16.2	14.2	13.2
Higher education 3 years or more	19.8	24.3	20.8	30.0	31.1	27.5
Post graduate education	1.4	4.4	2.2	0.8	2.6	1.5
Years of (using the mean +/- 1 st.dev):						
Undereducation	0.14	0.23	0.24	0.10	0.17	0.26
Required education	12.46	12.62	12.02	12.79	12.83	12.16
Overeducation	0.13	0.25	0.38	0.13	0.21	0.35
Years of (using the mode):						
Undereducation	0.69	0.85	0.76	0.55	0.70	0.69
Required education	12.41	12.59	11.56	12.82	12.82	11.96
Overeducation	0.62	0.80	1.08	0.59	0.72	1.08
Number of observations	1,265,344	54,591	125,621	1,165,324	59,470	133,242

Table 4 Returns to actual, required, over and undereducation, pooled OLS regressions, men

	Actual schooling, men			Under-, required and overeducation, men		
	Natives	Western immigrants	Non-Western Immigrants	Natives	Western immigrants	Non-Western Immigrants
Schooling	0.063** (0.000)	0.059** (0.000)	0.043** (0.000)			
UE_mode				-0.033** (0.000)	-0.036** (0.000)	-0.030** (0.000)
RE_mode				0.078** (0.000)	0.080** (0.000)	0.075** (0.000)
OE_mode				0.060** (0.000)	0.047** (0.001)	0.017** (0.000)
R-squared	0.433	0.427	0.410	0.464	0.468	0.503
Number of observations	5,413,151	239,869	420,554	5,413,151	239,869	420,554

Note: The models also include controls for age, age squared, marital status, having young children, years since migration (for immigrants), sector (five categories), municipality, birth region for immigrants (six categories) and year dummies. Standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 5 Returns to actual, required, over and undereducation, pooled OLS regressions, women

	Actual schooling, women			Under-, required and overeducation, women		
	Natives	Western immigrants	Non-Western Immigrants	Natives	Western immigrants	Non-Western Immigrants
Schooling	0.055** (0.000)	0.054** (0.000)	0.038** (0.000)			
UE_mode				-0.031** (0.000)	-0.032** (0.000)	-0.027** (0.000)
RE_mode				0.064** (0.000)	0.070** (0.000)	0.066** (0.000)
OE_mode				0.047** (0.000)	0.040** (0.000)	0.013** (0.000)
R-squared	0.436	0.413	0.405	0.462	0.458	0.506
Number of observations	6,929,785	347,853	543,752	6,929,785	347,853	543,752

Note: The models also include controls for age, age squared, marital status, having young children, years since migration (for immigrants), sector (five categories), municipality, birth region for immigrants (six categories) and year dummies. Standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

paper is however mostly descriptive and do not estimate the ORU-model or analyze state dependence in overeducation.

The returns to required schooling can be interpreted as the return to schooling for correctly matched workers. When required schooling is entered linearly into the regression there is in practice no difference between natives and immigrants. This means that given that workers are correctly matched, the return to education is about the same. This result is consistent with findings presented in (Chiswick and Miller 2008).

The returns to undereducation are about the same for natives and both groups of immigrants; for each year of undereducation, wages drop by 3 to 4 percent. There is however a large difference in the returns to overeducation. They are substantially smaller for non-Western immigrants than for natives; for each year of overeducation wages increases by 6.0 percent for natives but only by 1.7 percent for immigrants. Similar results have been found for Denmark (Nielsen 2011) and Canada (Wald and Fang 2008). As for Western immigrants, the estimated returns to overeducation are again between that of native and non-Western immigrants; for each year of overeducation wages increases by 4.7 percent.

One of the regularities is that the estimate for required education is larger than the estimate for overeducation but both are positive. This means that overeducated workers earn more than correctly matched workers in the same kind of jobs but less than correctly matched workers with the same education. However, our estimations reveal that the difference in the estimates is much larger for immigrants than for natives. Overeducated immigrants earn only slightly more than correctly matched workers in the same kind of jobs but much less than correctly matched workers with the same education. Large wage differences between the jobs that correctly matched and mismatched immigrants with the same education have, and compressed wage distributions and no individual wage bargaining within occupations could be the explanation.

