IFAU – INSTITUTE FOR
LABOUR MARKET POLICY
EVALUATION

Neighbourhoods, economic incentives and post compulsory education choices

Lars Lindvall

WORKING PAPER 2009:11

The Institute for Labour Market Policy Evaluation (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market policies, studies of the functioning of the labour market, the labour market effects of educational policies and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P.O. Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

Papers published in the Working Paper Series should, according to the IFAU policy, have been discussed at seminars held at IFAU and at least one other academic forum, and have been read by one external and one internal referee. They need not, however, have undergone the standard scrutiny for publication in a scientific journal. The purpose of the Working Paper Series is to provide a factual basis for public policy and the public policy discussion.

ISSN 1651-1166

Neighbourhoods, economic incentives and post compulsory education choices[a]

by

Lars Lindvall[b]

May 17, 2009

Abstract

There are large differences in income and education levels, unemployment and ethnic composition between neighbourhoods. An interesting question is whether a neighbourhood's characteristics affect the behaviour of its residents. This paper investigates neighbourhood effects on youths' post primary education choice. Besides including usual variables the paper also includes neighbourhood specific economic incentives. Estimating linear probability models as well as multinomial logit models using Swedish register data, covering the county of Stockholm and the years 1988–1992, I find that both neighbourhood characteristics and economic incentives affect the choice. For the latter the results are quite clear although the size of the effect is small: an increase in the expected income of an alternative increases the probability that this alternative is chosen. For the neighbourhood variables the results differ to some extent depending on the model. The proportion of individuals with at most compulsory education in a neighbourhood does however seem to have a negative effect on applying for a university preparatory programme. The proportion of immigrants in a neighbourhood tend to have a positive effect on immigrants' probability to apply for a university preparatory programme.

Keywords: Neighbourhoods, economic incentives, educational choice
JEL-codes: I22, I20, R19

[a]Comments from Martin Lundin, Ryszard Szulkin and Olof Åslund, and seminar participants at Department of Economics, Uppsala University, and IFAU are greatly appreciated.
[b]Department of Economics, Uppsala University, e-mail: lars.lindvall@nek.uu.se.

Table of contens

1 Introduction

You do not have to travel far in a city to notice that neighbourhoods differ in many aspects. The type and state of houses, the cars parked in the street or the different languages you may hear could all tell you that a neighbourhood differs from another you visited. These differences come around since individuals sort themselves into neighbourhoods. The sorting can, for example, be due to preferences for amenities associated with a neighbourhood, preferences for living with some types of people or not to live with other types, or economic incentives (see e.g., Durlauf, 2004). Whatever the reason, a neighbourhood is generally a poor sample, i.e. not representative, of the whole population. There might be large differences between neighbourhoods in terms of, for example, income and educational level, unemployment rate and ethnic composition. In Sweden and many other Western countries, the effects of these neighbourhood differences are an issue of political and scholarly interest. Indeed, there is a large, too large to review here, and growing empirical literature investigating so called neighbourhood effects on individual social and economic outcomes.[1]

The understanding of neighbourhood effects is important for several reasons. From a policy perspective knowledge is important since it may tell whether policies ought to be neighbourhood/area based or more general. The most important reason is perhaps the self reinforcing nature of neighbourhood effects. For example, if growing up in a deprived neighbourhood has detrimental effects on individual outcomes these individuals are more likely to settle in deprived neighbourhoods later in life. This will lead to decreased social mobility and the cementing of segregation. Education is an important aspect of social mobility and is one of the many outcomes studied in the literature – often in form of educational attainment (e.g. years of schooling) or achievement (e.g. grades), risk of dropping out or chance of graduating.[2]

[1] See e.g. Atkinson and Kintera (2001), Borjas (1995), Brännström (2005), Crane (1991), Cutler and Glaeser (1997), Edin et al. (2003), Freidrichs and Blasius (2003), or Oreopoulos (2003) for some specific studies or the articles by Ellen and Turner (1997), Ginther et al. (2000) and Galster (2007), which include reviews.

[2] See e.g. Aaronson (1998), Ainsworth (2002), Cardak and McDonald (2004) and Vartanian and Gleason (1999). The determinants of educational outcomes has its own large literature, see e.g. Haveman and Wolfe (1995).

This paper consider education too. Departing from a random utility framework, using discrete choice models and Swedish register data this paper studies the impact of neighbourhoods on post compulsory education choice. Besides including usual neighbourhood variables this paper adds neighbourhood specific economic incentives, in terms of expected income, to the analysis and the litteratur.

There are already a number of Swedish studies on neighbourhoods and education. For example, Andersson and Subramanian (2006) find negative association between neighbourhood characteristics, such as financial resources and ethic concentration, and years of education. Similarly, Nordin (2006) finds that ethnic segregation has a detrimental impact on educational attainment. In a study of ethnic enclaves, Grönqvist (2006) arrive at this result too, as he shows that ethnic concentration impacts the educational attainment of immigrants negatively. However, Brännström (2008) analyses the impact of neighbourhood and school characteristics on upper secondary education (USE) achievements, i.e. final grades, and finds no direct evidence that neighbourhoods matter.

This paper differs from these studies in at least two important aspects. First, it studies educational choice rather than outcomes. The interest here is the determinants of who is applying for USE, rather than who is, for example, graduating from USE. Attainments and achievements are important and interesting, but it all starts with the post compulsory education choice, i.e. you have to apply for graduate. It is the first watershed towards higher education. Second, it includes economic incentives in terms of expected incomes. In economics the prevailing view, following Becker's (1964) seminal work, of educational attainment is that of human capital theory. Education is an investment comparable to any other investment. As such the return and the cost of additional education determine the decision whether to invest, i.e. to further one's education. One important aspect of the returns to education is future income. Indeed, empirical studies, such as Wilson (2001) and Wilson et al. (2005), reveal that individuals respond to economic incentives, in terms of future income, when making their educational choices.[3]

Human capital theory leads us to expect economic incentives to impact educational decisions. What about neighbourhoods? The literature propose several mechanism, for

[3]These two studies use a similar framework to the one used in this paper, although not from a neighbourhood perspective.

example, collective socialisation, social control, social capital and perception of opportunity, through which neighbourhoods may effect education (Ainsworth, 2002). The mechanisms often involve 'role models', which mediate behaviour, attitudes, values and information. Neighbourhood characteristics shape the type of role models individuals are exposed to outside the home. For example, with few positive role models in their neighbourhood, individuals may be less likely to learn important behaviours and attitudes that lead to success in school. Moreover, lack of positive role models may deflate the value of such behaviours and attitudes. Neighbourhoods may also provide an important base for the perception of the value of education. The value of an education is to a large extent dependent on future incomes. These cannot be known with certainty. Expectations about the future and how these expectations are formed becomes important (Manski, 1993b). One way for individuals to gather information and form expectations is to observe the outcome of older individuals who already made their choices (Manski, 1993a). If these role models are confined to the neighbourhood, neighbourhood specific economic incentives become important. Neighbourhoods may thus influence educational choice directly through, for example, attitudes and indirectly through economic incentives.

The paper begins with a outline of the Swedish school system and neighbourhoods, which aims to provide a brief institutional background. The rest of the paper is structured as follows: Section 3 describes and discusses the econometric framework of the study. Departing from a random utility framework two versions of the educational choice is modelled. First, a binary choice of whether to apply for a university preparatory programme using a linear probability model. The second version is a multinomial choice between discontinue (not applying), applying for a vocational programme and applying for university preparatory programme. The latter version is modelled with a multinomial logit model. This section also discusses the estimation of expected incomes. Following this discussion section 4 describes the data and the variables used in the analysis. Before section 6 concludes the paper, section 5 presents and discusses the empirical results. The estimation results show that neighbourhoods affect youths' post compulsory education choice. This is true after controlling for individual and parental characteristics as well as neighbourhood fixed effects. Neighbourhood specific economic incentives and neigh-

bourhood levels of education and immigration affects the probability to make different choices. For the neighbourhood variables the results differ to some extent depending on how the choice problem is formulated. The economic incentives, measured as expected incomes, have an effect on the choice regardless of model choice. An increase in the expected income, here estimated with neighbourhood wage regressions, of an alternative increases the probability that this alternative is chosen. This is a result consistent with human capital theory. The effect is, however, in some cases very small.

2 The Swedish school system and neighbourhoods

The Swedish school system has gone through substantial changes over the years, especially during the 1990s and a detailed account will not be given here, but rather a brief outline and some aspects relevant for later discussions. Readers interested in further details are referred to e.g. Björklund et al. (2005). According to the Education Act, all children and youths shall have equal access to education regardless of gender, where they live, or social or economic factors (Skolverket, 2008). To meet this end all education in the public school system is free and there is usually no charge to pupils nor their parents for teaching materials, school meals or transports. Before the changes to the system in the 1990s the government provided earmarked public funds for the financing of the education, although 'topping up' by municipalities were allowed. After the changes the municipalities has full financial responsibility for primary and secondary schools. Another change was the introduction of school choice, which requires municipalities to satisfy, subject to space limitations, parental preferences regarding school choices. The main principle of allocating individuals to schools is, however, still the residence principle.[4] Due to which a majority of individuals attends compulsory school in or close to the neighbourhood where they live.

Starting at the age of seven, school is compulsory for nine years and follows a national curriculum. The typical graduation age is hence 16. Even though voluntary, a vast major-

[4] Närhetsprincipen.

ity attends USE. In the 1980s about 80 percent of the pupils continued, and in the 1990s almost all (above 95 percent) continued (Landell et al., 2000). USE consists of several different study programmes. These programmes can be divided into vocational programmes (VP) and university preparatory programmes (UPP).[5] Like the rest of the Swedish school system these programmes where subjects of change in the early 1990s. One particular change is of interest here. In the eighties VP were two years long whereas UPP where three or four years long. Besides differing in length and curricula the VP did not give eligibility for university entry. In the early 1990s, however, a third year of mainly theoretical studies was added to the VP curricula. A completed programme now gave eligibility for university studies as VP graduates now fulfilled the general requirement for university studies. In practice this meant that some university programmes became available for VP graduates. The difference between the two tracks thus became smaller.

