ECONOMIC STUDIES

DEPARTMENT OF ECONOMICS
SCHOOL OF ECONOMICS AND COMMERCIAL LAW
GÖTEBORG UNIVERSITY
134

ESSAYS ON IMMIGRANTS' ECONOMIC INTEGRATION

Kerem Tezic

ISBN 91-88514-94-3
ISSN 1651-4289 print
ISSN 1651-4297 online

To my dear parents
and to Deniz

Abstract

This thesis consists of five papers, related to each other in terms of study-sample, study-subject or methods used.

The first paper is concerned with second-generation immigrants' educational attainments, using the Longitudinal Individual Data-set (LINDA), which gave us the possibility to examine changes over time, from ages 16-17 to 21-22 and to compare second-generation immigrants with a randomly-chosen matched control-group of native Swedes. Since Swedish youth are obligated to remain in school through grade 9, finishing at age 16, so the focus was on post-compulsory (upper-secondary) education before university. The data available allowed us to analyze the influence of parental-income on post-compulsory educational choices, and even to decompose the sources of the income (i.e., labour-income, asset-income, welfare-income, etc.). We were also able to include parental levels of education as possible determinants. Thus we could take into account the effects of parents as role-models, as postulated by socialization theory. We found differences in these effects, and thus in the educational outcomes, both between second-generation immigrants and native Swedes, and among groups of second-generation immigrants identified by geographic origin.

Again based on LINDA, the second paper focuses on the early labour-market experiences of second-generation immigrants in Sweden from age 16-17 in 1991 to age 25-26 in 2000. The initial experiences of new entrants into the labour-market can seriously influence later developments in their lives. Using transition-data analysis in a competing-risks framework, four different types of transitions into the labour-market were analyzed: The first two from either compulsory or post-compulsory education to various competing states; the last two from non-employment to work after either compulsory or post-compulsory education. Again a control-group of native Swedes was used for comparison. Parental characteristics not only influenced second-generation immigrants' prospects for continuing their education but also their later labour-market success. For all youths, regardless of ethnic background, parental education, occupation and income were vital. Other inter-generational transmission-channels such as ethnic capital and "neighborhood characteristics" were also important. The study verifies that finding a job was difficult for second-generation immigrants, especially for those from Africa, Latin America, and the Middle East.

The third paper focuses on the relationship between university education and employment during the first four years after graduation. The study-population from a survey conducted for Statistics Sweden (SCB) during the spring of 1999, consisted of individuals who graduated during 1994. The data allowed us to examine the graduates' demographic backgrounds, their educational fields and achievements, as well as their initial and second labour-market experiences, including their disposable incomes in 1998. There were differences between the sexes as well as between universities attended, regions of residence, and occupational orientations, with respect both to types of transition and earnings.

Again using LINDA, the fourth and fifth papers focus on arrival-cohort effects on the earnings of an unbalanced panel of 60,000-70,000 first-generation immigrants during 1990-2000, analyzed separately for men and women since their labour-market determinants were expected to be different. The econometric model used handled potential sample-selection bias by estimating the employment-and earnings-equations simultaneously while allowing for random effects in both, which allowed us to distinguish both age and cohort-effects. In the fifth paper a possible endogeneity-problem when using the husband's-earnings as a control variable was also corrected for by predicting their earnings and using them as an instrument in women's employment-and earnings-equations. As in the first and second papers, a matched control-group of randomly-selected native Swedish men (in the fourth paper) and women (in the fifth), was used. In terms of both employment-probabilities and earnings, there were considerable differences in terms of the marginal effects of some variables for immigrants with different geographic groups, and our findings were pessimistic for some of them especially Africans and Middle Easterners.

Keywords: Second-generation immigrants, educational attainments, early labour-market experiences, competing-risks, graduate employment, calibration, first-generation immigrants, sample-selection in panel data, random effects.

Acknowledgement

First of all I am indebted to Ali Cevat Tasiran for his excellent guidance, encouragement, and continous support. He always commented on whatever I gave to him and discussed many things with me very carefully. Ali has always been very generous with his time and knowledge. Suzanne Evans from the Department of Mathematics and Statistics, Birkbeck College, also read my papers carefully and I am grateful to her for valuable comments and suggestions.

I would also like to thank to Lennart Hjalmarsson and Gunilla Bornmalm-Jardelöw for their continuous support; they played a major role in the accomplishment of this thesis.

I am also grateful to financial support provided by the Bank of Sweden Tercentenary Foundation (Stiftelsen Riksbankens Jubileumsfond) and by the Evaluation Unit of National Insurance Board (Utvärderingsavdelningen av Riksförsäkringsverket, RFV).

In four of the five papers, I used the Longitudinal Individual Data-set (LINDA). For answers to all my questions, I am very grateful to Håkan Björk from Statistics Sweden, who helped me immediately whenever I called or emailed him. Leif Johansson from SCB provided me with the LINDA data-sets while Anna Demérus and Margaretha Säfström provided the figures I needed for the fourth and fifth papers; to all of them I am very grateful. For the third paper, help from Sixten Lundstöm in the form of calibration-weights and his book is greatly acknowledged.

I also want to thank my econometrics professor, Lennart Flood who guided me to my studies at Chalmers University of Technology and later had me as his assistant in econometrics courses. It was a pleasure to have him as a teacher as well.

William H. Greene was also very helpful with lots of excellent much-appreciated helps and suggestions regarding Paper 4 and 5.

Most of the time I worked at Södra Allégatan. I feel very lucky to have met both Dominique Anxo and Donald Storrie who always recieved me with an open door and a friendly manner; they are very special to me. I also thank them for the seminars organized at Södra Allégatan together with Henry Ohlsson and Katarina Katz. I shared many things with my colleagues and friends who stayed very late during many long nights at Södra Allégatan. My friends Florin Maican, Violeta Piculescu, Anton and Eugene Nivorozhkin, Constantin Belu, and Jorge Garcia provided great moral support. Short coffee breaks, to be able to talk a few minutes is such a relief. I would like to thank Hala Abou-Ali, Sten Dieden, Henrik Hammar, Katarina Nordblom, Håkan Eggert, Roger Wahlberg, Nizamul Islam, Alexis Palma, Klaus Hammes, Bengt Haraldsson, Fredrik Andersson, Martin Linde-Rahr, Marcus Eliasson and many others for help in many small but very useful ways. Rick Wicks' corrections, editorial suggestions, and comments were very useful and are greatly appreciated. Haldun Sonkaynar's corrections in the first paper were greatly appreciated. I also want to thank to Eva Jonason, Eva-Lena Neth, Jeanette Saldjoughi, and Gunilla Leander for all their administrative help, which was very important for me. Another big thanks goes to Ingvar Holmberg, Lars-Erik Peterson, and Margareta Westberg for having me as their assistant in various statistics courses, as well as for their always very friendly manner.

Ali Cevat Tasiran's wife Hulya called me very often and she gave me great moral support whenever I needed to talk. I hope and I am sure little and lovely Deniz will get much better future and she fully deserves it.

This year, the arrival of my friend and brother Alpaslan Akay made things much better for me. We had so much in common already, from my mother's house near the Bosporus. With his help, friendship, and solidarity, things became much easier, as we were able to share many cultural aspects of life. The long talks about existentialist philosophy, jazz culture, absurdness of life and the universal solitude of being human were unforgettable; I am sure we will continue to talk as long as we live and I am sure we will also work in the same field together.

My great friends, Tolga Ebevi, Faik Barutogullari, Anna Hjärne, Fahri Yilmaz, Britt-Marie Ingeby, Gurbet-Sirac Demiral, Hilmi Fayek, Mattias Barve and Camilla Carlestav were always besides me and I am very lucky to have such friends.

Finally, I dedicate this work to my parents. For me they mean everything. Having such human, kind, intellectual, honest, and modest academician-parents was the best that ever happened to me. I have learned a lot from them and I thank them a lot for being with me all the time.

Göteborg, Sweden, April 2004

Contents

Paper 5 "Arrival-Cohort Effects on the Incomes of Immigrant Women in Sweden" pp. 1-47.

Kerem Tezic

The Effect of Parental Income on the Post-Compulsory Education of Second-Generation Immigrants in Sweden

Kerem Tezic and Ali C. Tasiran*

Department of Economics, Göteborg University

April 2002

Abstract

Understanding the economic integration of minority ethnic communities requires an analysis of the educational process. Second-generation immigrant-youths' educational attainments were studied in comparison with those of similarly-aged native Swedes. Binomial-logit, grouped-regression and multinomial-logit models were applied to longitudinal data. Evidence was found for socioeconomic determinants of post-compulsory education and for parental influence on educational choices. Parental income affected second-generation immigrants' post-compulsory education and Swedes' choice of level of education while parental education was found to affect the choice of type of education in general. The geographical origin of second-generation immigrants mattered, with youths of Asian origin having a higher probability of continuing their education.

Keywords: Second-generation immigrants, educational choices, probability- and grouped-regression models.

J.E.L. Classification: C25, I21, J15.

*We thank Jerry Coakley for his valuable comments and Håkan Björk from Statistics Sweden for his continous help . We have also benefited from the comments of the participants at presentations at the Centre for European Labour Market (CELMS) in Göteborg; at the ILM conference "Employment, Unemployment, and Under-unemployment" at the Centre for International Labour Market Studies (CILMS) in Aberdeen; and at the Department of Economics, Växjö University. Any remaining errors are our own.

Contents

1 Introduction

The number of immigrants in Sweden increased significantly after World War II, from one percent in 1940 to about four percent in 1960 and seven percent in 1970 (Ekberg, 1994). Until the 1970s migration to Sweden was predominantly for economic reasons, and immigrants were predominantly from cultural and ethnic backgrounds somewhat similar to those of native Swedes,[1] either fellow Nordics or Europeans. But with political events in Chile in the 1970s; Poland, Iran, and Iraq in the 1980s; and former Yugoslavia, Somalia, and other parts of Africa in the 1990s; and a corresponding shift from economic to political immigration, the ethnic picture of Sweden started to change. The cultural and educational background of these more recent immigrants —largely refugees— were also different from those of previous groups.

The economic performance of Swedish immigrants has varied substantially, just as in other western countries (Ekberg and Gustavsson, 1995; Ekberg 1997; Ekberg and Rooth, 2002). In recent years, the integration of second-generation immigrants has gained increased attention, in policy discussions. They have been called a lost generation, and the general opinion of their economic future has been pessimistic. Empirical research on the integration of second-generation immigrants in Sweden is nevertheless limited, and the same is true for other countries. In the United States, in contrast to the voluminous literature analyzing the economic impact of immigrants, little is known about the labour-market performance of the second-generation (Borjas, 1993). The most important reason has been the lack of appropriate data-sets useful for the analysis.

Understanding the economic integration of minority ethnic communities requires analysis of the educational process. The poor labour-market performance of recent second-generation immigrants in Sweden has thus drawn policy-makers attention to education. To analyze the scope for policy to affect the educational environment in Sweden, this study focused on immigrant-children born in Sweden or who immigrated before age seven. We used Longitudinal Individual Data-set (LINDA), which gave us the possibility to examine changes over time and to compare second-generation immigrants with a control-group of native Swedes from the ages of 16-17 to 21-22. Since Swedish youth are obligated to remain in school through grade 9 finishing at age 16, the focus was on post-compulsory (upper-secondary) education.

[1] Similar tendencies have been observed in migration to Canada, the USA and Australia, where the majority of the immigrants were originally from Europe.

We were able to study the choices of students as "expected education" when they started this phase and as "realized education" when they finished it.

In earlier literature, the educational choices of second-generation immigrants (Gang and Zimmerman, 2000; Van Ours and Veeneman, 2001) have been examined separately from the economic conditions of their parents. Because of the nature of the data used, we were able to analyze the influence of parental-income on post-compulsory educational choices, and even to decompose the sources of the income (i.e., labour-income, asset-income, welfare-income etc.). We were also able to include the parental level of educations as a possible determinant. Thus we could take into account the effects of parents as role-models, as postulated by socialization theory. We found differences in these effects, and thus in the educational outcomes, between second-generation immigrants and native Swedes, and among groups of second-generation immigrants identified by geographic origin.

The next section describes theories regarding post-secondary education and the results from previous research. Section 3 describes the data, while section 4 develops the statistical models of different educational outcomes: the decision to continue with post-compulsory education, the length of pre-university education and the choice of type of post-compulsory education. The last section summarizes the results and draws conclusions.

2 Theories and evidence regarding post compulsory education

The theoretical framework used here is based on the concepts of human capital and household-production (Becker, 1965, 1975), and social capital (Chiswick, 1988; Borjas, 1992, 1994). In human-capital theory, investment in education, like investment in physical capital by firms, involves both costs and stream of benefits that accrue over a long period. Individuals make decisions about the amount they invest in education in order to maximize the present value of "profits", which is the difference between benefits and costs. The demand for investment-funds relates the marginal rate of return, and the supply of investment-funds relates the marginal cost, to the level of investment; the optimal choice occurs where supply equals demand. The implication of this model is that the total amount invested in education differs among individuals due to the differences in demand-conditions, supply-conditions or both. Family socioeconomic characteristics can affect the amount of educational investment by altering both.

In household-production theory, it is assumed that the household obtains utility from some underlying goods that cannot be bought in the market but are instead produced in the household from inputs of market-goods and "leisure" time. In this context, children's educational attainment can be viewed as an underlying good, produced with inputs of market-goods, and time that enters the household's utility-function indirectly, via future income. Children benefitting from greater parental inputs can be expected to attain and achieve higher productivity and income.

The extension of Becker's household-production theory to ethnic groups was made by (Chiswick, 1988), who incorporated additional "social-capital" inputs such as cultural preferences for education (see also Borjas, 1992), the relative desire for future vs. present consumption, and the parents' levels of education. Highly-educated minorities often seem to have a cultural taste or preference for education and to place higher value on future than on present consumption. Even if all children had equal access to financial resources and had similar genetic endowments, the family environment could thus result in different educational attainments.

These theoretical frameworks suggest the possible influence of family socioeconomic characteristics on the educational attainment of children. The effect of parental income has been the subject of much debate; e.g., the role played by the components of parental income is not clear in Becker's model as economists do not usually distinguish among them. In contrast, socialization-theory looks at the different components of parental income source rather than only at total income as a financial resource. Parents, above all, can serve as role-models and the parents' type of income can affect the child's level and type of education. For example, parents on welfare may induce increased dependency in their children by discouraging self-sufficiency, and this can limit educational achievement.

Another socioeconomic factor is the presence or absence of the father, which can influence not only family income but also the amount of parental time spent with children, thus affecting school performance. According to welfare-theory, the male role-model is important for the cognitive and educational attainment of children. According to Beller and Chung (1992), the presence of both parents reduces time-pressure although the presence of a step-father can complicate the college-entrance decision. Krein and Beller (1988), Astone and McLanahan (1991), and Beller and Chung (1992), all concluded that the educational effect of living in a single-parent family was negative, worse for males and increasing with duration.

During the last two decades, empirical research has focussed on young people's educational attainment and on the influence of family characteristics, to explain why some young adults succeeded and some did not. Based on several waves of PSID data, Alwin and Thornton (1984), McLanahan (1985), Hill and Duncan (1987), and Corcoran and Datcher (1989) all found family income to be statistically significant and positively associated with educational attainment. Using data based on a cohort of male high-school seniors (grade 12), Sewell and Hauser (1975) and later Sewell, Hauser and Wolf (1980), also found statistically significant positive effects of parental income on the level of completed education.

The level of parental education, measured typically by the number of years in school has also been emphasized in many studies of the inter-generational transmission of socioeconomic status. The evidence from Hill (1979), Haveman *et al.* (1991), and Manski *et al.* (1992) was similar: Parental education and mother's labour-market employment were statistically significant positive determinants of high school (grade 12) completion.

Contrary to the general findings in the literature Gang and Zimmerman (2000) in a study in Germany, found that the level of parental education played no role in the educational choices of foreign-born children. But Van Ours and Veeneman (2001), controlling for the level of parental education, found that differences between second-generation immigrants and natives in Holland vanished largely. In Denmark, Nielsen *et al.* (2001) focused on the second-generation immigrants' probability of obtaining a "qualifying education", meaning at least 18 months beyond the compulsory level (grade 9). The educational level of the parents was found to be statistically significant only for native Danes whereas having parents with several years of labour-market experience had a statistically significant positive effect for all groups. For second-generation immigrant-women, the parents' income was an important positive factor. In Sweden, Österberg (2000) found that the parent's level of education had a positive impact on the child's educational attainment and that reduced the negative effect from belonging to some ethnic groups. Higher parental education may correlate with more or higher-quality attention to children, resulting in their increased desire and capacity to continue their education.

3 The data

The institutional features of the Swedish educational system are described in Appendix A. In 1996, there were approximately 449,000 students in the upper-secondary schools in Sweden.[2]

The data set used in this study is from the 1991-1996 panel of the register-based Longitudinal Individual Data set (LINDA), which includes socioeconomic characteristics for individuals and their household members, and designed to be representative of the population for each year. The principal data-sources are the official Income Registers and Population and Censuses. The definition of the family differs between the Censuses and the Income Registers; the Census definition is based on whether individuals actually reside together, while the Income Registers use the tax definition.

Following Kossoudji (1989), the study sample consisted of 1106 second-generation immigrants, 16-17 years old in 1991 (532 female) either born in Sweden with at least one foreign-born parent or immigrated before age six. We followed their behaviour until 1996 when they reached age 21-22.

The second generation immigrants' geographical origins were determined from the father's country of birth (or if only the mother was foreign-born, from her country of birth) and categorized in seven groups. Nearly half were of Nordic origin, most of them Finnish. The rest came from Western countries, including the USA, Canada, Australia, New Zealand and the EU; from Eastern Europe; the Middle East; Asia; Africa; or Latin America.

A control-group of 1106 same-age native Swedes (16-17 years old in 1991, born in Sweden with both Swedish parents) was matched by county of residence. See Appendix B for the details.

Table 1 in Appendix C then provides descriptive statistics about them. The immigrant-parents were very slightly older on average, but the educational differences were substantial. Far more immigrant parents than native Swedes (28 percent vs. 15 percent) had not completed upper-secondary (high-school), whereas far more native Swedes than immigrant-parents (37 percent vs. 29 percent) had a university degree. Nearly half of the second-generation immigrants were from Nordic countries, followed by Westerners (17 percent), and Eastern Europeans (16 percent),

[2] Thirty percent went to vocational schools, providing at most two additional years of education, while the rest went to general schools longer than two years.

Latin Americans (9 percent), Middle Easterners and Asians (6 percent each) and Africans (1 percent).

Slightly more of the second-generation immigrant sample than the native-Swedish control group (52 percent vs. 50 percent) were male. About two-thirds of the second-generation immigrants had been born in Sweden; 43 percent had one Swedish parent. The second-generation immigrant had very slightly more siblings, but the two groups had almost identical proportions living in two-parent families (78 percent).

Total annual disposable family income was about the same for both groups, but for the native Swedes, father's and mother's labour-incomes, as well as asset incomes, were considerably higher, whereas for the immigrant-households welfare income was higher.

4 Statistical models of various educational outcomes

We modelled the probability, length and the type of pre-university continuing education beyond the compulsory level (grade 9). A binary-logit model was fitted to estimate the effects of the variables on the decision whether or not to continue with post-compulsory education. The level of completed education was analyzed using a grouped-regression model, and the type of education chosen was analyzed using a multinomial-logit model. Such separate estimations are more appropriate for modelling these decision processes than a simultaneous model would be because these decisions are sequential rather than contemporaneous. Furthermore, the length of schooling varies during the study period in upper-secondary school and threshold levels are known (see Section 4.2). We thus used a grouped-regression rather than an ordered-probit model to analyze the choice of completed pre-college educational level.

The explanatory variables are essentially the same in all three models (see Appendix C including Table 1). But because of high correlation between family income and some background variables they were not all used for all groups (i.e., second-generation immigrants, and the native Swedish control-group). For the second-generation immigrants, high correlation between parental age and family income led us drop parental age as an explanatory variable in all three models. And similarly for the native Swedish control-group we dropped parents' education in all three models.

4.1 Continuing on to post-compulsory education

The decision whether to continue with post-compulsory education was analyzed using a binary-logit model in which an underlying response-variable Y_i^* (for the yes or no decision for each individual i) can be defined for a $(1xk)$ vector of observable explanatory variables $\mathbf{x}_i$. by the statistical model

$$y_i^* = \mathbf{x}_i\boldsymbol{\beta}+e_i \tag{1}$$

which is related to the associated observable random-choice variable by

$$y_i = I_{(0,\infty)}(y_i^*) = I_{(0,\infty)}(\mathbf{x}_i\boldsymbol{\beta}+e_i) \tag{2}$$

where $I_A(Z)$ is the standard indicator-function for which $I_A(Z) = 1$ if $Z \in A$, and $I_A(Z) = 0$ otherwise. In the context of discrete choice (Quandt, 1966, and McFadden, 1976), we can represent the probability that $y_i = 1$ as

$$p_i = P(y_i = 1) = P(y_i^* > 0) = P(e_i > -\mathbf{x}_i\boldsymbol{\beta}) \tag{3}$$

If we assume the logistic distribution for the Cumulative distribution-function $F(\mathbf{x}_i\boldsymbol{\beta})$ then[3]

$$p_i = P(y_i = 1) = \frac{\exp(\mathbf{x}_i\boldsymbol{\beta})}{1+\exp(\mathbf{x}_i\boldsymbol{\beta})} \tag{4}$$

This is the binomial-logit model, the likelihood-function for which can be expressed as

$$\ln[L(\boldsymbol{\beta};\mathbf{y})] = \sum_{i=1}^{n}\left[y_i \ln\left(\frac{\exp(\mathbf{x}_i\boldsymbol{\beta})}{1+\exp(\mathbf{x}_i\boldsymbol{\beta})}\right) + (1-y_i)\ln\left(\frac{1}{1+\exp(\mathbf{x}_i\boldsymbol{\beta})}\right)\right] \tag{5}$$

[3] We can establish a functional linkage between p_i and $\mathbf{x}_i\boldsymbol{\beta}$ by assuming a logistic distribution for the unobservable noise-component e_i:

$$\begin{aligned} p_i &= P(e_i > -\mathbf{x}_i\boldsymbol{\beta}) = 1 - F(-\mathbf{x}_i\boldsymbol{\beta}) \text{ or} \\ &= P(-e_i < \mathbf{x}_i\boldsymbol{\beta}) = F(\mathbf{x}_i\boldsymbol{\beta}) \end{aligned}$$

4.2 Choosing the length of education

Although the m (short, medium, long)[4] outcomes are ordinal, the grouped regression model was used instead of an ordinal-probability model because the cut points did not need to be estimated. Even though we did not know the total number of years of education (pre-university) a youth has had, we knew the level of education and the cut points (see footnote above). The latent variable y_i^* was now used for the choice of length of education for each individual i with observable outcomes

$$y_i = m \qquad \text{if} \quad \mu_{m-1} \leq Y_i^* < \mu_m \tag{6}$$

for alternatives $m = 1, 2, 3$. The related likelihood-function was

$$\ln[L(\boldsymbol{\beta}; \mathbf{y})] = \sum_{i=1}^{n} \sum_{m=1}^{M} \ln \left[F(\mu_m - \mathbf{x}_i \boldsymbol{\beta}) - F(\mu_{m-1} - \mathbf{x_i} \boldsymbol{\beta}) \right] \tag{7}$$

4.3 Choosing the type of education

In an unordered multinomial discrete model for the choice of type of education there can be M nominal types, which the dependent variable y_i can take. Thus $y_{im} = 1$ indicates that a certain type of education was chosen and $y_{im} = 0$ means that this type was not chosen. The probability of choosing a certain type can be written as follows

$$P(y_i = m \mid \mathbf{x}_i.) = \frac{\exp(\mathbf{x}_i \boldsymbol{\beta}_m)}{\sum_{k=0}^{M} \exp(\mathbf{x}_i \boldsymbol{\beta}_k)} \tag{8}$$

A convenient normalization is to assume $\boldsymbol{\beta}_o = 0$, so that the likelihood function is

$$\ln[L(\boldsymbol{\beta}; \mathbf{y}, \mathbf{x_i}.)] = \sum_{i=1}^{n} \sum_{k=0}^{M} d_{im} \ln \left[\frac{\exp(\mathbf{x}_i \boldsymbol{\beta}_m)}{1 + \sum_{k=1}^{M} \exp(\mathbf{x}_i \boldsymbol{\beta}_k)} \right] \tag{9}$$

[4]The term short refers to at most 9 years of education (i.e., highest completed level is lower-secondary); medium refers to upper-secondary studies with at most two years of further study (i.e., 10-11 years of education); and long refers to three years or more upper-secondary education (i.e., 12 years or more).

where $d_{im} = 1$ if alternative m is chosen, and $d_{im} = 0$ otherwise.

5 Results

5.1 Continuing on to post-compulsory education

It is possible that, in the binary-logit models of the underlying response-variable y_i^* for the decision on whether to continue with post-compulsory education, the error-terms for each individual might be heteroskedastic. In that case, the logit-model would no longer be appropriate and parameter-estimates based on it would be inconsistent. We therefore tested three binary-logit model, (of second-generation immigrants, native Swedes, and together all) for heteroskedasticity as a consequence of parental income-variation among both second-generation immigrants and Swedes. We could not reject the null hypothesis of homoskedasticity at the 5 percent significance level. See Appendix D for the details and test-results.

Maximum-likelihood estimates of the binary-logit model for continuing after compulsory education are reported in Tables 2a and 2b in the Appendix.[5] Table 2c shows odds-ratio estimates for the significant parameters.

The effects of parents education were both statistically significant and positive (Table 2a in the Appendix). The odds of continuing with post-compulsory education were 2.1 times higher for second-generation immigrants whose parents had a university degree compared to those whose parents had not completed upper-secondary (Table 2c in the Appendix). For the combined sample as a whole, the results were similar. This result is consistent with the intergenerational transmission of human-capital hypothesis and with the results of Österberg (2000).

Geographical origin was not a statistically-significant variable except for Africans where it was quite negative; the odds of their continuing with post-compulsory education were 0.3 times smaller than those with Nordic origin.

Coming from a two-parent family did not give a significant effect for native Swedes, but it was significant and positive for second-generation immigrants for whom the odds of continuing education were 1.9 times higher than otherwise.[6] For

[5] Table 2b contains only the parameters of decomposed income: father's labour-income, mother's labour-income, welfare- income,etc. The other control variables are the same -except for family income-but their coefficients are not reported. We ran seperate regressions with decomposed income for the other two models also; and the results are in Tables 3b and 4f.

[6] It is impossible to tell from LINDA whether or not both parents were the "birth" parents,

second-generation immigrants, having a Swedish parent was significant gave a statistically significant and positive effect. The odds to continuing education were 1.7 times higher than for those having no Swedish parent.

For immigrants alone, total family income was not a significant variable, but for native Swedes and for the pooled sample it was.

When income is decomposed (Table 2b), father's labour-income was significant and positive for second-generation immigrants, while welfare-income was significant but negative for both immigrants and native Swedes. Both of these results are consistent with theory.

5.2 Choosing the length education

The grouped-regression results on the length of pre-university education are presented in Tables 3a and 3b in the Appendix.

The parents' educational level had a positive influence both for second-generation immigrants and for the pooled sample. Again this is consistent with previous studies and shows the link across generations, as in Coleman's (1988) view that the culture in which an individual is raised alters their opportunity-set and has significant effects on future behavior, including human capital formation and labour-market outcomes. Geographical origin was statistically significant for Eastern-Europeans and Asians, perhaps indicating a "cultural" preference for education.

Being male reduced the length of education for Swedes and for the pooled sample, in accordance with earlier results by Beller and Chung (1992).

As in the binary logit model of continuing education, coming from a two-parent family and having a Swedish positively influenced the length of pre-university education for second-generation immigrants.

As expected, family income was statistically significant factor in the length of education chosen, giving substantial support for the economic hypothesis that income, regardless of its source, is a crucial determinant of the educational attainment of children. However the specific labour-income of the father or mother is also believed to represent the role-model provided by that parent, as in studies by Sewell and Hauser (1975), Corcoran and Datcher (1981), Kiker and Condon (1981), Hill

i.e whether or not the parents had remarried, so "parent" includes step-parents and "sambos" (cohabiting).

and Duncan (1987), and Nielsen *et al.* (2001), which all of found positive effects of parental income on the education of their children. When sources of income were decomposed (Table 3b), father's labour-income had a positive effect for second-generation immigrants, while mother's labour-income was positive for Swedes and the pooled sample. Welfare-income and all "other" income were negative for Swedes and the pooled sample.

5.3 Choosing the type of education

An important issue in the use of the multinomial-logit models is the assumption of independence (of the response categories) from irrelevant alternatives, or IIA, which simply means that the ratio of the choice-probabilities of any two response-categories is not influenced systematically by any other alternative. We verified the independence of the choice-alternatives for types of education using a Hausman-type test-statistic, based on eliminating one or more alternatives from the choice-set to see if the underlying choice-behavior would be different. The test-results indicated that we could not reject the hypothesis that IIA holds, i.e., the types of upper-secondary education were genuine choices independent from each other. See Appendix E for details. There were six education types: general (or non-occupational); humanities (artistry, art, theater, religion, etc.); social sciences (economics, accounting etc.); technical, health related (medicine, nursery, psychology etc.); and service oriented. Estimation of the multinomial-logit model yielded results in Tables 4a-f in Appendix.

The first columns of the Tables show the estimation results of general education comparing to various occupationally-oriented educations for (reading across the tables) second-generation immigrants, native Swedes and the pooled sample. The second, third, fourth, and fifth columns of each section of Tables 4b-f show comparisons of one occupationally-oriented education against another.

Among all the different possible comparisons, statistically-significant results are sometimes found on the variables for parental age, parental education, geographic origin, sex, born outside Sweden, two-parent household and income (Tables 4a-e). When income is decomposed (Table 4f), statistically significant results are sometimes found on all categories except "other" income.

To go through some examples of the effects for second-generation immigrants, the odds of choosing general education vs. humanities were 2.2 (= $exp[0.794]$) times higher for males than for females, while for the pooled sample the odds of choosing technical vs. general education were 5.4 (= $exp[1.688]$) times higher for males than for females. Such statistically-significant positive effects for males showed up in the choice of technical vs. all other types of education for all three groups (immigrants, Swedes and pooled).

The level of parental education also generate many statistically significant positive results. For example the odds of second-generation immigrants choosing humanities vs. other educational programs ranged from 2.5 (= $exp[0.900]$) times higher (against social sciences) to 6 times higher (= $exp[1.799]$) (against service related education) if their parents had a university degree, compared to those whose parents had not completed upper-secondary education. For the full sample, the same odds ranged from 4.3 to 15.4.

Geographical origin was important in some comparisons but without a clear pattern.

For second-generation immigrants, parental income had a positive effect on the choice of general, social-science or technical, versus health related education and on social science vs. service-oriented education.

When the sources of income were decomposed (Table 4f in the Appendix), father's labour-income had a positive effect on the choice of general, humanities, social-science, or technical education for both immigrants and Swedes, while the influence of mother's labour-income was limited to Swedes in the choices of humanities vs. technical, health-related or service-oriented education.

6 Summary and conclusions

The intergenerational influence of immigrants on their children's education is important since education plays a crucial role in their integration in the Swedish labour market. This study focused mainly on parental influences on post-compulsory education in Sweden, and in particular on the impact of parental income on second-generation immigrants compared with native Swedes. A unique data set (LINDA), was used, yielding six successive years of information (1991-1996) on 1106 second-generation immigrants (initially 16-17 years old) and an equal Swedish control-group matched for age and region.

Three educational outcomes were analyzed: the decision whether to continue with post-compulsory education, using a binomial-logit regression; the length of education, using a grouped-regression and the type of post-compulsory education, using multinomial-logit regression.

Similar results over the different models suggest that the approach used has support from the data.

• Parental income was found to have positive effects on young people's continuing with post-compulsory education. Those with higher parental incomes were more likely to attend longer upper-secondary programs, and were more likely to chose programs aimed at continuing at the university level.

• Decomposing the sources of income, showed father's labour-income increased the probability of second-generation immigrants' continuing education and the length of the upper-secondary education chosen. On the other hand, mother's labour-income played a positive role for native Swedes' continuing education.

• Geographical origin matters: Students with Asian origin had higher probabilities of continuing education, while chances were low for students with African origin.

• Having a Swedish parent played a positive role in second-generation immigrants' decision to continue with upper-secondary education.

• Gender was important in the choice of type of education: Regardless of their geographical origins, males were more likely to choose technical education.

Thus we observed some intergenerational transmission of parental characteristics via the educational attainment of second-generation immigrants as well as native Swedes. This is in line with general findings in the literature. In general, the stronger labour-market position of the parents, the higher the probability of their children continuing with upper-secondary education and thus the higher the chance of their own success in the labour-market.

The next step, is to study second-generation immigrants' entry into the labour-market, which has been done and is reported in a companion paper.

References

[1] Alwin, D., and Thornton, A. (1984), "*Family Origins and the Schooling Process: Early versus Late Influence of Parental Characteristics,*" American Sociological Review, Vol. 49, pp. 784-802.

[2] Astone, N.M., and McLanahan, S.S. (1991), "*Family Structure, Parental Practices and High School Completion,*" American Sociological Review, Vol. 56, pp. 309-20.

[3] Becker, G.S. (1965), "*A Theory of the Allocation of Time,*" Economic Journal, Vol 75, pp.493-517.

[4] Becker, G.S. (1975), "*Human Capital*" University of Chicago Press,Chicago.

[5] Beller, A.H., and Chung, S.-S. (1992), "*Family Structure and Educational Attainment of Children,*" Journal of Population Economics, Vol. 5, No. 1, pp. 39-59.

[6] Borjas, G.J. (1992), "*Ethnic Capital and Intergenerational Mobility,*" Quarterly Journal of Economics, Vol. 107, No. 1, pp.123-50.

[7] Borjas, G.J. (1993), "*The Intergenerational Mobility of Immigrants,*" Journal of Labor Economics, Vol. 11, No, 1, pp. 113-35.

[8] Borjas, G.J. (1994), "*The Economics of Immigration,*" Journal of Economic Litterature, Vol. 32, December, pp. 1667-717.

[9] Card, D. (1995), "*Earnings, Schooling and Ability Revisited,* " Research in Labor Economics, Vol. 14, pp. 23-48.

[10] Chiswick, B. (1988), "*Differences in Education and Earnings across Racial and Ethnic Groups: Tastes, Discrimination and Investments in Child Quality,*" Quarterly Journal of Economics, Vol. 103, No. 3, pp. 571-97.

[11] Coleman, J.S. (1988), "*Social Capital in the Creation of Human Capital,*" American Journal of Sociology, Vol. 94 (supplement), pp. 95-120.

[12] Corcoran, M., and Datcher, L. (1981), "*Intergenerational Status Transmission and the Process of Individual Attainment,*" In Five Thousand American Families: Patterns of Economic Progress, edited by M.S. Hill, D.H. Hill, and J.N. Morgan, University of Michigan Survey Research Center, Ann Arbor.

[13] Datcher, L. (1989), "*Family Background and School Achievement Among Low Income Blacks,*" Journal of Human Resources, Vol. 24, No. 3, pp. 528-44.

[14] Davidson, R., and MacKinnon, J. (1993), "*Estimation and Inference in Econometrics,*" Oxford University Press, New York.