To control for unobserved heterogeneity we estimate panel data models with individual fixed effects (see Table 6 for men and Table 7 for women). Although this is an

Table 6 Returns to actual, required, over and undereducation, panel data models with individual fixed effects, men

	Actual schooling, men			Under-, required and overeducation, men		
	Natives	Western immigrants	Non-Western Immigrants	Natives	Western immigrants	Non-Western Immigrants
Schooling	0.025** (0.000)	0.019** (0.001)	0.020** (0.001)			
UE_mode				-0.024** (0.000)	-0.022** (0.001)	-0.022** (0.001)
RE_mode				0.028** (0.000)	0.023** (0.001)	0.024** (0.001)
OE_mode				0.025** (0.000)	0.015** (0.001)	0.015** (0.001)
R-squared	0.333	0.347	0.345	0.334	0.348	0.348
Number of observations	5,413,151	429,587	422,451	5,413,151	429,587	422,451

Note: The models also include controls for age, age squared, marital status, having young children, years since migration (for immigrants), sector (five categories), municipality, birth region for immigrants (six categories) and year dummies. Standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

improvement over the previous model it only addresses the problem with unobservables that are constant over time. Identification in this model will come from job changes rather than from changes in years of schooling. Leuven and Oosterbeek (2011) argues that many other changes might happen related to the job change and that these might be unobserved and affect wages.

Our results from the pooled cross sectional model and the fixed effects models are very similar: the returns to actual schooling are lower among immigrants, the returns to undereducation and required education are about the same while the returns to overeducation are substantially lower for immigrants.

Table 7 Returns to actual, required, over and undereducation, panel data models with individual fixed effects, women

	Actual schooling, women			Under-, required and overeducation, women		
	Natives	Western immigrants	Non-Western Immigrants	Natives	Western immigrants	Non-Western Immigrants
Schooling	0.017** (0.000)	0.017** (0.001)	0.014** (0.001)			
UE_mode				-0.018** (0.000)	-0.018** (0.001)	-0.017** (0.001)
RE_mode				0.020** (0.000)	0.019** (0.001)	0.018** (0.001)
OE_mode				0.015** (0.000)	0.016** (0.001)	0.007** (0.001)
R-squared	0.401	0.380	0.387	0.402	0.380	0.391
Number of observations	6,929,785	355,518	545,849	6,929,785	355,518	545,849

Note: The models also include controls for age, age squared, marital status, having young children, years since migration (for immigrants), sector (five categories), municipality, birth region for immigrants (six categories) and year dummies. Standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Bauer (2002) argues that if there is a negative correlation between ability and overeducation, i.e. workers with lower ability would be more likely of being overeducated; we would therefore expect that the estimates in the pooled OLS model underestimate the returns to overeducation. This is not consistent with our findings since the estimates of the returns to overeducation, both for natives and immigrants, is lower in the specification where we control for unobserved heterogeneity. It is however interesting to note that the estimate for immigrants is much less affected by the inclusion of individual fixed effects than the corresponding estimate for natives.

A conclusion to be drawn from these results is that immigrants lose more from being overeducated in terms of lower wages. This is in contrast to the results for the US where immigrants gain more from being undereducated compared to natives Chiswick and Miller (2008).

It is worth stressing that the return to undereducation is about the same for natives and both groups of immigrants. That is, immigrants do not lose from being undereducated any more than natives. This is the reason for our focus on the persistence and state dependence in overeducation in the following section.

To summarize, we find that non-Western immigrants face a higher risk of being overeducated once they enter the labor market and lose more from being incorrectly matched. The outcomes for Western immigrants are not as bad, but still worse than that of natives. But is this a short run or long run phenomenon? We turn to this question in the next section.

6 State dependence in overeducation

Mavromaras et al. (2012) make a distinction between simple persistence and state dependence where the former can be interpreted as the duration of time an individual stays overeducated while the latter refers to the fact that previous overeducation has a direct causal effect on future overeducation. That is, state dependence is present if we find an effect of lagged overeducation on future overeducation once we have controlled for background factors that caused overeducation in the first place.