The neighbourhood situation is quite different in Sweden (and Western European countries) compared to the US. The levels of ethnic and socioeconomic segregation is usually lower in Swedish cities (Andersson, 2000). Moreover, areas with high poverty and different forms of deprivation are, also in contrast to the US experience, not located in the inner cities. Instead it is typically large scale public housing estates located in suburbs that show signs of deprivation. It is generally not possible to classify neighbourhoods as black, hispanic, etc. as is often done with American neighbourhoods. The number of different nationalities in neighbourhoods with large concentration of immigrants is very large and effectively stops such classification.

3 Econometric framework

When finishing compulsory education individuals have a veritable smorgasbord of options and opportunities to choose among. They could discontinue their education and start working or do something else, or they could stay in school. If they decide to stay in school there are several different programmes to choose among. Their choice problem may be modelled in several different ways. To keep things simple, two different versions

[5] Yrkesförberedande linje and teoretisk linje.

of the problem will be considered here. First, the plethora of possible choices are boiled down to the binary choice of whether to apply for a UPP. This formulation of the problem is interesting because it is a first watershed whether to study at a university. At the same time, the choices of individuals not applying for a UPP are treated the same even though there may be substantial differences. This motivates a second version in which the choice set is expanded into three alternatives – discontinuing education, applying for a VP and applying for a UPP. Both versions of the decision can be analysed departing from a random utility framework and similar formulations, to the one that follows, may be found throughout the literature.[6]

Consider the following linear utility function, giving the expected life time utility of choosing alternative s for individual i

$$U_i^s = \beta^s X_i + \delta Y_i^s + \varepsilon_i^s \tag{1}$$

where the first right hand side term is the consumption of an educational good or the non-pecuniary value of choosing alternative s. X denotes individual specific characteristics, e.g., individual, family or neighbourhood characteristics, that do not vary over the alternatives. Note that the variables have alternative specific parameters, β^s. If this where not the case the effects would not be identifiable in estimations as they would cancel out for different alternatives (see e.g Greene, 1997). It is in this first term the mechanisms mentioned above may come into play, i.e. neighbourhood characteristics affects the non-pecuniary value of education. The second term is the expected life time income, Y_i^s, following a choice, i.e. the alternative's pecuniary value or consumption of other goods. The income varies over alternative and have the same effect on utility, δ. Finally, the last term, ε_i^s, is an alternative specific random term.

Given the utilities of two alternatives, an utility maximising individual will choose alternative s over alternative k if

$$U_i^s - U_i^k = (\beta^s - \beta^k)X_i + \delta(Y_i^s - Y_i^k) + \varepsilon_i^s - \varepsilon_i^k > 0, \tag{2}$$

[6]See e.g., Wilson et al. (2005) for a educational choice application or Wolfe et al. (2007) for a similar econometric framework applied to youths' childbearing decision.

i.e. the utility from alternative s is larger than the utility from alternative k.[7] The choice thus depends on the difference in non-pecuniary value between the alternatives and the difference in expected income. At the time of the decision the expected income is not known to the individuals, however. Instead they must form expectations about the income generated by the alternatives.

There are several ways in which individuals may form these expectations. One way is to get information from role models or reference groups (Manski, 1993a,b). Individuals could, for example, look to their parents, friends' or classmates' parents, or other individuals in their surroundings. Here expectations are assumed to be formed using the neighbourhood as reference group. When forming their expectations individuals are thus assumed to observe the outcomes of older individuals in their neighbourhood who have already made the post compulsory education choice. The choice of neighbourhoods as reference group could always be contested. Most individuals' parents, at least some of their friends' and classmates' parents will be included, however. Moreover, in order to get information on all alternatives a wide reference group may be needed.

The expected income is here modelled with means of neighbourhood and time specific wage equations. The wage equations are, due to zero wage income observations, modelled with Tobit models (see e.g. Wooldridge, 2002). We have the following model;

$$Y_o = \begin{cases} BX_o + \varepsilon_o & \text{if } BX_o + \varepsilon_o > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where Y_o is the yearly wage income of older individual o, ε_o is a normally distributed random element of individual income and $X_o = D_o, D_o a_o, D_o a_o^2$, where D includes dummy variables for sex, immigrant background, and educational level, and a is age.[8] The Tobit estimates allow income streams to be predicted for different educational, combinations of sex, immigrant background. The present value of the income streams are used as the expected income of a specific alternative.

[7] It is clear from this equation that if $\beta^s = \beta^k$ the effect of X_i cannot be estimated.

[8] *Table A-2*, which presents estimations for three neighbourhoods and years, displays the structure of the wage equations.

Outlined above is, basically, a structural model containing a utility equation and an expected income equation. Here the expected income will be estimated first and used as a explanatory variable in the estimation of the probability model resulting from the utility equation. It is hence assumed that the random terms in the two equations are independent. Identification is possible due to the non-linear nature of the expected incomes and variation across neighbourhoods. Furthermore, variables appearing in both equations will not have any clear interpretation. A third equation, a neighbourhood selection equation, could also be added to a perceived structural model. Dependence between the selection equation and the utility equation will give rise to the selection problem that riddle neighbourhood studies. If, for example, the selection into neighbourhoods is based on some unobservable characteristics, any neighbourhood variables in the analysis will become endogenous, and hence render any estimates faulty, if the unobservable characteristics also affect the educational choice.[9] The selection equation will not be modelled here, but it must nevertheless be considered.

It is useful to consider the choices, the timing and the decision makers involved. The first choice is the choice of neighbourhood, which is probably made by the individuals' parents. This choice is hence affected by the parents' preferences and characteristics. Individuals may, however, effect the choice directly or indirectly by their preferences or characteristics. The second choice is the post compulsory education choice, which is made by the individuals under more or less influence from their parents either direct or indirectly. Individual as well as parental preferences and characteristics will thus affect this choice. The two choices may hence be intertwined and the selection issue present. Leaving out any characteristics affecting both choices when estimating the utility function parameters renders biased estimates. It thus becomes important to control for individual and parental characteristics that can be expected to affect both decisions. Data availability sets limits to which variables that can be included, however. One way to mitigate the selection issue is to include neighbourhood fixed effects. These take all neighbourhood characteristics, observed and unobserved, that are constant over time into account and should, if the selection is based on these characteristics, mitigate the selection bias.

[9] See e.g. Durlauf (2004) for a more lengthy discussion of problems and solutions

Depending on the number of alternatives and the assumption about the random term in the utility equation several different models may be fitted into this framework. Consider first the binary choice. This type of choice is often modelled with a logit or probit model, and sometimes with a linear probability model (LPM) (see e.g., Wooldridge, 2002). Here the latter is used. The three models often produce similar results and the LPM is chosen for its simplicity. It is simpler to estimate, using ordinary least squares (OLS), with large samples. Including neighbourhood fixed effects is straightforward using neighbourhood dummies. The LPM is also easier to interpret, especially with neighbourhood fixed effects. A major drawback with the LPM estimated with OLS is the possibility to get predicted probabilities outside the unit interval. This may lead to biased and inconsistent estimates.

When there are more than two alternatives, as in the second version of the choice, the LPM does not fit the bill. An often used model for multinomial choice situations, which will also be used here, is the multinomial logit (MNL) (see e.g., Wooldridge, 2002). The estimation of the MNL is straightforward save the case of neighbourhood fixed effects. When it comes to fixed effects in non-linear models in general, the incidental parameter problem presents challenges for unbiased estimation of parameters, which is also the case for the MNL (see e.g., Greene, 2002). These results are, however, based on fixed T asymptotics, i.e. the number of observations (often time periods) used to identify each fixed effect is held constant as the number of fixed effects increases towards infinity. For large T, this problem disappears. It is, however, unclear of which order T must be to produce unbiased estimates.[10] Here neighbourhood fixed effects are included in the MNL by means of neighbourhood dummies. The number of observations used to estimate each effect is the number of individuals in each neighbourhood, which should be sufficiently large. For the reminder of the paper, neighbourhood dummies will be used synonymously with neighbourhood fixed effects.

[10] E.g., based on Monte Carlo simulations Katz (2001) finds conditional and unconditional fixed effects estimation of a binary logit to be equivalent for $T \geq 16$, whereas Greene (2002) finds the bias decreasing rapidly with T larger than 3.

4 Data and descriptive statistics

The data used herein come from the IFAU database, which is a register based database covering the whole Swedish population between ages 16 and 64. Two different data sets will be used in the analysis. The primary data set covers most of the individuals finishing compulsory education in the county of Stockholm, from here on referred to as just Stockholm, during the years 1988 to 1992.[11] Availability of data determines the start year. The end year is due to the changes of USE in the early 1990s. Most notable was that all programmes were made three years long and gave eligibility for university entry upon completion. Restricting the analysis to this time period makes estimation of expected incomes straightforward as there is a clear difference between the two tracks. Furthermore, older individuals, used in the expected income estimation, will have educational attainment according to the old regime. These older individuals are part of a secondary data set, which is used to derive some of the variables used. The secondary data set covers the whole population of the county of Stockholm aged 16 to 64 and the same time period as the primary data set.

There are some qualifiers for inclusion of individuals in the primary data set. First, the individuals must be 16 years old when they finish compulsory education. This is the normal age for graduation. Second, the individual must reside in the same SAMS (Small Area Market Statistic) as at least one of their parents. SAMS are a division of Sweden, made by Statistics Sweden, into around 9,000 homogenous housing areas and will herein be used as a neighbourhood identifer. From hereon the terms neighbourhood and SAMS will be used synonymously. Restricting the analysis to individuals living in the same SAMS as at least one parent may lessen the selection problem to some extent. Even though it is not certain, it is quite probable that these individuals are still living at home. Living at home implies that they have not yet sorted themselves into neighbourhoods. Third, individuals must have observations for all variables to be used. Besides missing data, often pertaining

[11] Even though the database contains data for the whole of Sweden this studie is restricted to the county of Stockholm for a number of reasons. The issue studied could be perceived as a big city phenomenon. The other major cities, Malmö and Göteborg, are located in different parts of the country. Restricting the analysis the to the county of Stockholm implies that all the individuals face the same labour market conditions. The county of Stockholm is large enough to give observations from all kinds of neighbourhoods.

to information on fathers, another source of exclusion is the small size of some SAMS. For the smallest SAMS expected incomes cannot be estimated. Finally, in order to get stable expected income estimates a minimum neighbourhood population of 200 is required to be part of the primary data set.[12] Moreover, to estimate with neighbourhood dummies there is a need for larger SAMS, e.g., at least two observations with different choices is needed in the LPM case.

