Graduate migration, self-selection and urban wage premiums across the regional hierarchy*

Kent Eliasson[a,b] and Olle Westerlund[a]

[a] Umeå School of Business, Economics and Statistics, Umeå University, Sweden
[b] Swedish Agency for Growth Policy Analysis, Östersund, Sweden

Abstract

We use Swedish longitudinal population register data on university graduates and estimate the effect of migration on earnings. Migration between regional labour markets is used to identify static and dynamic agglomeration effects on earnings. Heterogeneity in effects is examined by individuals' position in the ability distribution and by origin-destination size categories of regional labour markets. The results indicate that the effect of upward migration (from smaller to larger labour markets) on earnings is positive throughout. Downward migration (from larger to smaller labour markets) is generally associated with negative or no convincing signs of positive effects on earnings. The estimates indicate positive short-term urban wage premiums (UWP) for all origin-destination flows of upward migration, especially high UWP for in-migration to the Stockholm labour market region. The UWP of upward migration is positive also for movers in the lower end of the ability distribution, but it is substantially higher for high ability migrants. We also find evidence of a positive dynamic UWP of migration to Stockholm from the other regions, particularly for high ability migrants.

Keywords: Urban wage premium, human capital, migration, agglomeration economies, ability

JEL classification: R12, R10, R23, J61, J24

* Email-addresses: kent.eliasson@tillvaxtanalys.se, olle.westerlund@umu.se. The authors would like to thank participants at the labour economics seminar at the Department of Economics, Umeå University, for valuable comments. The usual disclaimer applies. Corresponding author: kent.eliasson@tillvaxtanalys.se

1. Introduction

Empirical research generally confirms theories of positive causal effects of geographical concentration of economic activities (agglomeration) on productivity, wages, and growth (e.g. Combes and Gobillon 2015). Recent trends of continued spatial concentration into larger and more population dense labour markets have spurred increased interest among social scientists and policy makers (OECD 2018, Iammarino et al. 2019). Economic and social polarization between large urban and more sparsely populated regions is also a major issue in the public debate. Increased knowledge on current magnitudes, trends and underlying mechanisms for concentration is vital for design of policy.

The theoretical mechanisms driving the economic advantages of concentration (agglomeration economies) can be summarized into three categories: matching, learning and sharing (Duranton and Puga 2004). Empirical research confirms a variety of drivers of agglomeration economies within these categories. Job matching efficiency is higher in larger and thicker labour markets (e.g. Andini et al. 2013), learning and innovation are enhanced in agglomerations (e.g. Glaeser 1999, Gordon and McCann 2005), and sharing a common base of suppliers (including labour market pooling) is an advantage for firms (e.g. Rosenthal and Strange 2001).

Concentration of human capital is assumed to play a major role in generating static and dynamic effects on productivity (Moretti 2011), i.e. an urban wage premium (UWP). Findings differ between studies regarding magnitudes and interpretation of estimated effects. A major issue in research on agglomeration economies is to disentangle the effect of agglomeration from self-selection effects. Relatively higher wages in urban labour markets may not only derive from increased economic efficiency of spatial concentration, but may partially reflect non-random sorting of labour with higher market productivity into metropolitan areas and larger cities (Combes et al. 2008).

The observed concentration of the highly educated in urban labour markets may partially be driven by relatively higher returns to education in agglomerations. The UWP can emerge as a region-specific constant appearing at the time of labour force entry, which is a static effect that may result from more efficient job matching, for example (Wasmer and Zenou 2002, Wheeler 2001, Abel and Deitz 2015). The UWP may also increase over time as a dynamic effect due to positive learning and external effects, e.g. as a result of working and residing in environments with a high concentration of knowledge capital and a rich supply of education.

The purpose of this paper is to estimate the effects of migration on gross labour earnings among university graduates. Migration between regional labour markets is used to identify both static and dynamic agglomeration effects on earnings. In particular, we examine the heterogeneity in effects based

on the size of the origin and destination regional labour markets, by grouping such labour markets in categories by size. Migration from smaller to larger labour markets (upward migration) is distinguished from migration in the opposite direction, i.e., from larger to smaller labour markets (downward migration).

The analysis is confined to university graduates because of the important role of human capital for growth and the relatively high mobility of the highly educated, and because theory and evidence indicate an especially high urban premium for this group (e.g. Venables 2011, Korpi and Clark 2019).

We use Swedish longitudinal population register data to estimate both short-run and medium-run effects on earnings. Sweden provides a good setting for this type of study. It is a market economy characterized by continued strong tendencies of urbanisation and a growing urban-rural divide, and the main characteristics of recent trends in internal migration flows are similar as compared with most other developed economies. In addition, the Swedish population registers provide highly informative data. This is especially important because of the importance of controlling for non-random selection by individuals' abilities and skills into different types of labour markets. We go beyond previous studies on UWP in various respects that are further discussed in the literature review. In particular, we examine heterogeneity in effects of migration for downward as well as upward migration between all size category combinations of regional labour markets and for university graduates in different positions within the ability distribution.

Our results are by and large consistent with the hypothesis that a larger regional labour market is associated with a positive UWP. The estimated earnings effect of upward migration across the regional hierarchy is positive throughout. The estimated effect increases with the individual's high school grades, but migrants with low grades also receive a positive UWP from upward migration. The hypothesis of increased productivity by the size of regional labour markets is also confirmed by the estimated effects of downward migration. Migration from larger to smaller labour markets is generally associated with negative or insignificant effects on earnings. The hypothesis of a positive static UWP is generally supported by our data. Our estimates also indicate a positive dynamic UWP of upward migration and a negative dynamic UWP of downward migration. However, this is confined to the migration exchange between the largest labour market (Stockholm) and other regions. We find strong evidence of self-selectivity on ability in migration decisions.

The remainder of this study follows a common set-up of empirical papers in economics, i.e. sections in the following order: previous studies, data, method, results, and a summary with discussion.

2 Previous studies

Labour productivity is higher in agglomerations according to findings in a vast number of studies (Duranton and Puga 2004). Estimates of the urban wage premium without adjustment for workers' characteristics and residential self-selection indicate double digit differences in wages between large and smaller regional labour markets. For example, Combes et al. (2008) find around 35 percent higher wages in Paris relative to medium-sized cities; similar findings are reported in Glaeser and Maré (2001). In terms of wage elasticities with respect to population-related measures of regional market size, estimates of the UWP without adequate control for self-selectivity hover around 5–6 percent (Melo et al. 2009).

With adjustment for residential and other types of self-selectivity, studies indicate an UWP ranging from about 2 to 6 percent (e.g. Di Addario and Pattachini 2008, Mion and Naticchioni 2009, Lehmer and Möller 2010, Anderson et al. 2014, De la Roca and Puga 2017). Combes et al. (2008) find that 40–50 percent of aggregate regional wage differentials in France are accounted for by regional sorting of labour on observed and unobserved skills. Strong effects of selectivity and small or insignificant UWP:s are indicated in Gould (2007).

De la Roca and Puga (2017) use Spanish data and report estimates between 2 and 4 percent; they conclude that the higher estimate may not depend entirely on spatial sorting of abilities but may at least partially represent a dynamic UWP.

Carlsen et al. (2016) estimate UWP for the seven largest functional labour markets in Norway and find a combined static and dynamic UWP of 17 percent in Oslo, with an initial effect of 7 percent and a dynamic effect of 10 percent after eight years of experience. The combined UWP is found to be positive across education levels but considerably lower for primary-educated workers than for the college educated. Wage elasticities with respect to regional population density range between 1.6 and 3 percent according to their OLS estimates.

Using Swedish data, Andersson et al. (2014) find that the effects of agglomeration on earnings are small in general but larger for workers in occupations characterised by non-routine tasks. They conclude that spatial sorting of labour is the main explanation of higher earnings in dense labour markets.

Migration between regional labour markets can be used to identify residential self-selectivity and for bias reduction in estimates of the UWP. Relocations between smaller and larger labour markets can also be used to distinguish between initial and dynamic UWP (e.g. Glaeser and Maré 2001). Migration over longer distances between functional labour markets are mostly undertaken by young people and the

propensity to migrate increases with educational attainment (e.g., Greenwood 1997, Machin et al. 2012). Another reason to examine migration in a regional productivity context is the systematic spatial sorting of skills through migration that is evident in most developed countries (Winters 2011 on US data; Faggian and McCann 2009 on data from Great Britain; van Venhorst et al. 2010 on data from the Netherlands; Haapanen and Tervo 2012, and Haapanen 2013 on Finnish data; Berck et al. 2016, and Tano et al. 2018 on Swedish data).

Similar to our study, Ahlin et al. 2014 use Swedish longitudinal population register data on university graduates. They report an initial UWP of around 5.7 percent and a dynamic UWP of 2.2 to 3.6 percent effect on wage growth for graduates remaining in the large urban areas; the higher estimate pertains to in-migrants from the countryside.

Also using population register data from Sweden and similar estimation strategies as Ahlin et al. (2014), Korpi and Clark (2019) study differences in UWP by education and use migration of young individuals (age 22-29) to identify initial and dynamic UWP. They find an initial income gain of 4.5-8.0 percent for migrants moving into urban or city regions. They also find a positive dynamic UWP of around 2 percent yearly earnings increase, but this result is mostly confined to in-migrants being employed in the three largest urban labour markets.