In this section we first describe (simple) persistence in overeducation by presenting the share that is still overeducated in 2002–2008 among those who were overeducated in 2001 in a balanced panel over the years 2001–2008 that will be used in the estimation of state dependence. See Table 8. The modal method has been used to define the educational norm.⁵ For native men, 67 percent are still overeducated in 2008, i.e. seven years after observed overeducation. For all groups of male non-Western immigrants, the persistence is even stronger. Around 80 percent of those overeducated in 2001 are still overeducated in 2008. For Western immigrants, between 71 and 75 percent are still overeducated in 2008. For all groups of women, outflow from overeducation is higher compared to their male counterparts, but also for women we find that the persistence in overeducation is higher among non-Western immigrants.

Since job changes, or lack thereof, are likely to be important elements of the persistence in overeducation we take a closer look at the extent to which workers change jobs. Looking at job changes between consecutive years 13.9 percent of native women compared to 14.8 percent among immigrant women change workplace from one year to the next. For men the corresponding figures are 13.7 percent (natives) and 13.9 percent

Table 8 Persistence in overeducation

Region	2002	2003	2004	2005	2006	2007	2008
Men							
Sweden	92.1	83.1	78.9	77.0	73.5	72.7	67.0
Other Nordic countries	93.5	85.7	81.7	80.0	77.4	76.8	70.9
EU15	93.9	86.9	80.8	79.5	77.1	76.2	71.6
Rest of Europe	96.5	93.1	91.3	90.4	88.3	87.6	84.2
Africa	95.3	91.4	89.2	87.3	85.8	83.9	81.9
North America	94.7	89.7	85.0	83.4	80.7	79.8	75.8
South America	94.2	89.4	86.9	85.7	83.6	82.5	80.0
Asia	95.5	90.0	86.4	85.4	83.1	82.7	79.2
Women							
Sweden	90.3	71.7	69.1	64.9	61.1	60.3	50.5
Other Nordic countries	91.6	73.2	70.5	67.2	64.8	63.8	54.5
EU15	91.6	76.6	73.8	71.3	67.9	66.6	60.8
Rest of Europe	94.6	86.6	84.7	82.5	78.7	78.1	73.8
Africa	94.6	86.8	84.8	83.1	80.3	79.1	75.7
North America	90.3	79.1	77.3	74.4	72.7	72.9	67.5
South America	95.1	84.9	83.2	81.8	78.5	77.8	72.8
Asia	94.3	82.4	80.3	78.4	75.5	75.2	69.1

The percentage among those who were overeducated in 2001 that was also overeducated in subsequent years. (**Mode** method to decide the norm within occupations – balanced panel 2001–2008).

(non-Western immigrants). Job changes occur to a similar degree for Western immigrants. Estimating the probability of changing workplace in a regression framework and controlling for years of schooling, age, being married, having young children, sector, and municipality, we find that non-Western immigrants are less likely to change workplace between two consecutive years.⁶

The question of how important job changes are for explaining state dependence is analyzed below but as a descriptive exercise we have looked at persistence in overeducation among job switchers only. A bit surprisingly, we see that persistence in overeducation is also high among those switching jobs. Among native men that were overeducated in 2001 and had switched job in 2008, 63 percent were still overeducated. Among non-Western immigrants this share is around 75 percent. This shows that changing job is no guarantee for improved matching on the labor market.

Next we turn to the results from the dynamic random effects probit models. For comparison we present the results from a random effects probit model without correcting for the initial conditions problem and without the Mundlak correction in the upper panel of Table 9 for men and in the corresponding panel in Table 10 for women. The results from the random effects probit model with both corrections for the initial conditions problem by inclusion of overeducation in period $t = 0$, in our case year 2001, and the individual means over age, number of children, years of schooling, and years since migration for immigrants are presented in lower panels of Tables 9 and 10.