Table 1 and *Table 2* display descriptive statistics at individual and neighbourhood level, respectively, for the variables used in the analysis. The variable definitions are gathered in *Table A-1*. Before turning to the variables it can be noted that the primary data set contains 69,558 observations (individuals) distributed over 669 neighbourhoods. The total number of neighbourhoods in the original database is 892 for the time period of interest. The smaller number in the primary data set is mainly due to the aforementioned reasons. The average individual resides in a neighbourhood with a population (aged 16 to 64) of 2,410. The average neighbourhood population is 1,424, however.

4.1 Dependent variable

Starting with the dependent variable, i.e. the choice variable to be analysed, it is based on the individuals' applications for USE the year they finished compulsory education as follows; (i) discontinue – did not apply for any programme, (ii) VP – first choice is a two year programme, (iii) UPP – first choice is a three (or four) year programme. The applications are submitted during the last semester of compulsory education and the point of time of analysis is hence in the end of compulsory education. Individuals that discontinue, according to the classification above, may in fact have deferred their education. They could apply for USE later in their life. Around one quarter of the primary data set did not apply for any programme, one fifth applied for a VP and the rest (about 53 per cent) applied for a UPP. For the LPM the first two, discontinue and VP, are considered as one alternative. For both models the lowest alternative is used as base alternative, i.e. vocational/discontinue or 'not applying for a UPP' for the LPM and discontinue for the MNL.

[12] Removing or decreasing the population requirement does not affect the estimation result in any significant way. The time it takes to estimates the MNL with neighbourhood dummies increases a lot, however.

Table 1: Individual level descriptive statistics.

Variable	Mean	Standard deviation	Min	Max
Dependent variable				
Discontinue	0.2581			
Vocational programme	0.2066			
University preparatory programme	0.5351			
Individual variables				
Sex	0.4885			
Immigrant background	0.1189			
GPA	3.3169	0.6684	1	5
Expected income (binary choice, 1,000,000 SEK)				
Discontinue/Vocational	2.6077	0.6018	0.6742	21.7409
University preparatory	3.1720	0.9201	1.0454	26.4514
Difference in expected income (binary choice)				
University preparatory - Discontinue/Vocational	0.5643	0.4999	-1.7904	12.8225
Expected income (multinomial choice, 1,000,000 SEK)				
Discontinue	2.2971	0.6210	0.1136	21.7191
Vocational	2.9772	0.7064	0.1518	23.4190
University preparatory	3.4104	1.0955	0.8775	30.0815
Difference in expected income (multinomial choice)				
Vocational - Discontinue	0.6801	0.4527	-4.3201	7.1291
University preparatory - Discontinue	1.1133	0.7449	-3.7243	25.2683
Parental variables				
Father's income (1,000,000 SEK)	0.2043	0.1545	0	5.6838
Father's education				
Compulsory or less	0.2543			
Vocational	0.2242			
University preparatory	0.2107			
University	0.3107			
Mother's income (1,000,000 SEK)	0.1271	0.0736	0	1.783
Mother's education				
Compulsory or less	0.2288			
Vocational	0.3333			
University preparatory	0.1036			
University	0.3343			
Neighbourhood variables				
Population	2,410.81	2,181.319	203	12,288
Education[a]	0.0231	0.2925	-0.6861	1.1490
Immigration (normalised)[b]	-0.0558	0.4789	-0.7671	2.2329
Number of observations	69,558			

[a] The proportion of neighbours with immigrant background. Normalised using the yearly total average.
[b] The proportion of neighbours with not more than compulsory education. Normalised using the yearly total average.

Other, alternative, dependent variables could be which programme individuals are admitted to or which they graduate from. Even thought these are interesting in their own right they are affected by several factors, e.g., the individuals' grades – if their good enough to get admitted – and various events during the education. Applications are more likely, even though they may also be affected by grades and external events, to be the outcome of an optimisation process and thus expressing an optimal choice. Furthermore, the application for USE is the first watershed for higher studies, i.e. if there is no application for a UPP there is no graduation from a UPP.

Table 2: Neighbourhood level descriptive statistics, selected variables.

Variable	Mean	Standard deviation	Min	Max
GPA	3.3058	0.2113	2.54	4
Expected income (binary choice)				
Discontinue/Vocational	2.6298	0.35710	1.6142	7.3171
University preparatory	3.1875	0.4158	2.0308	8.5389
Difference in expected income (binary choice)				
University preparatory - Discontinue/Vocational	0.5576	0.2308	-0.1823 2.1769	
Expected income (multinomial choice)				
Discontinue	2.3045	0.3872	1.3001	7.0783
Vocational	3.0223	0.4066	2.0193	8.0857
University preparatory	3.5229	0.5488	2.1028	9.7166
Difference in expected income (multinomial choice)				
Vocational - Discontinue	0.7177	0.2930	-0.2976	2.2917
University preparatory - Discontinue	1.2183	0.4550	0.1662	5.4038
Neighbourhood				
Population	1,424.63	1,493.972	208.2	11,521.8
Education[a]	0.0172	0.2939	-0.6500	0.9882
Immigration[b]	-0.1162	0.3784	-0.7199	2.1877
Number of Neighbourhoods	669			

[a] The proportion of neighbours with immigrant background. Normalised using the yearly total average.
[b] The proportion of neighbours with not more than compulsory education. Normalised using the yearly total average.

4.2 Explanatory variables

There are two types of explanatory variables in the analysis; variables that vary over individuals and variables that vary over neighbourhoods and time. The former include individual and parental characteristics as well as expected income. The latter are neighbourhood characteristics.

4.2.1 Individual and parental characteristics

Individual and parental characteristics can be expected to influence individuals' educational choice by, in terms of equation 1, affecting the value of the educational good. They may also be expected to influence in which neighbourhood individuals reside and are therefore important in order to take some account for selection. Included are three individual variables; sex, immigrant background and grade point average (GPA). The first two are dummy variables taking the value one if the individual is male and if the individual has an immigrant background, respectively. Individuals are defined as having an immigrant background if they or both their parents are born outside Sweden. The variable thus includes both first and second generation immigrants. Whether this is an appropriate aggregation of individuals with immigrant background may always be debated. There is empirical evidence that suggests differences in educational outcomes between first and second generation immigrants (Szulkin and Jonsson, 2007). The time of arrival in Swe-

den also seems to matter for first generation immigrants (Böhlmark, 2008). Given that this issue is not the primary interest here and that only 12 per cent of the individuals in the primary dataset have immigrant background no further division will be made. Anyway, close to half of the individuals in the primary dataset are males. The GPA is the average of the final grades in compulsory education, i.e. average grade in the 9^{th} grade. The grade scale goes from one (lowest) to five (highest) in increments of one. The GPA in the primary data set covers the whole range with an average just above 3.3. The neighbourhood average GPA (*Table 2*) is just above 3.3 too. The GPA can be seen as a measure of ability, which could affect the value of education. Moreover, the GPA can also affect the individuals' choices directly since the admittance to different programmes are based on the GPA. For example, individuals with low GPA could refrain from applying for a UPP, although it would be optimal, in the belief that they will not get admitted.

In addition to the individual variables four parental characteristics are included in the analysis, the (wage) income and education level of both parents. Both the income and the educational level of parents are important since these can be expected to affect the sorting into neighbourhoods as well as the value of the educational good. The income variables are a two year, the year of the observation and the preceding year, average. In the primary data set the average incomes are around 240,000 SEK and 127,000 SEK for fathers and mothers, respectively. The parents' income show a large span in the data ranging from 0 SEK to almost 5.7 million SEK for fathers and 1.8 million SEK for mothers. The education level of the parents is divided into four dummy variables indicating the highest level attained; (i) compulsory education or less, (ii) VP, (iii) UPP, and (iv) university. The last category is used as base and is hence left out of the estimations. For both fathers and mothers about one forth has only compulsory education or less. The share of fathers with VP is about the same as the share with UPP, a couple of percentage units above 20. The rest of the fathers, some 30 per cent, has university education. For mothers both VP and university education have a share of one third whereas only 10 per cent of the mothers has attained UPP.

4.2.2 Expected income

Using the secondary data set, save the individuals in the primary data set, the expected income is calculated in two steps. First, Tobit models (equation 3) are estimated for each neighbourhood and year. The structure of the wage equation estimated can be seen in *Table A-2*, which presents estimations for three neighbourhoods and years. About 3,300 wage equations are estimated for each version of the choice.[13] The second step generates, based on the Tobit estimates, expected income for the alternatives in the model. The formulation of the Tobit model allows for income predictions for different educational levels and combinations of sex, immigrant background and age to be made. This results in predicted income for each education level, s, and year after 16 years of age, t, $\hat{Y}^s_{i,t}$. The discounted sum over 48 years of predicted incomes gives the expected income for each education level, i.e.

$$\hat{Y}^s_i = \sum_{t=1}^{48} \frac{\hat{Y}^s_{i,t}}{(1.02)^t}$$

where s is the same as the parental education levels; compulsory or less, VP, UPP, and university. In the MNL the two lowest levels correspond to the expected income of the corresponding alternatives. In the LPM the maximum of the two lowest levels corresponds to the expected income of discontinue/VP or 'not applying for a UPP'. The last two levels are used to derive the expected income of the UPP alternative. Upon completion of a UPP the individuals are eligible to apply for university, which has a value (Comay et al., 1973). The expected income of UPP is hence not only the expected income from this level, but includes the option value of university studies too. This option value is here included by letting the largest of the two expected incomes be the expected income of UPP, i.e.

$$Y^{UPP}_i = \max\{\hat{Y}^{UPP}_i, \hat{Y}^{university}_i\} = \hat{Y}^{UPP}_i + \max\{0, \hat{Y}^{university}_i - \hat{Y}^{UPP}_i\}$$

where the last term on the right hand side is the option value of university studies.[14]

[13] There are 669 neighbourhoods in the primary data set observed over five years, which gives a total of 3,345 regressions to be estimated for each version. All neighbourhoods do not have observations for all years, however. This reduced the number of regressions estimated.