Our contribution to previous research lies mainly with four aspects. First, we study heterogeneity in the effects of migration on earnings for upward as well as for downward migration across different origin-destination size categories of functional regional labour markets and by university graduates position within the ability distribution. Second, we do not condition the results on post-migration events, such as e.g. sector of employment, type of job, or size of establishment, because we consider these as potential drivers of the matching and learning effects of agglomeration and thus likely mediators of the urban wage premium. Third, we do not condition samples on labour force participation (e.g. we do not exclude observations with zero earnings), and we do not use sample restrictions causing any truncation of earnings – the variable measuring outcome of interest. The probability of employment (job finding rate) and better chances of optimizing the number of working hours may in fact be major economic drivers of labour migration. Fourth, we use propensity score matching instead of the commonly used fixed-effect estimator for reasons outlined in Section 4.

3. Data

The analysis is based on detailed longitudinal population register data on the Swedish population from Statistics Sweden. The dataset includes yearly information on all individuals 32 years or younger who graduated from at least three years of university education during the period 2001–2010. For those who are no older than 32 when they graduate, we have information on important covariates such as grade point average (GPA) from upper secondary school and parental background in terms of parents' education and earnings.[1] We will focus on graduates in the following fields of education: social sciences; business and law; science; engineering, manufacturing and construction. We have excluded graduates in education and health and welfare because a very large share of graduates in these fields work in the public sector, where wage formation is to a lesser extent determined by market outcomes. There are very few graduates in the remaining fields of education that are not included.

The regional dimension in the analysis is based on a grouping of 69 functional local labour market areas (LMA) into four types of regions. The LMA:s are defined on the basis of commuting patterns between Sweden's 290 municipalities in 2015 and constitute economically integrated regions where most people tend to both live and work. The LMA:s are grouped into the following four types of regions: very large regions (Stockholm LMA, with a population of about 2.6 million), large regions (Göteborg and Malmö LMA:s, with a population of 1.3 and 1.1 million, respectively), medium-sized regions (19 LMA:s, with a population between 100,000 and 300,000), and small regions (47 LMA:s, with a population under 100,000). The LMA:s in the medium-sized category typically include the regional administrative centres and contain the universities/university colleges located outside Stockholm, Göteborg and Malmö LMA:s. With a few exceptions, the LMA:s in the group small regions include neither regional administrative centres nor university colleges.

Based on these four types of regions, we define stayers and migrants by comparing each individual's region of residence at age 17 (approximately one year prior to the earliest possible entry into university education) with the region of residence two years after graduation.[2] Consequently, a stayer is a person who resides in the same type of region at both points in time, whereas a migrant is a person who moves from one type of region to another type during this time span (e.g. moves from a LMA in the category small regions to a LMA in the category medium-sized regions).[3] To make the definition of stayers and

[1] During the period 2001–2010, approximately 80 percent of all university education degrees were awarded to students 32 years of age or younger.
[2] A very large share of migration among university graduates take place during this time span. Expanding the window for the definition of migration to e.g. five years after graduation increases the number of migrants by less than 2 percent.
[3] Note that moves between LMA:s belonging to the same type of region do not count as migration (e.g. a move from a LMA in the category small regions to another LMA in the same category is not considered as migration).

movers as distinct as possible, and thereby facilitate the interpretation of our estimated effects, we do not allow any return or repeat migration across the regional levels during the two year time span after graduation.[4] However, note that we do not condition on possible future migration taking place more than two years after graduation. We treat stayers' possible subsequent migration, and migrants' possible subsequent return or repeat migration, as part of the causal effect we seek to identify.

The dependent variable in the analysis is annual gross labour earnings. The earnings measure includes no income transfers such as e.g. unemployment benefits. Data on earnings are not top coded and our sample is not conditioned on level of earnings.[5] That is, labour force participation and labour supply are considered as post-migration outcomes. Therefore, observations of zero or low levels of earnings are included in our estimations. We do not adjust for variation in regional prices. Instead, we agree with the argument in De la Roca and Puga (2017) that the focus should be on nominal earnings when trying to estimate the productive advantages of regions or cities. Nominal earnings reflect how much more firms are willing to pay similar or identical workers in larger regions compared to smaller regions due to the productive gains of larger regions. Regional differences in cost of living may very well affect workers' choice of location, but that does not change the fact that firms can only pay higher nominal earnings in larger regions if larger regions provide productive advantages. Without regional variation in productivity, firms in high wage locations wouldn't be able to compete with firms in low wage locations. Because we use data on several cohorts, earnings of the graduates and parents are adjusted with the national CPI to 2016 prices.

Our econometric approach is based on different propensity score matching methods (presented in Section 4) and a highly informative dataset. The covariates included in the propensity score models measure/indicate: gender, age, country of birth, field of education in high school, high school GPA, parental background at age seventeen in terms of parents' education, earnings, and country of birth, local labour market area and municipality type at age seventeen, family situation the year before graduation, number of years of university education, field of university education, a quality ranking of the attended university, and graduate cohort.[6] Some of the above covariates are standard in the UWP literature. But information such as high school grades, parental background, and university rank is unobserved in many empirical studies.

[4] It turns out that the incidence of return and repeat migration is rather rare, so the imposed restriction is not costly in terms of lost observations.

[5] See e.g. Chay and Powell (2001) for a discussion of potential problems with censored data.

[6] Country of birth is a dummy for born in Sweden or abroad. Field of education in high school cover three broad practical and theoretical programs. Municipality type refers to a classification of each municipality in a given LMA as either core, adjacent or fringe. Family situation is indicated by being married and having children. Field of university education is measured at ISCED 2-digit level. For quality of attended university, we partition all universities/university colleges into five quintiles based on enrollment selectivity in terms of high school GPA.

We estimate the size of the UWP by comparing the earning of stayers with the earnings of comparable migrants for each origin-destination combination. This is done for six upward migration flows (stayers in small regions compared to migrants from small regions to medium-sized regions, and so on) and for six downward migrations flows (stayers in very large regions compared to migrants from very large regions to large regions, and so on). Table 1 present the number of university graduates for each origin-destination combination. The dataset includes about 100,000 graduates. The diagonal in the destination matrix gives the number of stayers in each type of region. Of the graduates who start in LMA:s in the category small regions, 24 percent are classified as stayers (3,410/14,294). The share of stayers increases with the size of regions and reach 88 percent for the Stockholm region (very large). The table clearly reveals that upward migration flows dominate downward flows. Medium-sized regions and in particular small regions experience large net out-migration of graduates, whereas large regions and especially Stockholm experience net in-migration of graduates. From the table, we can conclude that there is a considerable redistribution of university graduates over time from smaller regions towards larger regions. This finding is in line with previous studies on migration among university graduates (see e.g. Faggian et al. 2007, Faggian and McCann 2009, Venhorst et al. 2011, Haapanen and Tervo 2012).

Table 1 University graduates distributed by region of residence at age 17 and two years after graduation

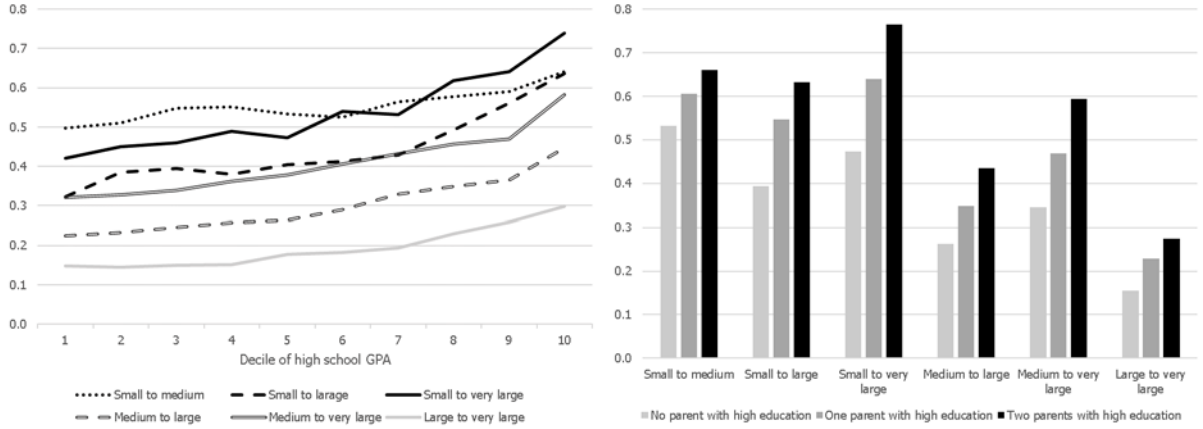
Region of residence at age 17		Region of residence two years after graduation			
		Small	Medium-sized	Large	Very large
Small	14,294	3,410	4,235	2,700	3,949
Medium-sized	38,034	1,679	17,201	7,377	11,777
Large	22,495	495	2,067	16,073	3,860
Very large	25,497	346	1,728	1,030	22,393

Figure 1 gives a flavour of the residential sorting going on based on two of our ability-related variables. The figure display the share of upward migrants distributed on deciles of high school GPA and parents' level of education for each origin-destination combination.[7] There is an almost monotonic increase in the share of university graduates moving upwards in the regional hierarchy from the first to the tenth decile of the grade distribution. This pattern is especially evident at the top of the grade distribution and tends to be more pronounced the larger the difference in the size of the origin-destination regions. For the origin-destination combination small to very large, 74 percent of the graduates in the top decile are migrants compared to slightly above 40 percent in the lower end of the grade distribution. The figure also reveals a distinct positive sorting on parents' level of education. The share of migrants increases

[7] For each origin-destination combination, the share of migrants from region *i* to region *j* (m_{ij}) is calculated as: m_{ij}=migrants from region *i* to region *j*/(stayers in region *i* + migrants from region *i* to region *j*). With this specification, Figure 1 shows how migrants from region *i* to region *j* differs on our ability-related attributes compared to stayers in region *i* (e.g. how migrants from LMA:s in the category small regions to LMA:s in the category large regions differs compared to stayers in LMA:s in the category small regions).

with the parents' level of education for each origin-destination combination. These findings are in line with Ahlin et al. (2018) who also find significant positive sorting into urban regions on high school grades and parents' level of education. Overall, these descriptive statistics indicate systematic residential sorting on ability-related variables that has to be accounted for when trying to identify the UWP by comparing the earnings of stayers and migrants.