In all models and for both natives and immigrants we find indications of a very high degree of state dependence in overeducation. In the model with controls for initial overeducation the estimates are about 2.1 for natives, 2.2 for Western immigrants and around 2.4 for non-Western immigrants. Using a similar method and Australian survey

Table 9 State dependence in overeducation, men

	Random effects probit, 2002 – 2008						
	Natives	Western immigrants			Non-Western immigrants		
		III	I	II	III	I	II
Overeducation t-1	3.193*** (0.004)	3.381*** (0.012)	3.381*** (0.012)	3.377*** (0.021)	3.596*** (0.010)	3.595*** (0.010)	3.619*** (0.018)
Job change	0.091*** (0.004)	0.033** (0.016)	0.032** (0.016)	0.108*** (0.023)	-0.088*** (0.013)	-0.089*** (0.013)	0.002 (0.019)
Region of birth	-	-	Yes	Yes	-	Yes	Yes
Occupation	Yes	-	-	Yes	-	-	Yes
Number of observations	4,800,915		194,683		292,424		
Number of individuals	708,673		28,759		43,664		
Insig2u	-0.303*** (0.010)	-13.737*** (3.870)	-13.739*** (3.874)	-0.351*** (0.054)	-14.209*** (3.650)	-14.210*** (3.659)	-0.342*** (0.044)
Random effects probit, 2002–2008 – Wooldridge estimator with Mundlak correction							
	Natives	Western immigrants			Non-Western immigrants		
		III	I	II	III	I	II
Overeducation t-1	2.130*** (0.005)	2.604*** (0.026)	2.603*** (0.026)	2.227*** (0.029)	2.760*** (0.023)	2.761*** (0.023)	2.391*** (0.024)
Overeducation t = 0	2.703*** (0.013)	1.422*** (0.052)	1.422*** (0.052)	2.884*** (0.074)	1.577*** (0.052)	1.571*** (0.051)	2.699*** (0.059)
Job change	0.110*** (0.005)	0.046** (0.019)	0.045** (0.019)	0.119*** (0.028)	-0.086*** (0.015)	-0.087*** (0.015)	-0.010 (0.022)
Region of birth	-	-	Yes	Yes	-	Yes	Yes
Occupation	Yes	-	-	Yes	-	-	Yes
Means over time	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	4,800,915		194,942		293,021		
Number of individuals	708,673		28,762		43,670		
Insig2u	0.7092 (0.0016)	-0.722*** (0.069)	-0.721*** (0.069)	0.984*** (0.041)	0.3323 (0.0139)	-0.704*** (0.063)	0.6855 (0.0074)

Note: The random effects probit models also include controls for: age, age squared, number of children, marriage, years of schooling and years since migration for immigrants, sector (five categories), county (21 categories), birth region for immigrants (six categories), and occupation (nine categories).

The random effects probit models using the Wooldridge estimator and Mundlak corrections include means over time for age, age squared, number of children, years of schooling and years since migration. Standard errors in parentheses.

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

data over six waves, the highest estimate of lagged overeducation that is reported by Mavromaras and McGuinness (2012) is 1.13, i.e. substantially below the estimates we find. Although there are many factors to consider when making cross-country comparisons, it has been argued that overeducation might be more self-persistent in countries with greater labor market rigidity (Verhaest and Van der Velden, 2013).

Even if it is difficult to compare estimates across groups we can say something about the state dependence among natives as compared to immigrants. First, we note that the standard errors of the estimates are small which makes the 95 percent confidence interval rather narrow. As a result, the confidence intervals of the same estimate for the two groups are not overlapping making the estimates of the lagged dependent variable for natives and immigrants being significantly different from each other. It might thus