[14] Whether the individuals completing a UPP attends university is a another problem than the one studied here. It may be modelled in a similar fashion, however. The individual will attend university if utility, including pecuniary and non-pecuniary rewards, of doing so is larger than not attending.

The number of years used in the discounted sum is such that the last age included is the same as the oldest individuals in the database. Due to this choice, pension payments are not considered in the expected income although they also depend on educational choices. Given the other choice made, a discount rate of two per cent, income such far into the future will have little impact on the expected income, however. Both the number of years and the interest rate are arbitrarily set, but this is of lesser importance. Moreover, the estimated income expectations are likely to be biased and it is of course doubtful that individuals estimate Tobit models when they form their expectations. The aim is not to exactly depict the expectations formations, but rather to get a proxy that captures differences in expected incomes between neighbourhoods, alternatives and individuals.

The estimated expected incomes, for both the binary and the multinomial version, follows what one may expect as the average expected income is increasing in the level of the alternatives. For example, in the binary case the average expected incomes are 2.6 million SEK for discontinue/VP and 3.2 million SEK for UPP. This correspond to a constant yearly income of 85,000 SEK and 104,000 SEK, respectively. For the multinomial case the figures are similar. On the neighbourhood level the numbers are somewhat larger. The difference in expected income which enters the estimations is calculated by subtracting expected income of the base alternative, i.e. discontinue/VP for the LPM and discontinue for the MNL, from the expected income of the alternative. If education is profitable, in terms of income, this difference should be positive. This seems to be the normal case as the average differences are all positive. There are, however, some individuals with a negative difference.

4.2.3 Neighbourhood characteristics

As discussed above there are many mechanisms through which neighbourhoods may affect the educational choice. In the framework used here neighbourhoods may, besides affecting the expected income, affect the value of the educational good. Neighbourhoods differ in many aspects, e.g., type of housing, amenities and location. Here two different neighbourhood characteristics, education level and immigration level, are included in the model. Here these levels will be measured as proportions of neighbourhood inhabitants (between 16 and 64) with immigrant background and low education level (compulsory or

less), respectively. The neighbourhood proportions are normalised with the yearly Stockholm proportion, i.e.

$$I_{jt} = \frac{x_{jt} - \overline{x}_t}{\overline{x}_t}$$

where x_{jt} is the proportion in neighbourhood j in year t and $\overline{x}_t$ is the proportion of Stockholm in year t. The range of the measures are $I_{jt} \in \left[-1, \frac{1-\overline{x}_t}{\overline{x}_t}\right]$. The normalisation facilitates interpretation later and neighbourhoods are characterised how they fare compared to a average neighbourhood which would be the result if immigrants and poorly educated individuals would be evenly distributed. Moreover, a neighbourhood with a constant proportion over the years will get a varying measure if proportion of Stockholm changes. These variables will be referred to as 'neighbourhood immigration' and 'neighbourhood education'. A neighbourhood with a measure of $I_{jt} = 0$, would in some sense be 'non-segregated'.[15] $I_{jt} > 0$ implies that the neighbourhood has, e.g., a larger than average proportion of immigrants and could be considered segregated. A negative measure would also convey segregation, but in the other direction. The neighbourhood immigration measure will in the analysis be interacted with individual immigrant background. The effect of living in a neighbourhood with low or high immigration may depend on whether oneself has immigrant background, i.e. whether the individual is living among peers.

In the primary sample the average individual resides in a neighbourhood where the proportion with immigrant background is six percentage units less (-0.06) than the proportion of Stockholm, which is around 0.24. The the average neighbourhood has 12 percentage units lower proportion than Stockholm. Moreover, there is a wide spread in the data as the immigration level for the neighbourhoods with the smallest proportion of immigrants is 72 percentage units less than the proportion of Stockholm, whereas it is almost 220 percentage units larger for the neighbourhoods highest concentration of immigrants. The variation for the education level, which is based on an proportion of around 0.26, is considerably narrower, -0.68 to 1.15 for individuals and somewhat narrower on neighbourhood basis (-0.65 to 0.99). Moreover, the average individual resides in a somewhat less educated neighbourhood (0.02), with a proportion that is two percentage units higher

[15] There is a myriad of different measures of segregation (see e.g., Massey and Denton, 1988). These measures are not applicable here since they are primarily constructed for comparing different metropolitan areas.

than the proportion of Stockholm. The average neighbourhood is also somewhat less educated (0.02).

Besides entering on their own, the neighbourhood variables are part of interaction terms that are included in the models. The neighbourhoods are measured along two dimensions, education and immigration. The effect of each of these could be quite different depending on the value of the other. For example, living in a poorly educated neighbourhood may have a different effect depending on the level of immigration. To take such possibilities into account an interaction term between neighbourhood variables enters the models.

5 Estimation results

Both the LPM and the MNL estimated without neighbourhood dummies show statistically significant effects of the neighbourhood variables and the difference in expected income on the post compulsory education choice. These estimates are most likely to suffer from selection bias, however. If neighbourhood dummies are included in the models the number of statistically significant variables decreases. Given the short time span used and that we can expect some inertia in neighbourhood change this result is not surprising. Some neighbourhood effects nevertheless survive the inclusion of the dummies. The marginal effect (ME) of one neighbourhood variable will in most cases depend on whether the individual has immigrant background as well as the level of the other neighbourhood variable. The difference in expected income have a positive effect on the probability of applying for a UPP as well as applying for a VP in the MNL. These effects are present regardless of whether the models are estimated with neighbourhood dummies. The inclusion of neighbourhood dummies should mitigate the selection bias. The estimates may still be affected by selection, however. The estimation results will now be discussed in more detail. First the LPM results, followed by the MNL results. To improve readability the number of tables in the text containing estimation results is kept low. Some estimation results, mainly different sensitivity specifications, are therefor deferred to appendix.

Table 3: Linear probability model OLS estimates, without and with neighbourhood dummies. Dependent variable: 1=applied to university preparatory programme and 0=otherwise.[a]

	(1)	(2)
Neighbourhood education[b]	-0.168***	-0.075*
	(0.008)	(0.045)
Neighbourhood immigration[c]		
×1	-0.025***	0.054
	(0.007)	(0.045)
×Immigrant background	0.062***	0.051***
	(0.015)	(0.015)
Neighbourhood education[b] ×Neighbourhood immigraton[c]		
×1	0.125***	0.104*
	(0.013)	(0.056)
×Immigrant background	-0.066***	-0.036
	(0.024)	(0.026)
Difference in expected income	0.019***	0.013**
	(0.005)	(0.005)
Sex (=1 if male)	-0.020***	-0.023***
	(0.005)	(0.005)
Immigrant background (=1 if immigrant background)	0.008	0.003
	(0.006)	(0.006)
GPA	0.34***	0.341***
	(0.002)	(0.002)
Father's income	0.08***	0.055***
	(0.011)	(0.011)
Father's education (University left out)		
Compulsory or less	-0.088***	-0.084***
	(0.005)	(0.005)
Vocational programme	-0.070***	-0.064***
	(0.005)	(0.005)
University preparatory programme	-0.032***	-0.030***
	(0.005)	(0.005)
Mother's income	0.049**	0.045**
	(0.022)	(0.023)
Mother's education (University left out)		
Compulsory or less	-0.050***	-0.048***
	(0.005)	(0.005)
Vocational programme	-0.043***	-0.041***
	(0.004)	(0.004)
University preparatory programme	0.004	0.003
	(0.006)	(0.006)
Neighbourhood dummies		√
No observations	69,588	69,588
R^2	0.326	0.339

[a] Significance levels 0.10, 0.05 and 0.01 are denoted by *, **, and ***, respectively. Robust standard errors in parenthesis. All estimations include a constant and time dummies.
[b] The proportion of neighbours with not more than compulsory education. Normalised using the yearly total average.
[c] The proportion of neighbours with immigrant background. Normalised using the yearly total average.

5.1 Linear probability model

Table 3 displays the OLS estimation results for two different specifications of the LPM.[16] Specification (1) includes all the explanatory variables but no neighbourhood dummies. This specification has a fairly high degree of explanatory power with a R^2 of 0.326. Half of this power comes from the GPA, however.[17] Moreover, all the included variables, save immigrant background and mother's education UPP, are statistically significant. The

neighbourhood variables may, however, suffer from bias due to the selection in to neighbourhoods. Before going deeper into the estimates this possibility is evaluated.

One simple strategy to gauge the problem is to check the selection on observable variables. If the estimates are sensitive to the exclusion of variables that can be expected to influence both the neighbourhood and post compulsory education choice it is likely that unobserved characteristics also play an important role. Here the parental variables and GPA are such variables. If these variables are removed from the estimation of the model the estimated coefficients increase.[18] This indicates selection on the parental variables and GPA. If this is the case, it is also likely that there are selection on unobservable variables. In addition, it may be noted, that the direction is, unfortunately, away from zero, i.e. any effects are overstated.

Another strategy to evaluate the possible bias, following Glaeser (1996), is to re-estimate the model using a sub sample of recent movers. Recent movers have made their selection of neighbourhood later and should, if there is a selection issues, exhibit a stronger selection effect. If neighbourhood characteristics are more important for this group it is suggestive of results driven by selection. If the neighbourhood characteristics are less important for this group it may be an indication that neighbourhood characteristics do affect the choice. Here the case is the latter. Estimation of the model with only recent movers gives smaller and fewer statistically significant coefficients.[19]

Taken together the two results do not give any clear picture of the selection issue. Even if they would, they only facilitate an evaluation of the problem. One possibility to mitigate any selection bias is to include neighbourhood dummies, as in *Table 3* specification (2). The dummies control for everything that is constant over time in each neighbourhood,

[16] A major drawback with the LPM estimated with OLS is predicted probabilities outside the unit interval. This may lead to biased and inconsistent estimates. A remedy to this problem is to use a sequential least square (SLS) procedure (Horrace and Oaxaca, 2006). This is an iterative procedure; Starting with the OLS estimates observations with predictions outside the unit interval are trimmed from the data. The model is then re-estimated using the trimmed data and OLS. Based on these estimations the data is trimmed again and the model re-estimated. This is repeated until no predictions are outside the unit interval and the SLS estimates are obtained. Estimating the LPM with SLS generates here, see *Table A-4*, somewhat larger (in absolute terms), estimates. The OLS estimates thus seem to be biased towards zero. Given the direction of the bias this issue is not pressing since OLS simply seems to underestimate any effects.