[15] Edin, P.-A., Lalonde, R.J., and Åslund, O. (2000), "*Emmigration of Immigrants and Measures of Immigrant Assimilation: Evidence from Sweden,*" Working Paper Series, Vol. 13, Uppsala University.

[16] Ekberg, J. (1994), "*Economic Progress among Immigrants in Sweden,*" Scandinavian Journal of Welfare, Vol. 3, pp. 148-57.

[17] Ekberg, J., and Gustavsson, P. (1995), "*Invandrare på Arbetsmarknaden (Immigrants in the Labor Market),* " SNS Förlag, Stockholm.

[18] Ekberg, J. (1997), "*Hur är Arbetsmarknaden för den Andra Generationens Invandrare (How is the Labor Market for Second-generation Immigrants)*?" Arbetsmarknad & Arbetsliv, Årg. 3, No. 1, pp. 5-16.

[19] Ekberg,J., and Rooth, D.-O. (2002) "*Unemployment and Earnings for Second-Generation Immigrants in Sweden,*" forthcoming, Journal of Population Economics.

[20] Gang, I.N., and Zimmermann, K.F. (2000), "*Is Child Like Parent? Educational Attainment and Ethnic Origin,* " Journal of Human Resources, Vol. 35, No. 3, pp.550-69.

[21] Haveman, R., Wolfe, B., and Spaulding, J. (1991), "*Childhood Events and Circumstances Influencing High School Completion,*" Demography, Vol. 28, No. 1, pp. 133-57.

[22] Hill, M.S., and Duncan, G. J. (1987), "*Parental Family Income and the Socioeconomic Attainment of Children,*" Social Science Research, Vol. 16, pp. 39-73.

[23] Hill, R.C. (1979), "*Capacities, Opportunities and Educational Investments: The Case of High School Dropout,*" Review of Economics and Statistics, Vol. 61, No. 1, pp. 9-20.

[24] Kiker, B.F., and Condon, C.M. (1981), "*The Influence of Socioeconomic Background on the Earnings of Young Men,"* Journal of Human Resources, Vol. 16, No. 1, pp. 94-105.

[25] Kossoudji, S.A. (1989), "*Immigrant Worker Assimilation: Is It a Labour Market Phenomenon*? " Journal of Human Resources, Vol. 24, No. 3, pp. 494-527.

[26] Krein, S.F., and Beller, A.H. (1988), "*Educational Attainment of Children From Single-Parent Families: Differences by Exposure, Gender and Race,*" Demography, Vol. 25, No. 2, pp. 221-34.

[27] Manski, C. F., Sandefur, G.D., McLanahan, S., and Powers, D. (1992), "*Alternative Estimates of the Effect of Family Structure during Adolescence on High Scholl Graduation,*" Journal of the American Statistical Association, Vol. 87, No. 417, pp. 25-37.

[28] McFadden, D. (1976), "*Quantal Choice Analysis: A Survey,*" Annals of Economic and Social Measurement, Vol. 5, pp. 363-90.

[29] McLanahan, S. (1985), "*Family Structure and Reproduction of Poverty,*" American Journal of Sociology, Vol. 90, No. 4, pp. 873-901.

[30] Nielsen, H. , Rosholm, M. , Smith, N. , and Nusted, L. (2001), " *Intergenerational Transmissions and the School-to-Work Transition of Second Generation Immigrants,*" Discussion Paper No 296 The Institute for the Study of Labour IZA Bonn.

[31] Quandt, R. (1966), "*Probabilistic Theory of Consumer Behavior,*" Quarterly Journal of Economics, Vol. 70, pp. 507-36.

[32] Sewell, H.W. and Hauser, R. M. (1975) " Education, Occupation and Earnings: Achievement in the Early Career", Academic Press, New York.

[33] Sewell, H.W., Hauser, R.M., and Wolf, W.C. (1980), "*Sex, Schooling and Occupational Status,*" American Journal of Sociology, Vol. 86, No. 3, pp. 551-83.

[34] Stewart, M. (1983), "*On Least Squares Estimation when the Dependent Variable is Grouped,*" Review of Economic Studies, Vol. 50, pp. 737-53.

[35] Sweetman, A., and Dicks, G.(1999), "*Education and Ethnicity in Canada*: *An Intergenerational Perspective,*" Journal of Human Resources, Vol. 34, No. 4, pp. 668-96.

[36] Van Ours J., and Veeneman, J. (2001), "*The Educational Attainment of Second-Generation Immigrants in the Netherlands,*" Discussion Paper No, 297, The Institute for the Study of Labour IZA Bonn.

[37] Willis, R.J. and Rosen, S. (1979), "*Education and Self-Selection,*" Journal of Political Economy, Vol. 87 No. 5, pp. 7-36.

[38] Österberg, T. (2000), "*Economic Perspectives on Immigrants and Intergenerational Transmissions,*" Economic Studies 102. Thesis. Göteborg University. Sweden.

Appendices

Appendix A. The institutional features of the Swedish educational system.

Swedish children attend primary school for six years from the year in which they turn seven, and continue with three or more of compulsory secondary school. Upon completion of this nine years of compulsory education, most students continue immediately on to upper-secondary school ("gymnasium"=American "high-school"). There are several areas of study at this level, the most important distinction being between general and vocational education. Upper-secondary vocational schools provide at most two years of further study. Most upper-secondary general education lasts for three years and prepares students for further study at universities or other institutions of higher education. Some upper-secondary educations such as nursing and engineering, last four years.

Appendix B. The control-group of native Swedes.

The control-group of native Swedes obtained from the same nationally representative data base (LINDA). In first stage, a sample of second-generation Immigrant youth sample is drawn from all second-generation immigrants. Let n_{SG} denotes the number of second-generation young immigrants and N_{SG} the number of all second-generation immigrants in LINDA 1991. Then the sampling fraction for second-generation young immigrants is:

$$f_{SG} = \frac{n_{SG}}{N_{SG}}. \tag{10}$$

In the second stage, units in the second-generation sample were grouped into disjoint cells according to their age and residential municipality areas in a table. Say there are C cells, $c = 1, ..., C$ in the table. Then the total of young second-generation immigrants can be written as

$$n_{SG} = \sum_{c=1}^{C} n_{c,SG} \tag{11}$$

The calculated sampling fraction for each cell is:

$$f_{c,SG} = \frac{n_{c,SG}}{N_{c,SG}} \tag{12}$$

We needed the same number of units in the Swedish sample as in the second-generation immigrant sample, $n_{SG} = n_{SW}$. The total number of individuals in the Swedish sample can also be grouped into the same age and residential municipality areas as in the second-generation sample. This is

$$n_{SW} = \sum_{c=1}^{C} n_{c,SW} \tag{13}$$

To obtain the number of Swedish young people in each cell of the Swedish table, which is equal to the number of second-generation immigrants in each cell of the second-generation table, we utilized the sampling-fraction information of second-generation immigrants in each cell, $f_{c,SG}$. In the sampling procedure, the number of each age and municipality-cell for second-generation immigrants was taken as a base for Swedes. The total number of Swedes in each cell is:

$$n_{c,SW} = f_{c,SG}\ N_{c,SW} \tag{14}$$

In order to get twin groups, we used the sampling-fraction information of second-generation young immigrants in each cell, $f_{c,SG}$, to draw the same number of Swedes as their twin (control) group of second-generation immigrants. In order to do this we kept all those sampled individuals who were 16-17 years old in 1991 and not dropped from the data set until 1996. There were about ten times more such Swedish youths, N_{SW}. The 1106 second-generation immigrants were divided into 50 cells depending upon whether they were 16 or 17 when they finished compulsory education, and depending on which of the 25 Swedish "counties" (län) they resided in. The native Swedes aged 16-17 and who similarly remained in LINDA through 1996 were then also divided in 50 cells. From each of the Swedish cells the same number of individuals were then randomly chosen as there were second-generation immigrants in the corresponding cells using random number generator. We thus ended up with a control group consisting of 1106 same-aged native Swedes matched by county of residence.

Appendix C. Variables.

Three educational outcomes were analyzed: whether the individual continued with post-compulsory education or not (i.e., upper-secondary after completing lower-secondary); the level of completed education; and the type of higher education.

"Years of Education" is common for studies in the United States, but not the case in Sweden because Statistics Sweden reports only the level of schooling completed, and upper-secondary programs are of varying length from less than 2, to 4 years.

The explanatory variables available and which we thought might affect educational outcomes were:

• Parental age, taken as the age of the head of the household, since, for those children living with only one parent, age would be missing value for the other.

• Parental education, taken as the highest completed level of either parent. Again since many children were living with only one parent and information was not available about the other one, it would have been difficult to analyze the effects of mothers' education and fathers' education separately. Information about the levels of education completed in the country of origin vs. those completed in Sweden (or elsewhere) was not available, making analysis based on such distinctions impossible. Educational levels were categorized as

- lower-secondary at most: had not completed upper-secondary (high-school);

- completed upper-secondary (high-school);

- university degree.

• Geographic origin: as explained in the test.

• Individual variables

- Sex;

- Whether born in Sweden;

- Whether or not either parent was Swedish;

- Whether or not two "parents" in the household (including step-parents and "sambos", i.e., cohabiting)

• Economic variables

- Total annual disposable income; family income averaged over the two-year period 1990-1991 when the child was 15-16 and 16-17 (using LINDA 1990); this was also decomposed as

- Father's labour-income;

- Mother's labour-income;

- Asset-income;

- Welfare-income (child-allowance, housing-allowance, general welfare, etc);

- Other income (sum of all income from sources other than parent's labour income, asset income, and parents' welfare income).

Table 1: Descriptive Statistics

	Second-generation Immigrants (n=1106)		Native Swedes (n=1106)		All (n=2212)	
	mean	std dev	mean	std. dev	mean	std. dev
Parental age (maximum if 2)	46.659	6.459	45.857	5.351	46.258	5.943
Parental education						
Lower-secondary	0.277	0.448	0.155	0.362	0.216	0.411
Upper-secondary	0.436	0.496	0.474	0.500	0.455	0.498
University degree	0.288	0.453	0.371	0.483	0.329	0.470
Mother working	0.769	0.423	0.816	0.387	0.792	0.406
Father working	0.858	0.352	0.932	0.256	0.895	0.306
Geographical Origin						
Nordic (not incl. Swedish)	0.498	0.500			0.249	0.432
Swedish					0.500	0.500
Western countries	0.170	0.376			0.085	0.279
Eastern-Europe	0.158	0.365			0.079	0.270
Middle-East	0.058	0.234			0.029	0.168
Asia	0.056	0.230			0.028	0.165
Africa	0.014	0.119			0.007	0.085
Latin-America	0.045	0.207			0.023	0.149
Male	0.519	0.500	0.498	0.500	0.509	0.500
Born outside Sweden	0.323	0.468			0.836	0.370
At least one Swedish parent	0.426	0.495			0.713	0.452
Two-parent family	0.788	0.409	0.783	0.413	0.785	0.411
Number of siblings	1.877	1.017	1.752	0.835	1.815	0.932
Parental income (log)	12.252	0.416	12.343	0.382	12.297	0.402
Father's labor-income[a]	13.899	12.290	18.449	13.523	16.174	13.117
Mother's labor-income[a]	9.745	7.386	11.968	7.098	10.857	7.327
Asset-income[a]	0.862	4.230	1.168	3.862	1.015	4.052
Welfare-income[a]	0.263	1.218	0.053	0.412	0.158	0.916
Other income[a,b]	6.574	6.132	3.962	4.566	5.268	5.560

[a] in tens of SEK per year

[b] Sum of all incomes other than parents' labor income,asset income and welfare income averaged over the period when the child was 15-16 (or 16-17) years old.

Table 2a: Binary-logit estimates of the probability of post-compulsory education

	Second-generation Immigrants (n=1106)		Native Swedes (n=1106)		All (n=1106)	
Constant	-8.349	(2.922)	-9.729***	(0.262)	-8.349**	(1.269)
Parental age (maximum if 2)						
Parental education			0.581***	(0.156)	0.188*	(0.099)
Upper secondary	0.428**	(0.199)			0.418***	(0.161)
University degree	0.741***	(0.259)			0.624***	(0.199)
Swedish					0.183	(0.214)
Western countries	0.032	(0.260)			0.077	(0.261)
Eastern-Europe	0.039	(0.250)			0.020	(0.244)
Middle-East	0.112	(0.374)			0.159	(0.374)
Asia	0.733	(0.475)			0.739	(0.462)
Africa	-1.071**	(0.479)			-1.032*	(0.573)
Latin America	0.024	(0.421)			-0.030	(0.420)
Male	-0.076	(0.172)	-0.235	(0.216)	-0.155	(0.134)
Born outside Sweden	-0.062	(0.208)			0.027	(0.896)
At least one Swedish Parent	0.512**	(0.225)			0.554**	(0.216)
Two-parent family	0.661***	(0.213)	-0.056	(0.292)	0.369	(0.173)
Number of siblings	0.003	(0.090)	-0.169	(0.133)	-0.044	(0.077)
Economic Variable						
Parental Income	0.234	(0.249)	0.841**	(0.342)	0.377**	(0.203)
AIC	957.933		668.712		1647.300	
SC	962.942		673.720		1653.002	
-2log L	955.933		666.712		1645.300	

Note: * = significant at 10 percent; ** = significant at 5 percent; *** = significant at 1 percent;
The standard errors are reported in parantheses
The reference variables included in the constant are: lower-secondary education, nordic, female, born in Sweden, one-parent family, no Swedish parent.

Table 2b: Binary-logit estimates of the probability of post-compulsory education

	Second-generation immigrants (n=1106)		Native Swedes (n=1106)		All (n=2212)	
Father's labor income	0.024**	(0.012)	0.014	(0.013)	0.019**	(0.008)
Mother's labor income	0.002	(0.014)	0.026	(0.019)	0.011	(0.012)
Asset income	0.056	(0.056)	0.053	(0.051)	0.054	(0.037)
Welfare income	-0.126**	(0.062)	-0.581***	(0.204)	-0.174***	(0.056)
Other income	-0.016	(0.018)	-0.045	(0.024)	-0.031**	(0.014)

Note: Other control variables are as shown in table 2a (except parental income)

Table 2c: Odds Ratio Estimates (for significant estimates from Tables 2a and 2b)

	Second-generation immigrants	Native Swedes	All Sample
Parent education			
Upper-secondary	1.534		1.518
University degree	2.097		1.866
Africa	0.343		1.356
At least one Swedish parent	1.668		1.741
Two-parent family	1.936		
Parental income(log)		2.319	1.458
Father's labor income	1.025		1.019
Welfare income	0.882	0.559	0.841

Appendix D. Heteroscedasticity test results.

Family income might show heteroskedasticity since there might be greater variation in continuing education among low-income families than in high ones, and so the disturbance-variance might not be constant across observations. The statistical model is the same as above

$$Y_i^* = \mathbf{x}_i\boldsymbol{\beta}+e_i \tag{1}$$

with the error-term e_i now distributed as

$$e_i \sim N\left(0, \exp(Z_i\gamma)^2\right) \tag{2}$$

where Z_i is a row-vector of log-family-income. In this model, to ensure that both $\boldsymbol{\beta}$ and γ are identifiable, Z_i must not include a constant term. The probability that $y_i = 1$ can now be written as

$$\begin{aligned} p_i &= P(y_i = 1) = P\left(\frac{e_i}{\exp(Z_i\gamma)} > \frac{-\mathbf{x}_i\boldsymbol{\beta}}{\exp(Z_i\gamma)}\right) = 1 - F\left(\frac{-\mathbf{x}_i\boldsymbol{\beta}}{\exp(Z_i\gamma)}\right) \\ &= P\left(\frac{-e_i}{\exp(Z_i\gamma)} < \frac{\mathbf{x}_i\boldsymbol{\beta}}{\exp(Z_i\gamma)}\right) = F\left(\frac{\mathbf{x}_i\boldsymbol{\beta}}{\exp(Z_i\gamma)}\right) \end{aligned} \tag{3}$$

and the log-likelihood function for this heteroskedastic model is

$$\ln[L(\boldsymbol{\beta};\mathbf{y})] = \sum_{i=1}^{n}\left\{y_i \ln F\left(\frac{\mathbf{x}_i\boldsymbol{\beta}}{\exp(Z_i\gamma)}\right) + (1-y_i)\ln\left[1 - F\left(\frac{\mathbf{x}_i\boldsymbol{\beta}}{\exp(Z_i\gamma)}\right)\right]\right\} \tag{4}$$

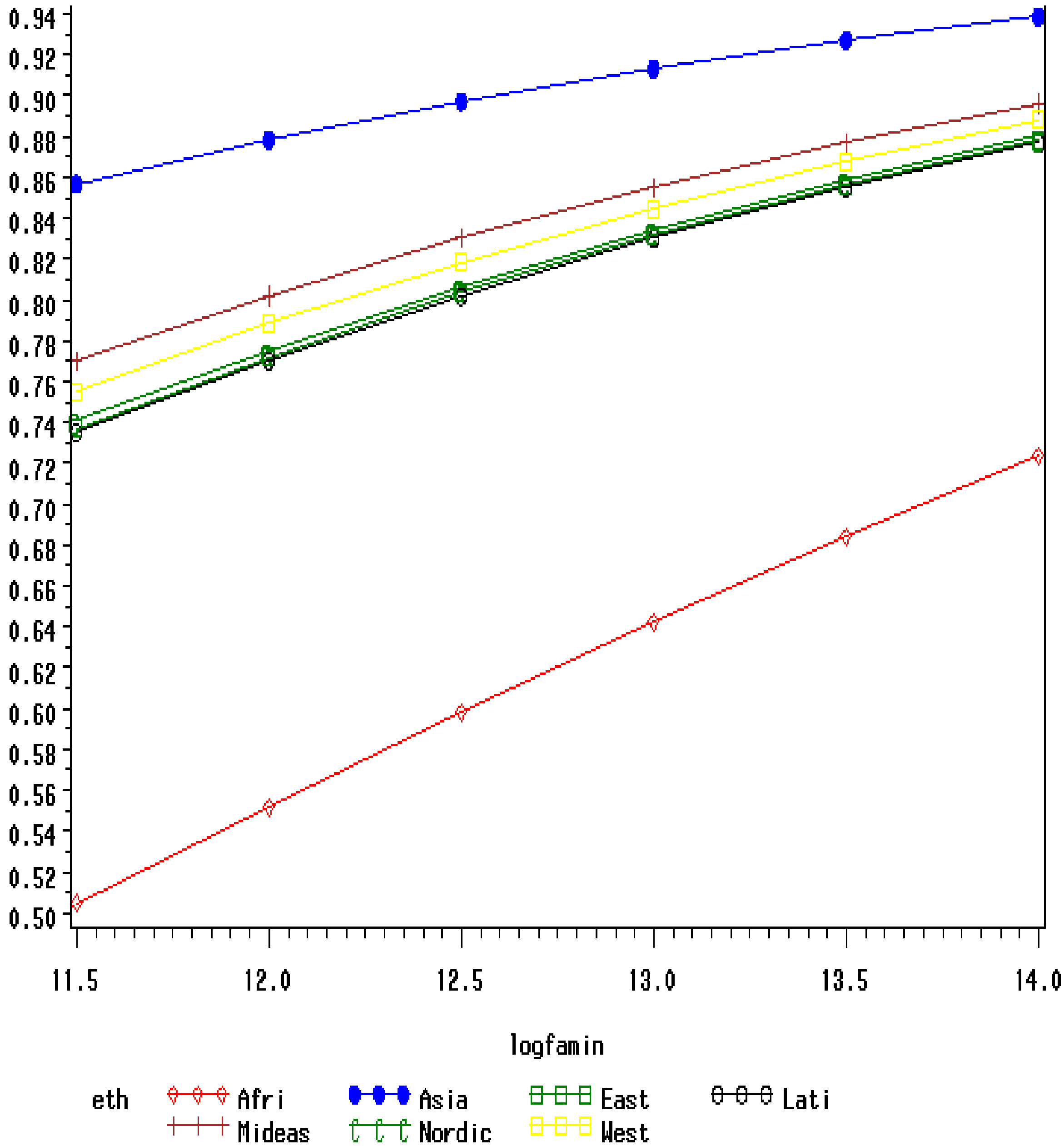
Now we can test for heteroskedasticity of log-family-income by testing the null hypothesis of homoskedasticty, $\gamma = 0$, with a likelihood- ratio statistic

$$LR = -2[\ln L_{Homoskedastic} - \ln L_{Heteroskedastic}] \tag{5}$$

The test results below indicate no heteroskedasticity as a consequence of parental-income variation among second generation immigrants or Swedes. We cannot reject the null hypothesis of homoskedasticity at 5 percent significance levels.

Testing for Heteroscedasticity

Binary Logit Models	γ	Likelihood Ratio Test
Second-generation immigrants	-0.274 (0.298)	2.556
Native Swedes	-0.337 (0.354)	3.610
All	-0.037 (0.439)	0.175



Predicted Probabilities of post-compulsory education based on family income, by geographical origin

Figure 1

(Reference group: Second-generation immigrant, parent with university degree, born in Sweden, two-parents family, one-parent Swedish, parent age=mean value, number of siblings=mean value.)

Table 3a: **Grouped-regression estimates of the length of education**

	Second-generation immigrants (n=1106)		Native Swedes (n=1106)		All (n=1106)	
Constant	4.657	(3.772)	-12.365***	(5.683)	-4.897	(3.929)
Parental age (maximum if 2)			0.742***	(0.188)	0.373***	(0.123)
Parental education						
Upper-secondary	0.821***	(0.264)			0.810***	(0.194)
university degree	1.661***	(0.321)			1.559***	(0.225)
Swedish					0.288	(0.220)
Western Countries	0.320	(0.310)			0.312	(0.293)
Eastern-Europe	0.707**	(0.319)			0.706**	(0.304)
Middle-East	0.333	(0.493)			0.246	(0.467)
Asia	1.147**	(0.514)			1.013	(0.492)
Africa	-1.272	(0.888)			-1.251*	(0.842)
Latin-America	0.122	(0.549)			0.060	(0.522)
Male	-0.262	(0.216)	-0.675***	(0.203)	-0.479***	(0.147)
Born Outside Sweden	-0.053	(0.271)			0.043	(0.255)
At least one Swedish parent	0.686***	(0.265)			0.680***	(0.240)
Two-parent family	0.687**	(0.290)	0.126	(0.284)	0.386*	(0.203)
Number of siblings	-0.083	(0.111)	-0.024	(0.133)	-0.039	(0.087)
Parental Income	0.527*	(0.319)	1.174***	(0.329)	0.549**	(0.234)

Note: * = significant at 10 percent; ** = significant at 5 percent; *** = significant at 1 percent;
The standard errors are reported in parantheses
The reference variables included in the constant are: lower-secondary education, nordic, female, born in Sweden, one-parent family, no Swedish parent.

Table 3b: Estimates of Educational Level Attained (Grouped Regression)[a]

	Second-generation immigrants (n=1106)		Native Swedes (n=1106)		All (n=2212)	
Father's Labor Income	0.030***	(0.011)	0.009	(0.010)	0.012	(0.008)
Mother's Labor Income	0.007	(0.018)	0.070***	(0.015)	0.031**	(0.012)
Asset Income	0.032	(0.028)	0.044	(0.026)	0.039**	(0.021)
Welfare Income	-0.143	(0.094)	-0.853***	(0.282)	-0.223***	(0.008)
All Other Income	-0.027	(0.022)	-0.064**	(0.026)	-0.046***	(0.016)

[a] Other control variables are as shown in table 3a

Table 4a: Multinomial-logit regression results on the type of education chosen, with base category human sciences, for second-generation immigrants, native Swedes, and the pooled sample

	Second-generation immigrants (n=931)	Native Swedes Twin Group (n=1010)	All (n=1941)
	General education	General education	General education
Parental age (maximum if 2)		-0.743*	-0.226
Parental education			
Upper-secondary	-0.496*		- 0.686**
University degree	-0.974**		-1.526***
Swedish			-0.109
Western countries	0.138		0.159
Eastern-Europe	-0.110		-0.066
Asia and Middle East	0.582		0.566
Africa	0.466		0.511
Latin-America	-0.050		0.034
Male	0.794***	0.834***	0.833***
Born outside Sweden	-0.189		-0.158
At least one Swedish parent	-0.308		0.145
Two-parent family	-0.198	0.324	-0.023
Parental Income (log)	0.306	-1.051**	0.046

Note 1: * = significant at 10 percent; ** = significant at 5 percent; *** = significant at 1 percent;
Note 2: The reference variables included in the constant are: lower-secondary education, nordic, female, born in Sweden, one-parent family, no Swedish parent.

Table 4b: Multinomial-logit regression results on the type of education chosen, with base category social sciences, for second-generation immigrants, native Swedes, and the pooled sample

	Second-generation immigrants (n=931)		Native Swedes Twin Group (n=1010)		All (n=1941)	
	General education	Human sciences	General education	Human sciences	General education	Human sciences
Parental age (maximum if 2)			-0.005	0.739	0.033	0.259
Parental education						
Upper-secondary	-0.047	0.448			-0.103	0.584*
University degree	-0.073	0.900**			-0.075	1.451***
Swedish					-0.368**	-0.259
Western Countries	-0.094	-0.231			-0.119	-0.278
Eastern-Europe	0.056	0.166			0.044	0.110
Asia and Middle East	-0.407	-0.989*			-0.473*	-1.038
Africa	1.631	1.165			1.597	1.086
Latin-America	0.234	0.284			0.275	0.241
Male	0.129	-0.665**	0.026	-0.808***	0.086	-0.748***
Born outside Sweden	0.074	0.264			0.049	0.208
At least one Swedish parent	0.203	0.511			0.121	0.219
Two-Parent family	-0.739***	-0.541	-0.218	-0.542	-0.451	-0.429
Parental Income (log)	-0.315	-0.620	-0.340	0.711*	-0.289	-0.335

Table 4c: **Multinomial-logit regression results on the type of education chosen, with base category technical sciences, for second-generation immigrants, native Swedes, and the pooled sample**

	Second-generation immigrants (n=931)			Native Swedes Twin Group (n=1010)			All (n=1941)		
	General education	Human sciences	Social sciences	General education	Human sciences	Social sciences	General education	Human sciences	Social sciences
Parental age (maximum if 2)				-0.285*	0.459	-0.280	-0.049	0.177	-0.082
Parental education									
Upper-secondary	0.021	0.517	0.069				-0.044	0.642*	0.059
University degree	0.217	1.191***	0.290				0.131	1.656***	0.205
Swedish							-0.161	-0.052	0.207
Western countries	0.515**	0.378	0.609**				0.507***	0.348	0.626**
Eastern-Europe	0.216	0.327	0.161				0.220	0.286	0.176
Asia and Middle East	0.815	0.233	1.221***				0.793**	0.227	1.265***
Africa	0.727	0.261	-0.904				0.741	0.231	-0.856
Latin-America	0.042	0.092	-0.192				0.057	0.023	-0.218
Male	-1.551***	-2.345***	-1.680***	-1.838***	-2.672***	-1.863***	-1.688***	-2.521***	-1.773***
Born outside Sweden	0.069	0.259	-0.004				0.097	0.255	0.047
At least one Swedish parent	0.108	0.416	-0.096				0.118	0.216	-0.002
Two-parent family	-0.448**	-0.249	0.291	-0.235	-0.559	-0.017	-0.327**	0.124	0.371
Parental Income (log)	-0.112	-0.418	0.202	-0.129	0.922**	0.211	-0.139	-0.185	0.150

Table 4d: **Multinomial-logit regression results on the type of education chosen, with base category health sciences, for second-generation immigrants, native Swedes, and the pooled sample**

	Second-generation immigrants				Native Swedes Twin Group				All (n=1941)			
	General education	Human sciences	Social sciences	Technical sciences	General education	Human sciences	Social sciences	Technical sciences	General education	Human sciences	Social sciences	Technical sciences
Parental Age					-0.034	0.709*	-0.029	0.251	0.096	0.322	0.064	0.146
Parental education												
Upper-secondary	0.269	0.765	0.317	0.248					0.083	0.770	0.186	0.127
University degree	0.449	1.423***	0.522	0.232					0.473*	1.998***	0.547*	0.342
Swedish									0.030	0.139	0.398	0.191
Western countries	0.847**	0.709	0.940**	0.331					0.789**	0.630	0.909**	0.282
Eastern-Europe	0.911*	1.021*	0.855*	0.695					0.892*	0.959*	0.849*	0.673
Asia and Middle East	-0.032	-0.614	0.375	-0.847*					-0.187	-0.752	0.286	-0.979**
Africa	0.551	0.085	-1.080	-0.176					0.542	0.031	-1.056	-0.200
Latin-America	0.342	0.392	0.108	0.300					0.364	0.330	0.089	0.307
Male	1.102***	0.308	0.973***	2.653***	0.672**	-0.162	0.646**	2.509***	0.873***	0.039	0.787***	2.561***
Born Outside Sweden	0.235	0.425	0.162	0.166					0.291	0.449	0.243	0.194
At least one Swedish parent	-0.421	-0.114	-0.625	-0.529					-0.445	-0.347	-0.566*	-0.564*
Two-parent family	-0.897**	-0.698	0.158	-0.449	-0.485	-0.809*	-0.267	-0.250	-0.698**	-0.675**	-0.247	-0.371
Parental Income (log)	0.725**	0.420	1.040***	0.838**	0.482	1.533***	0.821**	0.610	0.526**	0.480	0.815***	0.665**

Table 4e: Multinomial-logit regression results on the type of education chosen, with base category service-related education for second-generation immigrants, native Swedes, and the pooled sample

	Second-generation immigrants (n=931)					Native Swedes Twin Group (n=1010)				
	General education	Human sciences	Social sciences	Technical sciences	Health sciences	General educ vs	Human sciences	Social sciences	Technical sciences	Health sciences
Parental age (maximum if 2)						0.085	0.829*	0.090	0.370	0.112
Parental education										
Upper-secondary	-0.358	0.138	-0.311	-0.379	-0.627					
University degree	0.825	1.799***	0.898	0.608	0.376					
Swedish										
Western countries	0.616	0.478	0.709	0.101	-0.231					
Eastern-Europe	0.194	0.304	0.138	-0.023	-0.717					
Asia and Middle East	-0.053	-0.636	0.354	-0.868	-0.022					
Africa	0.031	-0.435	-1.600	-0.696	-0.520					
Latin-America	0.969	1.019	0.735	0.927	0.627					
Male	0.067	-0.727*	-0.062	1.618***	-1.034*	0.869**	0.035	0.844**	2.707***	0.197
Born outside Sweden	-0.463	-0.274	-0.537	-0.533	-0.699					
At least one Swedish parent	0.016	0.324	-0.187	-0.091	0.438					
Two-parent family	-0.778*	-0.579	-0.039	-0.330	-0.440	0.201	-0.124	0.419	0.436	0.686
Parental Income (log)	0.585	0.280	0.899**	0.697	-0.140	-0.077	0.974	0.263	0.052	-0.559

"continued"

Table 4e: Multinomial-logit regression results on the type of education chosen, with base category, service-related education

	All (n=1941)				
	General education	Human sciences	Social sciences	Technical sciences	Health sciences
Parental age (maximum if 2)	-0.164	0.061	-0.198	-0.115	-0.261
Parent education					
Upper-secondary	0.235	0.922**	0.339	0.280	0.153
University degree	1.212***	2.737***	1.287***	1.081***	0.739*
Swedish	-0.139	-0.031	0.228	0.021	-0.170
Western	0.539	0.380	0.658	0.032	-0.251
Eastern-Europe	0.074	0.140	0.030	-0.145	-0.819
Asia	-0.247	-0.812	0.226	-1.039*	-0.060
Africa	-0.064	0.575	-1.662	-0.806	-0.606
Latin-America	1.016	0.981	0.740	0.958	0.651
Male	0.416*	-0.418	0.330	2.104***	-0.456
Born outside Sweden	-0.620	-0.462	-0.700*	0.717*	-0.911*
At least one Swedish parent	-0.188	-0.090	-0.309	-0.307	0.258
Two-parent family	-0.263	-0.241	0.187	0.063	0.434
Parental Income (log)	0.141	0.095	0.430	0.280	-0.385

Table 4f: Partial multinomial-logit regression results on the type of education chosen, using decomposed income, all base categories

	Second-generation immigrants (n=931)					Native Swedes Twin Group (n=1010)				
base category human sciences	General education	Human sciences	Social sciences	Technical sciences	Health sciences	General education	Human sciences	Social sciences	Technical sciences	Health sciences
father's labour-income	-0.004					-0.012				
mother's labour-income	0.012					-0.022				
asset-income	-0.017					-0.073				
welfare-income	0.292					0.414				
other income	0.021					0.014				
base category social sciences										
father's labour-income	-0.010	-0.006				0.001	0.012			
mother's labour-income	-0.003	-0.016				-0.019	0.003			
asset-income	-0.025	-0.008				-0.043	0.030			
welfare-income	0.016	-0.276				0.431	-0.982			
other income	-0.001	-0.022				0.032	0.018			
base category technical sciences										
father's labour-income	-0.003	0.001	0.007			-0.005	0.007	-0.006		
mother's labour-income	0.002	-0.010	0.005			0.013	0.036**	0.032**		
asset-income	-0.016	0.001	0.009			-0.028	0.044*	0.015		
welfare-income	0.185*	-0.107	0.169			0.321	-0.093	-0.110		
other income	-0.001	-0.002	0.001			0.045*	0.031	0.013		
base category health sciences										
father's labour-income	0.030*	0.034*	0.041*	0.033*		0.026*	0.038**	0.025*	0.031**	
mother's labour-income	0.007	-0.005	0.011	0.005		0.027	0.049**	0.046**	0.014	
asset-income	0.081	0.098	0.106	0.097		-0.043	0.029	-0.001	-0.015	
welfare-income	-0.012	-0.304	-0.028	-0.197		0.472	-0.942	0.041	0.151	
other income	-0.021	-0.042	-0.020	-0.021		0.054	0.040	0.022	0.009	
base category service related education										
father's labour-income	0.020	0.024	0.030	0.023	-0.010	-0.008	0.004	-0.009	-0.003	-0.034
mother's labour-income	-0.003	-0.015	0.001	-0.004	-0.010	0.040	0.062**	0.059**	0.027	0.013
asset-income	0.004	0.021	0.029	0.020	-0.077	-0.059	0.014	-0.015	-0.030	-0.015
welfare-income	0.008	-0.284	-0.007	-0.177	0.020	0.984	-0.430	0.553	0.663	0.512
other income	0.015	-0.006	0.016	0.015	0.036	0.022	0.008	-0.010	-0.023	-0.032

[a] Other control variables are shown on tables 4a-e

"continued"

Table 4f: Partial multinomial-logit regression results on the type of education chosen using decomposed income, all base categories

	All (n=1941)				
base category human sciences	General education	Human sciences	Social sciences	Technical sciences	Health Sciences
father's labour-income	-0.002				
mother's labour-income	0.006				
asset-income	0.001				
welfare-income	0.389				
other income	0.021				
base category social sciences					
father's labour-income	-0.003	-0.001			
mother's labour-income	-0.012	-0.019			
asset-income	-0.032	0.007			
welfare-income	0.044	-0.344			
other income	0.013	-0.007			
base category technical sciences					
father's labour-income	-0.006	-0.004	-0.002		
mother's labour-income	0.007	0.001	0.019*		
asset-income	-0.019	0.020	0.013		
welfare-income	0.204**	-0.184	0.159		
other income	0.020	-0.001	0.006		
base category health sciences					
father's labour-income	0.024**	0.027**	0.030***	0.029*	
mother's labour-income	0.014	0.008	0.027*	0.007	
asset-income	-0.015	0.025	0.018	0.005	
welfare- income	0.035	-0.353	-0.009	-0.169	
other income	0.011	-0.010	-0.002	-0.009	
base category service related education					
father labor income	-0.004	-0.002	-0.001	0.002	-0.029*
mother labor income	0.004	-0.003	0.016	-0.003	-0.011
assetincome	-0.032	0.008	0.001	-0.012	-0.017
welfare income	0.045	-0.344	0.434	-0.160	0.001
other income	0.010	-0.011	-0.003	-0.009	-0.001

[a] Other control variables are shown on tables 4a-e

Appendix E. Hausman test results.