Table 10 State dependence in overeducation, women

	Random effects probit, 2002 – 2008						
	Natives	Western immigrants			Non-Western immigrants		
	III	I	II	III	I	II	III
Overeducation t-1	2.725*** (0.0033)	3.057*** (0.010)	3.057*** (0.010)	2.891*** (0.016)	3.298*** (0.009)	3.299*** (0.009)	3.101*** (0.014)
Job change	0.0774*** (0.0038)	0.054*** (0.013)	0.054*** (0.013)	0.106*** (0.018)	-0.051*** (0.012)	-0.052*** (0.012)	0.022 (0.016)
Region of birth	-	-	Yes	Yes	-	Yes	Yes
Occupation	Yes	-	-	Yes	-	-	Yes
Number of observations	4,655,054		223,213		300,290		
Number of individuals	693,242		33,682		46,057		
Insig2u	-0.4198 (0.0091)	-12.415*** (2.327)	-12.416*** (2.327)	-0.493*** (0.045)	-14.026*** (3.824)	-14.030*** (3.820)	-0.455*** (0.040)
Random effects probit, 2002–2008 – Wooldridge estimator with Mundlak correction							
	Natives	Western immigrants			Non-Western immigrants		
	III	I	II	III	I	II	III
Overeducation t-1	2.253*** (0.004)	2.555*** (0.019)	2.555*** (0.019)	2.285*** (0.020)	2.739*** (0.019)	2.739*** (0.019)	2.428*** (0.018)
Overeducation t = 0	1.148*** (0.007)	0.863*** (0.032)	0.863*** (0.032)	1.376*** (0.037)	1.003*** (0.034)	1.005*** (0.034)	1.383*** (0.034)
Job change	0.108*** (0.004)	0.073*** (0.015)	0.073*** (0.015)	0.138*** (0.019)	-0.037*** (0.013)	-0.038*** (0.013)	0.049*** (0.017)
Region of birth	-	-	Yes	Yes	-	Yes	Yes
Occupation	Yes	-	-	Yes	-	-	Yes
Means over time	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	4,655,054		223,392		300,712		
Number of individuals	693,242		33,685		46,061		
Insig2u	-0.155*** (0.008)	-1.211*** (0.067)	-1.212*** (0.067)	0.038 (0.036)	-1.151*** (0.065)	-1.156*** (0.065)	-0.007 (0.033)

Note: The random effects probit models also include controls for: age, age squared, number of children, marriage, years of schooling and years since migration for immigrants, sector (five categories), county (21 categories), birth region for immigrants (six categories), and occupation (nine categories).

The random effects probit models using the Wooldridge estimator and Mundlak corrections include means over time for age, age squared, number of children, years of schooling and years since migration. Standard errors in parentheses.

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

seem reasonable to conclude that state dependence is a more severe problem among non-Western immigrants than among natives but that it exists and is fairly large also for natives. The estimates for Western immigrants are between those for natives and non-Western immigrants.

State dependence is present if overeducation in an earlier period has an independent effect on current overeducation. By taking into account the initial conditions problem we hope to be able to distinguish this from the heterogeneity between workers that makes some more likely to be overeducated due to unobservable factors such as ability. Since we have found that state dependence is present and stronger among non-Western immigrants this would mean that immigrants are more negatively affected in the future by having been overeducated earlier in their careers. This is different from

saying that immigrants on average are more likely to have, to the researcher, unobservable qualities that increase the risk of being overeducated.

Looking at the estimate for having changed a job we find interesting results. First, we note that the inclusion of job changes does not significantly change the estimate of lagged overeducation suggesting that there is a high degree of state dependence even after having controlled for job changes. Second, the effect of changing job on overeducation differs between natives and immigrants. For natives and Western immigrants, changing jobs seems to be correlated with an increasing probability of being overeducated while for immigrants we find that changing jobs is correlated with a lower probability of being overeducated.

In specification II and III for immigrants we successively add region of birth and occupation. Note that occupation is already added in specification III for Natives. As discussed earlier, the quality of education in the home country of the immigrant does not need to be the same as the corresponding Swedish education. To investigate whether the higher persistence of overeducation of immigrants just reflects a different quality of education, lack of language proficiency or cultural factors, we add region of birth effects in specification II. The estimate on overeducation remains unchanged. Even if the education obtained is of the same quality, due to entry restrictions in certain occupations, it may take time for immigrants to get their credentials validated. As a partial test of this hypothesis, we control for both region of birth and occupation in the last specifications of the model in Tables 9 and 10. Doing so reduces the estimate on overeducation from about 2.6 to 2.2. for Western immigrants and from 2.8 to 2.4 for non-Western immigrants. Thus, imperfect transferability of human capital across countries has greater merit as explanation but even after controlling for region of birth and occupation, persistence in overeducation for immigrants significantly exceeds that for natives.