Figure 1 Share of upward migrants distributed on deciles of high school GPA and parents' level of education



4. Econometric approach

Our identification strategy can be described using the Rubin causal model (Rubin 1974) and rural-to-urban migration as a generic case. Let Y_1 represent the potential earnings of moving from a rural to an urban region and Y_0 the potential earnings of staying in the rural region. Furthermore, let T represent a binary treatment which in our case is moving or staying ($T = 1$ when moving from a rural to an urban region, and $T = 0$ when staying in the rural region). The parameter of main interest is the average treatment effect on the treated: $ATT = E(Y_1 - Y_0|T = 1) = E(Y_1|T = 1) - E(Y_0|T = 1)$. In our context, ATT is the average effect on earnings of moving from a rural to an urban region rather than staying in the rural region, for those individuals who actually move to an urban region. The fundamental evaluation problem is that we only observe Y_1 or Y_0 for each individual, but never both. Given that the conditional independence assumption holds (Dawid 1979), we can use the earnings of matched stayers as an approximation of the unobserved counterfactual, i.e. what movers would have earned had they chosen to stay: $E(Y_0|X, T = 1) = E(Y_0|X, T = 0)$. Identification of ATT also relies on a common support or overlap condition, which for ATT can be expressed as $\Pr(T = 1|X < 1)$. This condition prevents X from being a perfect predictor of treatment status and thus ensures that for all values of X there are observations of both stayers and migrants.

We use several different propensity score matching methods. The propensity score is the probability of receiving treatment conditional on the covariates X: $\Pr(T = 1|X)$. Rosenbaum and Rubin (1983) show that, if the treatment and control groups have the same distribution of propensity scores, they also have the same distribution of all covariates in X, no matter what the dimension of X. We can therefore implement the conditional independence assumption by matching on the propensity score instead of matching on all covariates in X.

The suggested identification strategy requires that X include all confounding factors that affect both the treatment and the outcome. For the strategy to be plausible, this clearly requires having access to highly informative data. We believe that the data set available for this study goes a long way in meeting this requirement. Among many other things, our data include a large potential comparison group, detailed information on ability-related variables such as high school grades and parents' education and earnings. Having access to direct measurement of ability-related variables is especially important given that we focus on young college graduates, with little or no pre-treatment labour market experience, which makes the information content of pre-treatment earnings less useful for corrections of selection bias.[8] We therefore consider the commonly used individual fixed-effects estimator, where estimates of the urban wage premium primarily are identified by comparing migrants' pre- and post-migration earnings, as a less suitable identification strategy. Another major advantage of using matching, instead of commonly applied parametric regression-based methods to estimate the UWP is that matching allows for explicit and transparent examination of the degree of overlap in the distribution of covariates in the treatment group and the control group. When using regression-based methods, it is much more difficult to assess the degree of overlap in the distribution of covariates for the two groups; this drawback, in combination with functional form assumptions, makes conventional regression techniques sensitive to potential extrapolation bias. A number of authors, including Heckman et al. (1998), Rubin (2001), Ho et al. (2007), and Imbens (2015), emphasize that incorrect functional form assumptions can generate substantial bias in the estimates, especially when the distribution of covariates in the treatment group and the control group are far apart.

Finally, note that, in order to avoid post-treatment bias in the estimation of the UWP, it is important that X do not include any variables that may have been affected by the treatment (Rosenbaum 1984, Ho et al. 2007). In our case, we do not match on covariates measured after migration. This means that outcomes such as, e.g. getting a highly skilled job after moving from a rural to an urban region are considered as part of the causal effect of the treatment.

[8] Carlsen et al. (2016) find that regional sorting is driven by the young and college educated. Differences in unobserved abilities is more important in early stages of workers' careers, according to the authors' interpretations of results.

Given the rich data available for this study and the applied propensity score matching methods, we think that the bias in estimates due to residential sorting is substantially reduced. As is the case for all studies based on non-experimental data, influence of unobserved heterogeneity may still be a remaining source of bias.

5. Results

In this section, we begin by reporting descriptive statistics for the sample of stayers and migrants used in the analysis. This is followed by a presentation of the probit propensity score estimates of selection into migration. These estimates provide information on the potential selectivity of migration in upward and downward migration flows. We then provide some balancing diagnostics that indicate whether our identification strategy has been successful in generating comparable stayers and migrants. The final sub-section contains *ATT* estimates of the urban wage premium.

5.1 Descriptive statistics

Table A1 in the Appendix reports covariate means for the different origin-destination combinations. The first four columns in Panel A compare stayers in small regions with migrants from small regions to larger regions. Family ties in terms of being married or having children are less common among migrants. Migrants also tend to be positively selected in terms of high school GPA and parents' education and earnings. The differences in the means between stayers and migrants for these covariates increases with the size of the destination region. The table also shows that a larger share of migrants has graduated from universities at the upper end of the university quality distribution. Less than 40 percent of the stayers in small regions have graduated from universities in the top two quintiles of the quality distribution. The corresponding figures for migrants from small regions to large or very large regions are 67 and 73 percent, respectively. The remaining columns in Panel A compare stayers in medium-sized and large regions with migrants moving upward from these regions. In these cases as well, family ties are less common among migrants and the migrants are positively selected in terms of high school GPA, parental background and university quality. Panel B in Table A1 presents comparisons between stayers and downward migrants. The general picture is that the differences in covariate means are smaller. The differences that do exist indicate that migrants tend to be negatively selected on several ability-related attributes. If we compare downward migrants from very large and large regions with stayers in these regions, the migrants have lower high school GPA, are less likely to have graduated from universities at the upper end of the university quality distribution, and tend to have parents with lower earnings.

5.2 Propensity score estimates

Table 2 reports the probit estimates of the propensity score models for upward and downward migration in the regional hierarchy. The presentation is restricted to some of our ability-related attributes. Tables A2 and A3 in the Appendix report estimates for a larger set of covariates included in the propensity score models. The positive selectivity for upward migration (Panel A) on high school GPA, parents' education, and university quality stand out clearly and are consistent for all origin-destination pairs. Parents' earnings may be partially correlated with their education, but the results still indicate a positive selection on parents' earnings for migration to the very large labour market (Stockholm). Besides a potential nature-nurture association with individual ability, this may reflect the need for financial back-up from parents to buy housing in the Stockholm region.

Table 2 Probit propensity score estimates for upward and downward migration

Panel A: Upward migration	Small to medium	Small to large	Small to very large	Medium to large	Medium to very large	Large to very large
High school GPA	0.030***	0.025***	0.034***	0.024***	0.025***	0.049***
	(0.008)	(0.009)	(0.008)	(0.005)	(0.004)	(0.005)
One parent with high education	0.165***	0.253***	0.218***	0.126***	0.147***	0.142***
	(0.043)	(0.050)	(0.045)	(0.024)	(0.021)	(0.027)
Two parents with high education	0.274***	0.296***	0.402***	0.250***	0.281***	0.198***
	(0.069)	(0.076)	(0.066)	(0.033)	(0.028)	(0.032)
Parents' earnings	-0.193**	0.127	0.167**	0.067	0.209***	0.165***
	(0.077)	(0.092)	(0.076)	(0.043)	(0.037)	(0.036)
4-year university	0.249***	0.255***	0.269***	0.046**	0.108***	0.186***
	(0.037)	(0.044)	(0.040)	(0.023)	(0.020)	(0.029)
University quality Q5	-0.153**	1.030***	1.179***	1.063***	1.274***	0.422***
	(0.067)	(0.069)	(0.066)	(0.044)	(0.039)	(0.058)
Panel B: Downward migration	Very large to large	Very large to medium	Very large to small	Large to medium	Large to small	Medium to small
High school GPA	-0.016**	-0.001	-0.024**	0.007	-0.014	-0.008
	(0.007)	(0.007)	(0.011)	(0.006)	(0.011)	(0.007)
One parent with high education	0.084**	-0.069**	0.038	0.069**	0.010	0.060*
	(0.037)	(0.034)	(0.056)	(0.032)	(0.052)	(0.035)
Two parents with high education	0.148***	0.061	0.074	0.068	0.034	0.120**
	(0.042)	(0.038)	(0.067)	(0.042)	(0.071)	(0.052)
Parents' earnings	-0.011	-0.095**	-0.168**	-0.128**	-0.051	-0.083
	(0.037)	(0.048)	(0.069)	(0.053)	(0.107)	(0.066)
4-year university	0.116***	0.201***	0.148**	0.207***	0.205***	0.102***
	(0.042)	(0.036)	(0.063)	(0.035)	(0.058)	(0.033)
University quality Q5	-0.194**	-0.949***	-0.744***	-0.387***	-0.615***	0.005
	(0.087)	(0.064)	(0.096)	(0.058)	(0.076)	(0.063)

Notes: See Section 3 for definition of covariates included in the propensity score models. Tables A2 and A3 in the Appendix report estimates for a larger set of covariates. Robust standard errors in parentheses. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively.

The probit estimates of the propensity score for downward migration (Panel B) indicate in most cases insignificant or negative selectivity on variables measuring individual ability, parental background and university quality. Seemingly at odds with expectations, the coefficients on longer university education are positive and significant. However, this is in line with most research on internal migration – higher educational attainment is generally associated with higher propensity of interregional migration.