The Hausman test for IIA was conducted on the six choices of education type: general (or non-occupational); humanities (art, theater, religion); social sciences (economics, accounting, etc.), technical; health sciences (medicine, nursing, psychology, etc.) and service- oriented. The Hausman test-statistic for the null hypothesis (IIA holds) was $H_0 : \hat{\beta}_R - \hat{\beta}_U = 0$, and the alternative hypothesis was $H_1 : \hat{\beta}_R - \hat{\beta}_U \neq 0$, the IIA does not hold. The test-statistics can be written as

$$H^{IIA} = (\hat{\beta}_R - \hat{\beta}_U)' \left[\widehat{Var}(\hat{\beta}_R) - \widehat{Var}(\hat{\beta}_U) \right]^{-1} (\hat{\beta}_R - \hat{\beta}_U) \tag{1}$$

which if IIA is true, is asymptotically distributed as chi-square with degrees of freedom equal to the rows in the restricted model . Significance values of H^{IIA} indicate that the IIA assumption has been violated. Hausman and McFadden (1984) note that test-statistics can be negative when the difference-matrix is not positive semi-definite in finite-sample applications, and conclude that a negative value for H^{IIA} is evidence that IIA holds. The test-statistic for service-oriented education was 4.032 and for technical education was 7.892. These two positive values are smaller than any reasonable critical value for a $\chi^2_{(5)}$ which means that when we add service-oriented or technical education to the other choices, we cannot reject the hypothesis that IIA holds, i.e., the types of upper-secondary education were genuine choices independent from each other.

Testing for Independence from Irrelevant Alternatives (IIA test)

	Service oriented	Health sciences	Technical sciences	Social sciences	Human sciences	General education
$\chi^2_{(5)}$	4.032	-269.845	7.892	-15.826	-314.149	-687.557
Skipped observations	94	172	604	514	149	693
P-values	0.750	0.900	0.950	0.975	0.990	0.995
Tabulated Values of $\chi^2_{(5)}$	6.63	9.24	11.07	12.83	15.09	16.75

Early Labour-Market Experiences of Second-Generation Immigrants in Sweden

Kerem Tezic and Ali C. Tasiran
Department of Economics, Göteborg University

April 2003

Abstract

The first experiences of new entrants into the labour-market may be decisive for later developments in their life. Their success in the labor-market can be measured by their time spent in employment and their returns from human-capital investment such as wages. We therefore focused on four different types of transitions into the labor-market: The first two are from compulsory and post-compulsory education to various competing states; the last two are from non-employment to work after compulsory or post-compulsory education.

Our results reveal that parental resources not only affected second-generation immigrants' continuing education but also their later labour-market success. For all young people, regardless of their ethnic background, parental capital in the form of educational attainment, occupation and income was vital. Ethnic capital and neighboring characteristics such as other intergenerational transmission-channels were also important. The study verifies that finding job is difficult for second-generation immigrants, especially those coming from Africa, Latin America and Middle East. The significance of an unobserved-heterogeneity parameter may indicate discrimination.

Keywords: Second-generation immigrants, intergenerational transmissions, duration-models, competing-risks, unobserved heterogeneity.

J.E.L. Classification: C23, C24, C41.

Contents

1 Introduction

One might think that "second-generation" immigrants, born in the new country to immigrant parents, would be better off than their parents in terms of economic and social integration since they go to school in the new country, and thus have better opportunities to learn the language, make friends with the native population, and develop country-specific human capital. But what about the empirical facts? What kind of early labour-market experiences have second-generation immigrants had after finishing compulsory education? This paper attempts to find answers to these questions.

Until the early 1970s, immigration to Sweden was mostly labour-force migration from other European countries. But since then, political upheavals elsewhere and immigration policy have caused the character of immigration to change. The number of refugees and tied movers has increased, regardless of the state of the labour market.

Parallel with deterioration in the labour-market during the 1970s, unemployment increased in Sweden among immigrants. During the 1980s the employment-intensity among immigrants continued to fall, despite improvements in the labour-market from the mid-1980s onwards, see Ohlsson (1975), Ekberg and Gustafsson, (1995). Immigration reached its highest levels during the extreme boom of the late 1980s, but with the recession at the beginning of the 1990s, unemployment among immigrants increased further. Nevertheless, immigrants have stayed in Sweden, built families, and raised children conventionally called "second-generation immigrants".

Despite the fact that studies about immigrants are abundant, they have mainly concerned economic assimilation based on cross-sectional studies. There are few studies about second-generation immigrants, either in the United States or in Europe, perhaps because of lack of data. The ones that exist are again based on cross-sectional data, which do not permit to making strong causal interpretations. Longitudinal research has certain advantages, such as measuring change and establishing temporal order, but there are offsetting difficulties since it requires more advanced techniques.

The purpose of this paper is to analyze the early labour-market experiences of second-generation immigrants in Sweden based on a longitudinal data set (LINDA) for the period 1991-2000. The early experiences of new entrants into the labour-market may help determine developments later in their life. Youths 16-25 may encounter many important changes, such as completing their education, getting their first job, building a family and becoming a parent. All of these factors will affect their later labour market behaviour and success, such as time spent in employment, and returns from human-capital investment such as wages.

The study is focused on four different types of transitions into the labour-market: The first two are from compulsory and post-compulsory education to various competing states; the last two are from non-employment to work after compulsory or post-compulsory education. The first transition was modelled by fitting a multinomial-logit model since the waiting-time after compulsory education was the same for all individuals and there was no time-dependency. The transition from post-compulsory education to work, non-employment or the military was modelled in a competing-risks hazard-framework by using semi-parametric Cox and parametric duration-dependence distributions. We also look to the transitions to work after non-employment after compulsory and post-compulsory education. Based on the results of an information-matrix test, we fitted hazard-models to both of the transitions —transition from post-compulsory education to various states and transition from non-employment after post compulsory education to work— taking into account inter-individual heterogeneity.

The data we use differ from other second-generation studies in many aspects. We use yearly observations from the Swedish register based Longitudinal Individual Data (LINDA). The data set is rich about the socioeconomic variables. Besides the heterogenous feature of the second-generation immigrants are taken into account by grouping them into similar geographical units. The second-generation immigrants are all from the same age cohorts. They are 16/17 in 1991 and 25/26 in 2000. This prevents us from average age differences across ethnic groups which can affect the results. A twin native group is also randomly chosen which helps us to make reliable comparisons.

Our results reveal that parental resources not only affected second- generation immigrants' continuing education but also their later labour-market success. For all young people, regardless of their ethnic background, parental capital in the form of educational attainment, occupation and income were vital. Ethnic capital and neighboring characteristics such as other intergenerational transmission-channels were also important. The study verifies that finding job is difficult for second-generation immigrants, especially those coming from Africa, Latin America and Middle East. The significance of an unobserved-heterogeneity parameter may indicate discrimination.

Next section discusses previous research about second-generation immigrants, both in the US and in Europe. Section 3 discusses the theoretical issues while Section 4 describes the data and discusses sample-selection issues are described. Section 5 analyses the modelling issues. Section 6 presents the results and finally last section offers a conclusion.

2 Previous research about second-generation immigrants

2.1 Studies in the United States

Europe and the United States, and especially Sweden and United States, differ in many respects, including differences in social policies and provisions of the welfare state. An important goal of Swedish immigrant policy has been to assure the equality of immigrants, who should have the same employment and income opportunities, although this hasn't always worked out in practice. Even post-compulsory education is tuition-free for all in Sweden and there is a fairly generous transfer system, including low-cost medical care. Nevertheless, both US studies and those from other European countries may be relevant to Swedish experience. Using US census-data to analyze contemporary immigrant-flows, Chiswick (1977), Carliner (1980) and Borjas (1993 and 1994) come to different conclusions about the experience of the sons of immigrants in the US labour market. Using data from the 1/1000 sample of the 1970 Census of Population, Chiswick found that the sons of immigrants had a 5 percent earnings-advantage over white male of native-born Americans. Both Chiswick and Carliner concluded that characteristics associated with positive selection in the immigrant-population had been transmitted to their children. Analyzing the inter-generational mobility experienced by immigrants in the context of economic model of immigration, Borjas (1993) questioned Chiswick's and Carliner's results, since their studies were based on cross-sectional data. Borjas examined the earnings of immigrants and second-generation immigrants across decennial censuses from 1940 to 1970. Assuming that first-generation immigrants in the 1940 census were the parents of second-generation immigrants in the 1970 census, he found that, although there was some regression toward the mean, the average earnings of the second-generation of any given ethnic group were strongly influenced by the earnings of the corresponding first-generation group. He postulated that the immigrant characteristics in the source countries which determine migrant-selection had introduced differentials in skills between ethnic groups which resulted in persistent earnings-differentials. On the other hand, Card, DiNardo and Estes (1998), found that, despite the fact that an increasing fraction of today's second-generation immigrants were the grand children of the formerly lowest-paid immigrant groups, they now tended to have higher wages than long-time natives.

2.2 Studies in Europe

2.2.1 Educational attainment

In Europe, most of the research about second-generation immigrants has been done in Germany, the Netherlands, Denmark, and Sweden. Since understanding the economic integration of ethnic communities requires understanding of the human capital process, educational attainment has been one of the main research topics.

Gang and Zimmerman (2000), using information for 1984 from the German Socioeconomic Panel, explored three aspects of educational attainment: total years of education, schooling-level, and vocational training. Their sample consisted of 4678 individuals, 17-38 years old, 3895 of whom were Germans while 783 were second-generation immigrants. They found that parents' education had no effect on the educational attainment of their children, but that ethnicity mattered. They concluded that the effects of the parents' education had been completely depreciated, and that the inter-generational transfer of human capital disappeared with the shock of immigration.

Riphahn (2001), using five pooled microcensuses (a one percent random sample of the population) from 1989, 1991, 1993, 1995 and 1996, estimated completed degrees and school-attendance of second-generation immigrants in Germany. The sample consisted of 55,570 natives and 3,627 second-generation immigrants. By focusing on the development over time using an ordered-probit model, she found that second-generation immigrants did not assimilate to the native education-level and that the attainment-lag did not diminish over time. She interpreted this by pointing to the changing country of origin of second-generation immigrants.

Van Ours and Veeneman (2001) explored the extent to which differences in educational attainment between second-generation immigrants and natives exist in the Netherlands. Using a 1998 nationwide survey where the second-generation immigrants consisted of Turks, Moroccans, Surinamese, and Antilleans, they analyzed the probabilities of continuing education and of the level of education attained, with binary-probit and ordered-probit models, respectively. Their main conclusion was that, controlling for parental education-level, differences between second-generation immigrants and natives largely vanished.

Nielsen, Rosholm, Smith, and Nusted (2001), using Danish panel-data originating from administrative registers covering the period 1985-1997, focused on the probability of obtaining a "qualifying education", defined as education lasting at least 18 months beyond the compulsory level. Their results indicated that, for males the number of years since the parents immigrated to Denmark had a statistically significant positive effect, whereas for females the parents' income was an important factor affecting them positively. The education of parents was found to be statistically significant only for ethnic Danes, while having parents with several years of labour-market experience had a statistically significantly positive effect for all groups.

Österberg (2000), using the Swedish Income Panel, where 98 percent of the second-generation immigrants had European background, analyzed level of educational attainment using an ordered probit model. She found that the parent education had a positive effect on the child's educational level which decreased the negative effect from belonging to certain ethnic groups.

2.2.2 Early labour-market experiences

Van Ours and Veeneman (2002) using the same Netherlands data as before, investigated the early-labour market experiences of second-generation immigrants. Using a simple probit-model to model of the probability that individuals had a job while still at school, they found that it was an age-related phenomenon, rather than depending on the level of education or on "neighborhood" characteristics. They also focused on the probability of having a job conditional on being out of school. The results of their bivariate-probit model pointed to the importance of the level of education for females in finding a job. For both males and females, Turkish and Moroccans had greater difficulty in finding a job. Again using a bivariate-probit model, they found that the probability of having a steady job increased with work experience, while for males education mattered. Turkish and Surinamese females were less likely to hold a steady job. Regarding the amount of hourly wages, taking possible selectivity into account, they found that education had a statistically significant effect for males, while ethnicity did not matter.

Nielsen *et al.* (2001), focused on three aspects of the early-labour market experiences of second-generation immigrants in Denmark: time from leaving school until the first ordinary job; duration of the first unemployment spell; and hourly wages. In terms of data-structure and estimation-methods, this study was the closest to our work in this paper. Using a proportional-hazards model with a piecewise constant baseline-hazard for the waiting time, they found that parental capital had strong effects on entry into the labour market for females, whereas ethnic capital and "neighborhood" variables were also important for males. The education of parents had a negative effect, which they interpreted as a perverse effect on the probability of taking a first job.

Ekberg (1997), conducted the first study about second-generation immigrants in Sweden. His descriptive analyses, based on 1994 labour-market data from Statistics Sweden, showed that most were born before 1970, of European origin, had integrated well into the labour-market with hardly any difference between them and the native Swedes. The situation was completely different for those born after 1970, for whom there were also differences between ethnic groups. Månsson and Ekberg (2000), found a similar result.

Using Swedish survey data from 1988, Schröder and Wilhelmsson (1998) analyzed labour-market position of second-generation immigrants seven years after graduation, where labour market position was categorized as; working; unemployed; studying; out of labour force; or other. Estimation results based on a multinomial-logit model indicated that second-generation immigrants, especially those coming from Asia, Africa, or Latin America, had a disadvantage in the labour-market. Second-generation immigrants who migrated after compulsory school had a higher probability of being unemployed compared to those who immigrated earlier or were born in Sweden, and the differences persisted even other variables such as language proficiency and other Sweden-specific human capital were controlled.

Ekberg and Rooth (2002), divided 1998 data from National Labour Market Board

(AMS) and Statistics Sweden into two subsets and focused on the probability of being unemployed for 25-40 years old individuals who were part of the labour-force and on their earning levels. Probit-regression for the first outcome (being unemployed) and OLS for the second, indicated that the outcomes differed greatly between different groups of second- generation immigrants, and compared to native born Swedes. The pattern was similar to that in the parent generation. The outcome was more favorable if one parent was born in Sweden.

3 Determinants of early labour-market experiences and related hypotheses

It is not clear what hypotheses to use for the early labour-market experiences of second-generation immigrants. Early works point to the inter-generational transmission process, whereby the weak labour-market position of first-generation immigrants would be transmitted to their children through various channels: parental capital, or direct effects from the parents; ethnic capital; and neighborhood effects. This was also the main point of Coleman (1988) where he asserted that both social capital in the family and social capital in the community play roles in the creation of human capital in the rising generation.

However, one has to be careful in defining these entities. For example, family background also includes financial capital and human capital, approximately measured by the parents' education which contributes to the cognitive environment of the child. Family social capital is different from either of these. Family human capital may be irrelevant if it is not complemented by social capital embodied in family relations, which in turn can depend on the number of parents present, the number of siblings in the family, parental time devoted to the children, etc.

In immigration literature, Borjas (1995) extended the notion of social capital to include ethnic capital and its interaction with neighborhood effects. Ethnic capital may include linguistic skills and external effects from the average human capital of the ethnic group, as a whole while neighborhood effects include the negative impact on youths growing up in poor neighborhoods. According to Borjas, residential segregation and ethnic external effects may be linked since ethnic capital includes the socioeconomic background of the neighborhood.

In our data parental characteristics are rich allowing us to disaggregate the sources of family income, such as father's labour-income, mother's labour-income, family asset-income and family welfare-income. Such an approach, used also by Hill and Duncan (1987), enabled us to test various human capital hypotheses such as the role-model hypothesis and the welfare-dependency hypothesis, which have implications for the success and social integration of the child. To test neighborhood effect, we constructed a variable called "ethnic concentration", see Nielsen *et al.* (2001), which is the percentage of first-generation immigrants in the municipality.

4 The data and sample-selection

The data-set used in this study is the 1991-2000 panel of the register-based Longitudinal Individual Data-set (LINDA), database of about 300,000 individuals and their household members, see Edin and Frederiksson (2001). The core of the data is the annually-available Income Registers and Population Censuses. The database contains individuals' socioeconomic characteristics such as education-related variables and labour-market related variables, such as employment-status, labour-income, etc. Family members are only included in the sample as long as they stayed in the family.

Our sample, following Kossoudji, (1989), consisted of second-generation immigrants who were either born in Sweden, or immigrated before six years of age and who were 16-17 year old in 1991. We included children who immigrated before age six since they started school with native Swedes and had the opportunity to learn the language thoroughly of the migrated country at a very early period of their life cycle. In other words, they have access to the migration country's specific human capital. Those who are above 18 years of age were not included since most leave their parental home and it is then impossible to identify whether or not they are second-generation immigrants. In 1991, there were 1106 second-generation immigrants and we followed these individuals until they reached age 25-26 in 2000. The geographical origin of second-generation immigrants was determined from the father's country of birth (or if only the mother was foreign-born, it was determined from hers). We also randomly chose a Swedish statistical twin —control— group consisting of 1106 similarly aged individuals. We followed these individuals until they reach 25-26 years of age by 2000.

We have analyzed four sets of labour-market experiences. The first was the transition from compulsory school to various states: continuing their education, finding a job, or non-employment. Among the 2212 individuals in both groups, 1941 continued their education while 162 found a job and 109 were non-employed during the first year after compulsory education (Table 1 in the Appendix). None of the 109 individuals returned to school during the following 10 years of the study period while the overwhelming majority of the 162 individuals who found a job continued to work. Next we analyzed the transition from continuing education to work, non-employment, or the military. Finally we looked at transitions from non-employment to work or the military either after compulsory education or after continuing education.

An inconvenience with such data registered annually is that alternative outcomes can occur in the same year. To deal with this, we gave priority to educational status. After the completion of compulsory education, sometimes we had to look at many variables in the data in order to decide what happened next and after that.

Among the 1106 second-generation immigrants, just over half were males, to just under half of the native Swedes (Table 2 in the Appendix). Only about a third of the second-generation immigrants had been born outside Sweden. Three quarters lived in two-parent families, essentially the same as native Swedes. The number of siblings was slightly higher for the immigrants. The average age of the "head of household" (the father if two parents were present) was almost the same for second-generation immigrants and for native Swedes (about 46).

We classified parent's education according to the highest level attained by at least one parent: compulsory level; high-school level; or university degree. Considerably more immigrant parents had completed only ninth grade (27.7 percent) than native Swedish parents (15.5 percent). On the other hand, considerably more native Swedish parents (37.1 percent) than immigrant parents (28.8 percent) had a college or university degree. The percentages who had completed high-school were more similar though slightly higher for native Swedes.

Nearly half of the second-generation immigrants were of Nordic origin, most of them Finnish. The rest can be classified in six groups coming from industrialized western countries, including USA, Australia, Canada and the EU; Eastern Europe; Middle-East; Asia; Africa; and Latin America.

The ethnic concentration, was only slightly higher for second-generation immigrants than for native Swedes.

Native Swedes had only slightly more total parental income than did second generation immigrants. Native Swedes' father's labour income, mother's labour income, and asset income were all substantially higher, whereas the immigrant's parents' welfare income was somewhat higher, though not enough to make up the difference. A somewhat mysterious category of "other income"[1] was substantially higher for immigrants however.

For the first transition, considering all 2212 second-generation immigrants and native Swedes combined, of the 109 who transited to non-employment after compulsory education, a larger proportion (≈ 67 percent) were male, whereas of those who transited to work 46.2 percent were male (Table 3 in the Appendix). Those who transited to non-employment were also disproportionately born outside of Sweden (32 percent, versus 15-18 percent of those who transited to work or continued their education). A higher proportion of those who continued their education came from two-parent family (77 percent, versus about 60 percent of those who transited to work or to non-employment). A higher proportion also had at least one Swedish parent (74 percent, versus 58.6 percent of those who transited to work and 42.2 percent of those who transited to non-employment). The parents of those who continued their education had more education themselves, followed by those who transited to work. Those who continued their education also had substantially more mothers and fathers who were working, followed again by those who transited to work. Those who transited to non-employment had parents with the lowest educations and fewer of whom were working. With respect to geographical origin native Swedes tended disproportionately to continue their education or go to work. Second-generation Nordic immigrants tended disproportionately to go to work. Those who continued their education (followed by those who transited to work) also had the highest father's labour income, mother's labour income, and asset income. Conversely those who transited to non-employment (followed by those who transited to work) had the highest welfare income and the highest "other income".

Of the 1941 native Swedes and second-generation immigrants who initially continued their education, about a third later transited to work and another third to non-employment while some 262 went to military and the rest continued in school (Table 4 in the Appendix).

[1] All other income is the sum of all income from sources other than parents' labour income, asset income, and parents' welfare income, averaged over the period when the child was 15-16 years old.

All those who went to military were male, whereas females seem to have tended slightly towards non-employment. Proportionally more of those who transited to non-employment were also born outside of Sweden. Those from two-parent families tended disproportionately to continue their education or to go to work. Those with at least one Swedish parent went disproportionately into the military, followed by continuing their education or going to work. Again those whose parents had highest educations tended to continue their educations, while those whose parents had the lowest educations tended disproportionately to transit to non-employment. Those whose mothers were working tended again disproportionately to continue their education or to go to work as did those whose father were working. With respect to geographical origin native Swedes tended disproportionately to go into the military, followed by going to work or continuing their education. Nordics tended slightly towards non-employment and away from continuing their education. Those from Western countries tended disproportionately towards continuing their education and away from the military. Those from the Middle-East, Africa and Latin America tended disproportionately towards non-employment and away from work. Those who continued their education had the highest father's labour income, followed by those who went to work and those who joined military. Those who continued their education also had the highest mother's labour-income, followed by those who went to work. Those who transited to non-employment had the highest parents' asset and welfare income (probably two different groups).

5 Statistical modelling

Transiting from compulsory education to various states is analyzed using multinomial-logit regression —since there was no time-dependency— , the later transitions from continuing education or from initial non-employment after compulsory-education to various states are modeled in a competing-risks framework, and transition from non-employment after post-compulsory education to work is modeled in a single-destination framework. This means not only that we analyzed the duration of the non-occurrence of the event of interest —except transitions from compulsory education—, but we also distinguished between different types of transitions. We assumed a priori that the occurrence of each type of transition had a different causal structure. The same covariates might be relevant but each transition could have an independent set of parameters, see Kalbfleisch and Prentice (1980).

When time T is continuous and measured precisely so that there are no ties such continuous time survival procedure could be adopted: let x be a vector of covariates. In a competing-risks framework, a cause-specific or type-specific model can be represented by

$$h_j(t;\mathbf{x}) = \lim_{dt\to 0} dt^{-1} P(t \leq T < t+dt, \quad J=j \mid T \geq t,\ \mathbf{x}) \tag{1}$$

where $j = 1, ..., m$ and $t > 0$; $h_j(t;\mathbf{x})$ denotes the instantaneous risk of experiencing a transition of type j in the time interval $(t \leq T < t+dt)$ given that no transition occurred before $T = t$. The overall hazard-rate can be obtained by summing the transition specific hazard rates, that is,

$$h(t;\ \mathbf{x}) = \sum_{j=1}^{m} h_j(t;\ \mathbf{x}) \tag{2}$$

The overall survivor function is

$$S(t;\ \mathbf{x}) = P(T > t \mid \mathbf{x}) = \exp\left(-\int_0^t h(u;\ \mathbf{x})\ du\right)$$

if we substitute in the transition specific hazard rates then we obtain

$$\begin{aligned} S(t;\ \mathbf{x}) &= \exp\left(-\int_0^t \sum_{j=1}^{m} h_j(u;\ \mathbf{x})\ du\right) \\ &= \prod_{j=1}^{m} \exp\left(-\int_0^t h_j(u;\ \mathbf{x})\ du\right) \end{aligned} \tag{3}$$

The density-function for the time until a type j transition is then

$$\begin{aligned} f_j(t;\mathbf{x}) &= \lim_{dt\to 0} dt^{-1} P(t \leq T < t+dt,\ J=j \mid \mathbf{x}) \\ &= h_j(t;\ \mathbf{x})\ S(t;\ \mathbf{x}) \end{aligned} \tag{4}$$

It must be noted, however, that $f_j(t;\ \mathbf{x})$ is not the density-function of the duration-time. In particular

$$\int_0^{\infty} f_j(t;\ \mathbf{x})\ dt = P(J = j \mid \mathbf{x}) = \boldsymbol{\pi}_j(\mathbf{x}) \tag{5}$$

where $\boldsymbol{\pi}_j(\mathbf{x})$ is the probability of transition into the jth state $j = 1, ..., m$, given the covariate-vector $\mathbf{x}$, with the relationship

$$\sum_{j=1}^{m} \boldsymbol{\pi}_j(\mathbf{x}) = 1 \tag{6}$$

If $t_{j1} < t_{j2} < ... < t_{jn_j}$ represents the n_j uncensored durations until the transition j, then the likelihood-function may be rewritten as

$$L = \prod_{j=1}^{m} \prod_{k=1}^{n_j} h_j(t_{jk};\ \mathbf{x}_{jk}) \prod_{i=1}^{n} S_j(t_i; \mathbf{x}_i) \tag{7}$$

where $\mathbf{x}_{jk}$ is the covariate of an individual with the observed noncensored duration t_{jk} and

$$S_j(t_i;\ \mathbf{x}_i) = \exp\left(- \int_0^{t_i} h_j(u;\ \mathbf{x}_i)\ du \right) \tag{8}$$

The likelihood-function may be divided into the product

$$L = \prod_{j=1}^{m} L_j \qquad \text{with} \qquad L_j = \prod_{k=1}^{n_j} h_j(t_{jk};\ \mathbf{x}_{jk}) \prod_{i=1}^{n} S_j(t_i;\ \mathbf{x}_i) \tag{9}$$

The L_j-factors may be further rearranged as

$$L_j = \prod_{i=1}^{n} [h_j(t_{jk};\ \mathbf{x}_{jk})]^{\delta_{ij}} \quad S_j(t_i;\ \mathbf{x}_i) \tag{10}$$

$$\text{with } \delta_{ij} = \begin{cases} 1 \text{ if for individual } i \text{ a transition to state } j \text{ occurs at time } t_i \\ 0 \text{ otherwise} \end{cases} . \tag{11}$$

The log-likelihood function $\ln L = \sum_{j=1}^{m} \ln L_j$ can be maximized separately for each transition type $j = 1, ..., m$, given that the transition-specific hazard rates $h_j(t \mid \mathbf{x})$ are dependent upon the parameter-vector θ_j, where the θ's have no common components. In particular, a parametric model $h_j(t; \mathbf{x}, \theta_j)$ can be specified for the type-specific hazards (see, Cox and Oakes, 1984).

For the transitions from continuing education to various states, as well as later transitions from non-employment after compulsory or post-compulsory education, we first plotted the smoothed non-parametric hazard functions against time and the hazard functions displayed non-monotonic curves (Figures 1-4 in the Appendix). The log-logistic model was ideal in catching the turning-points in these cases. We then fitted log-logistic model[2] in the competing-risks framework discussed above since in the case fitting a proper parametric distribution function, one can obtain more efficient estimates than estimates of a semi-parametric model. In addition to that, we also fitted Cox proportional-hazards model.[3] Since Cox's method does not require some particular probability distribution to represent survival times, is considerably more robust. It can also accommodate both discrete and continuous measurement of transition times. The cause-specific hazard functions mentioned above, can be modeled by using Cox model in the following way:

$$h_j[t; \mathbf{x}_i] = h_{0j}(t) \exp[\mathbf{x}_i' \beta_j], \qquad j = 1, \ldots\ldots, m, \tag{12}$$

where $\mathbf{x}_i$ is a vector of covariates, h_{0j} and β_j are the baseline hazards and the regression coefficients respectively which vary arbitrarily over the m transition types. As before, let $t_{j1} < t_{j2} < \ldots < t_{jn_j}$ represents the n_j uncensored durations until the transition j, $j = 1, \ldots\ldots, m$. The corresponding partial likelihood is

$$L(\beta_1, \ldots\beta_m) = \prod_{j=1}^{m} \prod_{i=1}^{k_j} \frac{\exp[x_i \beta_j]}{\sum_{l \in R(t_{ji})} \exp[\mathbf{x}_i' \beta_j]} \tag{13}$$

the arbitrary baseline hazard function has been eliminated and the resulting likelihood can be used for inferences about $\beta_j' s$.

The modelling issues above rely on the implicit assumption that the exogenous variables were measured without any error and that there were not any omitted variables in the model. In other terms, there was an implicit assumption that the error term in the model had white noise characteristics. If we have any omitted variable in the model the omission of such an effect can introduce important biases

[2]The log-logistic hazard function is $h(t) = \frac{\gamma\lambda(\gamma t)^{\gamma-1}}{1+(\gamma t)^{\gamma}}$ where $\lambda = \exp[-(\beta_0 + \beta_2 x_2 + \ldots + \beta_k x_k)]$ and the corresponding density function is
$f(t) = \frac{\lambda^{\frac{1}{\gamma}} t^{\frac{1}{\gamma - 1}}}{\gamma(1+(\lambda t)^{\frac{1}{\gamma}})^2}$

[3]When there were many ties, we used the approximation method proposed by Efron (1977).

on the estimates of the parameters of interest (Gourieroux, 1989). The results based on Information Matrix Test detected such unobserved heterogeneity in the case of transitions from continuing education to various states, as well as transition from non-employment after continuing education to work (see Appendix A.). As a result of that we introduced a gamma type unobserved heterogeneity term to the parametric and semi-parametric specifications mentioned above, which is the survival-data analog to regression models with random effects. Such an unobserved heterogeneity is a latent random effect that enters multiplicatively on the hazard function.

The estimated parameters of the transitions from compulsory education to various states, based on multinomial-logit model are reported in Section 6.1. The estimated parameters of the transitions from continuing education to various states, as well as transition from non-employment after continuing education to work, based on Cox proportional-hazard gamma-mixture and log-logistic hazard gamma-mixture[4] specifications are reported in Sections 6.2 and 6.4. The estimated parameters of the transitions from non-employment after compulsory education, based on Cox proportional-hazard and accelerated failure-time log-logistic specifications are reported in Section 6.3.

6 Results

6.1 Transitions from compulsory education

After the completion of compulsory education at age 16-17, everyone in the sample either continued their education, went to work, or transited to non-employment. Table 5 in the Appendix shows the estimated parameters from the multinomial-logit regression for those who transited to work, continued to higher education or transited to non-employment.

Having at least one Swedish parent made the odds of continuing education versus working 1.6 ($\approx \exp(0.496)$,looking at the first column) times higher, and the odds of continuing education versus being in non-employment state 1.8($\approx \exp(0.563)$[5], looking at the second column) times higher. Coming from a two-parent family is

[4]The observed log-logistic hazard with gamma mixture specification becomes $h(t) = \frac{\lambda\gamma(\lambda t)^{\gamma-1}}{1+(\lambda t)^{\gamma}}[1+\delta\log(1+(\lambda t)^{\gamma})]^{-1}$ where δ is the variance of the mixing gamma distribution.

[5]Between two contrasts, reversing the reference category causes reversing the sign.

also statistically significant and made the odds of continuing education versus being non-employed 1.6 times higher. Both results can be interpreted in the light of Coleman's (1988) theory of social capital where he postulates that social capital in the family plays a role in the creation of human capital in the raising generation. Parent education is also a discerning factor. Having parents with university education made the odds of continuing education versus working 2.0 ($\approx \exp(0.680)$ times, and the odds of continuing education versus being in non-employment 1.9 times higher than having parents with secondary education. Again these results confirm the earlier research. It can be interpreted in the light of intergenerational transmission process and are in accordance with those of Österberg (2000). Geographical origin mattered in the case of Asians (continuing education versus working), Middle-Easterns and Africans[6], and East-Europeans (transiting to non-employment versus continuing to higher education). The significance of parental income is at the 10 percent significance limit but when we disaggregated the source of parental income, father's labour income and welfare income were significant but affected the odds in the opposite way. Father's labour income increased the odds of continuing education versus both working and being in non-employment state whereas welfare income decreased it.

6.2 Transitions after continued education

Table 6 in the Appendix shows the results of the estimated parameters based on Cox proportional-hazard gamma-mixture and Log-logistic hazard gamma-mixture models. The results from both models are similar. The signs, as expected are the same in both models. One can see that having parents with higher education decreases the hazard of exit from continuing education to work, to non-employment and to military. This result is, in somewhat similar to that of Nielsen *et al.* (2001), with Danish data, which they interpreted as a "perverse effect" on the probability of entering the first job after leaving the educational system. In our case, the original state is continuing education, which one can interpret in the social-capital framework of the family. Highly-educated parents may be motivate their children, act as role-models.

Compared to the native Swedes, the hazard of transiting to work was higher for

[6]We put Middle-Easterns and Africans in the same category since Middle-Easterns did not transit to working state.

other Nordics, and for Asians, —although not statistically significant— but lower for everybody else, especially the Middle-Easterns, Africans and Latin Americans.

Father's labour-income and mother's labour-income similarly decreased the hazard of exit to work, joining the military and transiting to non-employment.

The statistical significance of the unobserved heterogeneity-parameter in the parametric case may indicate ethnic discrimination since we controlled nearly for all the individual and socioeconomic variables.

6.3 Transitions from non-employment (after compulsory education)

After compulsory education 109 individuals were initially non-employed within a year. The majority of them, 70 transited from non-employment to work, while 27 joined the military. The estimated parameters of the transition from non-employment, based on Cox proportional hazard and accelerated failure time log-logistic models are reported in Table 7 in the Appendix. The signs, as expected, are opposites in both models (since one of them modelizes the hazard, and the other waiting time. If the hazard is high, then transitions occur quickly and survival times are short). Coming from a two-parent family shortened the time in the non-employment state; the ratio of estimated hazard of transiting to work was 2 ($\approx exp(0.688)$, controlling for other covariates. The risk of exit non-employment state is lower for those from the Middle-East, Latin America or Africa. The neighborhood-effect (ethnic concentration) also prolonged the waiting time in the non-employment state. Again this result is similar to that of Nielsen, *et al.* (2001). Having at least one Swedish parent reduced the waiting time and increased the hazard of transiting to military.

6.4 Transitions from non-employment to work (after continued education)

The estimated parameters of the transitions from non-employment after continuing education to work, based on Cox proportional-hazard gamma-mixture and Log-logistic gamma-mixture specifications are reported in Table 8 in the Appendix.

The results are very robust for both specifications. Parents with more education again seem to have children who are more likely to transit to work faster. On the other hand, those from Africa, Middle-East or Latin America were considerably less likely to exit to work. This was also found by Schröder and Wilhelmsson (1998) and by Ekberg (1997 and 2002). Total parental income was statistically significant (subcategories of income were not) and had a positive effect the hazard of transiting to work. As before in both models, the parameter of unobserved heterogeneity was significant. This could indicate to discrimination in the labour market.