Comparing men and women, both simple persistence and state dependence is slightly lower among women than among men, which might be explained by the sample of women being more select than the male sample or that women have an outside option. But also for women we find a positive correlation between changing job and the probability of being overeducated for natives and Western immigrants while we find a negative correlation for non-Western immigrant women. For women, the persistence effect is reduced when controlling for occupation.⁷

7 Summary and conclusion

This paper studies educational mismatch among non-Western immigrants in Sweden. We first analyze differences in the incidence of overeducation and find that the overeducation is higher among all groups of immigrants than natives but the difference is larger for regions from which Sweden traditionally receives many refugees and tied movers. Concerning the incidence of undereducation, we find that it is highest among immigrants from the Nordic countries and lowest among immigrants from North America. The difference between natives and non-Western immigrants is smaller than the corresponding difference in overeducation. Turning next to the wage effects of educational mismatch, we find that on average the return to required schooling is about the same for natives and immigrants. However, the returns to overeducation are

substantially smaller for immigrants than for natives; immigrants lose much more from not being correctly matched than natives do. The return to undereducation is about the same for natives and both groups of immigrants.

However, there are some caveats to our analysis. The estimates of the returns to actual, required, under and overeducation presented in this paper should be interpreted with some care. There are several sources of selection that might influence our results. First, selection into over or undereducation is not random. It is difficult to determine in which direction the selection goes. Some have argued that there is negative correlation between ability and the probability of being overeducated since less able workers might require a higher education to perform a job for which the formal requirements actually are lower. But in the paper we argue that this might not be true for immigrant workers. Employers may see overeducation as a signal of ability and conscientiousness and prefer to hire immigrant workers with a higher education than what is actually needed to perform the job instead of hiring an immigrant worker with the required educational qualifications.

In the third part of the paper we analyze state dependence in overeducation, in particular if this appears to be a more severe problem among immigrants than among natives. We choose to focus on state dependence in overeducation since there are smaller differences between natives and immigrants in the incidence of undereducation and the return to undereducation is about the same for natives and immigrants. We find that there is a very high degree of state dependence in overeducation among natives as well as among immigrants, but it is stronger for immigrants than for natives. One hypothesis that may explain this result is that job mobility is lower for immigrants compared to natives in Sweden, and hence will result in a greater degree of state dependence in a spell of overeducation. However, job mobility in our sample is about the same for natives and immigrants and even after controlling for job changes in the regression, there is a higher degree of state dependence among immigrants. A novel and interesting result is that changing job seems to be associated with a higher probability of being overeducated for natives while immigrants that change jobs are less likely of being overeducated compared to those who do not change jobs. Another hypothesis we test is whether quality of education or imperfect transferability of human capital obtained in the home country could explain the greater persistence of overeducation among immigrants. Our results indicate that the latter explanation does indeed hold some merit, but only in part.

We will conclude by discussing some policy implications. If the explanation that immigrants are overeducated is because they have insufficient knowledge of the Swedish language or a lower quality of education then increased and improved education in Swedish could be one method to improve the situation. However, controlling for region of birth did not reduce persistence in overeducation. Our results rather point to the fact that immigrants' qualifications are not viewed the same as Swedish ones, and thus validation of credentials but also additional complementing education could be part of the solution. Such policies exist already but could be further developed. Compared to those born in Sweden immigrants on average have less information about the Swedish labor market and their networks may be less suited for finding a job corresponding to their qualifications. More intense assistance from the labor market administration may be part of a solution as well. All policy changes have of course to be evaluated.

Endnotes

¹For an overview of the literature on educational mismatch of immigrants, see Piracha and Florin (2012).

²Employees at Statistics Sweden with good knowledge of the 1990 census indicate, however, that many may have reported an education level higher than their acquired one.

³Carlsson and Rooth (2007) show that the degree of employer discrimination in the hiring process varies with the characteristics of the job and the workplace.

⁴We have also estimated the probability of being overeducated using the mean method to define the educational norm within occupation. All marginal effects are slightly smaller using the mean but the sign of the estimates are not affected by the way overeducation is defined.

⁵Using the mean method when defining the educational norm leads to a lower degree of persistence for all groups. The mean method implies that the norm is an interval for years of schooling and not an absolute number of years of schooling.

⁶The results are not presented in the paper but are available from the authors upon request.

⁷We have also estimated models for state dependence in undereducation (results available on request). Our focus in this paper is, however, to look more at overeducation, since immigrants in Sweden are hurt more by overeducation relative to natives, whereas they face the same penalties for undereducation. We find patterns of persistence also in undereducation.