Otherwise, there is virtually no evidence of a positive selection on abilities in the migration flows from larger to smaller regional labour markets.[9]

Overall, the estimates indicate systematic spatial sorting of abilities and concentration of human capital into larger labour markets. Given the strong interest in analysis of agglomeration of human capital and location choices of university graduates, the estimates show the importance of controlling for self-selectivity in migration decisions within the group of university graduates.

5.3 Balance diagnostics

Before turning to the estimates of the urban wage premium, it is important to assess whether the chosen identification strategy has been effective in generating comparable stayers and migrants. Figure A1 in the Appendix presents propensity score density plots for stayers and migrants before and after matching. Before matching, the density for the stayers lies well to the left of that of upward migrants in most cases. This indicates that many of the stayers have very low predicted probabilities of upward migration. The difference in the distribution of propensity scores for stayers and upward migrants also tends to increase with the difference in the size of the origin-destination regions. Concerning downward migration, the differences in the distributions for stayers and downward migrants before matching are much smaller compared to the case of upward migration and the predicted probabilities of migration also tend to be quite small. After matching, the distribution of propensity scores is identical for stayers and migrants in all cases of both upward and downward migration.

Figure 2 provides a graphical presentation of covariate balance in terms of standardized differences for selected covariates in the unmatched and matched samples.[10] The standardized difference of a covariate is defined as the difference of the sample means in the treatment and control group, scaled by the square root of the average of the sample variance in the two groups (Rosenbaum and Rubin 1985). In the applied literature, a standardized difference within the range +/– 0.1 is often considered as negligible (see, e.g. Austin 2009). This interval is indicated by the dotted vertical lines in the figure. Before matching, there is considerable imbalance between stayers and upward migrants on important ability-related covariates such as high school GPA, parents' education, and university quality, but also on covariates such as being married or having children (Panel A). There is also some covariate imbalance between stayers and downward migrants before matching, but the differences are generally less pronounced (Panel B). After matching, the figure demonstrates that the matched samples of stayers and migrants are very similar in

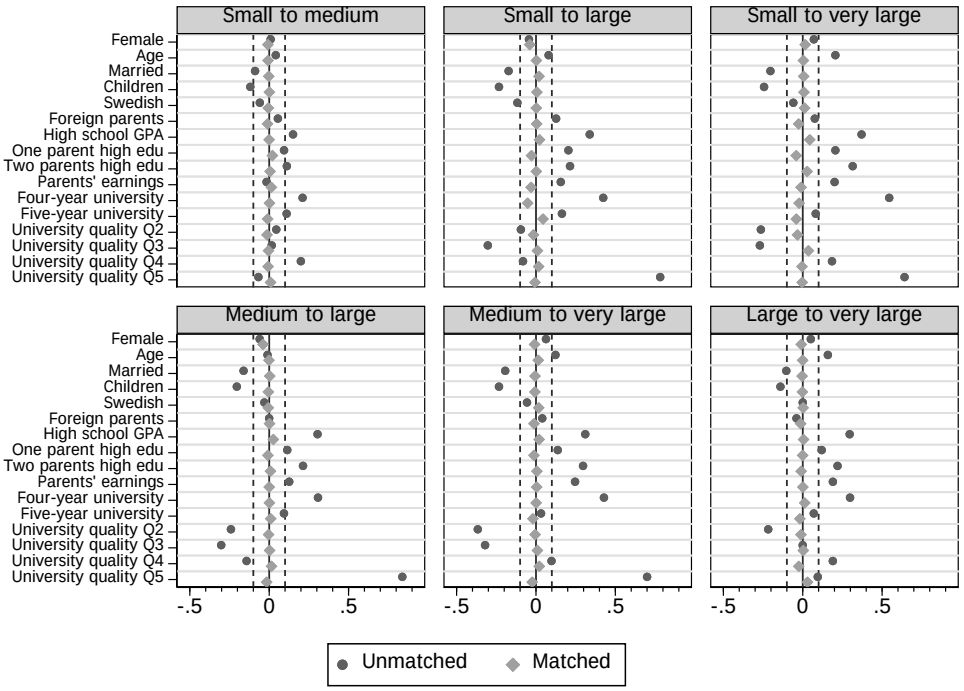
[9] Note that exact symmetry in upward/downward estimation results cannot be expected because samples of stayers and migrants vary in characteristics by each origin-destination category. Also, some individual traits are almost always associated with relatively higher probability of interregional migration, e.g. highly educated.

[10] To save space, the figure excludes standardized differences for field of education in high school, field of university education, local labour market area and municipality type at age seventeen, and graduate cohort.

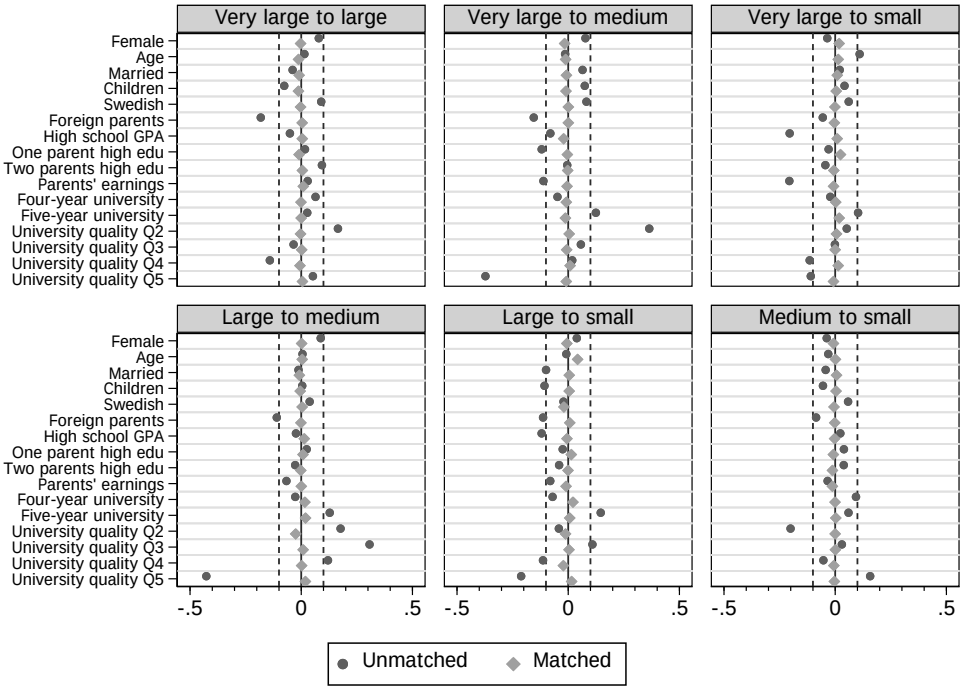
all cases. The standardized differences across all covariates used in the propensity score models, including the ones not shown in Figure 2, are well within the acceptable interval +/– 0.1.

Figure 2 Covariate balance in terms of standardized differences between stayers and migrants

Panel A: Upward migration



Panel B: Downward migration



Notes: The dotted vertical lines in the figure indicate standardized differences in the interval +/– 0.1. See Section 5.4 for details of the applied matching algorithm.

5.4 Estimated urban wage premiums

We begin by reporting the estimates of migrants' initial earnings premiums measured in the first year of observation in the new location. To check robustness, we present results using several different propensity score matching methods. We then proceed with reporting the estimates of the dynamic earnings premium over the following five years after migration. We close by examining heterogeneity in effects by rank of high school GPA.

Table 3 presents the estimated effects of migration on annual gross labour earnings the first year after migration (initial earnings premium). All estimates are based on propensity score matching using an Epanechnikov kernel with bandwidths based on cross validation with respect to the means of the covariates in the propensity scores.[11] Panel A in Table 3 shows estimates for migration upwards in the regional hierarchy, and panel B reports estimates for downward migration from larger to smaller regions. Column (1) contains matching estimates of migration on earnings with graduate cohort and local labour market area and municipality type at age seventeen as the only covariates. The overall pattern of estimates is that migration from smaller to larger labour markets is associated with an economically significant initial earnings premium. The premium ranges from SEK 13,732 up to 70,040 (corresponding to between 1,500 and 7,500 Euro) and increases the larger the difference in size between the out-migration region and the in-migration region. Migration in the opposite direction, from larger to smaller labour markets, is associated with completely different earnings outcomes as compared with upward migration estimates. The estimated initial earnings premiums of downward migration from the Stockholm region (very large) to large, medium-sized and small regions are negative and substantial in magnitude, ranging from SEK -18,243 to -46,511. There is also a negative premium of migrating from large to medium-sized regions. A relatively small but positive earnings premium is indicated for migrants starting in medium-sized labour markets and migrating to small regions.

Column (2) in Table 3 contains matching estimates of migration on earnings including the full set of confounders. Adjusting for spatial sorting based on demographic attributes, type of university education and the various ability-related variables clearly affect the estimated earnings premiums. The premiums of upward migration from small and medium-sized region to large and very large regions drop by between 11 and 39 percent. The estimated negative effects on earnings of downward migration from the Stockholm region drop by between 21 and 36 percent.[12] The estimated premium for upward migration using the full set of confounders ranges from SEK 10,654 up to 62,641. The estimated effects are largest for migration into the Stockholm region (very large). Compared to the earnings of comparable stayers,

[11] The reported results are robust with regard to alternative approaches for bandwidth selection, such as e.g. cross validation with respect to the dependent variable.

[12] Previous studies using fixed effect specifications to control for spatial sorting of migrants also find substantial reductions in the estimated elasticity of the earnings premium with regard to city size/density (e.g. Combes et al. 2008, Mion and Naticchioni 2009, and De la Roca and Puga 2017).

migrants into the Stockholm region receive an earnings premium of about 17–21 percent. The estimates of downward migration from the Stockholm region ranges from SEK -14,488 to -32,049, which corresponds to an earnings loss of about 4–9 percent compared to the earnings of comparable stayers in the Stockholm region. The estimates including the full set of confounders also indicate a small negative effect of downward migration from large to medium-sized regions and a relatively small but positive premium of migrating from medium-sized to small regions.