7 Summary and conclusions

We have analyzed the early labour-market experiences of second-generation immigrants in Sweden. A register based data set (LINDA) containing information on 1106 16-17 year-old second-generation immigrants and a similar Swedish control group (in terms of age and region), also 1106 individuals were followed for the period 1991-2000. Four types of labour-market outcomes were analyzed: Transitions after compulsory education; after continued education; and from non-employment after compulsory education and after continued education.

The alternative models used showed similar results:

• Parental resources; marital-status, education, occupation, and income, are not only affecting second-generation immigrants' continued education but also their labour-market success. For all young people, regardless of their ethnical backgrounds, parental capital in the form of parents' attained education, occupation and income is vital. Inter-generational transmission channels are thus still important, contrary to what we expected.

• Even after controlling for numerous individual, parental, socioeconomic variables, geographical origin was a major labour-market hindrance for second-generation immigrants from Africa, Middle-East and Latin America.

• The significance of an unobserved-heterogeneity parameter may indicate discrimination.

References

[1] Becker, G. S., and Tomes, N. (1986), "*Human Capital and the Rise and Fall of Families,*" Journal of labour Economics, Vol. 4, No. 3, pp. 1-39.

[2] Bloissfeld, H-P., Hamerle, A., and Mayer, K-U., (1989) "*Event History Analysis*" Hillsdale, NJ.

[3] Borjas, G. J. (1992), "*Ethnic Capital and Intergenerational Mobility,*" Quarterly Journal of Economics, Vol. 107, No. 1, pp. 123-50.

[4] Borjas, G. J. (1993), "*The Intergenerational Mobility of Immigrants,*" Journal of labour Economics, Vol. 11, No. 1, pp. 113-35.

[5] Borjas, G. J. (1994), "*Long-Run Convergence of Ethnic Skill differentials: The Children and Grandchildren of the Great Migration,*" Industrial and labour Relations Review, Vol. 47 (July), pp. 553–73.

[6] Borjas, G. J. (1995), "*Ethnicity, Neighborhoods, and Human-Capital Externalities,*" Vol. 85, No. 3, pp. 365-90.

[7] Carliner, G. (1980), "*Wages, Earnings and Hours of First, Second and Third Generation American Males,*" Economic Inquiry, Vol. 18, pp. 87-102.

[8] Chiswick, B. (1977), "*Sons of Immigrants. Are they at an Earning Disadvantage*?" The American Economic Review, Vol. 67, pp. 376-80.

[9] Coleman, J. S. (1988), "*Social Capital in the Creation of Human Capital,*" American Journal of Sociology, Vol. 94, (Supplement), pp. 95-120.

[10] Cox, D. R., and Oakes, D. (1984), "*Analysis of Survival Data,*", London: Chapman and Hall.

[11] Edin, P-A., and Frederiksson, P. (2001), "*LINDA-Longitudinal Individual Data for Sweden,*" Working Paper 2001:6, Department of Economics, Uppsala University.

[12] Efron, B. (1977), "*The Efficiency of Cox's Likelihood Function for Censored Data,*" Journal of the American Statistical Association, Vol. 76, pp. 312-19.

[13] Ekberg, J. (1994), "*Economic Progress among Immigrants in Sweden,*" Scandinavian Journal of Welfare, Vol. 3, pp. 148-57.

[14] Ekberg, J., and Gustafsson, B. (1995), "*Invandrare på Arbetsmarknaden (Immigrants in the Labour Market),*" SNS förlag, Stockholm.

[15] Ekberg, J. (1997), "*Hur är Arbetsmarknaden för den Andra Generationens Invandrare (How is the Labour Market for Second-Generation Immigrants)*?" Arbetsmarknad&Arbetsliv, Årg 3, No. 1, pp. 5-16.

[16] Ekberg, J., and Rooth, D-O. (2002), "*Unemployment and Earnings for Second-Generation Immigrants in Sweden,*" stencil, Högskolan i Kalmar.

[17] Gang, I. N., and Zimmermann, K. F. (2000), "*Is Child Like Parent? Educational Attainment and Ethnic Origin,*" Journal of Human Resources, Vol. 35, No. 3, pp. 550-69.

[18] Gourieroux, C. (1989), "*Econometrie des Variables Qualitatives,*" Paris: 2^e edition Economica.

[19] Hill, M. S., and Duncan, G. J. (1987), "*Parental Family Income and the Socioeconomic Attainment of Children,*" Social Science Research, Vol. 16, pp. 39-73.

[20] Kalbfleisch, J-D., and Prentice, R-L. (1980), "*The Statistical Analysis of Failure Time Data,*" New-York:Wiley.

[21] Kossoudji, S. A. (1989), "*Immigrant WorkerAssimilation: Is It a Labour Market Phenomenon?*" Journal of Human Resources, Vol. 24 No. 3, pp. 494-527.

[22] Lancaster, T. (1984), "*The Covariance Matrix of the Information Matrix Test,*" Econometrica, Vol. 52, No. 4, pp. 1051-1053.

[23] Månsson, J. and Ekberg, J. (2000), "*Second-Generation Immigrants in the Swedish Labour- Market,*" Working paper no. 4, School of Management and Economics, Växjö University.

[24] Nielsen, H., Rosholm, M., Smith, N., and Nusted, L. (2001), " *Intergenerational Transmissions and the School-to-Work Transition of Second Generation Immigrants,*" Discussion Paper No. 296, The Institute for the Study of Labour IZA Bonn.

[25] Ohlsson, R. (1975), "*Immigrants in the Labour Market,*" Dissertation. Dep. of Economic History, Lund University.

[26] Riphahn, R. T. (2001), "*Cohort Effects in the Educational Attainment of Second-Generation Immigrants in Germany: An Analysis of Census Data,*" Discussion Paper No. 291, The Institute for the Study of Labour IZA Bonn.

[27] Van Ours J., and Veeneman, J. (2001), "*The Educational Attainment of Second-Generation Immigrants in the Netherlands,*" Discussion Paper No. 297, The Institute for the Study of Labour IZA Bonn.

[28] Van Ours J., and Veeneman, J. (2002), "*From Parent to Child: Early labour Market Experiences of Second-Generation Immigrants in the Netherlands,*" Discussion Paper No. 649, The Institute for the Study of Labour IZA Bonn.

[29] White, H., (1982), "*Maximum likelihood estimation of misspecified models*", Econometrica, 50, pp. 1-26.

[30] Österberg, T. (2000), "*Economic Perspectives on Immigrants and Intergenerational Transmissions,*" Economic Studies 102., Thesis. Göteborg University.

Appendix A. Information matrix test and mixture models

The problem of unobserved heterogeneity, or the bias caused by not being able to include particular important explanatory variables has a larger impact in transition data models than in other types of regression models. Unobserved heterogeneity may introduce, among other things, downward bias in the duration effects, spurious effects of time-varying covariates, spurious time-covariate interaction effects, and dependence among competing risks and repeatable events. For transition data models testing for unobserved heterogeneity was a complex problem. A feasible approach was suggested by Lancaster (1984) and it is based in the information matrix (IM) test introduced by White (1982).

In a duration model, T be a random variable with probability density function

$$f(t \mid \theta_0,\ \theta_1), \tag{14}$$

where θ_0 is a scalar and θ_1 a vector of parameters. Let

$$g(\nu \mid \sigma^2), \tag{15}$$

be the probability density function of a random scalar variable V with mean zero and variance σ^2. Now consider the more general model for T with probability density function

$$f(t \mid \theta_0,\ \sigma^2,\ \theta_1) = \int_\nu f(t \mid \theta +\ \nu,\ \theta_1)\ \ g(\nu \mid \sigma^2)\ d\nu. \tag{16}$$

The parameter θ_0 that was fixed in the null model has in the generalized model with mean θ_0 and variance σ^2. To test the model we can do a score test of the null hypothesis $\sigma^2 = 0$ against the alternative hypothesis $\sigma^2 > 0$.

Chesher shows that for a large, flexible class of distributions $g(\cdot)$, and for a general choices of models $f(\cdot)$, the score function that is the basis of score test for testing the null that $\sigma^2 = 0$, is proportional to

$$\left[\frac{\partial \ln f}{\partial \theta}(t;\theta_0,\theta_1)\right]^2 + \frac{\partial^2 \ln f}{\partial \theta^2}(t;\theta_0,\theta_1). \tag{17}$$

which has expectation zero by the information matrix equality.[7]

Lancaster (1984) suggested a simple way to carry such tests. Under the null hypothesis this test statistic has a chi-squared distribution. Blossfeld, Hamarle and Mayer (1989) used a version of the same test statistic that had a normal distribution. The test statistic can be written

$$\tau = \frac{1}{2n} \sum_{i=1}^{n} \left[H^2(t_i \mid \mathbf{x}_i; \hat{\theta}) - 2\delta_i H(t_i \mid \mathbf{x}_i; \hat{\theta}) \right] \tag{19}$$

where $H(t_i \mid \mathbf{x}_i; \hat{\theta})$ are the maximum likelihood estimates of the integrated hazards for the model without heterogeneity and δ_i is the censoring indicator with $\delta_i = 1$ for uncensored and $\delta_i = 0$ censored observations. We report the values of this test statistic below. The results indicate that the assumption of unobserved heterogeneity holds for these transitions.

Transition from continuing school to work:	4.49
Transition from continuing school to non-employment:	5.27
Transition from continuing school to military service:	4.50

Transition from non-employment after continuing school to work: 2.16

[7] The score function has expectation zero at the true values of the parameters.

$$E\left[\frac{\partial \ln f}{\partial \theta}(t, \theta_0)\right] = 0.$$

It is based on the Information Matrix Equality and says that

$$I(\theta_0) = -E\left[\frac{\partial^2 \ln f}{\partial \theta \partial \theta'}(t, \theta_0)\right] = E\left[\frac{\partial \ln f}{\partial \theta}(t, \theta_0). \frac{\partial \ln f}{\partial \theta}(t, \theta_0)'\right]. \tag{18}$$

Note that this equality only holds at the true values of parameters.

Table 1: Transitions

from compulsory school (n=2212)	**to work**	**to continuing school**	**to non employment**	
frequency	162	1941	109	
percent	7.32	87.75	4.93	
from continuing school (n=1941)	**to work**	**still studying**	**to non employment**	**to military**
frequency	637	385	657	262
percent	32.82	19.84	33.85	13.50
from non employment (n=109) after compulsory school	**to work**	**to military**	**still not employment**	
frequency	70	27	12	
percent	64.22	24.77	11.01	
from non employment (n=657) after continuing school	**to work**	**still non employment**		
frequency	486	171		
percent	73.97	26.03		

Table 2: Descriptive Statistics

	Second Generation Immigrants (n=1106)		Swedes (n=1106)		All (n=2212)	
	mean	std dev	mean	std. dev	mean	std. dev
Individual variables						
Gender						
Male	0.519	0.500	0.498	0.500	0.509	0.500
Birth place						
Born outside Sweden	0.325	0.468			0.164	0.370
Family structure						
Two-parent family	0.752	0.432	0.747	0.435	0.750	0.433
Number of Swedish parent						
At least one Swedish parent	0.426	0.495			0.713	0.452
Number of siblings	1.877	1.017	1.752	0.835	1.815	0.932
Parental variables						
Parent education						
Secondary	0.277	0.448	0.155	0.362	0.216	0.411
High school	0.436	0.496	0.474	0.500	0.455	0.498
University degree	0.288	0.453	0.371	0.483	0.329	0.470
Parent working status						
Mother working	0.769	0.423	0.816	0.387	0.792	0.406
Father working	0.858	0.352	0.932	0.256	0.895	0.306
Geographical origin						
Swedish					0.500	0.500
Nordic (not incl. Swedish)	0.498	0.500			0.249	0.432
Western Countries	0.170	0.376			0.085	0.279
Eastern-Europe	0.158	0.365			0.079	0.270
Middle-East	0.058	0.234			0.029	0.168
Asia	0.056	0.230			0.028	0.165
Africa	0.014	0.119			0.007	0.085
Latin-America	0.045	0.207			0.023	0.149
Parental Age (maximum if 2)	46.659	6.459	45.857	5.351	46.258	5.943
Ethnic concentration in municipality	0.124	0.071	0.111	0.078	0.118	0.075
Economic variables						
Parental income (log) (annual)	12.252	0.416	12.343	0.382	12.297	0.402
Father's labour income[a]	13.899	12.290	18.449	13.523	16.174	13.117
Mother's labour income[a]	9.745	7.386	11.968	7.098	10.857	7.327
Asset income[a]	0.862	4.230	1.168	3.862	1.015	4.052
Welfare income[a]	0.263	1.218	0.053	0.412	0.158	0.916
All other income[a,b]	6.574	6.132	3.962	4.566	5.268	5.560

[a] in tens of SEK

[b] Sum of all incomes other than parents' labour income,asset income and welfare income averaged over the period when the child was 15-16 (or 16-17) years old.

Table 3: Descriptive statistics for first transitions after compulsory education

	First Transition to					
	Work n=162		Continuing Education n=1941		Non Employment n=109	
	mean	std dev	mean	std. dev	mean	std. dev
Individual Variables						
Gender						
Male	0.462	0.500	0.503	0.500	0.669	0.472
Birth place						
Born outside Sweden	0.179	0.384	0.154	0.361	0.321	0.469
Family structure						
Two-parent family	0.611	0.489	0.770	0.421	0.596	0.493
Number of Swedish parent						
At least one Swedish parent	0.586	0.494	0.740	0.439	0.422	0.496
Number of siblings	1.833	1.052	1.808	0.918	1.908	1.004
Parental variables						
Parent education						
Secondary	0.333	0.473	0.196	0.397	0.385	0.489
High school	0.451	0.499	0.459	0.498	0.404	0.492
University degree	0.216	0.412	0.346	0.476	0.211	0.409
Parent working Status						
Mother working	0.667	0.473	0.814	0.389	0.596	0.492
Father working	0.839	0.368	0.905	0.293	0.798	0.403
Geographical Origin						
Swedish	0.432	0.496	0.519	0.500	0.266	0.443
Nordic (not incl. Swedish)	0.388	0.489	0.239	0.426	0.229	0.422
Western Countries	0.074	0.263	0.084	0.278	0.110	0.314
Eastern-Europe	0.062	0.242	0.076	0.265	0.165	0.373
Middle-East			0.027	0.162	0.111	0.314
Asia	0.012	0.011	0.029	0.167	0.037	0.188
Africa	0.006	0.078	0.005	0.075	0.038	0.188
Latin-America	0.025	0.155	0.021	0.144	0.046	0.210
Parental Age (maximum if 2)	45.018	6.340	46.379	5.845	45.954	6.837
Ethnic Concentration in Municip.	0.128	0.076	0.116	0.075	0.129	0.060
Economic Variables						
Parental Income (log)	12.169	0.393	12.318	0.399	12.128	0.393
Father's labour income	11.419	9.158	16.997	13.174	8.585	6.927
Mother's labour income	9.311	6.429	11.145	7.343	8.019	7.451
Asset income	0.555	0.290	1.095	0.264	0.277	0.686
Welfare income	0.386	1.589	0.117	0.761	0.549	0.416
All other income	6.619	6.568	4.957	5.304	8.804	6.799

Table 4: **Descriptive statistics for second transitions of those who continued their education**

n=1941 obs	Transition to							
	Work		Higher Education		Non-Employment		Military	
	n=637		n=385		n=657		n=262	
	mean	std dev	mean	std. dev	mean	std. dev	mean	std. dev
Individual variables								
Gender								
Male	0.452	0.498	0.431	0.496	0.397	0.490	1.000	0.000
Birth place								
Born outside Sweden	0.116	0.321	0.122	0.328	0.221	0.415	0.126	0.332
Family structure								
Two-parent family	0.801	0.400	0.818	0.386	0.728	0.446	0.729	0.445
Number of Swedish parent								
At least one Swedish parent	0.766	0.424	0.784	0.412	0.662	0.473	0.805	0.397
Number of siblings	1.816	0.861	1.800	0.924	1.819	0.992	1.771	0.849
Parental variables								
Parent education								
Secondary	0.198	0.399	0.094	0.292	0.251	0.434	0.202	0.402
High school	0.484	0.499	0.343	0.475	0.464	0.499	0.553	0.498
University degree	0.319	0.466	0.564	0.497	0.285	0.452	0.244	0.430
Parent working status								
Mother working	0.841	0.366	0.852	0.336	0.770	0.421	0.802	0.400
Father working	0.920	0.272	0.932	0.251	0.883	0.322	0.885	0.319
Geographical origin								
Swedish	0.557	0.497	0.551	0.498	0.426	0.495	0.611	0.489
Nordic(not incl. Swedish)	0.235	0.425	0.205	0.404	0.265	0.442	0.229	0.421
Western Countries	0.075	0.264	0.112	0.315	0.094	0.293	0.042	0.201
Eastern-Europe	0.080	0.272	0.078	0.268	0.081	0.273	0.050	0.218
Middle-East	0.011	0.104	0.016	0.124	0.053	0.225	0.015	0.123
Asia	0.031	0.175	0.021	0.143	0.030	0.172	0.031	0.172
Africa[a]			0.005	0.072	0.012	0.110	0.004	0.062
Latin-America	0.009	0.097	0.013	0.113	0.038	0.191	0.019	0.137
Parental Age (maximum if 2)	45.958	5.576	47.117	5.710	46.496	6.309	46.023	5.368
Ethnic Concentration in Municip.	0.118	0.075	0.111	0.072	0.118	0.071	0.112	0.086
Economic variables								
Parental income (log)	12.348	0.392	12.394	0.399	12.264	0.405	12.265	0.377
Father's labour income[b]	17.841	12.795	20.381	14.433	14.596	12.556	15.992	12.416
Mother's labour income[b]	11.277	7.307	12.071	7.178	10.475	7.591	10.366	6.824
Asset income[b]	1.056	3.638	1.068	3.795	1.217	5.111	0.923	3.398
Welfare income[b]	0.071	0.004	0.121	0.008	0.175	0.043	0.080	0.044
All other income[b]	4.781	5.041	4.235	5.056	5.667	5.779	4.665	4.843

[a] In transition to work there was not any African originated individual.

Table 5:[a] **Multinomial-logit regression results for first transitions after compulsory education, with base category continuing school**

n=2212	Work (n=162)	Non Employment (n=109)
Constant[b]	4.721	-1.590
	(3.063)	(1.696)
Individual variables		
Male	-0.162	0.669^{0}
	(0.167)	(0.214)
Family structure		
Two-parent family	-0.234	-0.573^{2}
	(0.216)	(0.257)
At least one Swedish parent	-0.496^{6}	-0.563^{7}
	(0.266)	(0.317)
Parental variables		
High school education	-0.429^{3}	-0.428^{7}
	(0.199)	(0.242)
University degree	-0.680^{0}	-0.544^{7}
	(0.247)	(0.300)
Geographical Origin		
Nordic(not incl. Swedish)	0.238	0.135
	(0.258)	(0.358)
Western Countries	-0.211	0.669
	(0.368)	(0.409)
Eastern-Europe	-0.471	0.969^{1}
	(0.410)	(0.403)
Middle-East and Africa	-1.186	1.450^{0}
	(1.392)	(0.450)
Asia	-1.310^{9}	0.310
	(0.776)	(0.624)
Latin-America	-0.174	0.790
	(0.588)	(0.582)
Parental Age (maximum if 2)	-0.026^{8}	-0.004
	(0.015)	(0.017)
Ethnic Concentration in Municip.	1.936^{3}	0.491
	(0.921)	(1.385)

[a] numbers on the power of coefficients, are not powers, they indicate p-values in percents, i.e power 6 means p-value between 0.05 and 0.06...so on.
[b] The reference variables included in the constant are: female, one-parental family, no Swedish parent, secondary education, and Swedish.
* Individual variable, born outside Sweden was not significant, even in the sensitivity test, so we omitted this variable

Table 5: (cont.) Multinomial-logit regression results for first transitions after compulsory education, with base category continuing school

n=2212	Work (n=162)	Non-employment (n=109)
Economic Variables		
Parental Income (log)	-0.429[8]	-0.329
	(0.250)	(0.297)
Father's labour income[b]	-0.210[4]	-0.229[10]
	(0.105)	(0.140)
Mother's labour income[b]	-0.129	-0.081
	(0.146)	(0.179)
Asset income[b]	-0.027	-0.164
	(0.035)	(0.121)
Welfare income[b]	0.216[0]	0.097
	(0.075)	(0.073)
All other income[b]	0.013	0.039[4]
	(0.017)	(0.019)

[b] Separate estimates are obtained when the source of parental income disaggregated using the same models and the same control variables listed above (except log Family Income). The same methodology is used in the other outcomes listed in the tables below.

Table 6[a] : Estimated parameters for transitions after continued education

n=1941	to Work n=637		to Non-Employment n=657		to Military n=262	
	Cox Prop. hazard gamma mixture	Log-Logistic hazard gamma mixture	Cox Prop. hazard gamma mixture	Log-Logistic hazard gamma mixture	Cox Prop. hazard gamma mixture	Log-Logistic hazard gamma mixture
alpha Constant		-1.117		-0.561		-2.201^{1}
		(0.819)		(0.573)		(0.064)
Individual Variables						
Male[a]	-0.019	-0.028	-0.367^{0}	-0.014		
	(0.105)	(0.042)	(0.082)	(0.029)		
Born outside Sweden	-0.203	-0.064	0.242^{7}	0.066	-0.135	-0.027
	(0.190)	(0.082)	(0.133)	(0.045)	(0.254)	(0.049)
Family structure						
Two-parent family	0.261^{9}	0.101	-0.148	-0.004	-0.077	-0.033
	(0.153)	(0.062)	(0.111)	(0.038)	(0.169)	(0.039)
At least one Swedish parent[b]	0.079	0.037	0.094	-0.041		
	(0.179)	(0.073)	(0.128)	(0.044)		
Parental Variables						
High school education	-0.317^{3}	-0.106^{7}	-0.277^{1}	-0.066^{6}	-0.022	0.059
	(0.150)	(0.059)	(0.105)	(0.036)	(0.167)	(0.116)
University degree	-1.291^{0}	-0.628^{0}	-0.763^{0}	-0.218^{0}	-0.745^{0}	-0.732^{0}
	(0.168)	(0.065)	(0.123)	(0.044)	(0.203)	(0.139)
Geographical Origin						
Nordic(not incl. Swedish)	0.138	0.101	0.372^{2}	0.017	-0.018	-0.064
	(0.154)	(0.062)	(0.117)	(0.418)	(0.180)	(0.118)
Western Countries	-0.255	-0.069	0.211	-0.076	-0.823^{1}	-0.014
	(0.213)	(0.086)	(0.156)	(0.055)	(0.326)	(0.131)
East-Europe	-0.123	-0.077	0.208	-0.082	-0.486	-0.061
	(0.226)	(0.092)	(0.173)	(0.062)	(0.315)	0.123
Middle-East	-0.806^{6}	-0.834^{3}	0.832^{0}	0.027	-0.564	-0.176
	(0.484)	(0.396)	(0.231)	(0.086)	(0.559)	(0.164)
Asia	0.310	0.075	0.005	-0.098	-0.103	-0.067
	(0.355)	(0.143)	(0.269)	(0.089)	(0.427)	(0.141)
Africa	-1.046	-0.197	0.900^{2}	0.228	-0.479	-0.434^{5}
	(1.001)	(0.192)	(0.386)	(0.177)	(1.015)	(0.207)
Latin-America	-1.033^{6}	-0.357^{9}	0.597^{1}	0.223^{2}	-0.171	-0.068
	(0.549)	(0.216)	(0.253)	(0.097)	(0.512)	(0.151)

[a,b] Numbers on the power of coefficients, are not powers, they indicate p-values in percents, i.e power 6 means p-value between 0.05 and 0.06...so on.

The reference variables included in the constant are: female, one-parent family, no Swedish parent, secondary education, and Swedish.

Table 6: (cont). Estimated parameters for transitions after continued education (1941 obs)

n=1941	to Work		to Non-Employment		to Military	
	n=637		n=657		n=262	
	Cox Prop. hazard gamma mixture	Log-Logistic hazard gamma mixture	Cox Prop. hazard gamma mixture	Log-Logistic hazard gamma mixture	Cox Prop. hazard gamma mixture	Log-Logistic hazard gamma mixture
Parental Age	-0.034^{0}	-0.016^{0}	0.002	0.018	-0.004	-0.051^{8}
	(0.009)	(0.003)	(0.007)	(0.024)	(0.011)	(0.028)
Ethnic Conc.in Munic.	0.676	0.149	0.379	0.122	-0.306	0.134
	(0.566)	(0.280)	(0.537)	(0.207)	(0.864)	(0.144)
Economic Variables						
Parental Income (log)	0.145	0.006	-0.110	-0.050	-0.251	-0.064
	(0.129)	(0.066)	(0.123)	(0.045)	(0.192)	(0.051)
Father's labour income	-0.081^{4}	-0.066^{0}	-0.164^{0}	-0.029^{5}	-0.149^{2}	-0.130^{0}
	(0.039)	(0.022)	(0.042)	(0.015)	(0.066)	(0.042)
Mother's labour income	-0.054	-0.115^{0}	-0.132^{3}	-0.084^{0}	-0.331^{0}	-0.262^{0}
	(0.059)	(0.032)	(0.060)	(0.025)	(0.101)	(0.068)
Asset income	-0.007	0.008	0.009	0.015	-0.009	-0.017
	(0.010)	(0.006)	(0.008)	(0.028)	(0.018)	(0.012)
Welfare income	-0.103	-0.059	-0.039	-0.018	-0.111	-0.078
	(0.079)	(0.039)	(0.043)	(0.017)	(0.116)	(0.072)
All other income	0.009	0.003	0.003	0.001	-0.019	-0.012
	(0.008)	(0.005)	(0.008)	(0.006)	(0.015)	(0.009)
ln_gamma constant		1.114^{0}		1.883^{0}		2.498^{0}
		(0.052)		(0.053)		(0.067)
ln_delta constant		0.070		-1.837^{0}		3.166^{0}
		(0.245)		(0.066)		(0.081)
theta constant	1.260^{0}		0.112		0.005	
	(0.040)		(0.405)		(0.110)	

[a] " male" and "at least one Swedish parent" were omitted when we focused on the transition to the military, since all who joined military were male and almost all had at least one Swedish parent.

As before, the reference variables included in the constant are: female, one-parental family, no Swedish parent, secondary education, and Swedish.

Table 7: Estimated parameters for transitions from non-employment after compulsory education

n=109	To Work		To Military	
	70 obs.		27 obs.	
	Cox Proportional Hazard	Accel. Failure time Log-Logistic	Cox Proportional Hazard	Accel. Failure time Log-Logistic
Constant		-2.489		2.175
		(2.421)		(2.563)
Individual variables				
Male[a]	-0.373	0.038		
	(0.251)	(0.124)		
Born outside Sweden	0.284	-0.003	-0.425	0.221
	(0.279)	(0.124)	(0.543)	(0.176)
Family structure				
Two-parental family	0.688^{4}	-0.345^{3}	-0.122	0.053
	(0.343)	(0.165)	(0.546)	(0.187)
At least one Swedish parent	0.585	-0.302^{8}	0.555	-0.379^{5}
	(0.347)	(0.177)	(0.510)	(0.199)
Parental variables				
college education	-0.019	0.013	0.251	-0.109
	(0.287)	(0.145)	(0.454)	(0.152)
university degree	0.220	-0.062	-0.109	0.135
	(0.356)	(0.169)	(0.622)	(0.215)
Geographical Origin				
Mid-East, Africa and Latin America[b]	-2.047^{0}	0.388^{5}	-3.123^{0}	0.769^{1}
	(0.495)	(0.202)	(1.075)	(0.302)
Parental Age	-0.023^{10}	0.017^{8}	-0.017	0.004
	(0.020)	(0.010)	(0.031)	(0.011)
Ethnic Concentration in Municip.	-0.442	0.855^{0}	-0.108	0.391
	(0.347)	(0.193)	(0.430)	(0.257)

Table 7: (cont.) Estimated parameters for transitions from non-employment after compulsory education

	to work 70 obs.		to military 27 obs.	
	Cox Proportional Hazard	Accel. Failure time Log-Logistic	Cox Proportional Hazard	Accel. Failure time Log-Logistic
Economic Variables				
Parental Income (log)	-0.057	0.194	0.180	-0.091
	(0.376)	0.186	(0.609)	0.200
Father's labour income	-0.003	0.072	0.117	-0.025
	(0.136)	(0.078)	(0.206)	(0.073)
Mother's labour income	0.072	0.043	-0.086	0.064
	(0.221)	(0.100)	(0.310)	(0.118)
Asset income	-0.164	0.052	-0.189	0.036
	(0.155)	(0.063)	(0.205)	(0.072)
Welfare income	0.084	-0.023	-0.554	0.126
	(0.075)	(0.036)	(0.426)	(0.119)
All other income	-0.014	0.017	0.024	-0.101
	(0.022)	(0.011)	(0.033)	(0.266)
scale		0.334		0.271
		(0.033)		(0.040)

[a] As in the previous outcome, Individual variable being male is omitted when we focused on the transition to military, since all of the individuals who made transition to military were male.

[b] Due to the small number of observations, those from the Middle-East, Africa, and Latin America were combined as one group with all othere as the reference group. As before, The reference variables included in the constant are: female, one-parental family, no Swedish parent, and secondary education.

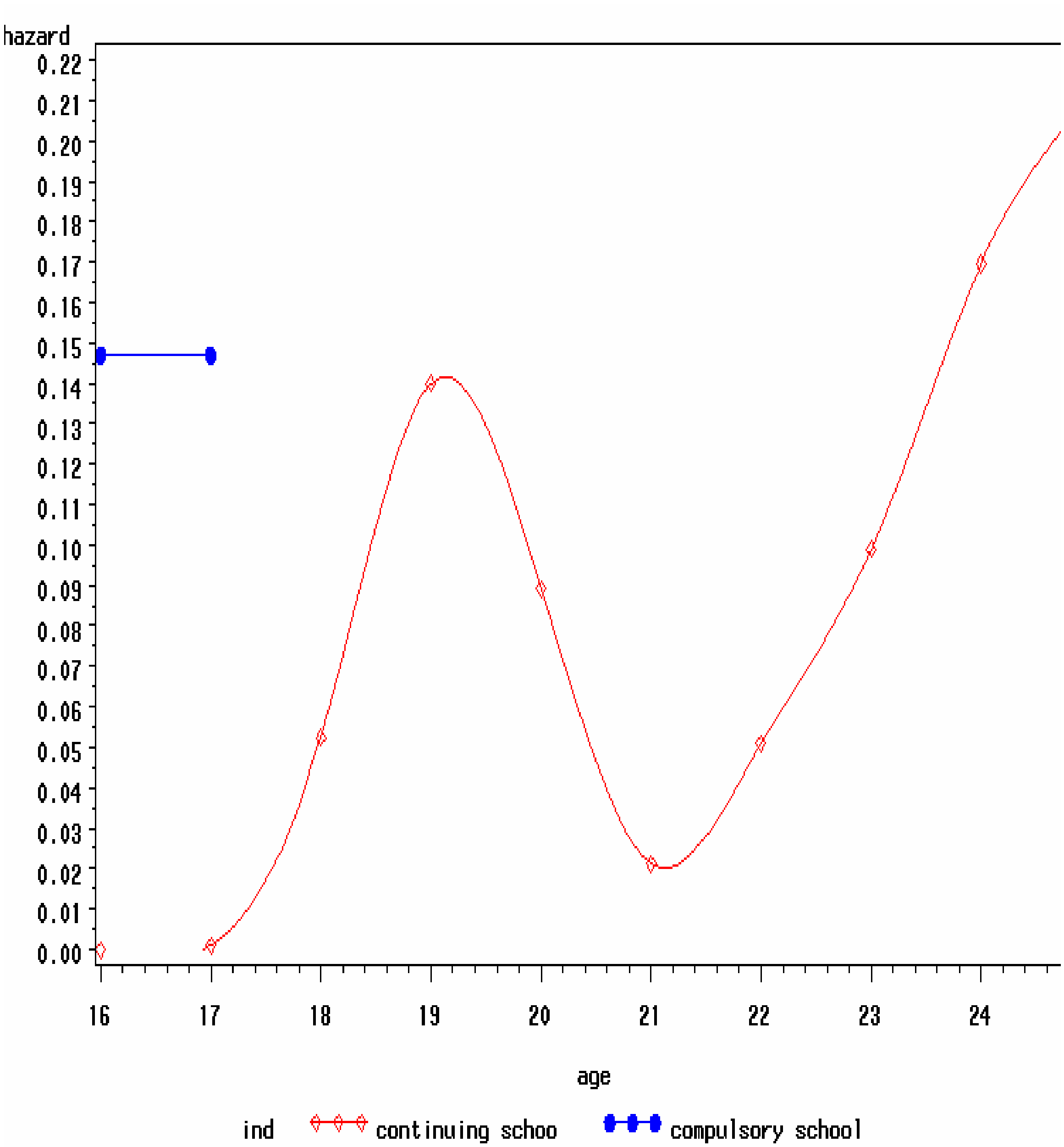
Table 8: Estimated parameters for transitions from non-employment after continued education

n=657	To Work	
	n=486	n=486
	Cox Proportional hazard gamma mixture	Log-Logistic hazard gamma mixture
Constant		-3.203^{1}
		(1.293)
Individual Variables		
Duration of higher education	0.011	0.002
	(0.048)	(0.024)
Male	0.116	0.080
	(0.132)	(0.064)
Number of siblings	-0.191	-0.094
	(0.079)	(0.035)
Birth place		
Born outside Sweden[a]	0.169	0.084
	(0.200)	(0.092)
Family structure		
Two-parent family	0.304^{8}	0.024
	(0.174)	(0.082)
At least one Swedish parent	-0.017	-0.113
	(0.194)	(0.091)
Parental Variables		
high school education	0.804^{0}	0.076
	(0.173)	(0.101)
university degree	0.614^{0}	0.232^{1}
	(0.205)	(0.085)
Geographical Origin		
Nordic(not incl. Swedish)	-0.278	-0.279^{0}
	(0.218)	(0.097)
Western Countries	-0.076	-0.095
	(0.276)	(0.123)
Eastern-Europe	-0.443	-0.451
	(0.295)	(0.133)
Middle-East	-2.675^{0}	-1.551^{0}
	(0.443)	(0.165)
Asia	-0.837	-0.523
	(0.471)	(0.227)
Africa	-2.838^{0}	-1.663^{0}
	(0.769)	(0.266)
Latin-America	-2.830^{0}	-1.761^{0}
	(0.541)	(0.200)

As before, the reference variables included in the constant are: female, one-parental family, no Swedish parent, secondary education, and Swedish .