Competing interest

The IZA Journal of Migration is committed to the IZA Guiding Principles of Research Integrity. The authors declare that they have observed these principles.

Acknowledgements

We want to thank Barry Chiswick, an anonymous referee and other seminar participants at the SULCIS Scientific Advisory Board Meeting, May 2011, the 2011 conference on Immigration and Labor Market Integration arranged by SULCIS, November 1–2, 2011, the conference on European Integration in Swedish Economic Research, May 22–25, 2012 and a seminar at Ratio September 25, 2013. Andersson Joona gratefully acknowledges funding from Jan Wallander's and Tom Hedelius stiftelse and the Swedish Council for Working Life and Social Research (FAS). Financial support from SULCIS is also gratefully acknowledged.

Responsible editor: Amelie F Constant.

Author details

¹Swedish Institute for Social Research, Stockholm University, SE-106 91 Stockholm, Sweden. ²Department of Economics and Business, Aarhus University, 8210 Aarhus V, Denmark.

Received: 3 December 2013 Accepted: 6 May 2014

Published: 28 May 2014

References

- Arai M, Bursell M, Nekby L (2010) The Inverted Gender Gap: Measuring the Intensity of Employer Priors against Men and Women with Arabic Names, unpublished paper. Stockholm University, Department of Economics,
- Arulampalam W, Stewart MB (2009) Simplified Implementation of the Heckman Estimator of the Dynamic Probit Model and a Comparison with Alternative Estimators. *Oxf Bull Econ Stat* 71(5):659–681
- Åslund O, Rooth D-O (2007) Do when and where matter? Initial labour market conditions and immigrant earnings. *Econ J* 117(518):422–448
- Battu H, Peter SJ (2002) To what extent are ethnic minorities in Britain overeducated? *Int J Manpow* 23(3):192–208
- Bauer TK (2002) Educational mismatch and wages: a panel analysis. *Econ Educ Rev* 21(3):221–229
- Böhlmark A (2003) Over- and Undereducation in the Swedish Labour Market: Incidence, Wage effects and Characteristics 1968–2000, unpublished paper, Swedish Institute for Social Research. Stockholm University,
- Bursell M (2007) What's in a name? A field experiment test for the existence of ethnic discrimination in the hiring process, SULCIS Working Paper, 2007:7. Stockholm University,
- Carlsson M, Rooth D-O (2007) Evidence of Ethnic Discrimination in the Swedish Labor Market using Experimental Data. *Labour Econ* 14(4):716–729
- Chevalier A (2003) Measuring Overeducation. *Economica* 70(279):509–531