[17] Excluding GPA from the model reduces R^2 to 0.158, see *Table A-3* specification (2).

[18] See *Table A-3*, columns (1) and (2).

[19] See *Table A-3*, specification (3).

both observable and unobservable variables. To the extent that the selection is based on such factors the neighbourhood dummies will mitigate the bias. Adding the dummies to the LPM renders the proportion with immigrant background (main effect) statistically insignificant. The interaction between the two neighbourhood variables interacted with immigrant background becomes statistically insignificant too. The significance level of some other variables also change although they are still significant at conventional levels. Even when controlling for neighbourhood fixed effects there are effects from our neighbourhood variables on the post primary education choice. The remaining coefficient estimates are similar to the without neighbourhood dummies estimates. The R^2 has increased somewhat, which is not surprising given the number of dummies included in the model.

Recalling that the dependent variable is one if the individual applied for a UPP and zero otherwise (discontinue/VP) the coefficients tell how much the probability of applying for a UPP, changes when the variables change, i.e. the ME. For both specifications in *Table 3* the probability of applying for a UPP increases with the income of the parents as well as the GPA. Having a father who is not university educated or a mother with either of the two lowest education categories decreases the the probability, however. These effects are all between three and nine per cent. For example, having a father with vocational education instead of university education would increase the probability of not applying for a UPP from, for example, 0.5 to 0.57.

Turning to the neighbourhood variables the results differ, as noted above, between the two specifications both in terms of magnitude and statistical significance. Since the estimates without neighbourhood dummies are likely to be biased the more conservative estimates with dummies are perhaps to prefer even though they may also be biased. Recalling that a neighbourhood variable of zero implies that a neighbourhood has the same proportion of immigrants or individuals with low educational level as the average proportion, the interpretation of the coefficients is straightforward.

For an individual living in a neighbourhood with the same level of immigrants as Stockholm the ME with respect to neighbourhood education is -0.075. An increase in the proportion of individuals with low education of the same size as the over all proportion

will hence reduce the probability to apply for a UPP with 7,5 percentage units. This effect is, however, increasing in neighbourhood immigration.[20] The marginal effect will be negative for neighbourhoods that have neighbourhood immigration less than 0.72, i.e. 72 per cent higher proportion of immigrants than Stockholm. Thus, increasing the proportion of poorly educated individuals in a neighbourhood will generally decrease the probability of applying for a UPP. In a neighbourhood with high proportion of immigrants the effect will be positive. For example, an individual living in a neighbourhood with a neighbourhood immigration of one, i.e. twice the proportion of Stockholm, the marginal effect would be 0.04.

The ME with respect to neighbourhood immigration depends both on neighbourhood education and whether the individual has immigrant background.[21] An individual with immigrant background has a ME of five per cent if living in a neighbourhood with the same proportion individuals with low education as Stockholm. This effect is increasing in neighbourhood education and will be negative for education values lower than -0.49, i.e. around half the proportion of Stockholm. For individuals with Swedish background an increase in the proportion of immigrants will decrease the the probability to apply for a UPP if they live in a neighbourhood with a lower proportion of individuals with low education compared to the average of Stockholm. If they live in a neighbourhood with higher proportion the effect will be positive.

The ME of difference in expected income is about one per cent. An increase in the expected income for an alternative would increase the probability for that alternative to be chosen. The change needed to affect the probabilities in any significant way is, however, quite large. For example, if the expected income for UPP increases with one million SEK (around one standard deviation), or if the yearly constant income increases with 32,000 SEK, the probability to apply for a UPP increases with one percentage unit. The effect is, although statistically significant, quite small.

[20] $\frac{\partial P(\text{UPP})}{\partial \text{Neigh. ecu.}} = -0.075 + 0.104 \times \text{Neigh. imm.}$.
[21] $\frac{\partial P(\text{UPP})}{\partial \text{Neigh. imm.}} = 0.051 \times \text{Imm. backg.} + 0.104 \times \text{Neigh. edu.}$

5.2 Multinomial logit model

Table 4 displays the MNL estimation results with all explanatory variables without and with neighbourhood dummies.[22] There are now two equations in the model. Comparing VP and UPP to discontinue, the base alternative. VP and UPP could be compared by subtracting one from the other. For reasons of identification variables that are constant over alternatives, which includes all variables but the difference in expected income, have alternative specific coefficients (see e.g Greene, 1997).

In specification (1), which is without neighbourhood dummies, both neighbourhood variables and their various interactions are statistically significant in all cases but one. In the case of UPP the interaction between immigrant background and the interaction between the neighbourhood variables is statistically insignificant. There thus seem to be neighbourhood effects on the post primary education choice. The difference in expected income is statistically significant to. The estimation results for the other variables are mainly statistically significant. Two notable exceptions are the parents' income in the VP case and mother's education in the UPP case. Moreover, the GPA coefficient is negative in the VP case, which is counter intuitive. This result may be driven by individuals that postpone education. This could also explain the positive coefficients of the parents' education in the VP equation. There may be several other explanations. Here these variables are included to control for selection and will not be discussed further. The MNL estimates may even with the included control variables be biased due to the selection into neighbourhoods. Applying the same strategies as for the LPM reveals that this may indeed be the case. Removing the parental variables and GPA causes the estimates to increase in magnitude.[23] Estimating the model with only recent movers increases the estimates too.[24]

Table 4 specification (2) adds neighbourhood dummies to the model. As in the LPM case, some of the neighbourhood variable entries loose their statistical significance. Here the loss of significance is limited to the interaction between the neighbourhood variables.

[22] The MNL with neighbourhood dummies has been estimated with different requirement on the minimum number of observations per neighbourhood. As the size increased the estimates where stable suggesting that the number of observations to estimate each neighbourhood effect is sufficient.

[23] See *Table A-5*, specification (1) and (2).

[24] See *Table A-5*, specification (3).

Table 4: Multinomial logit estimation results, without and with neighbourhood dummies. Dependent variable: post compulsory education choice – discontinue, vocational programme (VP) and university preparatory programme (UPP). Discontinue is used as base alternative.[a]

	(1)		(2)	
	VP vs Discontinue	UPP vs Discontinue	VP vs Discontinue	UPP vs Discontinue
Neighbourhood education[b]	1.159***	-0.237**	4.275***	3.218***
	(0.108)	(0.105)	(0.699)	(0.681)
Neighbourhood immigration[c]				
×1	-1.616***	-1.728***	-2.354***	-1.869**
	(0.106)	(0.103)	(0.784)	(0.77)
×Immigrant background	-1.654***	-1.117***	-1.894***	-1.454***
	(0.267)	(0.26)	(0.325)	(0.32)
Neighbourhood education [b] ×Neighbourhood immigration[c]				
×1	1.153***	2.121***	-2.282**	-0.947
	(0.172)	(0.169)	(0.901)	(0.901)
×Immigrant background	0.688*	0.109	0.320	-0.076
	(0.361)	(0.352)	(0.45)	(0.445)
Difference in expected income	14.876***		16.039***	
	(0.171)		(0.183)	
Sex (=1 if male)	1.067***	1.088***	1.208***	1.234***
	(0.046)	(0.045)	(0.053)	(0.052)
Immigrant background (=1 if immigrant background)	-0.078	0.152*	0.145	0.362***
	(0.085)	(0.081)	(0.102)	(0.099)
GPA	-1.134***	1.460***	-1.206***	1.498***
	(0.035)	(0.035)	(0.041)	(0.04)
Father's income	-0.212	0.766***	-0.663***	0.201
	(0.184)	(0.161)	(0.213)	(0.193)
Father's education (University left out)				
Compulsory or less	0.299***	-0.37***	0.311***	-0.337***
	(0.068)	(0.064)	(0.076)	(0.071)
Vocational programme	0.285***	-0.255***	0.312***	-0.192***
	(0.067)	(0.062)	(0.074)	(0.07)
University preparatory programme	0.107	-0.169***	0.057	-0.206***
	(0.066)	(0.059)	(0.072)	(0.066)
Mother's income	0.183	0.819**	-0.158	0.422
	(0.354)	(0.329)	(0.401)	(0.379)
Mother's education (University left out)				
Compulsory or less	0.437***	-0.002	0.397***	-0.03
	(0.066)	(0.062)	(0.073)	(0.07)
Vocational programme	0.398***	-0.007	0.377***	-0.004
	(0.059)	(0.055)	(0.065)	(0.061)
University preparatory programme	-0.024	0.008	-0.034	-0.007
	(0.081)	(0.073)	(0.089)	(0.081)
Neighbourhood dummies			√	
No observations	69,200		69,200	
Log likelihood	-27,805.45		-2565-3.14	

[a] Significance levels -0.10, -0.05 and -0.01 are denoted by *, **, and ***, respectively. All estimations include a constant and time dummies. For reasons of identification variables that do not vary over alternatives have alternative specific coefficients.
[b] The proportion of neighbours with not more than compulsory education. Normalised using the yearly total average.
[c] The proportion of neighbours with immigrant background. Normalised using the yearly total average.

In the VP equation the interaction with immigrant background becomes insignificant and in the UPP equation the main effect looses its significance. The estimated coefficients increase somewhat in magnitude except in the case of neighbourhood education. In the VP equation the coefficient increases threefold. In the UPP equation there is a sign change in addition to a large increase in the magnitude of the coefficient. Again, these estimates may also be biased, but probably less likely than the without neighbourhood dummies estimates. For the rest of the variables there are some small changes in the estimates and some significance levels. Most notable is the change of equation by father's income from UPP to VP.