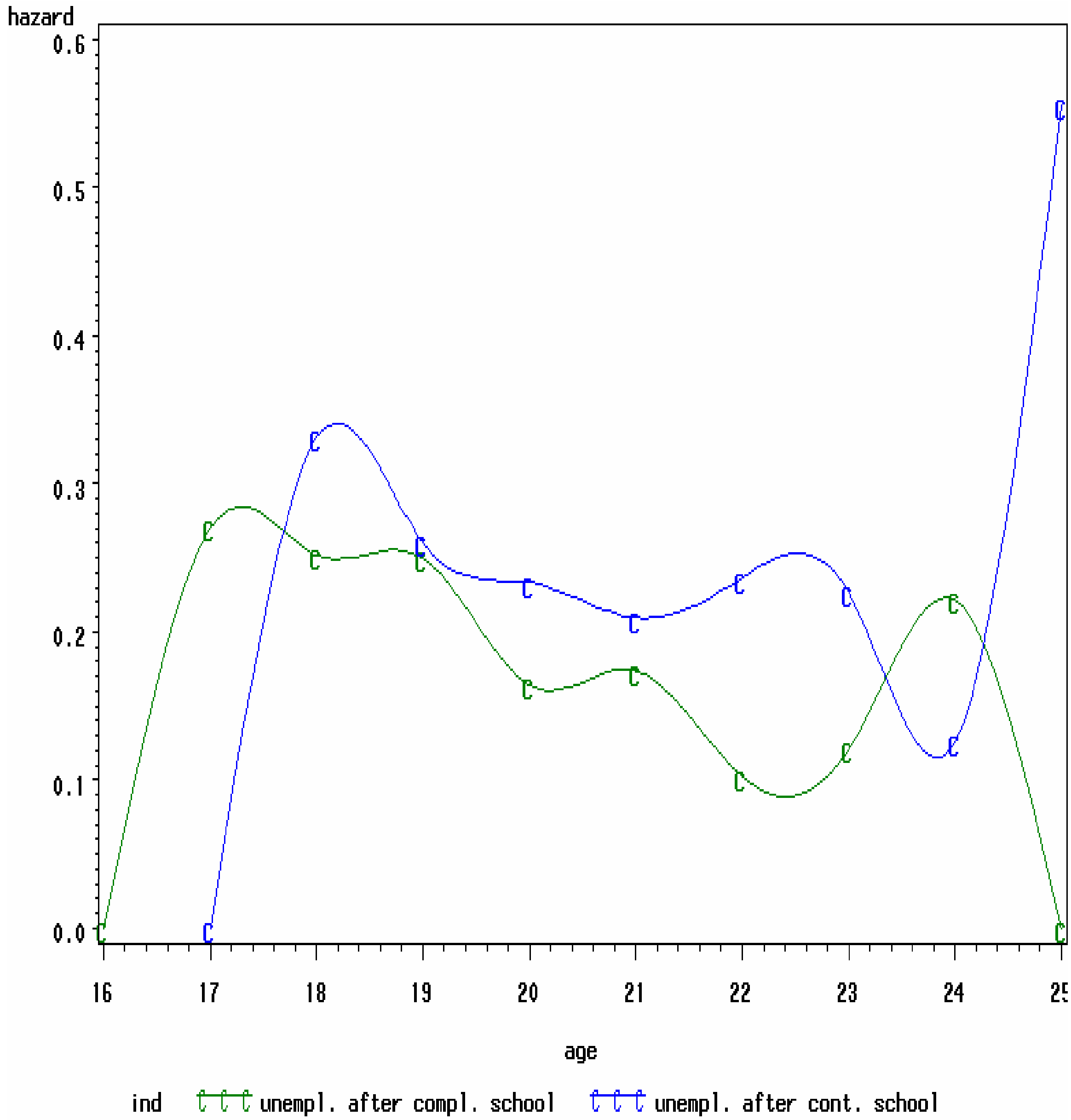
Table 8: (cont.) Estimated parameters for transitions from non-employment after continued education

n=657	To Work	
	n=486	n=486
	Cox Proportional hazard gamma mixture	Log-Logistic hazard gamma mixture
Parental Age	-0.016	-0.012^{7}
	(0.011)	(0.006)
Ethnic Concentration in Municip	0.052	0.009
	(0.134)	(0.045)
Economic Variables		
Parental Income (log)	0.566^{0}	0.275^{0}
	(0.201)	(0.097)
Father's labour income^{a}	0.018	0.034
	(0.042)	(0.032)
Mother's labour income^{a}	0.075	0.054
	(0.072)	(0.039)
Asset income^{a}	-0.003	-0.006
	(0.011)	(0.005)
Welfare income^{a}	-0.035	-0.046
	(0.067)	(0.029)
All other income^{a,b}	-0.004	0.032
	(0.010)	(0.059)
ln_gamma constant		1.199^{0}
		(0.077)
ln_delta constant		0.212
		(0.193)
theta	0.747	
	(0.275)	



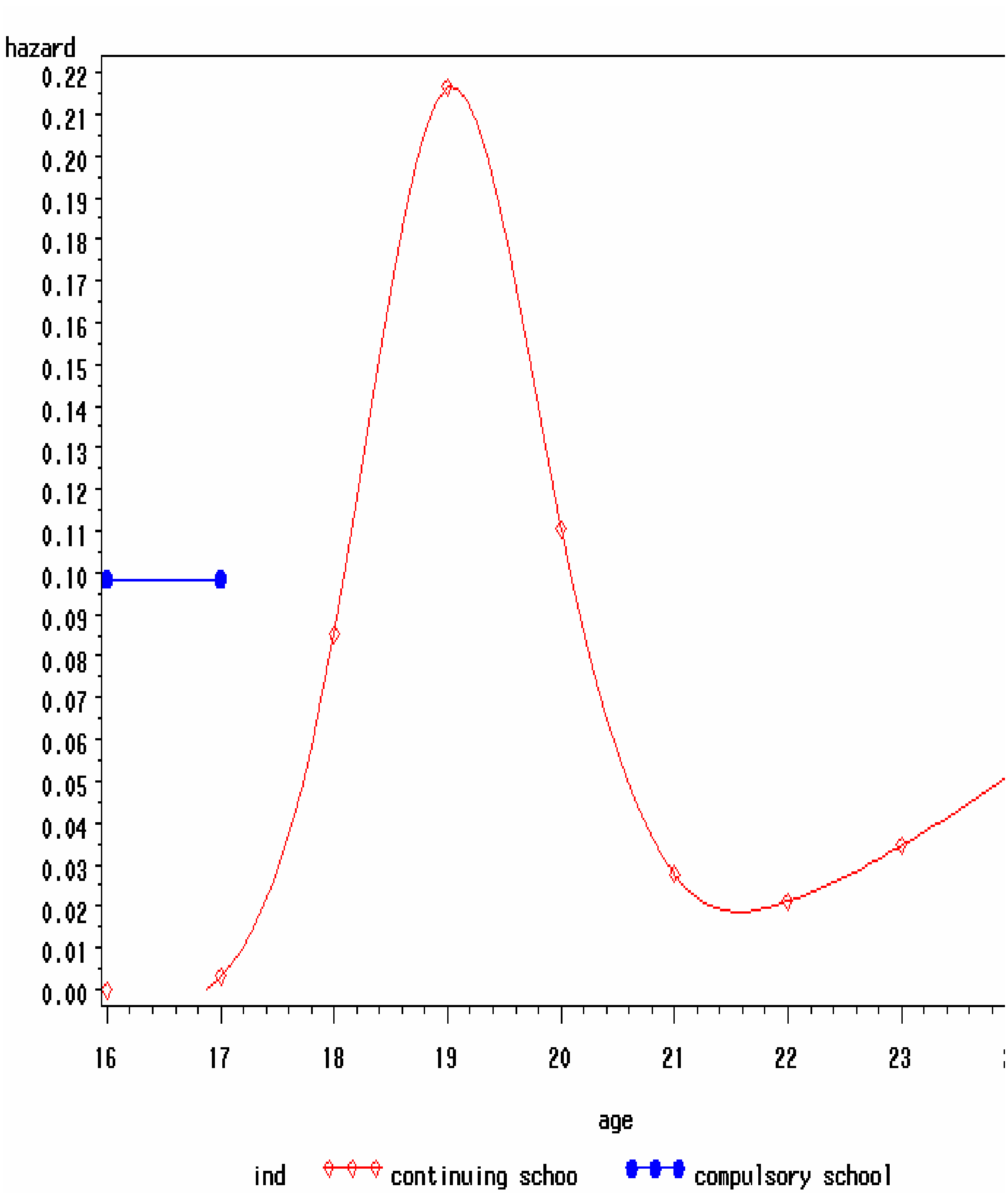
Smoothed non-paramteric plot : Estimated hazard of work after compulsory education and after some period of continuing education (life table estimator)

[Figure 1]



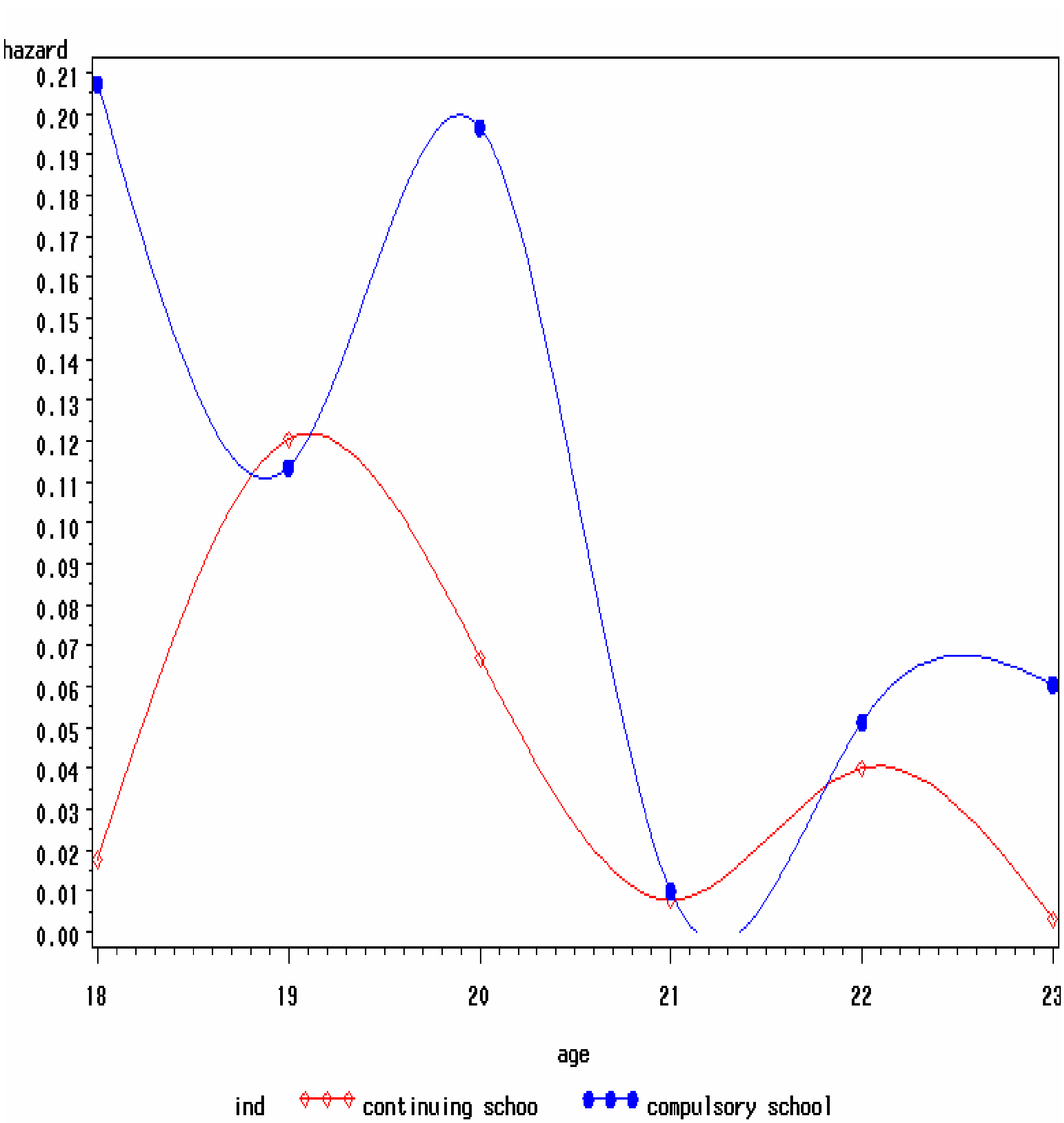
Smoothed non-parametric plot : Estimated hazard of transiting to work, after a period of non-employment (Life Table Estimator)

[Figure 2]



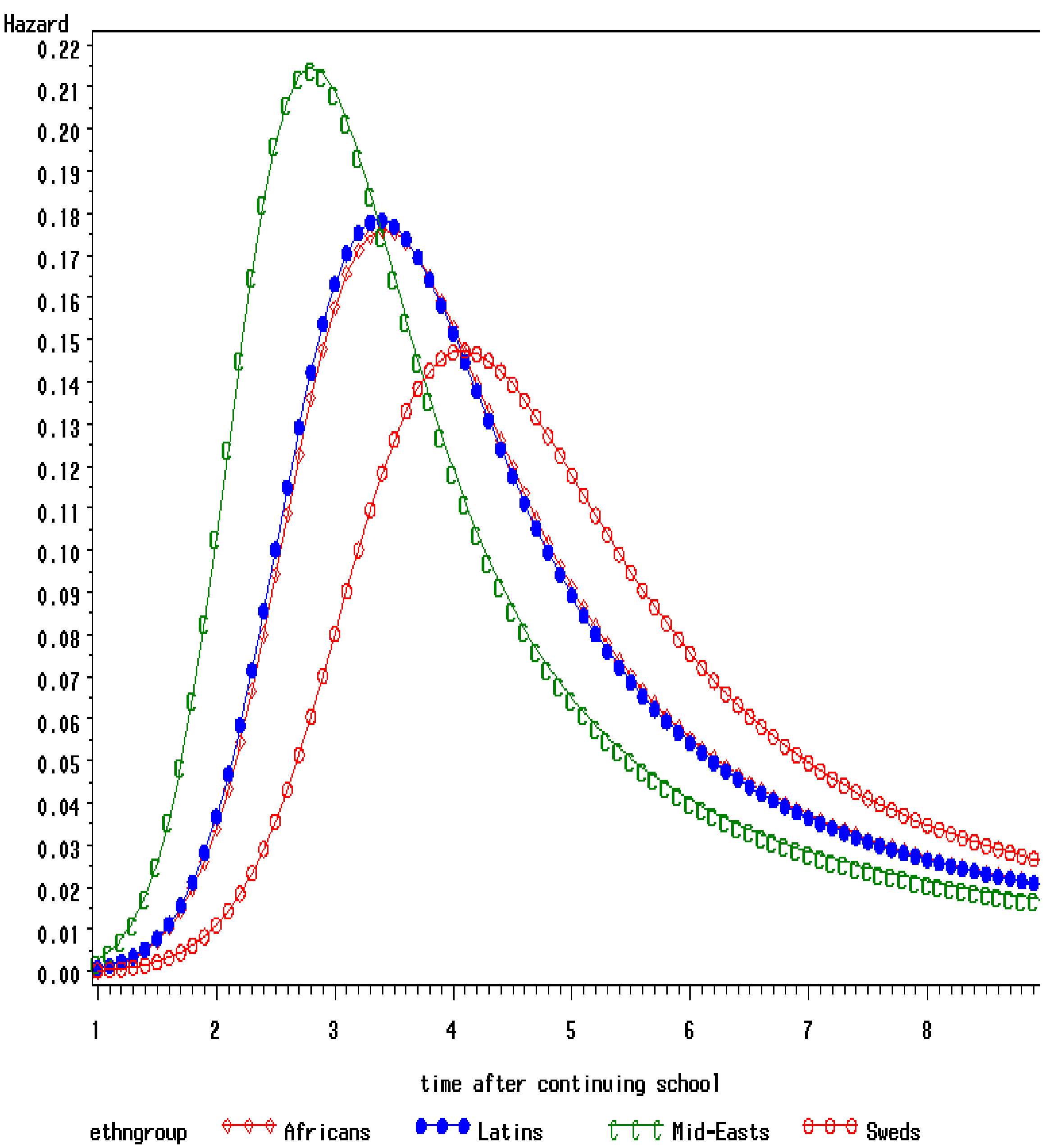
Smoothed non-parametric plot :Estimated hazard of non-employment after compulsory education and continuing education (life table estimator)

[Figure 3]



Smoothed non-parametric plot : Estimated hazard of military after compulsory education and continuing education (life table estimator)

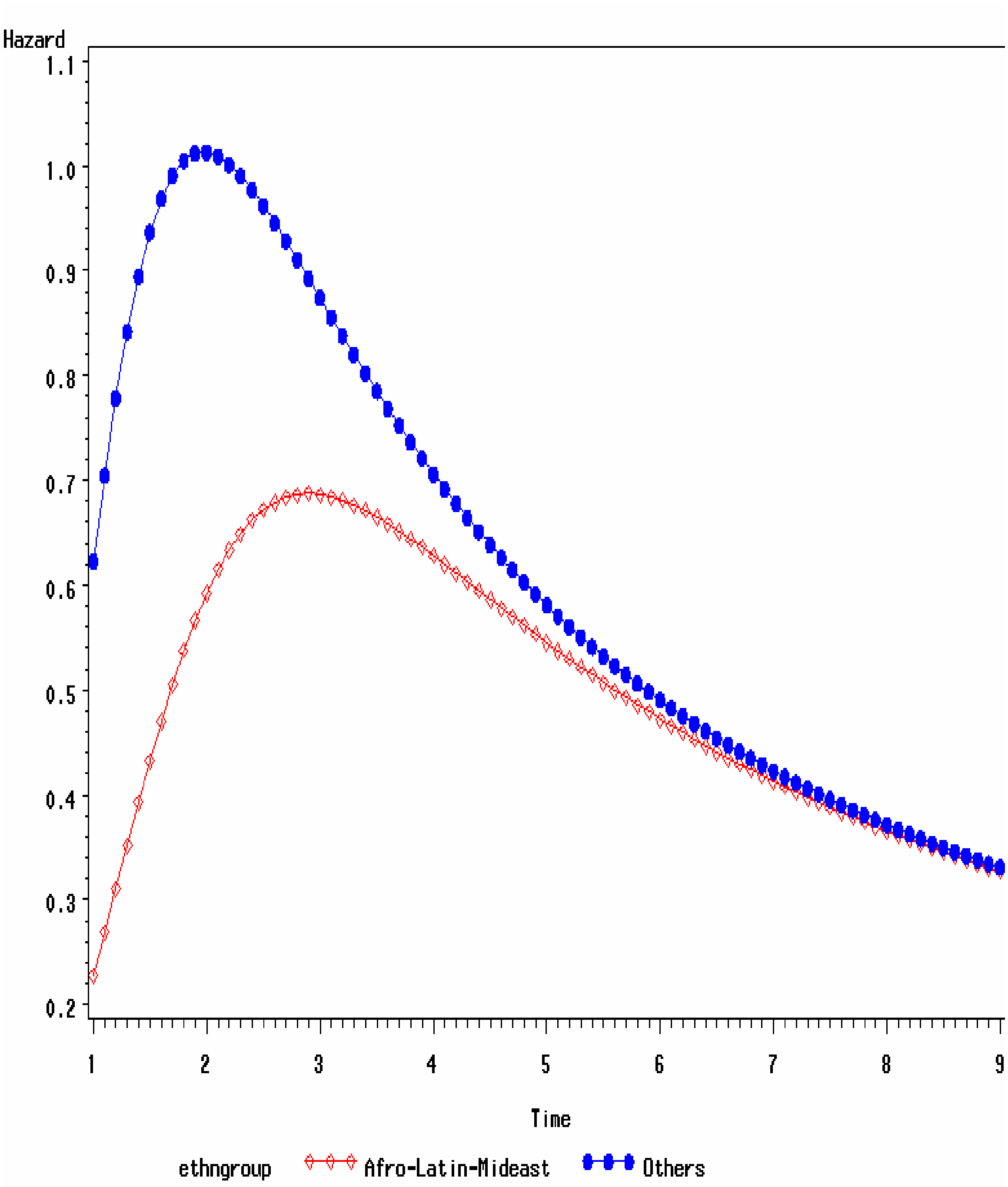
[Figure 4]



Parametric plot (based on log-logistic gamma-mixture parametric assumption) : Estimated hazard of transition from continuing education to non-employment.

Curves are plotted conditionally on these characteristics: male, two-parent family, born in Sweden, parents with university degree, family income=mean value

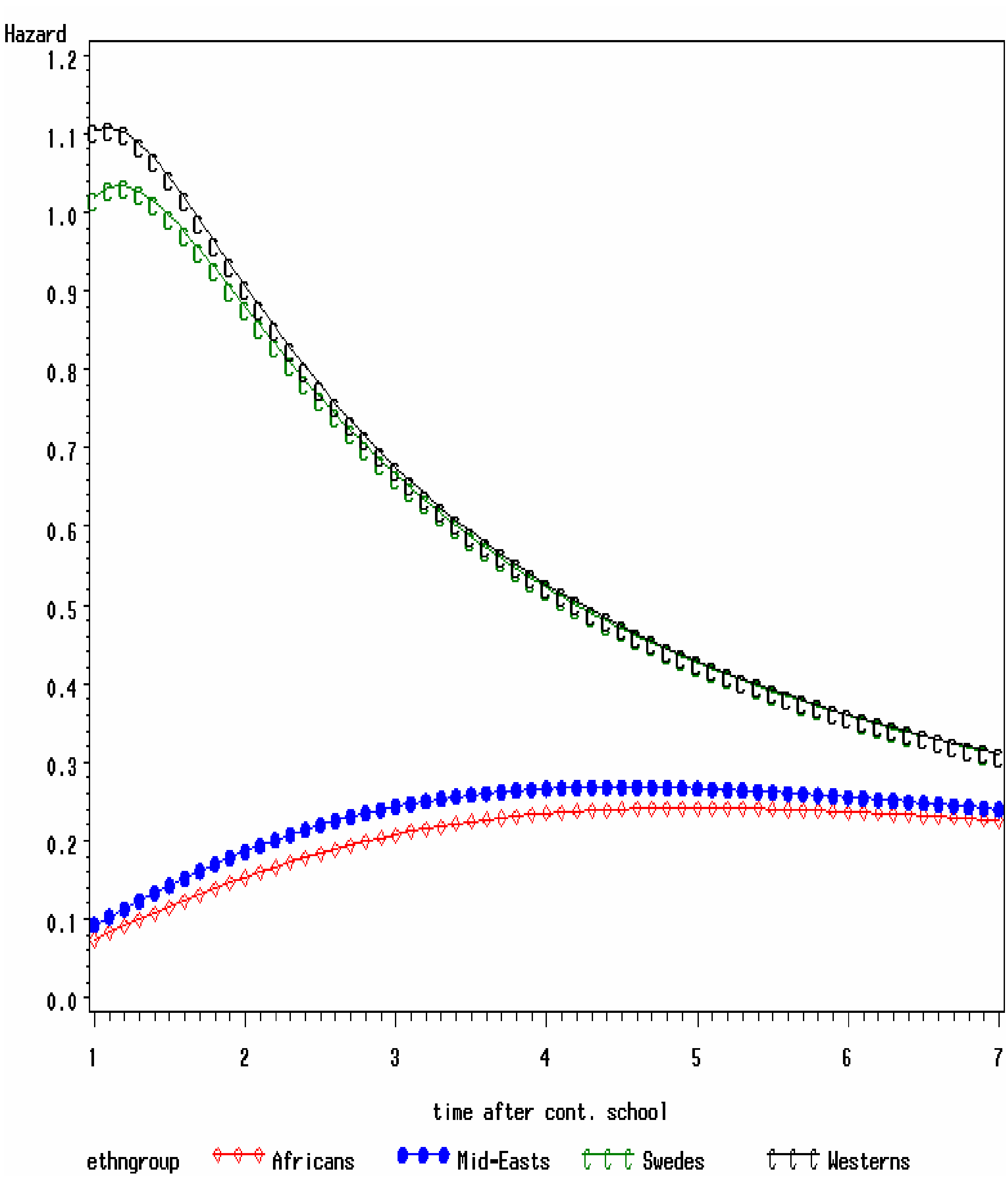
[Figure 5]



Parametric plot: Estimated hazard of transition from non-employment after compulsory education to work.

Curves are plotted conditionally on these characteristics: male, two-parent family, born in Sweden, parents with university degree, family income=mean value

[Figure 6]



Parametric plot: Estimated hazard of transition from non-employment after continuing education to work.

Curves are plotted conditionally on these characteristics: male, two-parent family, born in Sweden, parents with university degree, family income=mean value

[Figure 7]

An Economic Analysis of Graduate Employment in Sweden

Ali Tasiran, Kerem Tezic, Gunilla Bornmalm-Jardelöw*

Department of Economics, Göteborg University

January 2004

Abstract

This paper examines the new university-graduates' entry into the Swedish labour-market. The study-population consisted of 2598 individuals who took their first university-degree during 1994. Their subsequent transitions into various states were studied using Cox semi-parametric and parametric duration dependence distributions in a competing-risks framework. Transitions from post-graduate non-employment to subsequent work were separately examined. Then the annual earnings of graduates were investigated using Tobit-type models after correcting for selection-problems. Systematic differences with respect to age, sex, individual and university background and previous university financing showed up in all models.

Keywords: Graduate employment; calibration; competing-risks; labour market entry; annual-earnings

J.E.L classification: C41, I21, J23, J24.

*We thank Sixten Lundström from Statistics Sweden for providing calibration weights and for his continous help.

Contents

1 Introduction

This paper analyzes the new university graduates' entry into the Swedish labour-market. Graduation from university initiates an important period in students' lives: looking for a job; perhaps going through frequent job turnovers; possibly getting married and becoming a parent. Early experiences in the labour-market may have decisive influence on later life, and raise important policy issues: What factors are important in finding a job earlier? Why do some individuals remain non-employed longer? What factors determine earnings-differentials? And so on.

The focus in this paper is on the relationship between university education and employment during the first four years after university-graduation. The study-sample consisted of 2598 individuals who took their first university-degree during 1994. The data allowed us to examine the graduate's demographic backgrounds, their educational fields and achievements, as well as their early labour-market experiences. There were differences between the sexes as well as between universities attended, regions of residence, and occupational orientations, with respect both to types of transitions and earnings.

There has been a great deal of research done in this area, usually using search-models, see for instance Wolpin (1987), Jensen and Westerggard-Nielsen (1987), Eckstein and Wolpin (1995). We have used another approach, partly because we had no information about earnings before 1998, four years after graduation. We also wanted to model transitions into various states after graduation. Not all the graduates were job seekers (which is a very strong assumption used in many previous studies), we wanted to distinguish between different types of transitions and the causal-processes determining each.

The next section briefly discusses the data, methods and findings of previous research. Section 3 explains the data and variables used in the analysis; reports χ^2 test-results for high non-response rates; gives details of reweighting the sample observations and then describes the various transition-types and the earnings in different occupations. Section 4 presents the statistical methods used in the analysis, and empirical results are reported in Section 5: First, results are reported for transition-states after graduation; then for transitions from post-university non-employment into work; and finally Section 6 summarizes and draws conclusions.

2 Relation to previous research

School-to-work transitions have been explored by Jensen and Westerggard-Nielsen (1987), Wolpin (1987), Eckstein and Wolpin (1995), Bratberg and Nilsen (2000), and Nilsen *et al.* (2000). Wolpin (1987) estimated the structural parameters of a dynamic-programming model of individual search-behavior using a subsample of white males from 1979 U.S. NLS youth-cohort who were graduated from high school and who had not returned to school or entered military by the 1982 interview-date. The starting point was a standard discrete-time search-model, where the search-horizon was finite and the probability of receiving a job-offer was less than unity, with the cost of obtaining a job not subject to choice, and with a wage distribution known by the individual. Wolpin assumed that there were two possible states after graduation: working or searching while unemployed. The expected-duration of unemployment was very sensitive to changes in the wage-distribution, and reducing the cost of search increased it and the reservation-wage.

Jensen and Westerggard-Nielsen (1987) incorporated variable search intensity and employers' screening in their search-model of Danish data on new law-graduates who entered the labour-market during 1974-77. Assuming that the reservation-wage was unchanged over the period of analysis, they found that a higher reservation-wage increased the search-period and that the probability of receiving job-offers had a statistically significant positive influence on the reservation-wage. But since all the graduates were homogenous with respect to education level, it was not possible to assess the importance of education for the search-time.

Eckstein and Wolpin (1995) also used data from the 1979 NLS youth-cohort to investigate the joint distribution of the duration of the first post-graduation full-time job and the accepted wage for that job. They found that differences in unemployment-durations by race and level of schooling were due to differential rates at which job offers were accepted.

Bratberg and Nilsen (2000) simultaneously estimated search-time, hourly earnings, and job duration, using school-leavers aged 16-33 from the Norwegian register data-base. Their main findings reveal the importance of more education for getting a job quicker, at a higher, wage and keeping it longer. Such negative correlation between education and the incidence of unemployment was also pointed out by Nilsen *et al.* (2000). They also noted reservation-wage differences between males and females and the possibility of unobserved factors revealed by negative correlation between job-duration and wages.

But as mentioned above, the assumption that all graduates are job-seekers is very strong; there are other possible choices, such as post-graduate studies, child bearing, military service, etc. Each transition after graduation may also have its own causal interpretation. Our data reveal the event-history of each individual for five years after graduation allowing us easily to estimate dynamic aspects of the data. We also controlled for variables such as academic credentials upon entrance to university, the mode of financing university studies, and the university attended.

3 The data and the variables

3.1 The population and the original sample

This study was initiated by Statistics Sweden (SCB) at the request of Göteborg University in cooperation with 11 European countries, using the educational register of university-graduates who took their degree with minimum 120 credits during 1994. To ensure that graduates from certain lines of education could be reported separately, the population was divided into four strata; physicians, engineers, economists, and all others.

The total population consisted of 18,915 individuals. Unconstrained random-selection of 5500 of them and a response rate to a survey of 49.1 percent, yielded 2698 individuals.

The non-response rate was thus very high. Blom (2000) reports a drop-out analysis by comparing age, gender, and disposable annual-earnings distributions of the respondents with non-respondents. He did not find great differences between the groups. However, when we conducted a χ^2-test to judge whether or not the respondent-group had the same distribution function of annual earnings as selected group (the 5500 initially selected) we found an indication of different types of distributions (χ^2 =30 vs. 9.49=5 percent significance level for 5 different income categories); i.e., non-response was non-informative. We needed to correct non-response in the data and overall, we suspected that either some stigma effect was causing non-response or that high-earning people had not taken time to respond. So, we adjusted for non-response by reweighting the respondents' answers, thus calibrating the data.[1]

3.2 The variables

There were originally 2698 observations (responses to the survey). Then in order to obtain reweights for calibrating due to non-responses, Statistics Sweden recreated the sampling-frame which reduced the size of population from 18915 to 17981 because of changes in the registers. The number of observations in the sample after calibration was again 2698. This number fell to 2601 since there was no time-information for start and finish dates of some labour-market activities after graduation. However, three observations that were drawn after calibration technique were not in the original sample, so merging the original sample with the new sample including only original weights and calibrated weights gave a sample with 2598 observations.

Table A1 in the Appendix shows age frequency distribution of the graduates before and after calibration. Before calibration there was more concentration in the age-range 24-29, whereas calibration raised the frequency in the 30-42 range. This reflects the flatter age-distribution of university-students in Sweden compared to most other western countries, due to greater flexibility in admissions in Sweden and greater financial support.

Table A2 in the Appendix displays the frequency distribution of graduates by annual-income groups both before and after calibration. As we see, there are no big differences between before and after reweighted series, these are close to each other, except the last two groups of graduates. After calibration, we observe some decrease in the graduate group who earn between 235-310 thousand SEK while we observe some increase in the last group of graduates who earn more than 310 thousand SEK.

Table A3 in the Appendix gives the gender frequency distribution before and after reweighting, and there is no big difference between series.

Table A4 in the Appendix lists the independent variables used and their definitions.

Among the 2598 graduates over half were females (Table A5 in the Appendix). Only three percent were born outside Sweden. About 15 percent became parent

[1] See the Appendix A for estimation in the presence of non-response and the calibration approach.

during university studies. We classified parent's education according to the highest level attained by at least one parent. Less than half, 40 percent of the graduates had at least one parent with university diploma. Three quarters had university entrance with upper-secondary diploma while only five percent with university entrance exam. Nearly 20 percent worked either in foreign countries before university or studied in foreign countries during university.

3.3 The transition-types

In our sample we distinguished four different transitions made after university graduation in 1994: to employment; to non-employment; to further studies, including research or doctoral education; and other including child care and military service. 1760 who were initially non-employed after university-graduation, 93 .2 percent went to employment later during the 4-year study-period; 2.5 percent returned to school to further their educations; 3.3 percent made "other" transitions; and 1 percent remained non-employed[2] (Tables 1 and 2 below).

Table 1: First transitions after graduation (n=2598)

to employment	to non-employment	to further education	to other[1]	total
662	1760	124	52	2598
25.5%	67.7%	4.8%	2%	100%

[1] Other includes child care, military service, etc.

Table 2: Transitions from initial non-employment (n=1760)

to employment	continued non-employment	to further education	to other[1]	total
1642	17	44	57	1760
93.3%	1%	2.5%	3.2%	100%

[1] Other includes child care, military service, etc.

3.4 Annual earnings in different occupations

The data included 1998 annual-earnings information for 2324 of the 2598 individuals in the sample. The distribution of disposable income was concentrated between two-

[2] Informations were provided on monthly basis.

hundred and four hundred thousand SEK, slightly skewed to right (see Figure 1 in the Appendix).

Table 3 below shows the distribution of 1994 graduates' disposable income in 1998 (in 1,000 SEK) by educational field. Overall, about 28 percent of the graduates had less than 200,000 SEK disposable income while 8 percent had more than 400,000 SEK. About 7-8 percent of the engineers and doctors earned less than 200,000 SEK, compared to 17 percent of economists and 45 percent of "others". At the other extreme, about 12-14 percent of doctors, engineers and economists had more than 400,000 SEK compared to only about 8 percent for "others".

Table 3: Frequency distribution of 1994 graduates' disposable income in 1998 in (1,000 SEK) by educational field

	<100	100-199	200-299	300-400	>400	totals
economists	29	39	176	97	56	397
(percentages)	1.2	1.6	7.5	4.2	2.4	16.9
engineers	18	14	154	209	59	454
(percentages)	0.8	0.6	6.6	9	2.5	19.5
medical doctors	11	15	101	148	34	309
(percentages)	0.5	0.6	4.3	6.4	1.5	13.3
other	103	416	503	110	32	1164
(percentages)	4.4	18	21.6	4.7	13.8	50.3
totals	161	484	934	564	181	2324
(percentages)	6.9	20.8	40.2	24.3	7.8	100

Table 4 below shows the frequency distribution of 1994 graduates' disposable income in 1998 (in 1,000 SEK) by university. Graduates of "other" than seven major universities had the least: 33 percent had less than 200,000 SEK. On the other hand graduates of Royal Technical Institute had the most: 13 percent had more than 400,000 SEK.