- Chiswick BR, Miller PW (2008) Why is the payoff to schooling smaller for immigrants? *Labour Econ* 15(6):1317–1340
- Chiswick BR, Miller PW (2010a) An explanation for the Lower Payoff to Schooling for Immigrants. In: Ted MD et al (eds) *The Canadian Labor Market, in Canadian Immigration - Economic Evidence for a Dynamic Policy Environment*. McGill-Queen's University Press, Montreal and Kingston, pp 41–75
- Chiswick BR, Miller PW (2010b) Does the Choice of Reference Levels of Education Matter in the ORU Earnings Equation? *Econ Educ Rev* 29(6):1076–1085
- Chiswick BR, Miller PW (2010c) Educational mismatch: Are high-skilled immigrants really working at high-skilled jobs, and what price do they pay if they are not? In: Chiswick BR (ed) *High Skilled Immigration in a Global Labor Market*. American Enterprise Institute, Washington DC
- Chiswick BR, Miller PW (2010d) The Effects of Educational-Occupational Mismatch on Immigrant Earnings in Australia, with International Comparisons. *Int Migr Rev* 44(4):869–898
- Cuesta MB, Budría S (2012) Overeducation Dynamics and Personality. *Educ Econ* 20(3):260–283
- Dahlstedt I (2011) Occupational Match. Over- and Undereducation among Immigrants in the Swedish Labor Market. *J Int Migr Integr* 12(3):349–367
- Dolton P, Vignoles A (2000) The incidence and effects of overeducation in the U.K. graduate labor market. *Econ Educ Rev* 19(2):179–198
- Duncan GJ, Hoffman SD (1981) The incidence and wage effects of overeducation. *Econ Educ Rev* 1(1):75–86
- Eriksson S (2011) Utrikes födda på den svenska arbetsmarknaden, Bilaga 4 till Långtidsutredningen. p LU2011
- Fernández C, Ortega C (2008) Labor Market Assimilation of Immigrants in Spain: Employment at the Expense of Bad Job-Matches? *Span Econ Rev* 10(2):83–107
- Frenette M (2004) The overqualified Canadian graduate: the role of the academic program in the incidence, persistence, and economic returns to overqualification. *Econ Educ Rev* 23(1):29–45
- Green C, Kler P, Leeves G (2007) Immigrant overeducation: Evidence from recent arrivals in Australia. *Econ Educ Rev* 26(4):420–432
- Hartog J (2000) Overeducation and earnings: where are we, where should we go? *Econ Educ Rev* 19(2):131–147
- Heckman JJ (1981) Heterogeneity and state dependence. In: Sherwin R (ed) *Studies in Labor Markets*. Chicago Press, Chicago, pp 91–140
- Katz K, Österberg T (2013) Wages of childhood immigrants in Sweden – education, returns to education and overeducation, IFAU Working Paper 2013. p 8
- Korpi T, Tählin M (2009) Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000. *Labour Econ* 16(2):183–193
- Leuven E, Oosterbeek H (2011) Overeducation and Mismatch in the Labor market. In: Erik H, Finis W (eds) *Handbook of the Economics of Education*, vol 4. Elsevier Science, Amsterdam, pp 283–326
- Lianos TP (2007) Brain Drain and Brain Loss: Immigrants to Greece. *J Ethn Migr Stud* 33(1):129–140
- Mavromaras K, McGuinness S (2012) Overskilling Dynamics and Education Pathways. *Econ Educ Rev* 31(5):619–628
- Mavromaras K, Mahuteau S, Sloane P, Wei Z (2012) The persistence of overskilling and its effect on wages, National Vocational Education and Training Research and Evaluation Program, Research Report
- Mundlak Y (1978) On the Pooling of Time Series and Cross Section Data. *Econometrica* 46(1):69–85
- Nielsen CP (2011) Immigrant overeducation: evidence from Denmark. *J Popul Econ* 24(2):499–520
- Orme CD (2001) Two-step Inference in Dynamic Non-Linear Panel Data Models, unpublished paper, University of Manchester,
- Piracha M, Florin V (2012) Migrant Educational Mismatch and the Labor Market. In: Constant AF, Zimmermann KF (eds) *The International Handbook on the Economics of Migration*, vol 9. Edward Elgar Publishing, Cheltenham, UK, pp 176–192
- 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
- Rubb S (2003) Overeducation: A Short or Long Run Phenomenon for Individuals? *Econ Educ Rev* 22(6):389–394
- Sicherman N, Galor O (1990) A Theory of Career Mobility. *J Polit Econ* 98(1):169–192
- Statistics Sweden (2012). <http://www.scb.se>, 2013-09-05
- Stewart Mark B (2006) redprob- A Stata program for the Heckman estimator of the random effects dynamic probit model, *mimeo*, University of Warwick. <http://www2.warwick.ac.uk/fac/soc/economics/staff/academic/stewart/stata>
- Verdugo RR, Verdugo NB (1989) The impact of surplus schooling on earnings: some additional findings. *J Hum Resour* 24(4):629–643
- Verhaest D, van der Velden R (2013) Cross-country Differences in Graduate Overeducation. *Eur Sociol Rev* 29(3):642–653
- Wald S, Fang T (2008) Overeducated immigrants in the Canadian labour market: Evidence from the workplace and employee survey. *Canadian Public Policy* 34(4):457–479
- Wooldridge JM (2005) Simple solutions to the initial conditions problem in dynamic, non-linear panel data models with unobserved heterogeneity. *J Appl Econ* 20(3):39–54
- Zenou Y, Åslund O, Östh J (2010) How important is access to jobs? Old question – improved answer. *J Econ Geogr* 10(3):389–422

doi:10.1186/2193-9039-3-9

Cite this article as: Joona et al.: Overeducation among immigrants in Sweden: incidence, wage effects and state dependence. *IZA Journal of Migration* 2014 **3**:9.