The MNL is not as straightforward to interpret as the LPM. The marginal effects will not only depend on the values of all other variables, but also the probability of choosing the alternatives and may in fact be of a different sign than the coefficients (see e.g., Wooldridge, 2002). *Table 5* displays the average marginal effect (AME) for the variables of greatest interest based on the MNL with neighbourhood dummies.[25] The AME:s express the average change in the probability of choosing the different alternatives, i.e. VP and UPP. In addition, the table displays the statistical significance of the underlying individual ME:s within brackets. This is expressed as the percentage of ME:s that are; negative, insignificant and positive at five per cent significance level. Moreover, the AME:s are displayed for four groups based on sex and immigrant background.

Only nine out of 24 AME:s are statistically significant – neighbourhood immigration in the case of females with immigrant background and all difference in expected income AME:s. For all variables the individual ME:s may give some information of possible effects, however. Even though the AME:s are insignificant for neighbourhood education a majority of the the individual ME:s is positive in the VP case. An increase in the neighbourhood proportion of individuals with low education will thus increase the probability to apply for a VP for a majority of the individuals. Furthermore, no individuals have negative ME:s with respect to neighbourhood education. In the UPP case the distribution of individual ME:s are much more even and not much can be deduced from the estimations.

[25] Based on *Table 4* specification (2) ME:s have been calculated for all individuals. The AME:s are the average of these.

Table 5: Average marginal effects from multinomial logit with neighbourhood dummies.[a]

$\frac{\partial P(\text{VP})}{\partial x}$	Female		Male	
	Swedish backg.	Imm. backg.	Swedish backg.	Imm. backg.
Neighbourhood education[b]	0.0052	0.0041	0.0033	0.0025
	(0.0082)	(0.0084)	(0.0058)	(0.0071)
	[0,11,89]	[0,33,67]	[0,12,88]	[0,34,66]
Neighbourhood immigration[c]	-0.0059	-0.0084	-0.0038	-0.0056
	(0.0130)	(0.0123)	(0.0093)	(0.0094)
	[38,62,0]	[67,33,0]	[36,64,0]	[65,35,0]
Difference in expected income	0.0061***	0.0077***	0.0031***	0.0043***
	(0.0006)	(0.0006)	(0.0004)	(0.0005)
	[25,0,75]	[21,0,79]	[8,0,92]	[13,0,87]
$\frac{\partial P(\text{UPP})}{\partial x}$	Female		Male	
	Swedish backg.	Imm. backg.	Swedish backg.	Imm. backg.
Neighbourhood education[b]	0.0852	0.1590	0.0436	0.0949
	(0.0586)	(0.1036)	(0.0367)	(0.0792)
	[62,12,26]	[40,31,29]	[45,21,34]	[30,34,63]
Neighbourhood immigration[c]	-0.0520	-0.1276*	-0.0268	-0.0767
	(0.0561)	(0.0698)	(0.0364)	(0.0517)
	[22,61,17]	[33,23,44]	[25,59,16]	[38,24,38]
Difference in expected income	0.4506***	0.5854***	0.2343***	0.3533***
	(0.0127)	(0.0147)	(0.0082)	(0.0107)
	[25,0,75]	[21,0,79]	[8,0,92]	[13,0,87]

[a] Based on the estimates in *Table 4*. Significance levels 0.10, 0.05 and 0.01 are denoted by *, **, and ***, respectively. Standard errors in parenthesis. In brackets, percentage of individual marginal effects that are statistically significant, [-,0,+] at 0.05 significance level.
[b] The proportion of neighbours with not more than compulsory education. Normalised using the yearly total average.
[c] The proportion of neighbours with immigrant background. Normalised using the yearly total average.

A similar pattern occur for neighbourhood immigration, although with a different sign. No individuals have positive ME:s in the case of VP and a majority of individuals with immigrant background has negative ME:s. Thus an increase in the proportion of immigrants will reduce the probability to apply for a VP in general. For individuals with Swedish background about one third has negative ME:s. For females with immigrant background the AME is negative and statistically significant in the UPP case. The individual ME:s are rather evenly distributed, however. For the other groups the ME:s are evenly distributed too. Individuals with Swedish background seem to be less affected as a majority has insignificant ME:s.

The difference in expected income has positive and statistically significant AME:s all over. The effect is considerably smaller for VP, at most a half per cent, compared to UPP, at most 58 per cent. An one million SEK increase in the expected income will thus increase the probability to choose a VP with about a half percentage unit and 50 percentage units

to choose a UPP, respectively. For VP the effect is hence very small, in par with the linear probability model. For UPP the effect is, however, sizeable and even small changes to expected incomes may matter. A majority of the individuals, around three fourths of the females and around 90 per cent of the males, have positive ME:s. Moreover, females have twice the AME of males and individuals with immigrant background have larger AME then individuals with Swedish background. An increase in the expected income of an alternative increases (in average) the probability that this alternative is chosen, which is of course what can be expected.

6 Concluding remarks

The estimation results show that neighbourhoods affect youths' post compulsory education choice. This is true after controlling for individual and parental characteristics as well as neighbourhood fixed effects. Neighbourhood specific economic incentives and neighbourhood levels of education and immigration affects the probability to make different choices.

For the neighbourhood variables the results differ to some extent depending on how the choice problem is formulated. If the choice is formulated as a binary choice of whether to apply for a university preparatory programme the probability of applying for such a programme is generally decreasing in the neighbourhood proportion of individuals with at most compulsory education. For a neighbourhood with a high proportion of immigrants (somewhat less than twice the proportion of Stockholm) the effect is the opposite. One possible explanation for this is that the linear probability model works best around means. Another, more appealing, explanation is that education may be viewed as a way to escape neighbourhoods with high proportion of immigrants and individuals with low education.

The effect of immigrant concentration on the binary choice depends on whether the youth has immigrant background. For a youth with Swedish background living in a neighbourhood with higher proportion of individuals with low education that the average in Stockholm, the proportion of immigrants has a positive effect on the probability to apply

for a university preparatory programme. For more educated neighbourhoods the probability to apply for a university preparatory programme decreases when the proportion of immigrants increases. If the youth has immigrant background the effect of the proportion of immigrants on the probability to apply for a university preparatory programme is generally positive. This is in contrast to previous research on Swedish data. One possible explanation is that there is a discrepancy between applications (this paper) and outcomes (previous research). If this is the case is an interesting future research topic. Yet another interesting topic for the future is if there are differences between first and second generation immigrants. Here both generations have been included in immigrant background.

If the choice is formulated as a multinomial choice between not applying, applying for a vocational programme or applying for a university preparatory programme, the effect of the neighbourhood variables are less clear. In most cases the average marginal effects are statistically insignificant. The individual marginal effects do however point in certain directions. For a majority of the individuals the probability to apply for a vocational programme increases if the proportion of individuals with low education increases. The marginal effects on applying for a university preparatory programme are much more evenly distributed and not much can be deduced from the estimations. For the proportion of immigrants the effects on the probability to apply for a vocational programme are in general negative. An increase in the proportion of immigrants will hence reduce the probability to apply for a vocational programme. Again not much can be deducted regarding university preparatory programme applications. These results are in line with the linear probability model, however. The effects on vocational programme are the opposite of the general effects on university preparatory programme in the linear probability model.

Regardless of whether the problem is specified as a binary choice or as a multinomial choice economic incentives, measured as expected incomes, have an effect on the choice. An increase in the expected income, here estimated with neighbourhood wage regressions, of an alternative increases the probability that this alternative is chosen. This is a result consistent with human capital theory. The magnitude of the effects differ with the problem formulation, however, and is very small for some cases. One interesting results with respect to expected income is that females have considerably larger marginal effects

in the multinomial specification. Moreover, individuals with immigrant background also show larger marginal effects. The estimation of the expected incomes may however be challenged. It is of course questionable if youths collect information about different educational payoffs from their neighbours. There are other possible reference groups, such as parents and friends' parents. The path taken in this paper does however require a certain group size in order for the expected incomes to be estimable. As usual data effectively set constraints. Alternative approaches are left to future research. The difference in marginal effects between different groups is also an interesting future research topics.

References

Aaronson, D. (1998), "Using sibling data to estimate the impact of neighborhoods on children's educational outcome," *The Journal of Human Resources*, vol. 33, no. 4, pp. 915–974.

Ainsworth, J. W. (2002), "Why does it take a village? The mediation of neighbourhood effects on educational achievement," *Soical Forces*, vol. 81, no. 1, pp. 117–152.

Andersson, E. and S. V. Subramanian (2006), "Explorations of neighbourhood and educational outcomes for young Swedes," *Urban Studies*, vol. 43, no. 11, pp. 2013–2025.

Andersson, R. (2000), "Etnisk och socioekonomisk segregation i Sverige 1990–1998," in Fritzell, J. (ed.) "Välfärdens förutsättningar. Arbetsmarknad, demografi och segregation.", Nordsteds, Stockholm, pp. 223–266.

Atkinson, R. and K. Kintera (2001), "Disentangeling area effects: Evidence from deprived and non-deprived neighbourhoods," *Urban Studies*, vol. 48, no. 12, pp. 2277–2298.

Becker, G. S. (1964), *Human capital*, National Bureau of Economic Research.

Böhlmark, A. (2008), "Age at immigration and school preformance: A siblings analysis using register data," *Labour Economics*, vol. 15, pp. 1366–1387.

Björklund, A., M. Clark, P.-A. Edin, P. Fredriksson and A. Krueger (2005), *The Market Comes to Education in Sweden – An Evaluation of Sweden's Surprising School Reforms*, Russel Sage Foundation.

Borjas, G. J. (1995), "Ethnicity, neighborhoods, and human-capital externalities," *The American Economic Review*, vol. 85, no. 3, pp. 365–390.

Brännström, L. (2005), "Does neighbourhood origin matter? A longitudinal multilevel assessment of neighbourhood effects on income and reciept of social assistance in a Stockholm birth cohort," *Housing, Theory and Society*, vol. 22, no. 4, pp. 169–195.