Table 3 Estimation of initial earnings premium for upward and downward migration (SEK 1,000)

	(1) Raw epan pscore matching	(2) Full epan pscore matching	(3) Full exact and epan pscore matching	(4) Full bias-adjusted epan pscore matching
Panel A: Upward migration				
Small to medium	13.732***	10.654***	9.945***	10.367***
	(2.535)	(2.579)	(2.805)	(2.591)
Small to large	38.206***	23.946***	22.971***	21.267***
	(3.250)	(4.089)	(4.215)	(4.373)
Small to very large	70.040***	62.641***	63.701***	63.362***
	(3.223)	(4.090)	(3.860)	(3.616)
Medium to large	36.333***	22.109***	19.309***	19.196***
	(1.533)	(1.797)	(1.809)	(1.739)
Medium to very large	62.054***	54.336***	54.283***	53.842***
	(1.598)	(1.927)	(1.650)	(1.687)
Large to very large	62.470***	61.043***	60.063***	60.181***
	(3.233)	(3.409)	(3.009)	(3.109)
Panel B: Downward migration				
Very large to large	-18.243***	-14.488***	-14.785***	-14.739***
	(3.741)	(3.875)	(3.789)	(3.884)
Very large to medium	-43.695***	-32.049***	-31.456***	-32.426***
	(2.555)	(2.439)	(2.584)	(2.743)
Very large to small	-46.511***	-29.554***	-33.472***	-28.783***
	(5.865)	(6.563)	(6.078)	(6.022)
Large to medium	-14.531***	-3.966*	-2.771	-4.481**
	(2.410)	(2.317)	(2.693)	(2.094)
Large to small	-2.168	3.603	3.037	2.773
	(4.378)	(5.005)	(4.715)	(5.003)
Medium to small	12.775***	8.193***	8.158***	8.117***
	(2.760)	(2.505)	(2.541)	(2.347)

Notes: Bootstrap standard errors based on 200 replications in parentheses. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively.

To check the robustness of the estimated effects, the last two columns in Table 3 report estimates of migration on earnings based on the full set of confounders using propensity score matching in combination with other methods. In column (3), we combine exact matching on sex and field of university education with propensity score matching on the other covariates. The idea with this approach is to eliminate potential bias in the estimates due to remaining imbalance between migrants and stayers on covariates that, based on prior evidence, might be particularly important for the outcome of interest (see e.g. Rubin and Thomas 2000 and Ho et al. 2007). This approach is equivalent to blocking in a randomized experiment. In column (4), we follow Abadie and Imbens (2011) and combine propensity score matching with linear regression adjustment on all covariates. The aim of this approach is to reduce

potential bias in the estimates due to residual imbalance between migrants and stayers on all covariates. Note that regression adjustment on matched samples is far less sensitive to functional form assumptions than model-based adjustment on unmatched samples, where extrapolation bias can be a major concern (see e.g. Rosenbaum and Rubin 1983 and Ho et al. 2007). If we compare the estimates in columns (2) to (4), there are only minor differences in the results. The reported estimates of migration on earnings thus seem to be quite robust with regard to these alternative estimation methods. In the remaining analyses, we will stick with bias-adjusted propensity score matching.

Table 4 presents the estimated effects of migration on annual gross labour earnings the first year after migration (initial earnings premium) and up to four additional years after migration (dynamic earnings premium). Upward migration within the regional hierarchy is generally associated with a positive earnings premium (Panel A). Migrants from small, medium-sized, and large regions who locate in the Stockholm region (very large) receive a substantial earnings premium that increases over time. In absolute terms, the increase is between SEK 20,000 and 36,000. Compared to the earnings of comparable stayers, this corresponds to an additional increase in the earnings premium of about 3–8 percentage points from the initial to the fourth year after migration. For upward migration between the other origin-destination categories, the estimates indicate positive earnings premiums for all years, but there is no consistent pattern of a positive dynamic UWP.

The estimates in panel B generally indicate a negative or no premium following migration from larger to smaller labour markets. Modestly positive effects of migration are indicated for moves between medium and small labour markets. Migration from the Stockholm region (very large) to the other labour markets is associated with a negative earnings premium in all cases. With some exceptions, migration from the Stockholm region is also associated with a slightly growing negative earnings premium over time. That is, moving from the very large labour market to smaller ones is associated not only with an initial loss in earnings, but also an increased earnings loss the following 2-3 years after migration. There is no consistent pattern of negative dynamic effects of downward migration between the other regions.

Table 4 Estimation of initial and dynamic earnings premium for upward and downward migration (SEK 1,000)

Panel A: Upward migration	*t*	*t*+1	*t*+2	*t*+3	*t*+4
Small to medium	10.367***	11.953***	13.871***	12.224***	9.159**
	(2.591)	(3.118)	(3.062)	(3.680)	(3.929)
Small to large	21.267***	20.216***	23.914***	18.630***	9.120
	(4.373)	(4.639)	(5.561)	(5.637)	(6.250)
Small to very large	63.362***	74.233***	81.885***	90.833***	83.405***
	(3.616)	(3.743)	(5.003)	(6.073)	(6.724)
Medium to large	19.196***	20.110***	19.151***	18.611***	16.148***
	(1.739)	(2.016)	(2.351)	(2.502)	(3.086)
Medium to very large	53.842***	66.049***	71.864***	73.924***	79.918***
	(1.687)	(2.510)	(2.381)	(2.734)	(3.168)
Large to very large	60.181***	72.013***	81.855***	89.492***	96.059***
	(3.109)	(3.107)	(4.315)	(4.436)	(5.303)
Panel B: Downward migration					
Very large to large	-14.739***	-16.726***	-29.425***	-29.840***	-26.991***
	(3.884)	(4.810)	(5.490)	(6.568)	(6.591)
Very large to medium	-32.426***	-37.468***	-44.405***	-47.812***	-47.446***
	(2.743)	(3.390)	(3.644)	(3.942)	(4.632)
Very large to small	-28.783***	-30.597***	-33.474***	-47.472***	-47.974***
	(6.022)	(7.413)	(8.506)	(8.552)	(9.384)
Large to medium	-4.481**	-6.174**	-4.582	-9.042***	-3.067
	(2.094)	(2.637)	(2.972)	(3.300)	(3.577)
Large to small	2.773	2.045	-2.532	1.934	3.051
	(5.003)	(5.023)	(6.240)	(6.604)	(8.097)
Medium to small	8.117***	6.613**	9.045***	8.461***	11.271***
	(2.347)	(2.703)	(2.917)	(3.210)	(3.476)

Notes: Bootstrap standard errors based on 200 replications in parentheses. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively.

To examine heterogeneity in dynamic effects of upward migration by individuals' cognitive ability measured by high school grades, we estimate the earnings premium separately for the subsamples in the first and fifth quintile of the GPA distribution.[13] The estimates for the sample in the fifth quintile (Panel A in Table 5) show a similar pattern as for the full sample results in Table 4. As anticipated, the estimated earnings premiums are generally higher than indicated by the full sample estimates. For example, the estimated effect of migration from small regions to the Stockholm labour market (very large) amounts to around SEK 118,900 at year *t*+4 while the corresponding full sample result in Table 4 is around SEK 83,400. Again, the estimates suggest that upward migration to the Stockholm region yields the highest initial earnings premium, as well as a clearly positive dynamic UWP over the first five years after migration. For the other origin-destination migration flows, there is no consistent pattern of increasing earnings premium over time.

The estimates also suggest a positive earnings premium of upward migration for individuals with GPA in the first quintile of the GPA distribution (Panel B). The estimates signal a lower earnings premium than indicated by the full sample results and by the fifth quintile results. However, the estimated earnings premium for migration from small regions to the Stockholm labour market is substantial and increases

[13] The number of graduates moving from larger to smaller regions is too few to make a corresponding analysis for downward migration.

over time, although the time trajectory is flatter in this case. In-migration to the Stockholm region is generally associated with a fairly high earnings premium with some but not full consistent indications of a positive dynamic UWP.

Table 5 Estimation of initial and dynamic earnings premium for upward migration among high and low ability samples (SEK 1,000)

Panel A: High ability sample (High school GPA in 5th quintile)	*t*	*t*+1	*t*+2	*t*+3	*t*+4
Small to medium	8.187	23.307***	15.129*	24.915**	10.887
	(7.114)	(7.592)	(8.224)	(10.541)	(10.117)
Small to large	15.141*	18.821	25.566*	18.186	2.917
	(8.561)	(11.959)	(13.183)	(14.148)	(14.141)
Small to very large	74.311***	91.039***	103.683***	111.976***	118.917***
	(9.557)	(11.876)	(13.226)	(18.096)	(20.113)
Medium to large	20.176***	19.411***	19.222***	20.370***	17.933***
	(3.349)	(4.072)	(4.888)	(5.710)	(6.084)
Medium to very large	70.403***	87.846***	102.869***	107.648***	115.160***
	(3.905)	(5.204)	(6.087)	(6.817)	(6.964)
Large to very large	74.881***	87.262***	96.599***	111.800***	126.426***
	(6.933)	(6.794)	(9.532)	(8.849)	(11.695)
Panel B: Low ability sample (High school GPA in 1st quintile)					
Small to medium	18.343***	13.005	11.452	11.758	12.829
	(6.705)	(8.168)	(8.139)	(8.907)	(9.304)
Small to large	20.149**	12.605	5.928	6.409	16.271
	(10.001)	(10.200)	(10.557)	(12.109)	(12.952)
Small to very large	38.754***	40.342***	45.176***	48.570***	49.428***
	(7.931)	(9.804)	(10.099)	(12.655)	(12.783)
Medium to large	25.157***	21.293***	22.367***	16.577***	11.659*
	(4.072)	(4.564)	(5.137)	(6.388)	(6.909)
Medium to very large	51.278***	57.459***	59.606***	56.263***	56.024***
	(4.293)	(5.323)	(5.465)	(6.322)	(6.245)
Large to very large	46.888***	58.921***	76.669***	69.649***	78.617***
	(6.891)	(8.431)	(9.860)	(9.981)	(11.775)

Notes: Bootstrap standard errors based on 200 replications in parentheses. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively.