Table 4: Frequency distribution of 1994 graduates' disposable income in 1998 (in 1,000 SEK) by university

	<100	100-200	200-300	300-400	>400	total
Umeå	6	25	76	39	8	154
(percentages)	0.3	1.1	3.3	1.7	0.3	6.7
Uppsala	17	36	107	59	19	238
(percentages)	0.7	1.5	4.6	2.5	0.8	10.1
Stockholm	11	10	81	58	14	174
(percentages)	0.5	0.4	3.5	2.5	0.6	7.5
Karolinska Medical School	10	23	54	30	7	124
(percentages)	0.4	1	2.3	1.3	0.3	5.3
Royal Technical Institute	7	15	45	31	15	113
(percentages)	0.3	0.7	1.9	1.3	0.6	4.8
Göteborg	19	43	83	51	20	216
(percentages)	0.8	1.9	3.6	2.2	0.9	9.4
Chalmers University of Technology	1	28	55	28	12	124
(percentages)	0.04	1.2	2.4	1.2	0.5	5.3
Others	89	303	433	269	87	1181
(percentages)	3.8	13.3	18.6	11.6	3.7	51
Total	160	467	953	565	179	2324
(percentages)	6.9	20.8	40.2	24.3	7.8	100

Table 5 below displays the frequency distribution of 1994 graduates' disposable income in 1998 (in 1,000 SEK) by gender. About 13 percent of the graduate males had less than 200,000 SEK compared to 22 percent of graduate females. At the other extreme, nearly 14 percent of graduate males had more than 400,000 SEK compared to only three percent.

Table 5: Frequency distribution of 1994 graduates' disposable income in 1998 (in 1,000 SEK) by gender

	<100	100-200	200-300	300-400	>400	total
Male	45	98	395	389	147	1074
(percentages)	1.9	4.2	17	16.8	6.3	46.2
Female	116	386	539	174	35	1250
(percentages)	5	16.6	23.2	7.5	1.5	53.8
Total	6.9	484	934	563	1.5	2324
(percentages)	6.9	20.8	40.2	24.3	7.8	100

4 Econometric specifications

The economic analysis of university-graduate employment was done using two very broad sets of statistical models: *transition-models (both in a competing-risks- and single destination framework)* and *a selection-corrected Tobit-type of fixed-effects earnings-models*. In the first, the interest is on the waiting-time until the transition of interest, thus a dynamic analysis based on longitudinal data. In the second, earnings were subjected to static analysis based on cross-sectional data.

4.1 Transition-models in a competing-risks framework

Transition models in a competing risks framework can analyze waiting-times until multiple different events of interest where the occurrence of each may have a different causal structure. The same covariates may be relevant but each event has an independent set of parameters.

Given that T be a nonnegative random variable representing the length of a time period spent by an individual in a given state (it measures the length of time before graduation). Now, suppose that individuals under study can experience any one of m distinct transition types ($m = 4$ in our case and these are transitions to employment, non-employment, further education and others). For each individual we observe, possibly subject to right censoring, the time to transition and the type of transition. When time T is continuous and measured precisely so that there are no ties such continuous time survival procedure could be adopted: let x be a vector of covariates. In a competing-risks framework, a cause-specific or type-specific model can be represented by

$$h_j(t;\mathbf{x}) = \lim_{dt \to 0} dt^{-1} P(t \leq T < t + dt, \quad J = j \mid T \geq t, \ \mathbf{x}) \tag{1}$$

where $j = 1, ..., m$ and $t > 0$; $h_j(t;\mathbf{x})$ denotes the instantaneous risk of experiencing a transition of type j in the time interval $(t \leq T < t + dt)$ given that no transition occurred before $T = t$. The overall hazard-rate can be obtained by summing the transition specific hazard rates, that is,

$$h(t; \ \mathbf{x}) = \sum_{j=1}^{m} h_j(t; \ \mathbf{x}) \tag{2}$$

The overall survivor function is

$$S(t; \ \mathbf{x}) = P(T > t \mid \mathbf{x}) = \exp\left(- \int_0^t h(u; \ \mathbf{x}) \ du \right)$$

if we substitute in the transition specific hazard rates then we obtain

$$\begin{aligned} S(t; \ \mathbf{x}) &= \exp\left(- \int_0^t \sum_{j=1}^{m} h_j(u; \ \mathbf{x}) \ du \right) \\ &= \prod_{j=1}^{m} \exp\left(- \int_0^t h_j(u; \ \mathbf{x}) \ du \right) \end{aligned} \tag{3}$$

The density-function for the time until a type j transition is then

$$\begin{aligned} f_j(t;\mathbf{x}) &= \lim_{dt \to 0} dt^{-1} P(t \leq T < t + dt, \ J = j \mid \mathbf{x}) \\ &= h_j(t; \ \mathbf{x}) \ S(t; \ \mathbf{x}) \end{aligned} \tag{4}$$

It must be noted, however, that $f_j(t; \ \mathbf{x})$ is not the density-function of the duration-time. In particular

$$\int_0^\infty f_j(t; \ \mathbf{x}) \ dt = P(J = j \mid \mathbf{x}) = \boldsymbol{\pi}_j(\mathbf{x}) \tag{5}$$

where $\boldsymbol{\pi}_j(\mathbf{x})$ is the probability of transition into the jth state $j = 1, ..., m$, given the covariate-vector $\mathbf{x}$, with the relationship

$$\sum_{j=1}^{m} \boldsymbol{\pi}_j(\mathbf{x}) = 1 \tag{6}$$

If $t_{j1} < t_{j2} < ... < t_{jn_j}$ represents the n_j uncensored durations until the transition j, then the likelihood-function may be rewritten as

$$L = \prod_{j=1}^{m} \prod_{k=1}^{n_j} h_j(t_{jk};\ \mathbf{x}_{jk}) \prod_{i=1}^{n} S_j(t_i; \mathbf{x}_i) \tag{7}$$

where $\mathbf{x}_{jk}$ is the covariate of an individual with the observed noncensored duration t_{jk} and

$$S_j(t_i;\ \mathbf{x}_i) = \exp\left(- \int_0^{t_i} h_j(u;\ \mathbf{x}_i)\ du \right) \tag{8}$$

The likelihood-function may be divided into the product

$$L = \prod_{j=1}^{m} L_j \qquad \text{with} \qquad L_j = \prod_{k=1}^{n_j} h_j(t_{jk};\ \mathbf{x}_{jk}) \prod_{i=1}^{n} S_j(t_i;\ \mathbf{x}_i) \tag{9}$$

The L_j-factors may be further rearranged as

$$L_j = \prod_{i=1}^{n} [h_j(t_{jk};\ \mathbf{x}_{jk})]^{\delta_{ij}} \ \ S_j(t_i;\ \mathbf{x}_i) \tag{10}$$

$$\text{with } \delta_{ij} = \begin{cases} 1 \ \text{ if for individual } i \text{ a transition to state } j \ \text{ occurs at time } t_i \\ 0 \text{ otherwise} \end{cases}. \tag{11}$$

The log-likelihood function $\ln L = \sum_{j=1}^{m} \ln L_j$ can be maximized separately for each transition type $j = 1, ..., m$, given that the transition-specific hazard rates $h_j(t \mid \mathbf{x})$ are dependent upon the parameter-vector θ_j, where the θ's have no common components. In particular, a parametric model $h_j(t;\ \mathbf{x},\ \theta_j)$ can be specified for the type-specific hazards.

For the first transitions after graduation to various states, as well as later transitions from non-employment to work, we first plotted the smoothed non-parametric hazard functions against time and the hazard functions displayed non-monotonic curves. The log-logistic model was ideal in catching the turning-points in these cases. We then fitted log-logistic model[3] in the competing-risks framework discussed above since in the case fitting a proper parametric distribution function, one can obtain more efficient estimates than estimates of a semi-parametric model. In addition to that, we also fitted Cox proportional-hazards model. Since Cox's method does not require some particular probability distribution to represent survival times, is considerably more robust. It can also accommodate both discrete and continuous measurement of transition times. The cause-specific hazard functions mentioned above, can be modeled by using Cox model in the following way:

$$h_j[t; \mathbf{x}_i] = h_{0j}(t) \exp[\mathbf{x}_i' \beta_j], \qquad j = 1, ..., m, \tag{12}$$

where $\mathbf{x}_i$ is a vector of covariates, h_{0j} and β_j are the baseline hazards and the regression coefficients respectively which vary arbitrarily over the m transition types. As before, let $t_{j1} < t_{j2} < ... < t_{jn_j}$ represents the n_j uncensored durations until the transition j, $j = 1, ..., m$. The corresponding partial likelihood is

$$L(\beta_1, ...\beta_m) = \prod_{j=1}^{m} \prod_{i=1}^{k_j} \frac{\exp[x_i \beta_j]}{\sum_{l \in R(t_{ji})} \exp[\mathbf{x}_i' \beta_j]} \tag{13}$$

the arbitrary baseline hazard function has been eliminated and the resulting likelihood can be used for inferences about $\beta_j' s$.

4.2 Earnings-models

There was a problem with the data for estimating earnings-equations, since those who were not employed reported no earnings. To avoid sample-selection bias which can lead to inconsistent estimation of the behavioral parameters of interest, we had to "correct" the sample.

Given unknown population parameter-vectors β and α suppose that y^* and z^*

[3]The log-logistic hazard function is $h(t) = \frac{\gamma\lambda(\gamma t)^{\gamma-1}}{1+(\gamma t)^{\gamma}}$ where $\lambda = \exp[-(\beta_0 + \beta_2 x_2 + ... + \beta_k x_k)]$ and the corresponding density function is
$f(t) = \frac{\lambda^{\frac{1}{\gamma}} t^{\frac{1}{\gamma - 1}}}{\gamma(1+(\lambda t)^{\frac{1}{\gamma}})^2}$

are two latent variables based on exogenous variables $\mathbf{x}$ and $\mathbf{w}$, such that

$$\begin{aligned} y^* &= \boldsymbol{\beta}'\mathbf{x}+u \\ y &= 0 \quad if \quad y^* \leq 0 \\ y &= y^* \quad otherwise \end{aligned} \tag{14}$$

$$\begin{aligned} z^* &= \boldsymbol{\alpha}'\mathbf{w} + \mathbf{v} \\ z &= 1 \quad if \quad z^* > 0 \\ z &= 0 \quad if \quad z^* \leq 0 \end{aligned} \tag{15}$$

Values of y and $\mathbf{x}$ will only be observed when $z = 1$ and y is censored at 0. The residuals are assumed to be distributed according to

$$\begin{aligned} [u] &\sim N[0, \sigma_u^2] \\ [v] &\sim N[0, 1] \\ [u, v] &\sim N[0, 0, \sigma_u^2, 1, \rho] \end{aligned} \tag{16}$$

A log-likelihood can then be written as

$$\sum_{z=0} \log(P(z = 0)) + \sum_{z=1} \log(P(z = 1)\ f(y \mid z = 1) \tag{17}$$

The first term is exactly the same as the corresponding term in a probit- model for z by itself. Since u and v follow bivariate normal probability distribution one can write

$$z = \boldsymbol{\alpha}'\mathbf{w} + \rho(1/\sigma(y^* - \boldsymbol{\beta}'\mathbf{x})) + \varepsilon, \quad \varepsilon \sim N(0, (1 - \rho^2)) \tag{18}$$

consequently

$$P(z = 1) = \Phi\left(\frac{\boldsymbol{\alpha}'\mathbf{w} + \rho((\ y_{1t} - \boldsymbol{\beta}'\mathbf{x})/\sigma)}{(1 - \rho^2)^{1/2}}\right) \tag{19}$$

and finally the full likelihood function becomes

$$\sum_{z=0} \log(\Phi(-\alpha' w)) + \sum_{z=1} \log(\phi(\ y - \beta' x)/\sigma)) +$$
$$\sum_{z=1} \log\left(\Phi\left(\frac{\boldsymbol{\alpha}'\mathbf{w} + \rho((\ y - \boldsymbol{\beta}'\mathbf{x})/\sigma)}{(1-\rho^2)^{1/2}}\right)\right). \tag{20}$$

(see Gourieroux (1989)). The full likelihood function above can be estimated by maximum likelihood method.

5 Estimation results

5.1 First transitions after graduation

Tables A6a and b in the Appendix report the results of estimating parameters for initial transitions after graduation. Of the 2598 new graduates in 1994, 662 started working within one month. In the first outcome, we focus on the transition from graduation to different possible states. The states are defined as work, non-employment, further studies and all other categories. After graduation, from 2598 students, 662 became employed within one month, 1760 remained non-employed, 124 continued to further studies, 52 transited to “other” states. As expected, the parameter signs in the two models are inverse (since accelerated failure time estimates are in log-waiting time format, while the proportional hazards estimates are in log-hazard format).

Controlling for other covariates, the ratio of estimated hazard of transiting to non-employment for males, to the estimated hazard for females is 0.8 ($= exp[-0.215]$, looking at column 3). This means for males, the hazard of being non-employed is about 80 percent of females’ hazard. For each additional year of age the hazard of transiting to employment goes up by an estimated 1.8 percent ($100 * (exp[0.018] - 1)$, column 1). Possibly, by increasing ages graduates are becoming more goal-oriented. It is also true for the hazard of of transiting to “other”, such as child care. Being a parent during university reduced the hazard of transiting to employment and increased the hazard of transiting to non-employment. Having parents with university education reduced the hazard of transiting to non-employment.

Perhaps surprisingly, having studied in a foreign country either before or during university reduced the hazard of transiting to employment. The hazard of transiting to employment for those having studied in a foreign country during university was about 75 percent ($= 100 * exp[-0.283]$) of the hazard for those who have not studied abroad. Somewhat paradoxically the amount of student loans accumulated during university studies reduced the hazard of transiting to employment, to non-employment and increased the hazard of transiting to further studies.

The professionals analyzed, such as economists, engineers, and doctors had higher hazards of transiting to employment than "others".

5.2 Transitions from initial non-employment to work

Next we analyzed the second transitions of the 1760 individuals who initially remained non-employed after university graduation. At the second stage, we focused on the spell of non-employment after graduation. 1642 of them eventually found work during the study period. Table A7 (in the Appendix) shows the estimated effects on hazards and waiting times. Unlike the other analysis neither age nor the amount of student loans gave statistically significant results. On the other hand, being able to finance one's university education from one's own resources corresponded to increased waiting-time in non-employment.

Compared to those from "other" universities, those who graduated from Karolinska Medical School and Chalmers University of Technology had about 32 percent ($= 100 * (\exp[0.28] - 1)$) and those from Royal Technical Institute had about 21 percent ($= 100 * (\exp[0.197] - 1)$) greater hazard of leaving non-employment and transiting to work, while graduates of Göteborg had lower hazard of exiting non-employment.

5.3 The annual earnings of university-graduates

Table A8 (in the Appendix) shows estimated effects on the selection-corrected 1998 annual disposable income of the 1994 university graduates. Marginal effects on the probability part were very small[4] and had nearly no impact on total effects so the reported effects are based on regression part.

Males tended to earn about 28 percent more[5] than females, a not unexpected

[4] Selection-parameter ρ is not statistically significant.

[5] indirect effect was .125e-03

result. Student loans had positive and significant effect on earnings. Effects from universities of graduation were not statistically significant, but the professionals analyzed tended to earn more than "others": economists 42 percent; engineers 34 percent; and doctors 36 percent[6].

6 Summary and conclusions

We have examined the initial and secondary labour-market experiences of 2,598 graduates of Swedish universities in 1994, including their disposable incomes in 1998.

We initially analyzed first transitions after graduation: to work; further education; "other" including pregnancy, childcare, and the military; and remaining non-employed. A semi-parametric Cox proportional-hazards model was used in order to analyze the effects of a variety of individual, parental, socioeconomic, and academic variables on the hazards of these transitions; and a parametric log-logistic accelerated failure-time model was used to analyze the effects on waiting-times before transitions. The subsequent work-transitions of those who initially remained non-employed were then analyzed similarly. Finally, an earnings-equation was estimated to analyze the effects on annual disposable income four years after graduation, in 1998. The main findings were:

Age, being economist, engineer, or doctor, and having graduated from Karolinska Medical School increased the hazard of transiting to employment in the first transitions, while being a parent during university studies, having foreign experience decreased it.

For subsequent work-transitions of those who initially remained non-employed, compared to those from "other" universities, those who graduated from Karolinska Medical School, Chalmers University of Technology and those from Royal Technical Institute had higher hazards of transiting to employment. University entrance examination increased the hazard of transiting to employment as well.

As expected, gender was an important determinant in earnings differentials, even when other individual, socioeconomic and academic variables were controlled. The same can be said for professionals such as economists, engineers and medical doctors. Again, having a university entrance examination had a major impact on the earnings.

[6] Again indirect effects were very small: 732e-03; .593e-03; and 394e-03,

We understand that the function of university entrance examination becomes increasingly important not only for university entrance but also in the employment process as a screening tool used by employers. Those who take such an examination can be seen very flexible people who adapt better to the daily social structure. The increasing importance of university entrance examination in Sweden requires that its structure, role and consequences must be discussed formally.

A surprising result was the negative impact of foreign experience. Probably those who study or work abroad are not considered as goal-specific in the eyes of employers. On the other hand, the study loans might have played a motivating role for higher earnings and in order to pay them back.

References

[1] Blom, R. (2000), "*Technical Report: Higher Education and Graduate Employment in Sweden,*" Mail Survey Function Statistics, Sweden.

[2] Bratberg, E., and Nilsen, O. A. (2000), "*Transitions From School to Work and the Early Labour Market Experience,*" Oxford Bulletin of Economics and Statistics, Vol. 62, Special Issue, pp. 909-29.

[3] Eckstein, Z., and Wolpin, K. I. (1995), "*Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model,*" The Review of Economic Studies, Vol. 62, No.2, pp. 263-86.

[4] Gourieroux, C. (1989), "*Econometrie des Variables Qualitatives,*" 2^e edition, Economica, Paris.

[5] Jensen, P., and Westergard-Nielsen, N. C. (1987), "*A Search Model Applied to the Ttransition from Education to Work,*" The Review of Economic Studies, Vol. 54, No.3, pp. 461-72.

[6] Horowitz, J. L., and Manski, C. F. (1998), "*Censoring of Outcomes and Regressors due to Survey non-response: Identification and Estimation using weights and Imputations*," Journal of Econometrics, Vol. 84, pp. 37-58.

[7] Kalbfleisch, J-D., and Prentice, R-L. (2002), "*The Statistical Analysis of Failure Time Data,*" 2nd edition, Wiley, New York.

[8] Lundström, S., and Särndal, C. E. (2002), "*Estimation in the Presence of nonresponse,*" SCB Statistics Sweden.

[9] Nilsen, O. A., Risa, A. E., and Torstensen, A. (2000), "*Transitions from Employment among Young Norwegian Workers,*" Journal of Population Economics, Vol. 13, pp. 21-34.

[10] Stiglitz, J. E. (1975), "*The Theory of Screening: Education and the Distribution of Income,*" American Economic Review, Vol. 65, No.3, pp. 283-300.

[11] Särndal, C-E., Swensson, B., and Wretman, J. (1992), "*Model Assisted Survey Sampling,*" (Springer Series in Statistics), New York: Springer-Verlag.

[12] TSER (Targeted Socio-Economic Research) (2000), "*Final Report: Higher Education and Graduate Employment in Europe,*" Project funded by the European Community.

[13] Wolpin, K. I. (1987), "*Estimating a Structural Search Model: The Transition from School to Work,*" Econometrica, Vol. 55, pp. 801-17.

Appendix

A. Estimation in the presence of non-response

Non-response errors are the best known of the "non-sampling errors", which are not caused simply by the limitation of the investigation to a sample, rather than the entire population[7]. They may distort the representation of the true population and consequently distort inferences based on the observed data using standard methods.

For estimation under ideal conditions, consider the finite population of N elements $U\{1, ..., i, ..., N\}$,called the target-population. We could, for example estimate the total

$$Y=\sum_U y_k \tag{21}$$

where y_k is the value of the study variable, y, for the kth element.

Now suppose that s is a probability-sample of size n, drawn from the target-population U with probability $p(s)$. The inclusion- probabilities, known for all $i \in U$, are than $\pi_i = \sum_{s \ni i} p(s)$. Assuming that the design is such that $\pi_i > 0$ for all elements. Let $d_i = 1/\pi_i$ denote the design weight of element i. These weights are very important for computing estimators.

When the sampling design has been fixed, the inclusion probabilities π_i and the sampling design weights $d_i = 1/\pi_i$ are fixed, known quantities. Then the unbiased estimator of the total Y, is given by the Horvitz-Thompson estimator

$$\hat{Y}_{HT} = \sum_S d_i y_i \tag{22}$$

This estimator is unbiased for Y, under any sampling-design satisfying $\pi_i > 0$ for all elements i.

A wider and more efficient class of estimators are those that use auxiliary information explicitly at the estimation-stage. Some information may have been used at the design-stage. We can denote the auxiliary information vector by $\mathbf{x}$, and its value for individual i by $\mathbf{x}_i = (x_{i1}, ..., x_{ij}, ..., x_{iJ})$, a row vector with J components, where x_{ij} is the value for individual i of the j:th auxiliary variable. Suppose the population total, $\sum_{i=1}^{N} \mathbf{x}_i$, is accurately known. An estimator that uses this information is the *generalized regression estimator* (GREG- estimator), explained in Särndal,

[7] This part relies heavily on Lundström and Särndal (2002), and Särndal et al. (1992).

Swensson and Wretman (1992), Chapters 6 and 7, and given by

$$\hat{Y}_{GREG} = \hat{Y}_{HT} + \left(\sum_U \mathbf{x}_i - \sum_S d_i \mathbf{x}_i \right)' \hat{\mathbf{B}} \tag{23}$$

where $\hat{\mathbf{B}} = \left(\sum_U d_i c_i \mathbf{x}_i \mathbf{x}_i'\right)^{-1} \left(\sum_S d_i c_i \mathbf{x}_i y_i\right)$ is a vector of regression coefficients, obtained by fitting the regression of y on $\mathbf{x}$, using the data $(y_i, \mathbf{x}_i)$ for the elements $i \in s$. The data are weighted by $d_i c_i$, where the factor c_i is specified by the researcher. A simple choice is to take $c_i = 1$ for all i.

The GREG estimator is "almost unbiased". The bias, although not exactly zero, tends to zero with increasing sample size. The term $(\sum_U \mathbf{x}_i \; - \; \sum_S d_i \mathbf{x}_i)$ in the formula for $\hat{Y}_{GREG}$ can be viewed as a *regression adjustment* applied to the HT estimator, $\hat{Y}_{HT} = \sum_S d_i y_i$. The effect is an important reduction of the variance of $\hat{Y}_{HT}$, especially when there is a strong regression relationship between y and $\mathbf{x}$.

The estimator $\hat{Y}_{GREG}$ can be written as a linearly weighted sum of the observed values y_i,

$$\hat{Y}_{GREG} = \sum_S d_i g_i y_i \tag{24}$$

where the total weight given to the value y_i is the product of two weights, the design weight $d_i = 1/\pi_i$, and the weight, g_i, which depends both on the individual i and on the sample s of which i is a member as

$$g_i = 1 + c_i \left(\sum_U \mathbf{x}_i - \sum_S d_i \mathbf{x}_i \right)' \left(\sum_S d_i c_i \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \mathbf{x}_i \tag{25}$$

The value of g_i is near unity for a majority of individuals $i \in s$, and approaches unity as the sample gets larger.

In the case of non-response, assume that response in the sample is obtained for the elements in a set denoted r with size m. Full response implies that $r = s$. non-response implies that r is a proper set of s with size n. The non-response set is denoted $o = s - r$ with size $n - m$.

Suppose a vector $\mathbf{x}_i$ as follows: sex, age, residential area, citizenship, income, and marital status. It would be chosen if the survey had full response, so that $r = s$. A required input is the population total of the $\mathbf{x}_i$ -vector, $\sum_U \mathbf{x}_i$. Then the weights which will satisfy the calibration equation are

$$\sum_S d_i g_i \mathbf{x}_i = \sum_U \mathbf{x}_i \tag{26}$$

However, with non-response, values for y_i are available only for the elements i in the response set r, a subset of the sample s. Then the calibration estimator is, like the GREG estimator, formed as linearly weighted sum of the observed y_i values. It is defined by

$$\hat{Y}_W = \sum_r w_i y_i \tag{27}$$

where $w_i = d_i v_i$ with

$$v_i = 1 + c_i \left(\sum_U \mathbf{x}_i - \sum_r d_i \mathbf{x}_i \right)' \left(\sum_r d_i c_i \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \mathbf{x}_i \quad for \;\; i \in r \tag{28}$$

The principle behind the derivation of the formula is to minimize a function measuring the distance between the "old" weights, d_i, and the "new weights", w_i, subject to the calibration equation

$$\sum_r d_i v_i \mathbf{x}_i = \sum_U \mathbf{x}_i \tag{29}$$

Table A1: Age- frequencies before and after calibration

Age	Frequency Observed	Frequency after Calibration	Rounded off
.	5	.	.
21	2	1.950	2
22	15	14.131	14
23	77	76.777	77
24	176	151.923	152
25	281	239.459	239
26	366	297.806	298
27	369	306.897	307
28	329	313.412	313
29	237	216.954	217
30	140	143.161	143
31	95	108.152	108
32	81	89.539	90
33	67	81.446	81
34	48	60.279	60
35	35	35.715	36
36	37	43.490	43
37	16	26.222	26
38	25	35.605	36
39	29	39.150	39
40	22	36.839	37
41	16	26.632	27
42	25	35.258	35
43	8	11.271	11
44	22	40.509	41
45	14	25.441	25
46	15	32.384	32
47	11	23.610	24
48	4	6.528	7
49	5	10.664	11
50	5	9.273	9
51	7	12.976	13
52	2	4.834	5
52	2	6.902	7
54	2	5.020	5
55	1	2.857	3
57	2	7.710	8
58	2	6.476	6
59	1	2.455	2
61	1	5.749	6
62	1	2.543	3
	2598	2598.000	2598

Table A2: Distribution of graduates by annual income groups

Aincg[1]	**Frequency**	**Saincg**	**Saincgw**	**aincgc**	**aincgr**	**aincgwc**	**aincgwr**
0-85	489	489	3380.49	131.252	131	137.804	138
85-160	132	264	1794.46	70.860	71	73.150	73
160-235	799	2397	16398.99	643.377	643	668.497	668
235-310	661	2644	15574.37	709.674	710	634.882	635
310-	517	2585	18022.20	693.838	694	734.667	735

[1] in thousands of SEK

aincgr shows the theoretical distribution of graduates by income groups
aincgwr shows the empirical distribution of graduates after using calibartion reweights for income groups (The figures for the last two scores are close to each other)

Table A3: The gender-distribution of graduates before calibration and after calibration:malewc and femalewc (rounded as malewr and femalewr)

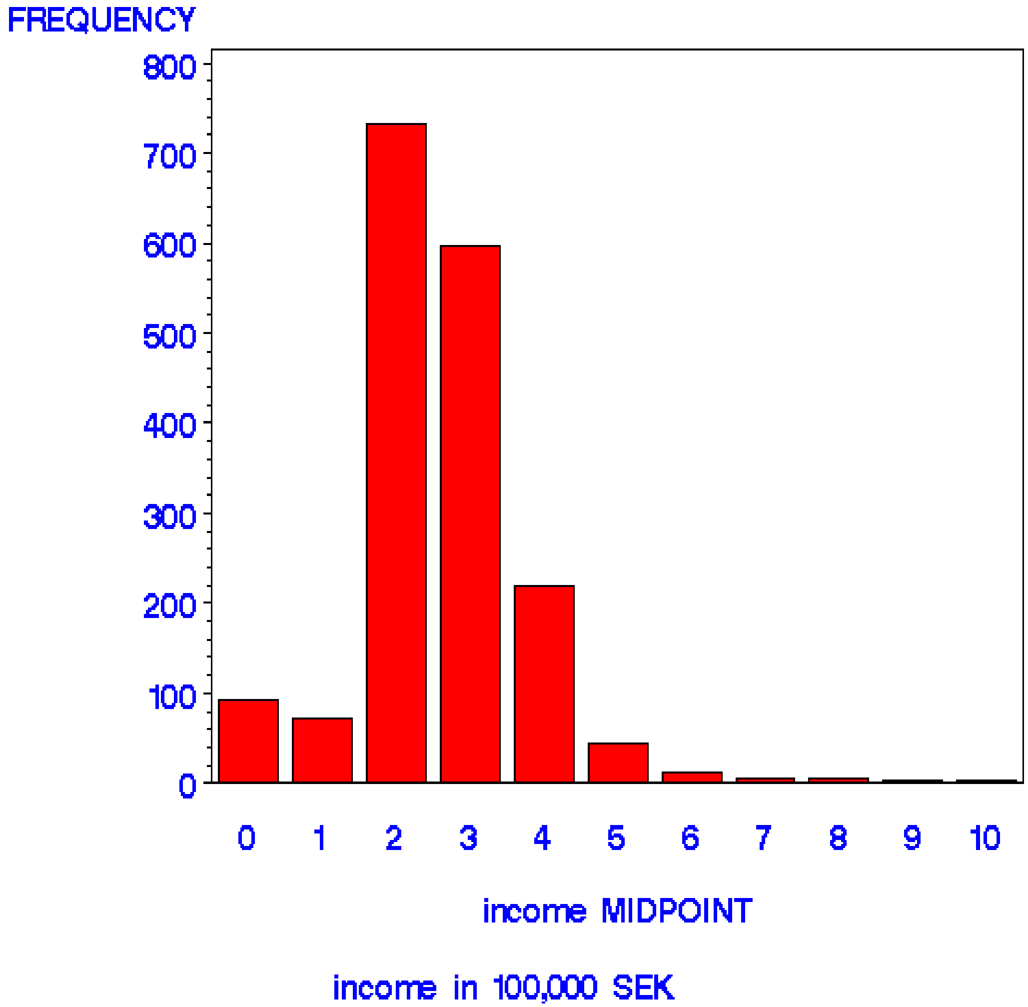
Sex	**Frequency**	**ssex**	**ssexw**	**malewc**	**malewr**	**femalewc**	**femalewr**
Male	1134	1134	1134	1130.53	1131	1467.47	1467
Female	1464	2928	0	1130.53	1131	1467.47	1467

Table A4: Definition of Variables

Background Variables	
Male	dummy variable=1 if male, 0 if female
Age at the start of university	Age at the start of university
Immigrant	dummy variable=1 if born as foreigner citizen; 0 if born Swedish
Parents with university education	dummy variable=1 if at least one parent had university degree
Having Child during University	dummy variable=1 if student was parent during university, 0 otherwise
University entrance with:	
upper-secondary diploma	University entrance with upper secondary diploma
college/university entrance exam.	University entrance with college/university examination
work experience	University entrance with working life experience
unknown	University entrance with unknown characteristics
Foreign experience:	
Worked before university	dummy variable=1 if prior work experience in foreign country(ies)
Studied before university	dummy variable=1 if prior studies in foreign country(ies)
Studied during university	dummy variable=1 if studied in foreign country(ies) during university
Financed university studies by:	
grant	Financed university studies with grant
work	Financed university studies by working
own resources	Financed university studies from own resources
parents	Financed university studies through parents
no one	Financed university studies by no one
other means	Financed university studies by other means
Student loan to finance studies	The amount of loan taken to finance studies, in SEK
University name:	
Umeå University	Umeå University
Uppsala University	Uppsala University
Stockholm University	Stockholm University
Karolinska Medical School	Karolinska Medical School
Royal Tecnical Institute	Royal Tecnical Institute
Göteborg University	Göteborg University
Chalmers University of Technology	Chalmers University of Technology
Other	Others
Graduated in as:	
economics/business management	Graduated as economist
engineer	Graduated as engineer
medical doctor	Graduated as medical doctor
other	Graduated as other
Logarithm of annual income	Logarithm of Annual Income in 1998

Table A5: Descriptive Statistics for the independent variables

	mean	std. dev.
Background Variables		
Male	0.436	(0.496)
Age at the start of University	24.132	(5.248)
Born Outside Sweden	0.033	(0.178)
Having Child during University	0.156	(0.363)
Parents with University education	0.371	0.376
University entrance with:		
upper-secondary diploma	0.748	(0.433)
college/university entrance examination	0.048	(0.215)
working life experience	0.089	(0.284)
with unknown characteristics	0.053	(0.225)
Foreign experience:		
Worked before university	0.191	(0.392)
Studied before university	0.294	(0.456)
Studied in during university		
Financed university studies by:		
grant	0.816	(0.387)
work	0.089	(0.285)
own resources	0.024	(0.152)
parents	0.036	(0.185)
no one	0.011	(0.103)
other means	0.022	(0.147)
Student loan to finance studies/100	1.653	(1.019)
University name:		
Umeå University	0.065	(0.245)
Uppsala University	0.104	(0.305)
Stockholm University	0.073	(0.259)
Karolinska Medical School	0.048	(0.213)
Royal Tecnical Institute	0.051	(0.219)
Göteborg University	0.094	(0.292)
Chalmers University of Technology	0.051	(0.219)
Other	0.514	(0.334)
Graduated in as:		
economics/business management	0.150	(0.356)
engineer	0.231	(0.421)
medical Doctor	0.136	(0.343)
other	0.483	(0.499)
Log Annual Income (1998)	11.638	4.239



The Distribution of 1994 graduates' disposable income in 1998

[Figure 1]

Table A6a: Estimated parameters for initial transitions after graduation, alternate models

N=2598 obs	to employment (n=662)		to non-employment (n=1704)	
	Cox proportional hazard	Accelerated failure time log-logistic	Cox proportional hazard	Accelerated failure time log-logistic
Constant*		3.863^{0}		3.024^{0}
		(0.126)		(0.100)
Background variables				
Male		0.423	-0.215^{0}	0.183^{0}
		(0.295)	(0.052)	(0.043)
Age	0.018^{2}	-0.947^{0}	-0.024^{0}	0.131^{0}
	(0.008)	(0.253)	(0.004)	(0.014)
Immigrant		-0.137		0.126
		(0.092)		(0.109)
Parents with university education		0.140	-0.127^{6}	0.288^{0}
		(0.388)	(0.068)	(0.064)
Parent during university	-0.662^{0}	0.182^{0}	0.281^{0}	-0.642^{0}
	(0.122)	(0.033)	(0.072)	(0.048)
University entrance with:				
upper-secondary diploma		0.067		
		(0.048)		
college/university entrance exam.		-0.208^{1}		
		(0.083)		
working experience		0.059		
		(0.063)		
unknown characteristics		0.000		
Foreign Experience				
Worked before		-0.029		
		(0.037)		
Studied before university	-0.153	0.061^{9}		0.072
	(0.109)	(0.036)		(0.053)
Studied during university	-0.283^{0}	0.097^{0}		0.073
	(0.098)	(0.033)		(0.055)
Financed university studies by				
other means		0.000		
grant		-0.043		
		(0.082)		
work		0.223^{0}		
		(0.081)		

"continued"

Table A6a: (cont.) **Estimated parameters for initial transitions after graduation, alternate models**

N=2598 obs	to employment (n=662)		to non-employment (n=1760)	
own resources		0.018		
		(0.133)		
parents		0.087		
		(0.108)		
no one		-0.041		
		(0.132)		
Student loan to finance studies	-0.107[1]	0.078[0]	-0.158[0]	0.257[0]
	(0.042)	(0.017)	(0.024)	(0.018)
University name:				
Umeå University		0.098[5]		
		(0.051)		
Uppsala University		0.029		
		(0.047)		
Stockholm University		0.008		
		(0.055)		
Karolinska Medical School		-0.163[0]		
		(0.045)		
Royal Tecnical Institute		-0.019		
		(0.052)		
Göteborg University		0.007		
		(0.047)		
Chalmers University of Technology		0.094		
		(0.060)		
other		0.000		
Graduated in as				
economist		-0.075[2]		0.337[0]
		(0.034)		(0.056)
engineer		-0.141[0]		0.411[0]
		(0.037)		(0.058)
medical doctor		-0.242[0]		0.353[0]
		(0.075)		(0.092)
other				
sigma		0.176[0]		0.445[0]
		(0.005)		(0.008)
Log Likelihood	-2621.430	-177.083	-8949.680	-2067.029