Brännström, L. (2008), "Making their mark: The effects of neighbourhood and upper secondary school on educational achievment," *European Sociological Review*, vol. 24, no. 4, pp. 463–478.

Cardak, B. A. and J. T. McDonald (2004), "Neighbourhood effects, preference heterogeneity and immigrant educational attainment," *Applied Economics*, vol. 36, pp. 559–572.

Comay, Y., A. Melnik and M. A. Pollatschek (1973), "The option value of education and the optimal path for investment in human capital," *International Economic Review*, vol. 14, no. 2, pp. 421–435.

Crane, J. (1991), "The epidemic theory of ghettos and neighborhood effects on dropping out and teenage childbearing," *The American Journal of Sociology*, vol. 95, no. 5, pp. 1226–1259.

Cutler, D. M. and E. L. Glaeser (1997), "Are ghettos good or bad?" *The Quarterly Journal of Economics*, vol. 112, no. 3, pp. 827–872.

Durlauf, S. N. (2004), "Neighborhood effects," in Henderson, J. V. and J.-F. Thisse (eds.) "Handbook of regional and urban economics," vol. 4, chap. 50, North-Holland Press, Amsterdam, pp. 2173–2242.

Edin, P.-A., P. Fredriksson and O. Åslund (2003), "Ethnic enclaves and the economic success of immigrants: Evidence from a natural experiment," *The Quaterly Journal of Economics*, vol. 118, no. 1, pp. 329–357.

Ellen, I. G. and M. A. Turner (1997), "Does neighborhood matter? Assessing recent evidence," *Housing Policy Debate*, vol. 8, no. 4, pp. 833–866.

Freidrichs, J. and J. Blasius (2003), "Social norms in distressed neighbourhoods: Testing the Wilson hypothesis," *Housing Studies*, vol. 6, no. 6, pp. 807–826.

Galster, G. (2007), "Should policy makers strive for neighborhood social mix? An analysis of the Western European evidence base," *Housing Studies*, vol. 22, no. 4, pp. 523–545.

Ginther, D., R. Haveman and B. Wolfe (2000), "Neighborhood attributes as determinants of children's outcomes: How robust are the relationships?" *The Journal of Human Resources*, vol. 35, no. 4, pp. 603–642.

Glaeser, E. (1996), "Discussion - comment on Katherine O'Regan and John Quigley, in this issue - Special issue: Earnings inequality," *New England Economic Review: Federal Reserve Bank of Boston*, vol. May/June, pp. 58–64.

Greene, W. (1997), *Econometric Analysis*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, third ed.

Greene, W. (2002), "The behavior of the fixed effects estimator in nonlinear models," Working Papers 02-05, New York University, Leonard N. Stern School of Business, Department of Economics.

Grönqvist, H. (2006), "Ethnic enclaves and the attainments of immigrant children," *European Sociological Review*, vol. 22, pp. 369–382.

Haveman, R. and B. Wolfe (1995), "The determinants of children's attainment: A review of methods and findings," *Journal of Economic Literature*, vol. 33, no. 4, pp. 1829–1878.

Horrace, W. C. and R. L. Oaxaca (2006), "Results on the bias and inconsistency of ordinary least squares for the linear probability model," *Economic Letters*, vol. 90, no. 3, pp. 321–327.

Katz, E. (2001), "Bias in conditional and unconditional fixed effects logit estimation," *Political Analysis*, vol. 9, no. 4, pp. 379–384.

Landell, E., O. Gustafsson and D. Grannas (2000), "Utbildningens omvägar – En ESO-rapport om kvalitet och effektivitet i svensk utbildning," Ds 2000:58, Finansdepartementet.

Manski, C. F. (1993a), "Adolescent econometricians: How do youth infer the returns to schooling?" in Clotfelter, C. T. and M. Rothschild (eds.) "Studies of supply and demand in higher education," The University of Chicago Press, Chicago, pp. 43–57.

Manski, C. F. (1993b), "Identification of endogenous social effects: The reflection problem," *The Review of Economic Study*, vol. 60, no. 3, pp. 531–542.

Massey, D. S. and N. A. Denton (1988), "The dimensions of residential segregation," *Social Forces*, vol. 67, no. 2, pp. 281–315.

Nordin, M. (2006), "Ethnic segregation and educational attainment in Sweden," *mimeo*, Department of Economics, Lund University.

Oreopoulos, P. (2003), "The long-run consequences of living in a poor neighborhood," *The Quarterly Journal of Economics*, vol. 118, no. 4, pp. 1533–1575.

Skolverket (2008), "The Swedish school system," http://www.skolverket.se/sb/d/354. 01/11/2008.

Szulkin, R. and J. O. Jonsson (2007), "Ethnic segregation and educational outcomes in Swedish comprehensive schools," Working paper 2007:2, SULCIS Stockholm University.

Vartanian, T. P. and P. M. Gleason (1999), "Do neighborhood conditions affect high school dropout and collage graduation rates?" *Journal of Socio-Economics*, vol. 28, pp. 21–41.

Wilson, K. (2001), "The determinants of educational attainment: Modeling and estimating the human capital model and education production functions," *Southern Economic Journal*, vol. 67, no. 3, pp. 518–551.

Wilson, K., B. Wolfe and R. Haveman (2005), "The role of expectations in adolescent schooling choices: Do youths respond to economic incentives?" *Economic Inquiry*, vol. 43, no. 3, pp. 467–492.

Wolfe, B., R. Haveman, K. Pence and J. A. Schwabish (2007), "Do youth nonmartial childbearing choices reflect income and relationship expectations?" *Journal of Population Economics*, vol. 20, no. 1, pp. 73–100.

Wooldridge, J. M. (2002), *Econometric analysis of cross section and panel data*, MIT Press, Cambridge, MA.

Appendix

Table A-1: Variable definitions

Variable	Definition
Dependent variable	
Binary choice	$= \begin{cases} 0 \text{ if first choice is a university preparatory programme (base alternative)} \\ 1 \text{ if discontinue, i.e. did not apply or if first choice is a vocational programme} \end{cases}$
Multinomial choice	$= \begin{cases} 0 \text{ if discontinue, i.e. did not apply} \\ 1 \text{ if first choice is a vocational programme} \\ 2 \text{ if first choice is a university preparatory programme (base alternative)} \end{cases}$
Individual variables	
Sex	1 if male, 0 if female
Immigrant background	1 if oneself or both parents are born outside Sweden
GPA	GPA when finishing grade 9 (possible grades 1. . . 5)
Expected income	the discounted sum of predictions from neighbourhood and year specific tobit regressions in 1,000,000 SEK
Difference in expected income	the difference in expected income between an alternative and the base alternative
Parental variables	
Father's (mother's) income	father's (mother's) average labour income last two years in 1,000,000 SEK
Father's (mother's) education	
Compulsory or less	1 if compulsory education or less
Vocational	1 if vocational programme
University preparatory	1 if university preparatory programme
University	reference group, more than upper secondary education
Neighbourhood variables	
Neighbourhood education	Proportion of neighbourhood population with a given year with at most compulsory education. Normalised using the overall proportion that year.[a]
Neighbourhood immigration	Proportion of neighbourhood population a given year with immigrant background. Normalised using the overall proportion that year.[a]

[a] The normalisation is made in the following way. $I_{jt} = \frac{x_{jt} - \bar{x}_t}{\bar{x}_t}$ where x_{jt} is the proportion in neighbourhood j year t and $\bar{x}_t$ is the proportion in the secondary data set year t.

Table A-2: Neighbourhood wage equation Tobit estimates for three neighbourhoods (SAMS) and years. Dependent variable wage income in 100 SEK.[a]

	SAMS 1170022 Year 1989	SAMS 1250015 Year 1990	SAMS 1910001 Year 1991
Sex	1119.585 (961.989)	304.345 (740.019)	1297.227** (603.093)
Immigrant background	-840.854 (1753.349)	-1472.433 (1083.654)	30.758 (757.713)
Compulsory or less	3010.075** (1266.227)	1299.088 (899.898)	1544.101** (708.952)
VP	2723.208** (1320.656)	506.071 (1181.805)	802.717 (907.590)
UPP	1491.363 (1572.229)	-1157.221 (1350.609)	1587.320 (1038.896)
Age	302.245*** (56.396)	74.750* (41.651)	182.689*** (31.274)
Age^2	-3.880*** (0.79)	-0.777 (0.58)	-2.165*** (0.406)
Sex × Immigrant background	132.731 (257.123)	94.292 (218.154)	11.239 (128.406)
Sex × Compulsory or less	6.705 (251.548)	-106.206 (197.036)	-160.491 (131.390)
Sex × VP	343.645 (288.077)	-244.238 (247.036)	-88.401 (176.432)
Sex × UPP	-414.011* (244.905)	-715.165*** (220.576)	-466.186*** (160.959)
Immigrant background × Compulsory or less	-209.057 (347.169)	-179.641 (300.063)	408.844*** (155.118)
Immigrant background × VP	-399.583 (384.911)	20.207 (354.688)	86.011 (203.853)
Immigrant background × UPP	-23.778 (324.113)	-422.084 (325.953)	197.860 (193.400)
Age × Sex	-108.019* (56.293)	-26.690 (44.741)	-85.161** (34.353)
Age × Immigrant background	46.223 (94.180)	94.218 (66.556)	-26.937 (41.509)
Age × Compulsory or less	-181.893** (77.501)	-48.638 (54.000)	-64.733 (40.191)
Age × VP	-193.532** (80.780)	-18.263 (69.699)	-28.059 (52.370)
Age × UPP	-86.339 (85.222)	89.322 (73.661)	-60.668 (56.501)
Age^2 × Sex	1.453* (0.777)	0.294 (0.609)	1.030** (0.449)
Age^2 × Immigrant background	-0.684 (1.237)	-1.358 (0.931)	0.265 (0.535)
Age^2 × Compulsory or less	2.480** (1.098)	0.644 (0.748)	0.766 (0.525)
Age^2 × VP	2.952** (1.147)	0.577 (0.959)	0.634 (0.697)
Age^2 × UPP	1.534 (1.112)	-0.708 (0.956)	1.005 (0.725)
Number of observations	282	565	1,059

[a] Significance levels 0.10, 0.05 and 0.01 are denoted by *, **, and ***, respectively. All estimations include a constant.