6 Summary and discussion

In this paper, we estimate the effect of migration on annual gross labour earnings among university graduates and examine heterogeneity in static and dynamic effects on earnings by origin-destination flows across the regional hierarchy. Of primary interest is the pattern of self-selectivity into migration and earnings outcomes in terms of an urban wage premium (UWP). This study contributes to previous research in this field by a detailed study of the UWP associated with upward migration as well as downward migration within the regional size hierarchy. Moreover, we do not condition samples on labour market outcomes, and we consider post-migration events (e.g. characteristics of jobs and workplaces at the new location) as part of the UWP. We also use an estimation method (propensity score matching) suitable for identification of the parameter of main interest, i.e. the treatment effect on the treated. The preference for the chosen method is also based on availability to highly informative data

and because we consider the most commonly used method (individual fixed effects) less suitable for estimation with data on young workers in the early stages of their career.

Although university graduates constitute a selective group by themselves, we find strong evidence of self-selection on ability in migration decisions within this group. Generally, the results indicate consistent positive self-selection on high school grades, parents' education, and quality ranking of universities into migration upward in the regional hierarchy (i.e. from smaller to larger regional labour markets). The findings of selectivity into migration downward within the regional size hierarchy exhibit a distinctly different pattern. In all, the magnitudes of migration flows, and the patterns of self-selectivity into upward and downward migration, point at a stronger geographical concentration and regional divergence of human capital than indicated just by regional distributions of workers' educational attainment.

The estimated treatment effect of upward migration on earnings is positive throughout, while downward migration is generally associated with negative or no convincing indication of positive effects on earnings. Estimates indicate positive static (short-term) UWP for all origin-destination flows of upward migration. The estimates for in-migration to the largest region (Stockholm) stands out with estimated static UWP ranging from 17 to 21 percent (compared to the earnings of comparable stayers). Upward migration into the medium-sized and large regions (outside Stockholm) is associated with more modest static UWP, around 3–7 percent. Downward migration flows from the largest regional labour market (Stockholm) to large, medium-sized and small labour markets are associated with quite large negative short-term effects on earnings. With minor exceptions, we find no strong indications of economically significant static UWP for the other origin-destination categories of downward migration.

We also find evidence of a positive dynamic UWP for in-migrants to the labour market region of Stockholm from all other size categories of labour markets. The estimates ranges between about 3–8 percentage points additional increase in the estimated earnings premium from the initial to the fourth year after migration. There are some indications of positive dynamic UWP for some other origin-destination size categories of upward migration, but the pattern is not consistent.

Out-migration from Stockholm to the smaller labour markets is associated with decreases in earnings, the initial UWP ranging from about -4 to -9 percent (compared to the earnings of comparable stayers in Stockholm). There is also some evidence of a dynamic (negative) UWP for downward migration from Stockholm.

Testing for heterogeneity in effects by the graduates' position in the ability distribution, we find substantially larger UWP of upward migration for individuals at the top of the high school grade (GPA)

distribution. However, migrants in all quintiles of the GPA distribution earn a positive UWP from upward migration. In-migration to the Stockholm region among individuals at the top of the GPA distribution yields the highest initial earnings premium, as well as a consistent and substantial increase in the premium over time.

Our UWP estimates are generally higher than in previous studies using migration as an identification strategy, although our estimates for in-migration to the Stockholm region are in the same ballpark as those reported by Carlsen et al (2016) for the Oslo region. One possible explanation is that we do not condition our sample on full-time employment, and we do not condition estimates on future events (after migration), such as changing jobs, being employed in certain types of jobs, occupations or types of firms. We consider all labour market outcomes after migration as potential components of UWP. Moreover, we estimate the treatment effects on the treated, which can differ from the average effects. Our empirical design also differs from most previous research in other respects. In this study, stayers (non-migrants) serve as comparisons, previous studies on Swedish data use migrants to countryside labour markets as reference category. Also, the regional set up of this study allows for heterogeneity in the effects of migration on UWP by the full set of specific origin-destination size-categories of regional labour markets.

Similarly to most empirical research in this field, our estimates do not identify specific mechanisms underlying the UWP. The positive initial UWP indicated for all origin-destination categories of upward migration is consistent with the hypothesis of more efficient job matching (e.g. Wasmer and Zenou 2002) in agglomerations in terms of higher job finding rates. The estimates of dynamic UWP of upward migration is consistent with the hypothesis of higher job match quality in larger labour markets, although the evidence on dynamic UWP is somewhat mixed for upward migration to regions other than Stockholm. Also, the heterogeneity in estimated effects of upward migration on UWP by migrants' position in the ability distribution is consistent with Venables (2011) – all individuals earn a positive UWP by working with highly skilled workers, but the highly skilled will benefit more. Upward migration is generally a move to labour markets with a higher density of highly skilled workers. Of course, there are alternative interpretations of our results that would be in line with implications from theory.

References

Abadie, A. and Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects, *Journal of Business & Economic Statistics*, vol. 29:1, 1–11.

Abel, J. R. and Deitz, R. (2015). Agglomeration and job matching among college graduates. *Regional Science and Urban Economics*, vol. 51:4, 14–2.

Ahlin L., Andersson M. and Thulin, P. (2014). Market thickness and the early labour market career of university graduates: an urban advantage?, *Spatial Economic Analysis*, vol. 9:4, 396–419.

Ahlin, L., Andersson, M. and Thulin, P. (2018). Human capital sorting: the "when" and "who" of the sorting of educated workers to urban regions, *Journal of Regional Science*, vol. 58:3, 581–610.

Andersson, M., Klaesson, J. and Larsson, J. P. (2014). The sources of the urban wage premium by worker skills: spatial sorting or agglomeration economies? *Papers in Regional Science*, vol. 93:4, 727–747.

Andini, M., de Blasio, G., Duranton, G. and Strange, W. (2013). Marshallian labour market pooling: evidence from Italy, *Regional Science and Urban Economics*, vol. 43:6, 1008–1022.

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples, *Statistics in Medicine*, vol. 28:25, 3083–3107.

Berck, P., Tano, S. and Westerlund, O. (2016). Regional sorting of human capital: the choice of location among young adults in Sweden, *Regional Studies*, vol. 50:5, 757–770.

Carlsen, F., Rattsø, J. and Stokke, H.J. (2016). Education, experience, and urban wage premium, *Regional Science and Urban Economics*, vol. 60, issue C, 39–49.

Chay, K. Y. and Powell, J. L. (2001). Semiparametric censored regression models, *Journal of Economic Perspectives*, vol. 15:4, 29–42.

Combes, P-P., Duranton, G. and Gobillon, L. (2008). Spatial wage disparities: sorting matters! *Journal of Urban Economics*, vol. 63:2, 723–742.

Combes, P-P. and Gobillon, L. (2015). The empirics of agglomeration economies, *Handbook of Regional and Urban Economics*, vol. 5, Amsterdam, North-Holland, 247–348.

Dawid, A. P. (1979). Conditional independence in statistical theory, *Journal of the Royal Statistical Society*, Series B, vol. 41:1, 1–31.

De la Roca, J. and Puga, D. (2017). Learning by working in big cities, *Review of Economic Studies*, vol. 84:1, 106–142.

Di Addario, S. and Patacchini, E. (2008) Wages and the city. evidence from Italy, *Labour Economics*, vol. 15:5, 1040–1061.

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies, in Henderson, V. and Thisse, J-F. (eds.), *Handbook of Regional and Urban Economics*, vol. 4, Amsterdam, North-Holland, 2063–2117.

Faggian, A. and McCann, P. (2009). Universities, agglomerations and graduate human capital mobility, *Tijdschrift voor economische en sociale geografie*, vol. 100:2, 210–223.

Faggian, A., McCann, P. and Sheppard, S. (2007). Human capital, higher education and graduate migration: an analysis of Scottish and Welsh students, *Urban Studies*, vol. 44:13, 2511–2528.

Glaeser, E. L. (1999). Learning in cities, *Journal of Urban Economics*, vol. 46:2, 254–277.

Glaeser, E. L. and Mare, D. (2001). Cities and skills, *Journal of Labor Economics*, vol. 19:2, 316–342.

Gordon, I. R. and McCann, P. (2005). Innovation, agglomeration, and regional development, *Journal of Economic Geography*, vol. 5:5, 523–543.

Gould, E. D. (2007). Cities, workers, and wages: a structural analysis of the urban wage premium, *Review of Economic Studies*, vol. 74:2, 477–506.

Greenwood, M. (1997). Internal migration in developed countries, in M R. Rosenweig and O. Stark (eds.), *Handbook of Population and Family Economics Vol 1B.* Amsterdam, Elsevier, 647–711.

Haapanen. M. (2013). The effect of a polytechnic reform on migration, *Journal of Population Economics*, vol. 26:2, 593–617.

Haapanen, M. and Tervo, H. (2012). Migration of the highly educated: evidence from residence spells of university graduates, *Journal of Regional Science*, vol. 52:4, 587–605.