[1] We couldn't use all variables in Cox-proportional hazard model due to its sensibility

* Variables with coeficients 0.000 denote reference categories

Table A6b: **Estimated parameters for initial transitions after graduation, alternate models**

N=2598 obs	to further studies[1] (n=124)		to other[2] (n=52)	
	Cox Proportional Hazard	Accelerated Failure Time Log-Logistic	Cox Proportional Hazard	Accelerated Failure Time Log-Logistic
Constant*		3.357[0]		4.335[0]
		(0.416)		(0.283)
Background variables				
Male		0.055		0.025[0]
		(0.064)		(0.115)
Age		0.005	0.136[3]	-0.033[0]
		(0.008)	(0.065)	(0.014)
Immigrants		0.024		0.161
		(0.109)		(0.480)
Parents with university education		-0.040		0.108[0]
		(0.076)		(0.097)
Parent during University		0.027		0.155
		(0.086)		(0.144)
University entrance with				
upper-secondary diploma		0.040		
		(0.094)		
college/university entrance exam.		0.192		
		(0.177)		
work experience		-0.036		
		(0.192)		
unknown characteristics		0.000		
Foreign experience				
Worked before university		-0.147		-0.064
		(0.122)		(0.098)
Studied before university		-0.030		0.047
		(0.091)		(0.106)
Studied during university		0.037	-0.936[2]	0.226[2]
		(0.071)	(0.408)	(0.102)
Financed University Studies by				
grant		0.339		
		(0.303)		
work		0.452		
		(0.317)		
own resource		0.261		
		(0.270)		
parents		0.183		
		(0.347)		

"continued"

Table A6b: (cont.) Estimated parameters for initial transitions after graduation, alternate models

n=2598 obs	to further studies (n=124)		to other (n=52)	
	Cox-proportional hazard	Accelerated failure time log-logistic	Cox-proportional hazard	Accelerated failure time log-logistic
no one		0.324		
		(0.332)		
other means		0.000		
Student loan to finance studies	0.256[1]	-0.074[6]		0.066[0]
	(0.105)	(0.038)		(0.043)
University name				
Umeå University	-0.798[9]	0.235	-1.494[1]	0.346[10]
	(0.483)	(0.155)	(0.616)	(0.209)
Uppsala University	-0.259	0.086	-1.690[0]	0.332[0]
	(0.283)	(0.114)	(0.543)	(0.088)
Stockholm University	0.247	-0.233[6]	0.048	0.005
	(0.408)	(0.125)	(0.559)	(0.133)
Karolinska Medical School	0.357	-0.009		
	(0.493)	(0.232)		
Royal Tecnical Institute	-0.128	0.022		
	(0.448)	(0.174)		
Göteborg University	0.053	-0.021	-1.313[5]	0.306
	(0.356)	(0.095)	(0.672)	(0.210)
Chalmers University of Technology	-0.142	0.194		
	(0.386)	(0.156)		
Other	0.000	0.000	0.000	0.000
Graduated in as				
economist		0.024		-0.129
		(0.112)		(0.505)
engineer		0.040		0.067
		(0.072)		(0.113)
medical doctor		0.268[6]		0.238
		(0.147)		(0.193)
other		0.000		0.000
sigma		0.142[0]		0.116[0]
		(0.012)		(0.018)
Log Likelihood	-331.1497	-7.685	-92.320	6.492

[1] Since Cox-Proportional hazard was very sensitive to some group of variables, we could not use all the variables in the estimations

[2] Other category includes child care, pregnancy, military service..etc

[*] As before variables with coeficients 0.000 denote reference categories

Table A7: Estimated parameters from initial non-employment after graduation to employment (n=1642)

n=1760

	Cox-proportional hazard[1]	Accelerated failure time log-logistic
Constant		1.470[0]
		(0.330)
Background Variables		
Male	0.034	-0.033
	(0.056)	(0.076)
Age	-0.003	0.003
	(0.006)	(0.007)
Immigrant	-0.009	0.292
	(0.153)	(0.217)
Parents with university education	-0.022	0.001
	(0.072)	(0.009)
Parent during university	-0.042	0.119
	(0.079)	(0.103)
University entrance with		
upper-secondary diploma	0.049	-0.123
	(0.081)	(0.103)
college/university entrance exam.	0.215[10]	-0.359[4]
	(0.132)	(0.175)
work experience	0.046	-0.045
	(0.112)	(0.144)
unknown characteristics	0.000	0.000
Foreign experience		
Worked before university	0.022	-0.078
	(0.067)	(0.092)
Studied before university	0.049	0.001
	(0.071)	(0.094)
Studied during university	0.015	-0.074
	(0.059)	(0.078)
Financed university studies by		
grant	-0.155	0.054
	(0.175)	(0.261)
work	-0.198	0.242
	(0.186)	(0.268)
own resource	-0.310	0.744[2]
	(0.236)	(0.327)
parents	-0.096	0.061
	(0.211)	(0.300)

"continued"

Table A7: (cont.) Estimated parameters from initial non-employment after graduation

n=1760	to employment (n=1642)	
	Cox Proportional Hazard	Accelerated Failure Time Log-Logistic
no one	-0.497[9]	0.781[3]
	(0.295)	(0.374)
other means	0.000	0.000
Student loan to finance studies	0.020	0.050
	(0.033)	(0.044)
University name		
Umeå University	-0.054	0.240[7]
	(0.106)	(0.132)
Uppsala University	-0.043	0.115
	(0.087)	(0.113)
Stockholm University	-0.167	0.180
	(0.114)	(0.134)
Karolinska Medical School	0.280[2]	-0.378[5]
	(0.127)	(0.196)
Royal Technical Institute	0.197[9]	-0.281[10]
	(0.118)	(0.170)
Göteborg University	-0.158[8]	0.394
	(0.091)	(0.116)
Chalmers University of Technology	0.281[2]	-0.296[10]
	(0.128)	(0.177)
other	0.000	0.000
sigma		0.804[0]
		(0.024)
Log Likelihood	-8729.736	-2799.925

[1] Since Cox-Proportional hazard was very sensitive to some group of variables, we could not use all the variables in the model.

Table A8: Estimated effects on the selection corrected of 1994 university graduates' annual disposable income (n=2598)

	Binary Probit	Earnings Equation
Constant	0.596	11.828[0]
	(0.389)	(0.120)
Background Variables		
Male	0.173	0.285
	(0.118)	(0.033)
Age	-0.001	0.002
	(0.008)	(0.003)
Immigrant	-0.131	-0.048
	(0.283)	(0.100)
Parents with university education	-0.097	
	0.135	
Parent during university	0.179	
	(0.143)	
University Entrance with		
unknown characteristics	0.000	0.000
upper-secondary diploma	0.181	0.048
	(0.143)	(0.051)
college/university examination	0.247	0.029[1]
	(0.250)	(0.079)
working experience	0.038	0.054
	(0.201)	(0.065)
Foreign experience	0.000	0.000
Worked before university	0.049	
	(0.132)	
Studied before university	-0.024	
	(0.134)	
Studied during university	-0.023	
	(0.123)	
Financed university studies by		
grant	0.188	
	(0.256)	
work	0.602[5]	
	(0.307)	
own resource	-0.044	
	(0.352)	
parents	0.385	
	(0.366)	

Table A8: (cont.) Estimated effects on the selection corrected of 1994 university graduates' annual disposable income (n=2598)

n=2500	Binary Probit	Earnings Equation
no one	-0.115	
	(0.403)	
other means	0.000	
Student loan to finance studies	0.052	0.039
	(0.061)	(0.015)
University name		
Umeå University	-0.076	0.063
	(0.191)	(0.078)
Uppsala University	-0.010	-0.013
	0.174	(0.046)
Stockholm University	0.182	-0.037
	(0.203)	(0.053)
Karolinska Medical School	0.071	0.090
	(0.249)	(0.068)
Royal Technical Institute	0.262	0.034
	(0.313)	(0.643)
Göteborg University	-0.017	-0.035
	(0.164)	(0.047)
Chalmers University of Technology	0.313	0.160
	0.304	(0.117)
Other	0.000	0.000
Graduated in as		
economist	1.012^{0}	0.418^{0}
	(0.261)	(0.067)
engineer	0.820^{0}	0.337^{0}
	(0.186)	(0.060)
medical doctor	0.545^{7}	0.360
	(0.304)	(0.086)
other	0.000	0.000
Sigma		0.637^{0}
		(0.006)
Rho		-0.004
		(0.329)
Log Likelihood	-2946.598	

Arrival-Cohort Effects on the Incomes of Immigrant Men in Sweden

Kerem Tezic

Department of Economics, Göteborg University

November 2003

Abstract

Arrival-cohort effects on the incomes of first-generation immigrant men in Sweden were analysed using eleven waves of panel-data. Employment-probabilities and earnings were estimated simultaneously in a random-effects model in order to control for individual effects and panel- selectivity due to missing earnings-information. The results indicate that labour-market outcomes differ considerably between immigrant-cohorts. And, although there has been some improvement with time spent in the adopted country, the economic integration is a long process.

Keywords: Immigrants, labour-market integration, unbalanced panel, sample-selection, random-effects.

J.E.L Classification: C33, J15, J61.

Contents

1 Introduction

Do immigrants earn less than native Swedes? If so, does this difference persist? Is the difference related to ethnic origin or arrival-cohorts? Do the earnings-gaps between native Swedes and immigrants get narrower with time spent in the adopted country, or do they persist? In other words, do immigrants catch up? What about labour-market differences between different immigrant groups? These questions are very important in a highly immigrated country such as Sweden, with a recent history of large immigrant flows from many corners of the world, and this study was conducted to try to find answers.

Like other European countries, Sweden had experienced some immigration and emigration already before industrialization, but the extent is not well known until the 1850s (Lundh and Ohlsson, 1994). From then until the 1930s, Sweden was characterized by net emigration. Figure 1 in the Appendix shows immigration to and emigration from Sweden since 1900.

During this period, most of the emigrants moved to the USA, mainly because of the difference in living-standard between the two countries. Nearly 1,2 million Swedes emigrated to the U.S. in total. Emigration fell during the Depression; probably the U.S. did not look so attractive during that period. Then immigration increased dramatically, first during the World War II, when Sweden was neutral and managed to stay out of the war thereby, and then after the war, when Swedish industry boomed and needed increased manpower to supply the needs of war-ravaged Europe. At first immigrants came mainly from northern and western Europe, but soon many where coming from southern Europe as well. During this period, the age-specific employment-intensity was often higher among immigrants than among native Swedes, and unemployment among immigrants was quite low (Ohlsson, 1975).

Starting from mid-70s, immigration largely switched from economic to political, partly due to decline in economic growth and industrial output, and because of resultant immigration restrictions. At the same time Sweden's liberal rules for political refugees led to a new influx of immigrants this time largely from non-European countries (at first Chile in the 1970s, later Iraq, Iran, Afghanistan and many African countries in the 1980s, and then from the former republics of Yugoslavia in the 1990s). Thus the composition of the immigrant-population by country of origin changed significantly.

During the same period, substantial deterioration began in the employment and earnings of immigrants relative to native Swedes (Figures 2a and 2b in the Appendix). This deterioration continued despite the boom in the Swedish economy during 1980s, and got worse during the 1990s. Probably not only supply-side, but also demand-side factors were responsible for this situation. A structural shift of the Swedish economy from industrial to service-oriented increased demand for employees with language and inter-personal skills, and culture-specific ability to deal with authorities and labour-market organizations. Such demand in informal competence made it difficult for immigrants to compete even if they had the same level of formal education.

The change in the type of immigration and the resulting ethnic composition of immigration and make it important to analyze the economic integration of immigrants and their children. Assimilation can be even more difficult than otherwise for the children if their parents were not only immigrants but low-income as well (Österberg, 2000).

This paper analyzes the arrival-cohort effects on the earnings of first-generation immigrant men using the register based Longitudinal Individual Data set (LINDA) during the period 1990-2000. The econometric model used, handles potential sample-selection bias by estimating the employment and earnings equations simultaneously while allowing for random effects in both which allows us to distinguish age and cohort-effects. In terms of both employment probabilities and earnings, our findings are pessimistic for some immigrant groups.

The next section describes previous studies and their relations to this one. Section 3 presents the data, while section 4 develops the model used and discusses econometric issues. Section 5 gives the estimation results, and Section 6 summarizes and draws conclusions.

2 Previous studies

2.1 In the USA and Canada

The literature on the economic integration of immigrants has been dominated by analysis of whether they experience rapid earnings growth over time and whether this leads to their catching up with the earnings of native-born workers within 10-15 years after arrival. Chiswick (1978), based on cross-sectional 1970 U.S. population-census data, suggested that new immigrants rapidly accumulated skills specific to the American labour-market. But a problem with that interpretation is that it was not possible to distinguish in Chiswick's study whether earnings of the typical immigrant had risen with time in the U.S. or whether the average quality of immigrant-cohorts had declined.

Borjas (1985) was questioned Chiswick's findings by arguing that the use of cross-sectional data at a single point in time to estimate immigrants' earnings-integration over time suffered from two major biases: First, if the average quality of recent immigrants were systematically lower than that of earlier ones, the cross-sectional data would overstate their earnings-growth. Second, there could be selection-bias due to the possible return of less-successful immigrants to their country of origin. Borjas thus used a "quasi-panel" of cross-sectional waves based on the 1970 and 1980 censuses, and his findings were considerably different from Chiswick's. He found much less earnings-integration, which he attributed to a decline in the quality of immigrants admitted to United States due to changes in immigration policies. Not surprisingly, he also found that the earnings-integration of later cohorts was slower. Borjas (1995) updated the previous study to include data from the 1990 census finding now that the decline in cohort-quality had slowed.

However, similar to Chiswick (1978), but using the "quasi-panel" method of Borjas (1985), LaLonde and Topel (1991, 1992) found evidence that even if average immigrant-quality, as measured by initial earnings, had declined, the new immigrants in fact assimilated rapidly and their long-run earnings-potential was much like the ethnically-similar native-born. Duleep and Regets (1996), even found that more recent cohorts of immigrants had experienced more rapid growth in earnings than did earlier ones.

The Canadian census is conducted every five years, which provides more frequent observations on immigrant-cohorts than are currently available for the U.S.. And while U.S. immigration policy has paid greater attention to the family reunification, Canada has attempted more to match the skills of immigrants with perceived shortages in the Canadian labour-market. Nevertheless, the findings of Baker and Benjamin (1994), based on the 1971, 1981, and 1986 censuses, gives a pessimistic picture of the immigrant-experience in the Canadian labour-market. While initial earnings were falling across successive immigrant cohorts, their rates of integration were also small. On the other hand, Grant (1999), using 1981, 1986, and 1991 censuses, was more optimistic. Initial earnings of new immigrants had stopped falling, and their subsequent integration rates was far better than that experienced by their predecessors.

But the quasi-panel approach does not capture individual effects, and it neglects the sample-selection problem. This study was able to control for unobserved individual heterogeneity by using panel-data estimation-techniques while also correcting for potential sample-selection bias.

2.2 In Europe including Sweden

Most European research about the economic integration of immigrants has been done in the Netherlands, Germany, Denmark, Norway, and Sweden, and the results have generally been pessimistic. Kee (1995) addressed wage-discrimination in Holland, using cross-sectional survey data. He corrected for sample-selection bias, estimated separate wage-equations for four different ethnic groups plus the native born, and then decomposed the wage-differences into those resulting from observed and unobserved variables. He found discrimination against Antilleans and Turkish immigrants but not against the two other non-native groups, Surinamese and Moroccans.

Husted *et al.* (2000), focused on employment and wage assimilation of male immigrants, using twelve waves of Danish panel-data. They found that immigrants assimilate partially to Danes but the assimilation process differed between refugees and non-refugees.

Longva and Raaum (2003), using "quasi panel techniques" based on 1980-1990 Norwegian population-censuses, found that non-OECD immigrants were at a disadvantage in the labour-market but there seemed to have been gradual improvement.

The first reported analysis of immigrant-earnings in Sweden is from 1967 (see Ekberg and Hammarstedt, 2002). More recent studies showed that there were differences among arrival-cohorts in terms of earnings-integration (Aguilar, Gustafsson, 1991); that earnings were lower for later-arriving cohorts (Ekberg, 1994); and that immigrants did not seem to be catching up with the native-born (Edin *et al.* 2000). Using 1970 and 1990 Swedish censuses, and focusing on Yugoslavian and Nordic non-refugee immigrants Bevelander and Nielsen (2001) first analyzed the determinants of employment-probability and then decomposed the part into explained by differences in observed characteristics from the unexplained part. Further, they made a detailed decomposition of the unexplained part showing which explanatory variables explained it. Unobserved characteristics seemed to be responsible for the decline of immigrants' employment-probabilities. Edin, Fredriksson, and Åslund (2003) analyzed the economic consequences of living in ethnic enclaves. Taking the recently changed Swedish policy concerning the initial location of refugees as exogenous, and focusing on two outcomes, earnings and employment-probabilities, they found evidence that living in ethnic enclaves caused significant improvements. Åslund and Rooth (2003) examined the effects of initial market conditions on the later performance of immigrants. Cohorts where most individuals had arrived before the recession of the early 1990s were 7-9 percentage points more likely to be employed ten years later, and had about 12-18 percent higher earnings than did individuals who arrived during the recession.

The empirical studies just discussed have given somewhat mixed results about the economic integration of immigrants. Researchers using similar data-sets have sometimes come to different conclusions, for two possible reasons: First is the type of data used, either cross-sectional or "quasi-panel". Such data are not really appropriate for analyzing either cohort-effects or unobserved individual-effects. Besides, age differences between comparison-groups of immigrants and native-born have been neglected, though certainly important for earnings-differences. Second, a problem in most of the earnings-integration studies, both in Sweden and elsewhere, is sample-selectivity due to unemployment, self-employment, and other reasons. Most earnings-comparisons of immigrants and the native-born have ignored this potential bias, something which this study hopes to avoid, as mentioned earlier. The data available for this study was also much more suitable to the analytic goals.

Based on what we have seen here and past studies, we can expect as a working hypothesis that some economic integration of immigrants may take place, but perhaps slowly, and arrival-cohort effects may be less important for immigrants from Nordic and Western countries than those from other regions. Other things equal, effects of other socioeconomic variables may differ substantially among different immigrant groups.

3 The data

The data used in this study consist of eleven annual waves (1990-2000) of the register-based nationally representative Longitudinal Individual Data-set (LINDA), which is a large panel of individuals and their household-members, updated each year. The principal data-sources are income registers and population censuses; Family members are included in the sample only as long as they stay in the household. LINDA includes a sub-panel of about 20 percent of the foreign-born population and the data are rich with individuals' socioeconomic characteristics. For more details see Edin and Frederiksson, (2001).

Analysis began with 33,568 male first-generation immigrants aged 18-55 in 1990, and followed them until 2000. The sample was restricted to male first-generation immigrants because the employment and earnings-conditions of immigrant women are considerably different and deserve a separate study. A Swedish control-group, of 33,568 similarly-aged men which was matched by county of residence was also selected, thus avoiding any bias in the analysis due to age-differences. Each additional year about 3000 more foreign-born male first-generation immigrant aged 18-55 were included in the panel with data from LINDA, and followed until 2000; and additional same-aged control group was also selected each year. By the year 2000, the unbalanced panel consisted of 525,689 observations of 69,041 first-generation male immigrants. Such data-work was cumbersome but worth the price since the large number of observations allowed analysis of geographical origin effects and the like which would not otherwise have been possible. All individuals were included except those who were self-employed. Immigrants' birth-places were classified as: Nordic not including Sweden; Western countries including the EU, the USA, Canada, Australia, and New Zealand; Eastern Europe; the Middle-East; Asia; Africa and Latin America.

Based on working-indicators in the data, an employment dummy was defined as 1 if the individual was employed, 0 if not.

The earnings-variable used in the study was calculated from the Tax Registers. The earnings were measured in thousands of SEK per year, adjusted using the consumer price-index to 2000 prices.

The key explanatory variables used were age; civil status; number of children at home; education level; the unemployment rate during the arrival year; birth place and arrival-cohort. No data on work-experience was available. In most U.S. based studies this is handled by calculating potential work experience as age minus years of schooling minus six. But in Sweden the education-data is given in terms of level, not years, so such a calculation would introduce severe measurement-error.

Table 1 in the Appendix shows the mean values for these variables, for both immigrants and native Swedes, for their first year in LINDA. Both the employment rate (79 percent vs. 46 percent) and earnings were considerably higher for native Swedes. On the other hand, immigrants' marriage or cohabitation rate was higher (47 percent vs. 37 percent). Slightly more immigrants lived in big cities (39 percent vs. 36 percent) and they had more children (0.56 vs. 0.46). The average native Swede was better educated than the average immigrant: About 72 percent of native Swedes had at least highschool education, compared to 60 percent for immigrants. The immigrant arrival-cohorts "before-1970", 1970-74, 1975-79, and 1980-84 all had 9-12 percent of the total, whereas 1985-89 had 18 percent and 1990-94 had almost 24 percent. Among others the Iran-Iraq war and various wars in former Yugoslavia occurred during the later periods. The most represented area of origin was the Nordic (24 percent) followed by Eastern Europe (21 percent), the Middle East (19 percent) and Western countries (15 percent). Asia, Africa, Latin America each had 6-8 percent.

The immigrant population was not homogenous, as can be seen in Table 2 (by geographical origin) and Table 3 (by arrival-cohort) in the Appendix. The employment rate and earnings were much higher for those coming from Nordic or Western countries (Table 2). Middle-Eastern and African immigrants were far less likely to be non-employed and had lower earnings if they were. The average Middle-Eastern or African immigrant was younger (as were the Asians and Latin Americans) and had more children (as did the Eastern Europeans followed by Asians and Latin Americans) than the Nordic and Western immigrants. Immigrants from Western countries had more education than all other groups (nearly 32 percent had a university degree), followed by Eastern Europeans. Despite the fact that Nordic immigrants, most of them from Finland, had low level of education, they had a higher employment-rate and earned more than all the others. Although not much older than the others (6-7 years), Nordic and Western immigrants had also predominantly arrived in earlier cohorts than had the others, and the Nordics had arrived when unemployment was much lower. In that respect, the Eastern Europeans arrived at the worst time. These descriptive statistics are generally in accord with the previous studies on immigrants to Sweden.

Looking at the statistics by arrival-cohort (Table 3 in the Appendix), earlier immigrants are much better established in the labour- market than are later ones, some of whom arrived when unemployment was very much worse (1990-2000). Again these results are in accord with earlier studies.

Finally, Table 4 (in the Appendix) shows the mean employment percentages and log of earnings for both immigrants and native Swedes by years from 1990 to 2000. While native Swedes' employment rates fell from almost 89 percent in 1990 to just below 80 percent in 1994, before recovering modestly to almost 83 percent by 2000, immigrants' rates fell drastically, from 67 percent in 1990 to 44 percent in 1994, before also recovering partially to almost 56 percent by 2000. Earnings for both groups generally increased throughout the period, however, with native Swedes consistently higher than immigrants.

4 Econometric specifications

Econometric model was chosen which both exploit the panel-aspect of data and correct for potential sample-selection bias. Sample-selection bias[1] can arise as a result of either self-selection by the individuals under investigation or sample-selection decisions made by data-analysts. Such sample-selectivity can be a major problem in cross-sectional as well as panel-data sets. It has been common way in many applied economic analyses of panel-data to study only the balanced sub-panel or the unbalanced panel without correcting for selectivity-bias.

One kind of selection-problem occurs when individuals do not disappear from the panel but certain variables are not observed for at least some time-periods. A well-known case is estimating earnings-equations using a panel of individuals, such as was done here. Some variables were observable for everyone in each time-period, but because some individuals did not work in some years, we cannot observe their earnings. Such selection may distort the representation of the true population and consequently distort inferences based on the observed data using standard methods.

[1] A simple sample-selection test, suggested by Nijman and Verbeek (1992) was also performed by adding the lagged selection indicator $r_{i,t-1}$, to the equation, estimating the model by fixed effects on the unbalanced panel and doing a t test for the significance of $r_{i,t-1}$. For all the groups, $r_{i,t-1}$ was significant.

Another big concern in empirical work is the presence of unobserved heterogeneity, otherwise known as "individual-effects". Heterogeneity across individuals may arise as a result of differences in individual preferences, characteristics, or endowments. Failure to account for such individual effects may result in biased and inconsistent estimates of the structural parameters (for more discussion of such topics, see for instance Matyas and Sevestre, 1995, ch.18; Kyriazidou, 1997 and Vella and Verbeek, 1999).

Due to the possibility of both sample-selectivity and unobserved heterogeneity, it was desirable to consider them simultaneously. This can be done in various ways. In our case, the random effects model suggested by Jensen *et al.* (2002) is specified and the model can be formulated as follows:

$$y_{it}^* = \beta' x_{it} + u_i + \epsilon_{it} \quad (1)$$

$$r_{it}^* = \gamma' z_{it} + v_i + \omega_{it} \quad (2)$$

$$r_{it} = \begin{matrix} 1 \ \text{ if } r_{it}^* > 0, \\ 0 \text{ otherwise} \end{matrix}$$

$$y_{it} = y_{it}^* * r_{it}$$

where i denotes the individual; t denotes the time period; y^* denotes earnings; x_{it} and z_{it} are row vector of exogenous variables; β and γ are column vectors of unknown parameters of interest; u_i and v_i are unobserved individual-specific effects; and ϵ_{it} and ω_{it} are idiosyncratic error terms. The observations for y_{it}^* may only be available if an unobserved latent variable r_{it}^*, measuring the extra benefits of being employed over not being employed, is non-negative. The following statistical assumptions are made for the idiosyncratic error terms:

$$\epsilon_{it} \sim N(0, \sigma_\varepsilon^2) \quad i = 1, ...N \quad t = 1, ...T \quad (3)$$

$$\omega_{it} \sim N(0, 1) \quad i = 1, ...N \quad t = 1, ...T \quad (4)$$

$$Corr(\epsilon_{it}, \omega_{it}) = \rho \quad (5)$$

The error-terms ϵ_{it} and ω_{it} are assumed to be non-autocorrelated. The likelihood of a single observation, conditional on the random effects, is

$$\begin{aligned}
L_{it}(\gamma,\beta,\sigma_\varepsilon^2,\rho \quad | \quad & u_i,v_i) = f(\epsilon_{it},\omega_{it} \mid u_i,v_i) && (6)\\
= \quad & \left[\int_{-\infty}^{-v_i-\gamma' z_{it}} \int_{-\infty}^{\infty} \phi_{\epsilon\omega}(\varepsilon,\omega) d\epsilon d\omega\right]^{1-r_{it}} \\
& \times \left[\int_{-v_i-\gamma' z_{it}}^{\infty} \phi_{\epsilon\omega}(y_{it}-u_i-\beta' x_{it},\omega) d\omega\right]^{r_{it}} \\
= \quad & \left[\int_{-\infty}^{-v_i-\gamma' z_{it}} \phi_{\omega}(\omega) d\omega\right]^{1-r_{it}} \\
& \times \left[\int_{-v_i-\gamma' z_{it}}^{\infty} \phi_{\omega|\epsilon}(\omega \mid y_{it}-u_i-\beta' x_{it}) \cdot \phi_{\epsilon}(y_{it}-u_i-\beta' x_{it}) d\omega\right]^{r_{it}} \\
= \quad & \left[\Phi_\omega(-v_i-\gamma' z_{it})\right]^{1-r_{it}} && (7)\\
& \times \left[(1-\Phi_{\omega|\epsilon}(-v_i-\gamma' z_{it} \mid y_{it}-u_i-\beta' x_{it})) \cdot \phi_\epsilon(y_{it}-u_i-\beta' x_{it})\right]^{r_{it}}
\end{aligned}$$

where the conditional distribution of $\omega \mid \epsilon \sim N(\frac{\rho\epsilon}{\sigma_\epsilon},(1-\rho^2)$. In this study, for the random effects and for their interactions with the idiosyncratic error terms, following specifications are made:

$$\begin{aligned}
v_i \quad &\sim \quad N(0,\sigma_v^2) && (8)\\
u_i \quad &\sim \quad N(0,\sigma_u^2) && (9)\\
\epsilon_{it},\omega_{it} \quad &\perp \quad u_i,v_i \qquad u_i \perp v_i && (10)
\end{aligned}$$

Thus, the individual-specific components in the selection equation and the equation of interest are assumed to be uncorrelated so that selectivity is assumed to show-up through the correlation of error terms ϵ_{it} and ω_{it}.

The contribution of the ith individual to the log likelihood , conditional on random effects is:

$$\begin{aligned} LogL_{it} &= \sum_{r_{it}=0} \log \Phi(-v_i - \gamma' z_{it}) + \\ & \sum_{r_{it}=1} \frac{-\log 2\pi}{2} - \log \sigma_\varepsilon - \frac{(y_{it} - u_i - \beta' x_{it})^2}{2\sigma_\varepsilon^2} + \\ & \log \Phi \left[\frac{(v_i + \gamma' z_{it}) + (\rho/\sigma_\varepsilon)(y_{it} - u_i - \beta' x_{it})}{\sqrt{1-\rho^2}} \right] \end{aligned} \tag{11}$$

The parameters of interest are estimated by maximum likelihood method.

The conditional mean function for the sample selection model is the same as in the case of cross-sectional data and is not changed by the presence of random effects:

$$E[y_{it} \mid x_{it}, r_{it=1}] = \beta' x_{it} + \rho \sigma_\varepsilon \sigma_\omega \frac{\phi(\gamma' z_{it})}{\Phi(\gamma' z_{it})} \tag{12}$$

where $\sigma_\omega = 1$ due to the normalization restriction.

5 Results

Tables 5 and 6 (in the Appendix) present the results and Tables 7 and 8 present the marginal effects for the joint estimation of the employment and earnings equations. Tables 5 and 7 (in the Appendix) show the coefficients and slopes for the employment-equation where the dependent variable indicates whether the individual is employed or non-employed. Because of the size of the sample, all parameters were statistically significant at the 5 percent level or most at 1 percent, except for those marked n.s (not significant). There were considerable differences in the magnitudes of the slopes within immigrant groups, and between immigrants and native Swedes but most standard results were confirmed for all groups (such as for age, married/cohabiting, having children at home and educational level). For example, for both immigrants and native Swedes, employment probabilities increased with age at a decreasing rate; the effect of age was weakest among Middle Eastern immigrants. Being married and having children increase the employment-probability though their magnitudes vary considerably for the different groups. The effect of being married/cohabiting was much larger for Nordic immigrants, native Swedes and Western immigrants than for other immigrant groups. For Eastern Europeans, Middle Easterners, Asians, Africans, and Latin Americans this effect was very close to zero. About the same can be said when it comes to number of children at home, though this time especially the Eastern Europeans, the Westerners and the Nordics were much further towards the high end. For both immigrants and native Swedes, having a university degree improved the employment probability but much less for the Middle Eastern, African and Latin American immigrants than the others. The unemployment rate during the year of arrival had a negative effect for all immigrant groups, especially for the Eastern Europeans. There was an immigration boom of Eastern Europeans in the mid to late-90s, which were high unemployment years, so the result is perhaps not surprising. On the other hand, as noted earlier, Eastern Europeans are one of the best educated groups of immigrants (Table 2). This finding is similar to that of Åslund and Rooth (2003), but in this case the effect shows up on employment probabilities instead of on earnings.

Cohort effects were similar for all the immigrant groups, with employment probabilities generally lower for later cohorts, as expected from previous research. Before 1980, Asians and Middle Easterners had the least negative (or even positive) effects; after that, Asians, and Eastern Europeans had the least negative cohort effects.

Figures 3,4 and 5 (in the Appendix) show predicted employment probabilities with respect to age for arrival cohorts ten years apart, starting with the most recent. In each the native Swedes have the highest probabilities as expected, but if the "immigrants catch up" hypothesis is correct one would expect to see the curves moving up in Figures 4 and 5 as the cohorts have been in Sweden longer. This is true for Nordics and perhaps Westerners and possibly somewhat for some others but the curves representing Africans seem to move the least.

Tables 6 and 8 in (the Apendix) show the coefficients and marginal (total) effects for the control variables in the earnings equation. The marginal effects consist of two components. There is the direct effect (β_ks) and the indirect effect (which is based on the probability of selection into the sample). As with employment probabilities, earnings rose with age (though at declining rate). The effect was strongest for native Swedes, Westerners, and Nordics, weakest for Middle Easterners, Africans and Latin Americans. There were small positive effects associated with being married or cohabiting for Westerners, Nordics, native Swedes and negative effects for Middle Easterners, Asians and Africans. Living in a big city was also slightly positive for

Nordics and native Swedes and slightly negative for the rest (with Middle Easterners and Eastern Europeans the most negative). The number of children at home showed positive effect for all groups, especially for Westerners, Nordics and Latin Americans.

Having finished high school (as highest education level) was no advantage for Westerners) but of some advantage to the other groups. Having a university degree was best for Nordics, Westerners and native Swedes with a university degree had predicted annual earnings about 35-37 percent higher than those with only lower-secondary education, Africans only about 20 percent higher. A weakness of the register-data is that it does not tell us where the immigrants got their education, which might reveal whether adopted-country education was more highly valued in the labour-market than country-of-origin education.

The pattern of arrival-cohort effects is different from and more mixed than employment probabilities. Compared to the pre-1970 cohort, Nordics and Westerners who came in 1970-74 were doing slightly worse, but generally each later cohort was doing better until, all other things equal, the last cohort was earning 9-16 percent more than the first. Possibly the changing country-composition of the Nordic and Western groups influenced this result, as Danes and Norwegians have increasingly taken the earlier place of Finns in Nordic immigration, while British and Germans have increasingly taken the earlier place of Greeks, Portugese, and Spanish among the Westerners. The 1970-74 cohort of all the other groups (except the Asians) was also doing worse than the pre-1970 cohort; and each later cohort was doing still worse especially the Asians and the Latin Americans. Of course this may be a sign of catch-up that earlier cohorts were doing better.

Figures 6-15 (in the Appendix) convey these results graphically. Figure 6,7 and 8 show predicted earnings[2] with respect to age for arrival cohorts ten years apart, starting again with the most recent. As just discussed, most recent cohorts of Nordics (also see Figure 9) and Westerners (also see Figure 10) were earning more than native Swedes, possibly a result of demand for highly-specialized immigrant labour. But in all cases, Nordics, Westerners and native Swedes were earning far more than all the other groups, and it is not at all clear that the other groups were catching up. Figures 11-15 do seem to show some small catch-up each of the groups (those who arrived earlier were doing slightly better than those who arrived

[2]Those are plots of predicted earnings for those who work, based on the expected values of earnings formula at the end of the section 'econometric specifications'.

later), most strikingly for Asians (figure 13) and Latin Americans (figure 15). But all remain far below the level of native Swedes.

Returning to Table 6, the estimates of the selectivity-parameter of ρ are lowest (in absolute value) for native Swedes, Nordics and Westerners, and highest for Africans and Middle Easterners. Apparently the unobserved factors influencing the employment-equations and earnings equations are negatively correlated.