Table A-3: Further linear probability model OLS estimation results, without neighbourhood dummies. Dependent variable: 1=applied to university preparatory programme and 0=otherwise.[a]

	(1)	(2)	(3)
Neighbourhood education[b]	-0.226***	-0.203***	-0.097***
	(0.007)	(0.008)	(0.026)
Neighbourhood immigration[c]			
×1	-0.044***	-0.034***	-0.024
	(0.007)	(0.008)	(0.028)
×Immigrant background	0.055***	0.114***	0.024
	(0.015)	(0.017)	(0.051)
Neighbourhood education[b] ×Neighbourhood immigration[c]			
×1	0.155***	0.128***	0.1**
	(0.013)	(0.014)	(0.05)
×Immigrant background	-0.065***	-0.123***	-0.073
	(0.024)	(0.027)	(0.091)
Difference in expected income	0.021***	0.017***	0.006
	(0.005)	(0.005)	(0.016)
Sex (=1 if male)	-0.025***	0.054***	-0.009
	(0.005)	(0.005)	(0.015)
Immigrant background (=1 if immigrant background)	0.009	-0.001	0.023
	(0.006)	(0.006)	(0.018)
GPA	0.368***		0.321***
	(0.002)		(0.009)
Parental variables		√	√
Sample	full	full	movers[d]
No observations	69,588	69,588	5,311
R^2	0.316	0.158	0.304

[a] Significance levels 0.10, 0.05 and 0.01 are denoted by *, **, and ***, respectively. Robust standard errors in parenthesis. All estimations include a constant and time dummies.
[b] The proportion of neighbours with not more than compulsory education. Normalised using the yearly total average.
[c] The proportion of neighbours with immigrant background. Normalised using the yearly total average.
[d] Movers are defined as individuals that have lived in a neighbourhood at most two years at the time of the observation.

Table A-4: Linear probability model sequential least squares estimates, without and with neighbourhood dummies.[a] Dependent variable: 1=applied to university preparatory programme and 0=otherwise.[b]

	(1)	(2)
Neighbourhood education[c]	-0.227***	-0.080
	(0.009)	(0.054)
Neighbourhood immigration[d]		
×1	-0.026***	0.024
	(0.008)	(0.054)
×Immigrant background	0.086***	0.073***
	(0.017)	(0.018)
Neighbourhood education[c] ×Neighbourhood immigration[d]		
×1	0.153***	0.166**
	(0.015)	(0.071)
×Immigrant background	-0.088***	-0.046
	(0.029)	(0.032)
Difference in expected income	0.024***	0.012*
	(0.006)	(0.007)
Sex (=1 if male)	-0.029***	-0.038***
	(0.005)	(0.006)
Immigrant background (=1 if immigrant background)	0.013**	0.007
	(0.007)	(0.007)
GPA	0.438***	0.451***
	(0.003)	(0.004)
Father's income	0.122***	0.102***
	(0.015)	(0.016)
Father's education (University left out)		
Compulsory or less	-0.096***	-0.092***
	(0.006)	(0.006)
Vocational programme	-0.078***	-0.071***
	(0.006)	(0.006)
University preparatory programme	-0.041***	-0.039***
	(0.005)	(0.005)
Mother's income	0.085***	0.08***
	(0.028)	(0.029)
Mother's education (University left out)		
Compulsory or less	-0.051***	-0.048***
	(0.006)	(0.006)
Vocational programme	-0.047***	-0.043***
	(0.005)	(0.005)
University preparatory programme	0.002	0.0002
	(0.007)	(0.007)
Neighbourhood dummies		√
No observations	57,352	55,877
R^2	0.258	0.264

[a] The sequential least square procedure contains the following steps; Estimate the LPM with OLS. Remove observations with prediction outside the unit interval. Repeat these two steps until no observations are removed. See Horrace and Oaxaca (2006) for more information.

[b] Significance levels 0.10, 0.05 and 0.01 are denoted by *, **, and ***, respectively. Robust standard errors in parenthesis. All estimations include a constant and time dummies.

[c] The proportion of neighbours with not more than compulsory education. Normalised using the yearly total average.

[d] The proportion of neighbours with immigrant background. Normalised using the yearly total average.

Table A-5: Further multinomial logit estimation results, without neighbourhood dummies. Dependent variable: post compulsory education choice – discontinue, vocational programme (VP) and university preparatory programme (UPP). Discontinue is used as base alternative.[a]

	(1)		(2)		(3)	
	VP vs Discontinue	UPP vs Discontinue	VP vs Discontinue	UPP vs Discontinue	VP vs Discontinue	UPP vs Discontinue
Neighbourhood education[b]	1.372*** (0.104)	-0.462*** (0.102)	1.231*** (0.106)	-0.131 (0.102)	1.600*** (0.39)	0.927** (0.392)
Neighbourhood immigration[c]						
×1	-1.549*** (0.105)	-1.817*** (0.103)	-1.610*** (0.105)	-1.714*** (0.101)	-2.365*** (0.432)	-2.617*** (0.434)
×Immigrant background	-1.621*** (0.265)	-1.148*** (0.259)	-1.689*** (0.264)	-1.031*** (0.255)	1.001 (1.004)	1.603* (0.971)
Neighbourhood education[b] ×Neighbourhood immigration[c]						
×1	1.042*** (0.171)	2.232*** (0.168)	1.281*** (0.169)	2.033*** (0.163)	1.951*** (0.614)	2.884*** (0.632)
×Immigrant background	0.665* (0.359)	0.119 (0.351)	0.623* (0.355)	-0.081 (0.343)	-2.899 (1.801)	-3.384* (1.756)
Difference in expected income	14.899*** (0.171)		14.834*** (0.17)		16.118*** (0.694)	
Sex (=1 if male)	1.099*** (0.046)	1.062*** (0.045)	0.996*** (0.044)	1.377*** (0.042)	0.891*** (0.17)	0.889*** (0.169)
Immigrant background (=1 if immigrant background)	-0.105 (0.084)	0.146* (0.081)	-0.055 (0.083)	0.083 (0.079)	0.075 (0.29)	0.600** (0.278)
GPA	-1.227*** (0.034)	1.565*** (0.034)			-0.853*** (0.13)	1.369*** (0.131)
Parental variables			√		√	
Sample	full		full		movers[d]	
No observations	69,200		69,200		5,280	
Log likelihood	-28,348.27		-34,513.08		-2,184.557	

[a] Significance levels -0.10, -0.05 and -0.01 are denoted by *, **, and ***, respectively. All estimations include a constant and time dummies. For reasons of identification variables that do not vary over alternatives have alternative specific coefficients.

[b] The proportion of neighbours with not more than compulsory education. Normalised using the yearly total average.

[c] The proportion of neighbours with immigrant background. Normalised using the yearly total average.

Publication series published by the Institute for Labour Market Policy Evaluation (IFAU) – latest issues

Rapporter/Reports

2009:1 Hartman Laura, Per Johansson, Staffan Khan and Erica Lindahl, ”Uppföljning och utvärdering av Sjukvårdsmiljarden”

2009:2 Chirico Gabriella and Martin Nilsson ”Samverkan för att minska sjukskrivningar – en studie av åtgärder inom Sjukvårdsmiljarden”

2009:3 Rantakeisu Ulla ”Klass, kön och platsanvisning. Om ungdomars och arbetsförmedlares möte på arbetsförmedlingen”

2009:4 Dahlberg Matz, Karin Edmark, Jörgen Hansen and Eva Mörk ”Fattigdom i folkhemmet – från socialbidrag till självförsörjning”

2009:5 Pettersson-Lidbom Per and Peter Skogman Thoursie ”Kan täta födelseintervaller mellan syskon försämra deras chanser till utbildning?

2009:6 Grönqvist Hans ”Effekter av att subventionera p-piller för tonåringar på barnafödande, utbildning och arbetsmarknad”

2009:7 Hall Caroline "Förlängningen av yrkesutbildningarna på gymnasiet: effekter på utbildningsavhopp, utbildningsnivå och inkomster"

2009:8 Gartell Marie ”Har arbetslöshet i samband med examen från högskolan långsiktiga effekter?”

2009:9 Kennerberg Louise ”Hur försörjer sig nyanlända invandrare som inte deltar i sfi?”

2009:10 Lindvall Lars ”Bostadsområde, ekonomiska incitament och gymnasieval”

Working papers

2009:1 Crépon Bruno, Marc Ferracci, Grégory Jolivet and Gerard J. van den Berg “Active labor market policy effects in a dynamic setting”

2009:2 Hesselius Patrik, Per Johansson and Peter Nilsson “Sick of your colleagues’ absence?”

2009:3 Engström Per, Patrik Hesselius and Bertil Holmlund “Vacancy referrals, job search and the duration of unemployment: a randomized experiment”

2009:4 Horny Guillaume, Rute Mendes and Gerard J. van den Berg ”Job durations with worker and firm specific effects: MCMC estimation with longitudinal employer-employee data”

2009:5 Bergemann Annette and Regina T. Riphahn “Female labor supply and parental leave benefits – the causal effect of paying higher transfers for a shorter period of time”

2009:6 Pekkarinen Tuomas, Roope Uusitalo and Sari Kerr “School tracking and development of cognitive skills”

2009:7 Pettersson-Lidbom Per and Peter Skogman Thoursie “Does child spacing affect childrens’s outcomes? Evidence from a Swedish reform”

2009:8 Grönqvist Hans “Putting teenagers on the pill: the consequences of subsidized contraception”

2009:9 Hall Caroline "Does making upper secondary school more comprehensive affect dropout rates, educational attainment and earnings? Evidence from a Swedish pilot scheme"

2009:10 Gartell Marie “Unemployment and subsequent earnings for Swedish college graduates: a study of scarring effects”

2009:11 Lindvall Lars “Neighbourhoods, economic incentives and post compulsory education choices”

Dissertation series

2009:1 Lindahl Erica “Empirical studies of public policies within the primary school and the sickness insurance”

2009:2 Grönqvist Hans “Essays in labor and demographic economics”