Heckman, J., Ichimura, H., Smith, J. and Todd, P. (1988). Characterizing selection bias using experimental data, *Econometrica*, vol. 66:5, 1017–1098.

Ho, D. E., Imai, K., King, G. and Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference, *Political Analysis*, vol. 15:3, 199–236.

Iammarino, S., Rodriguez-Pose, A. and Storper, M. (2019). Regional inequality in Europe: evidence, theory and policy implications, *Journal of Economic Geography*, vol. 19:2, 273–298.

Imbens, G. W. (2015). Matching methods in practice: three examples, *Journal of Human Resources*, vol. 50:2, 373–419.

Korpi, M. and Clark, W.A.V. (2019). Migration and occupational careers: the static and dynamic urban wage premium by education and city size, *Papers in Regional Science*, vol. 98:1, 555–574.

Lehmer, F. and Möller, J. (2010). Interrelations between the urban wage premium and firm-size wage differentials: a microdata cohort analysis for Germany, *The Annals of Regional Science*, vol. 45:1, 31–53.

Machin, S., Pelkonen, P. and Salvanes, K. (2012). Education and mobility, *Journal of the European Economic Association*, vol. 10:2, 417–450.

Melo, P. C., Graham, D. J. and Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies, *Regional Science and Urban Economics*, vol. 39:3, 332–342.

Mion, G. and Naticchioni, P. (2009). The spatial sorting and matching of skills and firms, *Canadian Journal of Economics*, vol. 42:1, 28–55.

Moretti, E. (2011). Local Labor Markets, in Ashenfelter, O. and Card, D. (eds.), *Handbook of Labor Economics*, vol. 4B, Amsterdam, North-Holland, 1237–1313.

OECD (2018). *Productivity and Jobs in a Globalised World – (How) can all regions benefit?* Paris, OECD.

Rosenbaum, P. R. (1984). The consequences of adjustment for a concomitant variable that has been affected by the treatment, *Journal of the Royal Statistical Society*, Series A, vol. 147:5, 656–666.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects, *Biometrika*, vol. 70:1, 41–55.

Rosenbaum, P. R. and Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score, *The American Statistician*, vol. 39:1, 33–38.

Rosenthal, S. and Strange, W. (2001). The determinants of agglomeration, *Journal of Urban Economics*, vol. 50:2, 191–229.

Rubin, D. B. (1974). Estimating causal effects of treatment in randomized and nonrandomized studies, *Journal of Educational Psychology*, vol. 66:5, 688–701.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation, *Health Services and Outcomes Research Methodology*, vol. 2:3–4, 169–188.

Rubin, D. B. and Thomas, N. (2000). Combining propensity score matching with additional adjustments for prognostic covariates, *Journal of the American Statistical Association*, vol. 95:450, 573–585.

Tano, S., Nakosteen, R., Westerlund, O. and Zimmer, M. (2018). Youth-age characteristics as precursors of power couple formation and location choice, *Labour Economics*, vol. 52, 98–111.

Venhorst, V., van Dijk, J. and van Wissen, L. (2010). Do the best graduates leave the peripheral areas of the Netherlands?, *Tijdschrift voor economische en sociale geografie*, vol. 101:5, 521–537.

Venhorst, V., van Dijk, J. and van Wissen, L. (2011). An analysis of trends in spatial mobility of Dutch graduates?, *Spatial Economic Analysis*, vol. 6:1, 57–82.

Venables, A. J. (2011). Productivity in cities: self-selection and sorting, *Journal of Economic Geography*, vol. 11:2, 241–251.

Wasmer, E. and Zenou, Y. (2002). Does city structure affect job search and welfare?, *Journal of Urban Economics*, vol. 51:3, 515–541.

Wheeler, C. H. (2001). Search, sorting, and urban agglomeration, *Journal of Labor Economics*, vol. 19:4, 879–899.

Winters, J. V. (2011). Why are smart cities growing? Who moves and who stays?, *Journal of Regional Science*, vol. 51:2, 253–270.

Table A1 Descriptive statistics (means) for stayers and migrants

Panel A: Stayers and upward migrants	Small region stayers	Small to medium region migrants	Small to large region migrants	Small to very large region migrants	Medium region stayers	Medium to large region migrants	Medium to very large region migrants	Large region stayers	Large to very large region migrants
Female	0.47	0.47	0.45	0.50	0.48	0.45	0.51	0.47	0.49
Age	25.18	25.29	25.36	25.67	25.45	25.42	25.72	25.42	25.76
Married	0.10	0.08	0.06	0.05	0.10	0.06	0.05	0.07	0.05
Children	0.08	0.06	0.03	0.03	0.08	0.03	0.03	0.05	0.02
Swedish	0.98	0.97	0.95	0.96	0.96	0.95	0.94	0.94	0.94
Foreign parents	0.03	0.04	0.06	0.05	0.06	0.07	0.08	0.09	0.08
High school GPA	14.91	15.25	15.69	15.77	14.99	15.70	15.72	15.53	16.22
One parent with high education	0.15	0.19	0.23	0.23	0.20	0.25	0.26	0.24	0.29
Two parents with high education	0.05	0.08	0.11	0.14	0.09	0.16	0.19	0.15	0.23
Parents' earnings (SEK 1,000)	448	445	482	497	476	507	542	520	592
4-year university	0.42	0.52	0.62	0.68	0.48	0.63	0.69	0.58	0.72
5-year university	0.02	0.04	0.05	0.03	0.03	0.05	0.04	0.03	0.05
University quality Q2	0.13	0.15	0.10	0.06	0.22	0.13	0.09	0.12	0.06
University quality Q3	0.25	0.26	0.13	0.14	0.27	0.15	0.14	0.08	0.08
University quality Q4	0.26	0.35	0.23	0.35	0.29	0.23	0.34	0.21	0.29
University quality Q5	0.12	0.10	0.44	0.38	0.09	0.43	0.37	0.49	0.54
Sample size	3,410	4,235	2,700	3,949	17,201	7,377	11,777	16,073	3,860

Panel B: Stayers and downward migrants	Very large region stayers	Very large to large region migrants	Very large to medium region migrants	Very large to small region migrants	Large region stayers	Large to medium region migrants	Large to small region migrants	Medium region stayers	Medium to small region migrants
Female	0.48	0.51	0.52	0.46	0.47	0.51	0.48	0.48	0.46
Age	25.69	25.73	25.68	25.93	25.42	25.46	25.47	25.45	25.38
Married	0.07	0.06	0.08	0.08	0.07	0.07	0.06	0.10	0.09
Children	0.04	0.03	0.06	0.06	0.05	0.05	0.04	0.08	0.06
Swedish	0.93	0.95	0.95	0.95	0.94	0.95	0.93	0.96	0.97
Foreign parents	0.13	0.08	0.09	0.11	0.09	0.06	0.07	0.06	0.05
High school GPA	15.84	15.71	15.64	15.35	15.53	15.47	15.21	14.99	15.05
One parent with high education	0.29	0.30	0.24	0.28	0.24	0.25	0.23	0.20	0.22
Two parents with high education	0.22	0.26	0.22	0.21	0.15	0.14	0.13	0.09	0.10
Parents' earnings (SEK 1,000)	641	648	591	568	520	502	494	476	467
4-year university	0.68	0.71	0.66	0.67	0.58	0.56	0.54	0.48	0.52
5-year university	0.03	0.03	0.06	0.05	0.03	0.06	0.06	0.03	0.04
University quality Q2	0.04	0.08	0.13	0.05	0.12	0.19	0.11	0.22	0.14
University quality Q3	0.10	0.09	0.13	0.10	0.08	0.19	0.12	0.27	0.28
University quality Q4	0.35	0.29	0.36	0.29	0.21	0.25	0.16	0.29	0.27
University quality Q5	0.48	0.50	0.30	0.43	0.49	0.29	0.38	0.09	0.15
Sample size	22,393	1,030	1,728	346	16,073	2,067	495	17,201	1,679

Notes: To save space, the table excludes descriptive statistics for field of education in high school, field of university education, local labour market area and municipality type at age seventeen, and graduate cohort. See Section 3 for additional definition of variables.

Table A2 Probit propensity score estimates for upward migration

	Small to medium	Small to large	Small to very large	Medium to large	Medium to very large	Large to very large
Female	0.013	0.071*	0.051	0.055**	0.054***	-0.072***
	(0.036)	(0.043)	(0.038)	(0.022)	(0.018)	(0.024)
Age	0.479***	0.614***	0.786***	0.403***	0.507***	0.822***
	(0.119)	(0.146)	(0.136)	(0.080)	(0.071)	(0.103)
Age square	-0.009***	-0.011***	-0.014***	-0.007***	-0.009***	-0.015***
	(0.002)	(0.003)	(0.003)	(0.002)	(0.001)	(0.002)
Married	-0.050	-0.173	-0.200**	-0.107**	-0.169***	-0.063
	(0.091)	(0.107)	(0.100)	(0.053)	(0.047)	(0.065)
Children	-0.220**	-0.441***	-0.709***	-0.383***	-0.540***	-0.424***
	(0.105)	(0.132)	(0.121)	(0.067)	(0.060)	(0.088)
Swedish	-0.267*	-0.196	-0.102	-0.160**	-0.173***	-0.137*
	(0.158)	(0.170)	(0.177)	(0.073)	(0.064)	(0.072)
Foreign parents	0.125	0.361***	0.465***	0.067	0.210***	0.035
	(0.131)	(0.140)	(0.145)	(0.059)	(0.051)	(0.059)
High school GPA	0.030***	0.025***	0.034***	0.024***	0.025***	0.049***
	(0.008)	(0.009)	(0.008)	(0.005)	(0.004)	(0.005)
One parent with high education	0.165***	0.253***	0.218***	0.126***	0.147***	0.142***
	(0.043)	(0.050)	(0.045)	(0.024)	(0.021)	(0.027)
Two parents with high education	0.274***	0.296***	0.402***	0.250***	0.281***	0.198***
	(0.069)	(0.076)	(0.066)	(0.033)	(0.028)	(0.032)
Parents' earnings	-0.193**	0.127	0.167**	0.067	0.209***	0.165***
	(0.077)	(0.092)	(0.076)	(0.043)	(0.037)	(0.036)
4-year university	0.249***	0.255***	0.269***	0.046**	0.108***	0.186***
	(0.037)	(0.044)	(0.040)	(0.023)	(0.020)	(0.029)
5-year university	0.332***	0.265**	0.012	0.006	-0.230***	0.146**
	(0.107)	(0.112)	(0.117)	(0.055)	(0.051)	(0.063)
University quality Q2	0.262***	0.281***	0.143*	-0.116**	-0.045	0.199***
	(0.060)	(0.076)	(0.078)	(0.045)	(0.040)	(0.064)
University quality Q3	0.287***	0.160**	0.409***	-0.110**	0.108***	0.474***
	(0.053)	(0.070)	(0.064)	(0.043)	(0.037)	(0.063)
University quality Q4	0.343***	0.475***	0.762***	0.193***	0.566***	0.446***
	(0.053)	(0.069)	(0.062)	(0.043)	(0.037)	(0.058)
University quality Q5	-0.153**	1.030***	1.179***	1.063***	1.274***	0.422***
	(0.067)	(0.069)	(0.066)	(0.044)	(0.039)	(0.058)
Constant	-6.667***	-9.767***	-11.567***	-7.133***	-7.873***	-13.237***
	(1.562)	(1.909)	(1.789)	(1.054)	(0.933)	(1.347)

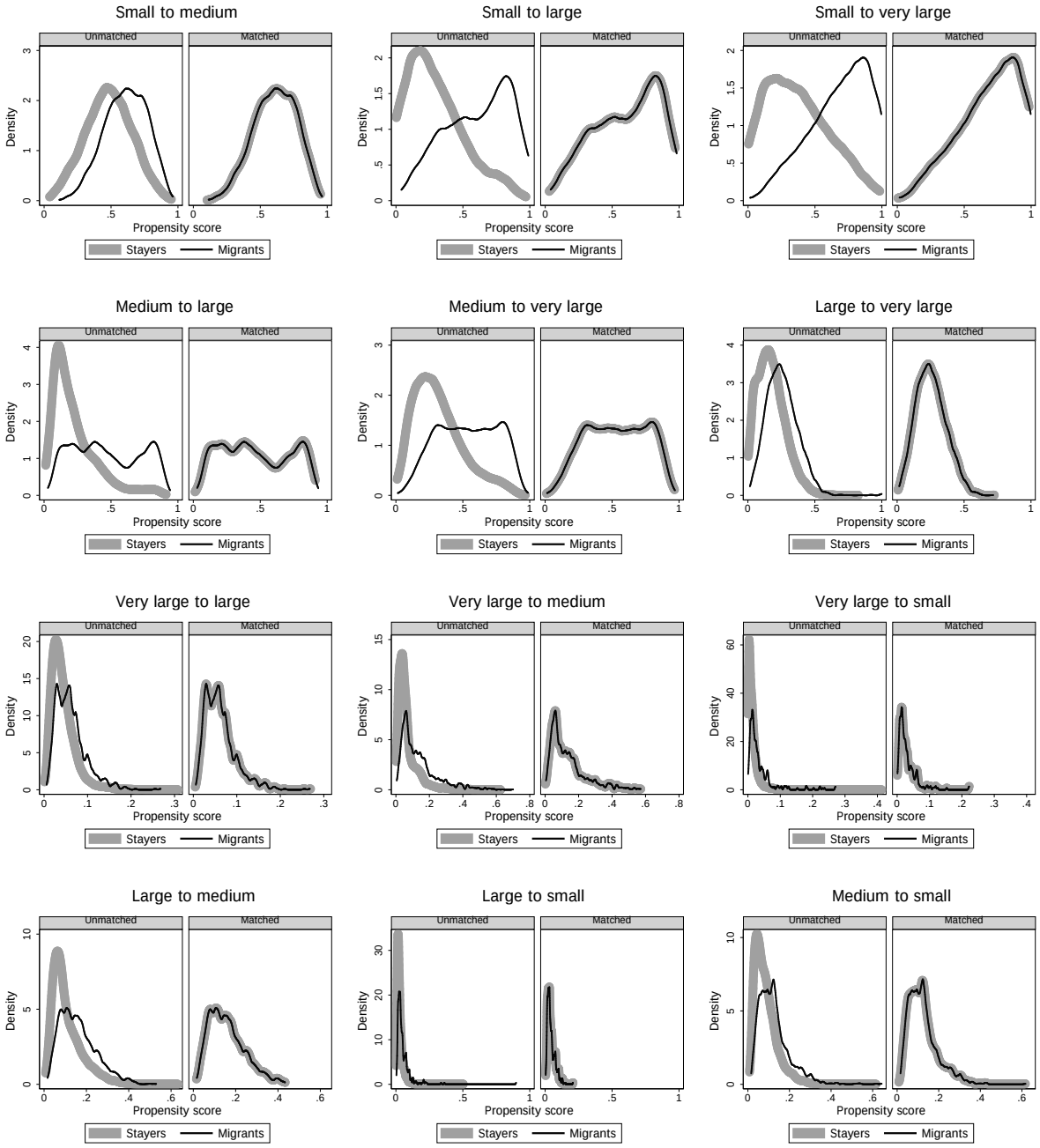
Notes: The specification of the propensity score models also include covariates for field of education in high school, field of university education, local labour market region and municipality type at age seventeen, and graduate cohort. See Section 3 for additional definition of covariates included in the propensity score models. Robust standard errors in parentheses. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively.

Table A3 Probit propensity score estimates for downward migration

	Very large to large	Very large to medium	Very large to small	Large to medium	Large to small	Medium to small
Female	0.147***	0.048	-0.011	-0.025	0.001	0.008
	(0.034)	(0.030)	(0.051)	(0.029)	(0.047)	(0.031)
Age	0.232	0.078	-0.106	0.143	-0.361**	0.141
	(0.143)	(0.119)	(0.194)	(0.109)	(0.163)	(0.111)
Age square	-0.004	-0.002	0.002	-0.003	0.007**	-0.003
	(0.003)	(0.002)	(0.004)	(0.002)	(0.003)	(0.002)
Married	0.074	0.068	-0.076	-0.072	-0.084	0.014
	(0.089)	(0.076)	(0.137)	(0.077)	(0.120)	(0.073)
Children	-0.244**	0.103	0.097	0.044	-0.237	-0.075
	(0.119)	(0.091)	(0.164)	(0.093)	(0.155)	(0.087)
Swedish	-0.060	-0.071	0.113	-0.181*	-0.507***	0.071
	(0.091)	(0.080)	(0.131)	(0.093)	(0.160)	(0.117)
Foreign parents	-0.247***	-0.289***	-0.022	-0.295***	-0.476***	-0.096
	(0.071)	(0.061)	(0.098)	(0.078)	(0.150)	(0.093)
High school GPA	-0.016**	-0.001	-0.024**	0.007	-0.014	-0.008
	(0.007)	(0.007)	(0.011)	(0.006)	(0.011)	(0.007)
One parent with high education	0.084**	-0.069**	0.038	0.069**	0.010	0.060*
	(0.037)	(0.034)	(0.056)	(0.032)	(0.052)	(0.035)
Two parents with high education	0.148***	0.061	0.074	0.068	0.034	0.120**
	(0.042)	(0.038)	(0.067)	(0.042)	(0.071)	(0.052)
Parents' earnings	-0.011	-0.095**	-0.168**	-0.128**	-0.051	-0.083
	(0.037)	(0.048)	(0.069)	(0.053)	(0.107)	(0.066)
4-year university	0.116***	0.201***	0.148**	0.207***	0.205***	0.102***
	(0.042)	(0.036)	(0.063)	(0.035)	(0.058)	(0.033)
5-year university	0.164*	0.536***	0.443***	0.535***	0.755***	0.278***
	(0.095)	(0.074)	(0.127)	(0.069)	(0.111)	(0.081)
University quality Q2	0.209**	0.165**	-0.522***	0.307***	-0.466***	-0.309***
	(0.103)	(0.072)	(0.136)	(0.056)	(0.083)	(0.060)
University quality Q3	-0.098	-0.350***	-0.532***	0.463***	-0.321***	-0.116**
	(0.095)	(0.069)	(0.109)	(0.057)	(0.084)	(0.054)
University quality Q4	-0.168*	-0.561***	-0.736***	0.033	-0.577***	-0.181***
	(0.088)	(0.064)	(0.099)	(0.058)	(0.084)	(0.058)
University quality Q5	-0.194**	-0.949***	-0.744***	-0.387***	-0.615***	0.005
	(0.087)	(0.064)	(0.096)	(0.058)	(0.076)	(0.063)
Constant	-4.892***	-2.017	0.077	-2.855**	3.892*	-3.260**
	(1.880)	(1.559)	(2.548)	(1.420)	(2.104)	(1.457)

Notes: The specification of the propensity score models also include covariates for field of education in high school, field of university education, local labour market region and municipality type at age seventeen, and graduate cohort. See Section 3 for additional definition of covariates included in the propensity score models. Robust standard errors in parentheses. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively.

Figure A1 Propensity score density plots for stayers and migrants before and after matching



Notes: See Section 5.4 for details of the applied matching algorithm.