6 Summary and conclusions

The register-based Longitudinal Data-set (LINDA) covering the period 1990-2000 was used to jointly estimate employment-probabilities and earnings for seven geographically based groups of first-generation immigrant men in cohorts from pre-1970 through 1995-2000, and for native Swedes. The problems of potential sample-selectivity bias and unobserved individual heterogeneity were handled by estimating the employment and earnings-equations simultaneously and by allowing for random effects in both equations. The panel structure of the data made it possible to control for unobserved heterogeneity and for cohort-effects due to country-of-origin or other changes in the composition of immigrant-groups.

The results show that there has been little catch-up for most immigrant groups. Nordic and Western immigrants may have largely caught up with native Swedes, though the effects may be masked by changes in the country-of-origin composition of the groups and by changing labour-demand. More recent cohorts of Nordic and Western immigrants men were actually doing better in the 1990s than were earlier cohorts. But most immigrant groups were catching up only slowly if at all in terms of employment-probabilities and earnings. Previous European studies which found similar employment and earnings-gaps have explained them as being due to labour-market discrimination, and it may not be possible to eliminate such gaps without taking strong preventive measures against that discrimination. Special policy-measures may be required too bring some immigrant groups more solidly into the labour-market.

References

[1] Aguilar, R., and Gustafsson, B. (1994), "*Immigrants in Sweden's labour Market During the 1980's*," Scandinavian Journal of Social Welfare, No. 3, pp. 139-47.

[2] Antelius, J., and Björklund, A.(2000), "*How Reliable are Register Data for Studies of the Return on Schooling? An Examination of Swedish Data,*" Scandinavian Journal of Educational Research, Vol. 44, No. 4, pp. 341-55.

[3] Baker, M., and Benjamin, D.(1994), "*The Performance of Immigrants in the Canadian labor Market,*" Journal of labor Economics, Vol. 12, No. 3 1994, pp. 369-405.

[4] Bevelander, P., and Nielsen, H.S. (2001), "*Declining Employment Success of Immigrant Males in Sweden: Observed or Unobserved Characteristics?*" Journal of Population Economics, Vol. 14, pp. 455-471.

[5] Borjas, G.J. (1985), "*Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants,*" Journal of Labor Economics, Vol. 3, No. 4, pp. 463-89.

[6] Borjas, G.J. (1987), "*Self-Selection and the Earnings of Immigrants,*" American Economic Review, Vol. 77, No. 4, pp. 531-53.

[7] Borjas, G.J. (1995), "*Assimilation, Changes in Cohort Quality, : What happened to Immigrant Earnings in the 1980s.,*" Journal of labor Economics, Vol. 13, No. 2, pp. 201-45.

[8] Chiswick, B.R. (1978), "*The effect of Americanisation on the earnings of Foreign-born Men,*" The Journal of Political Economy, Vol. 86, No. 5, pp. 897-921.

[9] Duleep, H.O., and Regets, M.C. (1996), "*The Elusive concept of Immigrant Quality : Evidence from 1970-1990,*" PRIP-UI-41 Working paper, The Urban Institute , Washington DC.

[10] "*Emmigration of Immigrants and Measures of Immigrant Assimilation*: *Evidence from Sweden,*" Working Paper Series, Vol. 13, Uppsala University.

[11] Edin, P.A., and Frederiksson, P. (2001), "*LINDA-Longitudinal Individual Data for Sweden,*" Working Paper 2001:6, Department of Economics, Uppsala University.

[12] Edin, P.-A., Fredriksson, P., and Åslund, O. (2003), "*Ethnic Enclaves and the Economic Success of Immigrants : Evidence from a Natural Experiment,*" Quarterly Journal Of Economics, Vol. 118, No. 1, pp. 329-57.

[13] Ekberg, J. (1994), "*Economic Progress Among Immigrants in Sweden,*" Scandinavian Journal of Social Welfare, Vol. 3, pp. 148-57.

[14] Ekberg J. (1999), "*Immigration and the Public Sector. Income Effects for the Native Population in Sweden,*" Journal of Population Economics, Vol. 12, pp. 411-30.

[15] Ekberg, J., and Gustafsson, B. (1995), "*Invandrare på Arbetsmarknaden (Immigrants in the labour-Market,*" SNS Förlag.

[16] Ekberg, J., and Hammarstedt, M. (2002), "*20 år med allt sämre arbetsmarknadsintegrering för invandrare,*" Ekonomisk Debatt 2002, Vol. 30, No. 4, pp. 343-53.

[17] Ekberg, J., and Rooth, D.-O. (2002), "*Unemployment and Earnings for Second-Generation Immigrants in Sweden,*" forthcoming, Journal of Population Economics.

[18] Grant, M.L. (1999), "*Evidence of New Immigrant Assimilation in Canada,*" The Canadian Journal of Economics, Vol. 32, No. 4, pp. 930-55.

[19] Hammarstedt, M. (2001), "*Disposable Income Differences between Immigrants and Natives in Sweden,*" International Journal of Social Welfare, Vol.10, No. 2, pp.117-26.

[20] Hausman, J.A., and Taylor, W.E. (1981), "*Panel Data and Unobservable Individual Effects,*" Econometrica, Vol. 49, No. 6, pp. 1377-398.

[21] Husted, L., Nielsen, H.S., Rosholm, M., and Smith, N. (2000), " *Employment and Wage Assimilation of Male First Generation Immigrants in Denmark* ", Discussion Paper No 101 The Institute for the Study of labor IZA Bonn.

[22] Jensen, P., Rosholm, M., and Verner, M. (2002), " *A comparison of Different Estimators for Panel Data Sample Selection Models* ", Working Paper, 2002-1, Dept. of Economics, Aarhus University

[23] Kee, P. (1995), "*Native-Immigrant Wage Differentials in the Netherlands : Discrimination?*" Oxford Economic Papers, Vol. 47, No. 2., pp. 302-17

[24] Kyriazidou, E. (1997), "*Estimation of a Panel Data Sample Selection Model,*" Econometrica, Vol. 65, No. 6, pp. 1335-364.

[25] Lalonde, R.J., and Topel, R.H. (1991), "*Immigrants in the American labor Market : Quality, Assimilation, and Distributional Effects,*" American Economic Review, Vol. 81, No. 2, pp. 297-302.

[26] Lalonde, R.J., and Topel., R.H. (1992), "*The Assimilation of Immigrants in the U.S. labor Market,*" NBER Working Paper Series, No. 3573, pp. 67-92.

[27] Longva, P., and Raaum, O. (2003), "*Earnings Assimilation of Immigrants in Norway: A Reappraisal,*" Journal of Population Economics, Vol. 16, pp. 177-93.

[28] Lundh, C., and Ohlsson, R. (1994), "*Från arbetskraftsimport till flyktning invandring,*" SNS Förlag.

[29] Matyas, L., and Sevestre, P. (1995), "*The Econometrics of Panel Data,*" A Handbook of Theory with applications, Kluwer Academic Publishers.

[30] Nielsen, H.S., Rosholm, M., Smith, N., and Husted, L. (2001), "*Qualifications, Discrimination, or Assimilation? An Extended Framework for Analysing Immigrant Wage Gaps,*" Discussion Paper No. 365, The Institute for the Study of labor IZA Bonn.

[31] Nijman, T., and Verbeek, M. (1992), "*Nonresponse in Panel Data: The Impact on Estimates of a Life Cycle Consumption Function,*" Journal of Applied Econometrics, Vol. 7, pp.243-257.

[32] Ohlsson, R. (1975), "*Immigrants in the labor Market,*" Dissertation, Ekonomisk-historia föreningen i Lund, Lund University, Sweden.

[33] Vella, F., and Verbeek, M. (1999), "*Two-Step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias,*" Journal of Econometrics, Vol. 90, pp. 239-63.

[34] Verbeek, M. (1990), "*On the Estimation of a Fixed Effects Model with Selectivity Bias,*" Economics Letters, Vol. 34, 267-70.

[35] Verbeek, M., and Nijman, T. (1992), "*Testing for Selectivity Bias in Panel Data Models,*" International Economic Review, Vol. 33, No. 3, pp. 681-703.

[36] Zabel, J.E. (1992), "*Estimating Fixed and Random Effects Models with Selectivity,*" Economics Letters, Vol. 40, 269-72.

[37] Åslund, O., and Rooth, D.-O. (2003), "*Do When and Where Matter? Initial labor Market Conditions and Immigrant Earnings,*" IFAU Institute For labor Market Policy Evaluation, Working Paper Series No. 7.

[38] Österberg, T. (2000), "*Economic Perspectives on Immigrants and Intergenerational Transmissions,*" Economic Studies 102. Thesis. Göteborg University, Sweden.

Appendix: Tables and Figures

Table 1: Mean characteristics of immigrants and native Swedes, first year in the data set LINDA (n=69,041) (Std. dev. in parantheses)

	Immigrants		Native Swedes	
	Mean	Std. dev.	Mean	Std. dev
Employed	0.461	(0.497)	0.792	(0.425)
Log of earnings	11.897	(0.619)	12.124	(0.538)
Age (first year in LINDA)	34.204	(11.142)	35.409	(11.210)
Age-squared	1294.6	(831.3)	1412.3	(952.2)
Married/cohabiting	0.467	(0.498)	0.370	(0.484)
Big city (>250,000)	0.388	(0.487)	0.364	(0.439)
Number of children at home	0.563	(0.917)	0.467	(1.044)
Education (highest level):				
Lower-secondary	0.399	(0.489)	0.278	(0.404)
Upper-secondary	0.384	(0.486)	0.502	(0.499)
University degree	0.217	(0.412)	0.220	(0.451)
Arrival cohort:				
before1970	0.124	(0.332)		
1970-74 (5 years)	0.094	(0.296)		
1975-79 (5 years)	0.115	(0.319)		
1980-84 (5 years)	0.094	(0.292)		
1985-89 (5 years)	0.182	(0.385)		
1990-94 (5 years)	0.239	(0.425)		
1995-2000 (6 years)	0.149	(0.355)		
Unemployment rate during arrival-year	3.668	(2.517)		
Birthplace:				
Nordic countries	0.245	(0.430)		
Western countries	0.148	(0.355)		
Eastern Europe	0.207	(0.405)		
Middle East	0.187	(0.391)		
Asia	0.083	(0.277)		
Africa	0.067	(0.249)		
Latin America	0.061	(0.240)		

Table 2: Mean characteristics of immigrants, first year in the data-set LINDA, by geographical origin (Std. dev. in parantheses)

	Nordic n=16,919	Western n=10,157	Eastern Europe n=14,316	Middle East n=12,912	Asia n=5,810	Africa n=4,658	Latin America n=4,269
Employed	0.685	0.554	0.347	0.303	0.350	0.304	0.487
	(0.465)	(0.498)	(0.473)	(0.454)	(0.474)	(0.473)	(0.499)
Log of earnings	12.050	12.018	11.897	11.577	11.648	11.609	(11.760)
	(0.562)	(0.682)	(0.590)	(0.577)	(0.583)	(0.589)	(0.554)
Age	37.539	36.080	35.387	31.979	29.135	30.298	30.440
	(11.447)	(11.600)	(11.558)	(9.494)	(9.479)	(9.198)	(10.185)
Age-squared	1540.2	1436.3	1385.8	1112.8	938.7	1002.6	1030.3
	(888.3)	(1080.4)	(869.7)	(674.4)	(625.5)	(630.1)	(675.3)
Married/cohabiting	0.369	0.443	0.520	0.580	0.443	0.483	0.393
	(0.483)	(0.497)	(0.493)	(0.499)	(0.496)	(0.499)	(0.488)
Big city (>250,000)	0.291	0.482	0.362	0.440	0.393	0.443	0.416
	(0.453)	(0.499)	(0.481)	(0.469)	(0.488)	(0.496)	(0.492)
Number of children at home	0.433	0.429	0.661	0.700	0.586	0.653	0.531
	(0.918)	(0.869)	(1.058)	(1.240)	(1.175)	(1.034)	(0.995)
Education (highest level):							
Lower-secondary	0.474	0.348	0.297	0.390	0.532	0.411	0.416
	(0.499)	(0.476)	(0.457)	(0.487)	(0.498)	(0.488)	(0.493)
Upper-secondary	0.379	0.335	0.472	0.368	0.288	0.410	0.376
	(0.485)	(0.472)	(0.499)	(0.482)	(0.453)	(0.491)	(0.484)
University degree	0.147	0.317	0.231	0.242	0.180	0.179	0.206
	(0.353)	(0.464)	(0.421)	(0.428)	(0.383)	(0.399)	(0.405)
Arrival Cohort:							
before 1970	0.284	0.161	0.117	0.015	0.013	0.021	0.011
	(0.452)	(0.371)	(0.326)	(0.125)	(0.118)	(0.144)	(0.108)
1970-74 (5 years)	0.171	0.114	0.106	0.026	0.057	0.051	0.044
	(0.377)	(0.318)	(0.309)	(0.158)	(0.232)	(0.220)	(0.206)
1975-79 (5 years)	0.162	0.109	0.059	0.085	0.158	0.079	0.192
	(0.369)	(0.312)	(0.236)	(0.284)	(0.365)	(0.273)	(0.394)
1980-84 (5 years)	0.070	0.091	0.077	0.103	0.140	0.081	0.175
	(0.255)	(0.288)	(0.267)	(0.304)	(0.348)	(0.270)	(0.380)
1985-89 (5 years)	0.102	0.128	0.106	0.333	0.229	0.230	0.303
	(0.303)	(0.333)	(0.307)	(0.471)	(0.419)	(0.421)	(0.459)
1990-94 (5 years)	0.104	0.154	0.395	0.262	0.284	0.386	0.156
	(0.304)	(0.359)	(0.488)	(0.437)	(0.451)	(0.486)	(0.362)
1995-2000 (6 years)	0.103	0.240	0.139	0.176	0.118	0.147	0.116
	(0.302)	(0.426)	(0.344)	(0.379)	(0.322)	(0.354)	(0.321)
Unemployment rate during arrival-year	2.796	3.794	4.757	3.727	3.323	3.863	3.084
	(2.065)	(2.524)	(2.896)	(2.390)	(2.177)	(2.473)	(2.076)

Table 3: Mean characteristics of immigrants, first year in the data-set LINDA, by arrival-cohort (Std. dev. in parantheses)

	before 1970 n=8,550	between 1970-74 n=6,681	between 1975-79 n=7,931	between 1980-84 n=6,498	between 1985-89 n=12,585	between 1990-94 n=16,483	between 1995-2000 n=10,313
Employed	0.699	0.726	0.608	0.590	0.546	0.154	0.278
	(0.469)	(0.456)	(0.492)	(0.494)	(0.497)	(0.360)	(0.448)
Log of earnings	12.188	11.978	11.927	11.841	11.688	11.551	(11.781)
	(0.510)	(0.522)	(0.559)	(0.564)	(0.563)	(0.690)	(0.828)
Age	47.045	37.566	33.961	32.401	30.592	30.382	33.217
	(10.938)	(10.390)	(10.679)	(9.763)	(8.950)	(9.348)	(9.694)
Age-squared	2332.9	1519.2	1267.4	1145.1	1016.1	1010.4	1197.4
	(973.1)	(747.9)	(719.1)	(652.8)	(601.4)	(639.8)	(737.4)
Married/cohabiting	0.536	0.462	0.411	0.405	0.424	0.485	0.507
	(0.498)	(0.497)	(0.492)	(0.491)	(0.494)	(0.499)	(0.499)
Big city(>250,000)	0.320	0.379	0.403	0.493	0.419	0.302	0.474
	(0.466)	(0.484)	(0.490)	(0.499)	(0.493)	(0.459)	(0.499)
Number of children at home	0.479	0.674	0.727	0.624	0.600	0.516	0.326
	(0.914)	(1.066)	(1.191)	(1.100)	(1.088)	(1.063)	(0.854)
Education (highest level):							
Lower-secondary	0.366	0.384	0.430	0.416	0.471	0.387	0.327
	(0.482)	(0.486)	(0.495)	(0.493)	(0.499)	(0.487)	(0.493)
Upper-secondary	0.478	0.464	0.424	0.429	0.363	0.354	0.265
	(0.499)	(0.498)	(0.494)	(0.494)	(0.481)	(0.478)	(0.441)
University degree	0.154	0.151	0.146	0.154	0.165	0.258	0.408
	(0.361)	(0.357)	(0.351)	(0.360)	(0.372)	(0.437)	(0.491)
Unemployment rate during arrival-year	1.743	2.074	1.861	2.806	2.012	5.571	7.203
	(0.345)	(0.481)	(0.271)	(0.549)	(0.491)	(2.473)	(0.956)

Table 4: Mean employment and log of earnings (in 1000 SEK), immigrants and native Swedes, 1990-2000 (Std. dev. in parantheses)

		Immigrants		Native Swedes	
		Employed	Log of earnings	Employed	Log of earnings
		Mean		Mean	
1990	n=33,568	0.674	11.861	0.887	12.138
		(0.478)	(0.561)	(0.364)	(0.485)
1991	n=36,509	0.618	11.894	0.876	12.150
		(0.491)	(0.563)	(0.369)	(0.509)
1992	n=39,262	0.546	11.930	0.844	12.155
		(0.499)	(0.581)	(0.394)	(0.522)
1993	n=42,930	0.473	11.915	0.800	12.131
		(0.498)	(0.601)	(0.423)	(0.534)
1994	n=47,852	0.442	11.916	0.796	12.141
		(0.494)	(0.627)	(0.425)	(0.546)
1995	n=50,405	0.455	11.935	0.812	12.163
		(0.496)	(0.626)	(0.415)	(0.531)
1996	n=51,315	0.456	12.005	0.811	12.226
		(0.497)	(0.624)	(0.417)	(0.531)
1997	n=51,591	0.464	12.028	0.806	12.244
		(0.497)	(0.632)	(0.421)	(0.541)
1998	n=56,808	0.519	12.080	0.819	12.303
		(0.499)	(0.625)	(0.415)	(0.546)
1999	n=57,038	0.553	12.097	0.822	12.332
		(0.499)	(0.621)	(0.413)	(0.544)
2000	n=58,411	0.558	12.134	0.827	12.370
		(0.496)	(0.619)	(0.411)	(0.552)

Standard deviations in parantheses

Table 5: Jointly estimated parameters for employment equation from the random-effects model of employment and earnings[1]

	Native Swedes	Nordics	Westerners	Eastern Europeans	Middle Easterners	Asians	Africans	Latin Americans
	n=523,873	n=132,422	n=69,382	n=106,914	n=95,743	n=46,455	n=38,119	n=36,654
constant	-5.410	-2.197	-3.347	-5.989	-3.435	- 4.181	-4.303	-4.717
	(.118e-01)	(.237e-01)	(.160e-01)	(.758e-02)	(.663e-02)	(.112e-01)	(.102e-01)	(.120e-01)
Age	0.360	0.224	0.245	0.335	0.176	0.247	0.273	0.277
	(.294e-01)	(.167e-02)	(.140e-2)	(.986e-03)	(.165e-02)	(.233e-2)	(.267e-02)	(.247e-02)
Age-squared	-0.0004	-0.0003	- 0.0003	-0.0004	-0.0002	-0.0003	-0.0004	-0.0003
	(.554e-05)	(.249e-04)	(.200e-04)	(.144e-04)	(.224e-04)	(.321e-04)	(.383e-04)	(.333e-04)
Married/cohabiting	0.544	0.542	0.308	0.052	0.092	0.125	0.065	0.148
	(.465e-02)	(.142e-01)	(.921e-02)	(.675e-02)	(.672e-02)	(.102e-01)	(.952e-02)	(.107e-01)
Big city (>250,000)	-0.036	-0.024[8]	-0.029	-0.341	-0.209	-0.146	-0.116	-0.136
	(.427e-02)	(.138e-01)	(.891e-02)	(.699e-02)	(.632e-02)	(.959e-02)	(.989e-02)	(.107e-01)
Number of children at home	0.084	0.111	0.146	0.134	0.069	0.061	0.052	0.095
	(.193e-02)	(.556e-02)	(.392e-02)	(.140e-01)	(.237e-02)	(.321e-02)	(.315e-01)	(.399e-02)
Education (highest level):								
Upper-secondary	0.479	0.421	0.582	0.568	0.251	0.516	0.371	0.419
	(.604e-02)	(.136e-01)	(.108e-01)	(.833e-02)	(.795e-02)	(.109e-01)	(.115-01)	(.123e-01)
University degree	0.888	0.887	0.904	0.754	0.554	0.744	0.456	0.523
	(.634e-02)	(.195e-01)	(.115e-01)	(.976e-02)	(.851e-02)	(.129e-01)	(.142e-01)	(.152e-01)
Arrival Cohort:								
1970-1974 (5 years)		-0.434	-0.676	-0.063	0.067[5]	0.116	-0.476	-0.087[6]
		(.206e-01)	(.197e-01)	(.157e-01)	(.331e-01)	(.403e-01)	(.467e-01)	(.471e-01)
1975-1979 (5 years)		-0.552	-0.893	-0.318	-0.009$^{n.s}$	-0.132	-0.641	-0.395
		(.205e-02)	(.204e-01)	(.178e-01)	(.287e-01)	(.379e-01)	(.459e-01)	(.421e-01)
1980-1984 (5 years)		-1.256	-0.967	-0.022$^{n.s.}$	-0.092	-0.264	-0.771	-0.462
		(.281e-02)	(.216e-01)	(.172e-01)	(.292e-01)	(.394e-01)	(.469e-01)	(.437e-01)
1985-1989 (5 years)		-0.920	-1.477	-0.238	-0.422	-0.722	-0.953	-0.557
		(.241e-02)	(.211e-01)	(.157e-01)	(.277e-01)	(.407e-01)	(.451e-01)	(.429e-01)
1990-1994 (5 years)		-1.213	-1.723	-0.668	-1.090	-1.353	-1.592	-1.100
		(.301e-01)	(.238e-01)	(.179e-01)	(.291e-01)	(.405e-01)	(.459e-01)	(.467e-01)
1995-2000 (6 years)		-1.413	-1.290	-0.619	-1.462	-1.341	-1.352	-0.855
		(.534e-01)	(.310e-01)	(.223e-01)	(.340e-01)	(.508e-01)	(.533e-01)	(.577e-01)
Unemployment rate during arrival-year		-0.033	-0.051	-0.144	-0.016	-0.032	-0.051	-0.044
		(.644e-02)	(.372e-02)	(.229e-02)	(.249e-02)	(.385e-02)	(.344e-02)	(.491e-02)
σ_v	1.794	1.975	1.915	1.495	1.156	1.318	1.218	1.248
	(.312e-02)	(.128e-01)	(.925e-02)	(.507e-02)	(.427e-02)	(.727e-02)	(.706e-02)	(.781e-02)

[1] Because of the size of the sample almost all parameters were statistically significant, at the 1 percent level except for those marked $^{n.s}$ or those marked on the power of coefficients which indicate p-values in percents, i.e power 6 means p-value between 0.05 and 0.06...so on.

Note: The reference variables are: single, small city, lower-secondary education, arrival cohort before 1970

Table 6: Jointly estimated parameters for earnings equation from the random-effects model of employment and earnings

	Native Swedes n=523,873	Nordics n=132,422	Westerners n=69,382	Eastern Europeans n=106,914	Middle Easterners n=95,743	Asians n=46,455	Africans n=38,119	Latin Americans n=36,654
constant	9.511	10.197	9.751	10.777	11.076	10.460	11.246	11.081
	(.301e-02)	(.149e-01)	(.143e-01)	(.129e-01)	(.173e-01)	(.179e-01)	(.223e-01)	(.204e-01)
Age	0.110	0.070	0.102	0.048	0.028	0.066	0.026	0.042
	(.164e-02)	(.764e-03)	(.760e-03)	(.416e-03)	(.865e-03)	(.109e-02)	(.117e-02)	(.106e-02)
Age-squared	-0.0009	-0.0006	-0.0009	-0.0004	-0.0002	-0.0006	-0.0001	-0.0003
	(.182e-05)	(.955e-05)	(.943e-05)	(.432e-05)	(.100e-05)	(.130e-05)	(.147e-05)	(.124e-04)
Married/cohabiting	0.036	0.042	0.049	0.009	-0.029	-0.054	$0.005^{n.s.}$	$-0.007^{n.s}$
	(.704e-03)	(.289e-02)	(.255e-02)	(.256e-02)	(.351e-02)	(.452e-02)	(.452e-02)	(.500e-02)
Big city (>250,000)	0.010	0.030	0.023	-0.007^{4}	-0.007^{3}	-0.008^{3}	-0.007^{10}	-0.016
	(.657e-03)	(.292e-02)	(.241e-02)	(.247e-02)	(.301e-02)	(.381e-02)	(.448e-02)	(.388e-02)
Number of children at home	0.008	0.022	0.024	0.010	$-0.001^{n.s}$	-0.008	$0.001^{n.s}$	0.018
	(.262e-03)	(.114e-02)	(.104e-02)	(.105e-02)	(.119e-02)	(.143e-02)	(.155e-02)	(.148e-02)
Education (highest level):								
Upper-secondary	0.130	0.074	$0.002^{n.s}$	0.055	0.043	0.077	0.036	0.053
	(.848e-03)	(.306e-02)	(.317e-02)	(.330e-02)	(.398e-02)	(.479e-02)	(.564e-02)	(.496e-02)
University degree	0.350	0.351	0.285	0.186	0.180	0.216	0.144	0.172
	(.896e-03)	(.386e-02)	(.312e-02)	(.347e-02)	(.391e-02)	(.489e-02)	(.634e-02)	(.148e-02)
Arrival Cohort:								
1970-1974 (5 years)		-0.066	-0.103	-0.073	-0.038	0.026^{7}	-0.063	-0.220
		(.415e-02)	(.486e-02)	(.455e-02)	(.143e-01)	(.146e-01)	(.181e-01)	(.147e-01)
1975-1979 (5 years)		-0.064	-0.076	-0.078	-0.028^{2}	$-0.009^{n.s}$	-0.080	-0.251
		(.427e-02)	(.496e-02)	(.529e-02)	(.127e-01)	(.137e-01)	(.178e-01)	(.137e-02)
1980-1984 (5 years)		-0.092	-0.039	-0.071	-0.089	-0.083	-0.111	-0.316
		(.571e-02)	(.542e-02)	(.519e-02)	(.130e-01)	(.144e-01)	(.184e-01)	(.142e-01)
1985-1989 (5 years)		-0.017	-0.021	-0.081	-0.073	-0.088	-0.069	-0.323
		(.495e-02)	(.513e-02)	(.484e-02)	(.123e-01)	(.143e-01)	(.177e-01)	(.139e-01)
1990-1994 (5 years)		0.107	0.032	-0.086	-0.095	-0.125	-0.114	-0.369
		(.632e-02)	(.629e-02)	(.573e-02)	(.131e-01)	(.196e-01)	(.180e-01)	(.154e-01)
1995-2000 (6 years)		0.233	0.202	-0.098	-0.084	-0.126	-0.110	-0.390
		(.120e-01)	(.858e-02)	(.800e-02)	(.163e-01)	(.197e-01)	(.220e-01)	(.194e-01)
σ_u	0.456	0.479	0.561	0.382	0.339	0.379	0.311	0.334
	(.315e-03)	(.149e-02)	(.132e-02)	(.113e-02)	(.144e-02)	(.181e-02)	(.207e-02)	(.179e-02)
σ_ε	0.314	0.345	0.390	0.432	0.513	0.445	0.492	0.441
	(.150e-03)	(.518e-03)	(.593e-03)	(.529e-03)	(.764e-03)	(.991e-03)	(.998e-03)	(.842e-03)
ρ	-0.205	-0.307	-0.357	-0.675	-0.792	-0.659	-0.783	-0.668
	(.252e-02)	(.831e-02)	(.586e-02)	(.393e-02)	(.321e-02)	(.402e-02)	(.307e-02)	(.453e-03)
Log-likelihood	-349008.7	-99987.5	-55620.8	-83484.3	-71862.9	-50443.6	-29823.3	-31324.9

Table 7: **Marginal effects in employment equation (reference Table 5)**[1]

		Native Swedes n=523,873	Nordics n=132,422	Westerners n=69,382	Eastern Europeans n=106,914	Middle Easterners n=95,743	Asians n=46,455	Africans n=38,119	Latin Americans n=36,654
Age	coefficient	0.360	0.224	0.245	0.335	0.176	0.247	0.273	0.277
	slope	0.120	0.082	0.075	0.127	0.063	0.099	0.108	0.095
Married/cohabiting	coefficient	0.544	0.542	0.308	0.052	0.092	0.125	0.065	0.148
	slope	0.181	0.182	0.094	0.019	0.033	0.049	0.027	0.050
Big city (>250,000)	coefficient	-0.036	-0.024[8]	-0.029	-0.341	-0.209	-0.146	-0.116	-0.136
	slope	-0.012	-0.008	-0.009	-0.131	-0.075	-0.059	-0.046	-0.047
Number of Children at home	coefficient	0.084	0.111	0.146	0.134	0.069	0.061	0.052	0.095
	slope	0.027	0.037	0.045	0.051	0.025	0.025	0.020	0.033
Education (highest level):									
Upper-secondary	coefficient	0.479	0.421	0.582	0.568	0.251	0.516	0.371	0.419
	slope	0.159	0.141	0.167	0.213	0.092	0.203	0.014	0.141
University degree	coefficient	0.888	0.887	0.904	0.754	0.554	0.744	0.456	0.523
	slope	0.293	0.299	0.244	0.262	0.208	0.284	0.180	0.164
Arrival Cohort:									
1970-74 (5 years)	coefficient		-0.434	-0.676	-0.063	0.067[5]	0.116	-0.476	-0.087[6]
	slope		-0.084	-0.235	-0.024	0.023	0.046	-0.178	-0.030
1975-79 (5 years)	coefficient		-0.552	-0.893	-0.318	-0.009$^{n.s}$	-0.132	-0.641	-0.395
	slope		-0.113	-0.319	-0.125	-0.003	-0.052	-0.234	-0.142
1980-84 (5 years)	coefficient		-1.256	-0.967	-0.022$^{n.s}$	-0.092	-0.264	-0.771	-0.462
	slope		-0.358	-0.349	-0.008	-0.032	-0.105	-0.274	-0.169
1985-89 (5 years)	coefficient		-0.920	-1.477	-0.238	-0.422	-0.722	-0.953	-0.557
	slope		-0.229	-0.526	-0.092	-0.149	-0.278	-0.348	-0.198
1990-94 (5 years)	coefficient		-1.213	-1.723	-0.668	-1.090	-1.353	-1.592	-1.100
	slope		-0.338	-0.606	-0.255	-0.336	-0.484	-0.554	-0.413
1995-2000 (6 years)	coefficient		-1.413	-1.290	-0.619	-1.462	-1.341	-1.352	-0.855
	slope		-0.434	-0.472	-0.242	-0.338	-0.424	-0.402	-0.328
Unemployment rate	coefficient		-0.033	-0.051	-0.144	-0.016	-0.033	-0.051	-0.044
in arrival year	slope		-0.006	-0.016	-0.055	-0.006	-0.014	-0.020	-0.015

[1]Because of the size of the sample almost all parameters were statistically significant, at the 1 percent level except for those marked $^{n.s}$ or those marked on the power of coefficients which indicate p-values in percents, i.e power 6 means p-value between 0.05 and 0.06...so on.

Table 8: **Marginal effects in earnings equation (reference Table 6)**

		Native Swedes	Nordics	Westerners	Eastern Europeans	Middle Easterners	Asians	Africans	Latin Americans
Age	direct effect	0.039*	0.028	0.030	0.016	0.016	0.023	0.016	0.016
	indirect effect	0.001	0.004	0.005	0.012	0.003	0.004	0.008	0.006
	total effect	0.040	0.032	0.035	0.028	0.019	0.027	0.024	0.022
Married/cohabiting	direct effect	0.036	0.042	0.049	0.009	-0.029	-0.054	$0.005^{n.s}$	$-0.007^{n.s}$
	indirect effect	0.001	0.014	0.019	0.008	0.027	-0.023	-0.016	0.021
	total effect	0.037	0.057	0.067	0.017	-0.002	-0.077	-0.011	0.014
Big city (>250,000)	direct effect	0.010	0.030	0.023	-0.007^{4}	-0.007^{3}	0.008^{3}	-0.009^{10}	-0.016
	indirect effect	-0.001	-0.001	-0.001	-0.057	-0.061	-0.027	-0.029	-0.020
	total effect	0.009	0.029	-0.022	-0.064	-0.067	-0.019	-0.038	-0.036
Number of Children at home	direct effect	0.008	0.022	0.024	0.010	-0.001	-0.008	0.001	0.018
	indirect effect	0.001	0.003	0.009	0.022	0.010	0.011	0.013	0.014
	total effect	0.009	0.025	0.033	0.032	0.009	0.003	0.014	0.032
Education (highest level):									
Upper-secondary	direct effect	0.130	0.074	$0.002^{n.s}$	0.055	0.043	0.077	0.036	0.053
	indirect effect	0.002	0.012	0.034	0.093	0.073	0.093	0.094	0.061
	total effect	0.132	0.086	0.036	0.148	0.115	0.170	0.130	0.114
University degree	direct effect	0.350	0.351	0.285	0.186	0.180	0.216	0.095	0.172
	indirect effect	0.003	0.019	0.080	0.114	0.101	0.128	0.111	0.071
	total effect	0.353	0.370	0.365	0.300	0.281	0.345	0.206	0.243
Arrival Cohort:									
1970-74 (5 years)	direct effect		-0.066	-0.103	-0.073	-0.038	0.026^{7}	-0.063	-0.220
	indirect effect		-0.014	-0.048	-0.010	0.019	0.021	-0.128	-0.013
	total effect		-0.080	-0.151	-0.083	-0.019	0.047	-0.191	-0.023
1975-79 (5 years)	direct effect		-0.064	-0.076	-0.078	-0.028^{2}	$-0.009^{n.s}$	-0.080	-0.251
	indirect effect		-0.019	-0.066	-0.055	-0.003	-0.025	-0.175	-0.062
	total effect		-0.083	-0.142	-0.133	-0.032	-0.034	-0.255	-0.313
1980-84 (5 years)	direct effect		-0.092	-0.039	-0.071	-0.089	-0.083	-0.111	-0.316
	indirect effect		-0.057	-0.073	-0.003	-0.027	-0.051	-0.213	-0.073
	total effect		-0.149	-0.112	-0.074	-0.116	-0.134	-0.324	-0.039

*Since the effect of age is not linear, the direct effect of age is different than its parameter value in table 6.

"continued"

Table 8: (cont.) Marginal effects in earnings equation (reference Table 6)

		Native Swedes	Nordics	Westerners	Eastern Europeans	Middle Easterners	Asians	Africans	Latin Americans
1985-89 (5 years)	direct effect		-0.017	-0.021	-0.081	-0.073	-0.088	-0.069	-0.323
	indirect effect		-0.036	-0.117	-0.040	-0.124	-0.142	-0.256	-0.086
	total effect		-0.053	-0.138	-0.121	-0.197	-0.230	-0.325	-0.409
1990-1994 (5 years)	direct effect		0.107	0.032	-0.086	-0.095	-0.125	-0.114	-0.369
	indirect effect		-0.053	-0.141	-0.113	-0.334	-0.272	-0.419	-0.192
	total effect		0.053	-0.109	-0.199	-0.429	-0.397	-0.533	-0.561
1995-2000 (6 years)	direct effect		0.233	0.202	-0.098	-0.084	-0.126	-0.110	-0.390
	indirect effect		-0.069	-0.105	-0.113	-0.475	-0.291	-0.394	-0.150
	total effect		0.164	0.097	-0.211	-0.131	-0.559	-0.504	-0.540

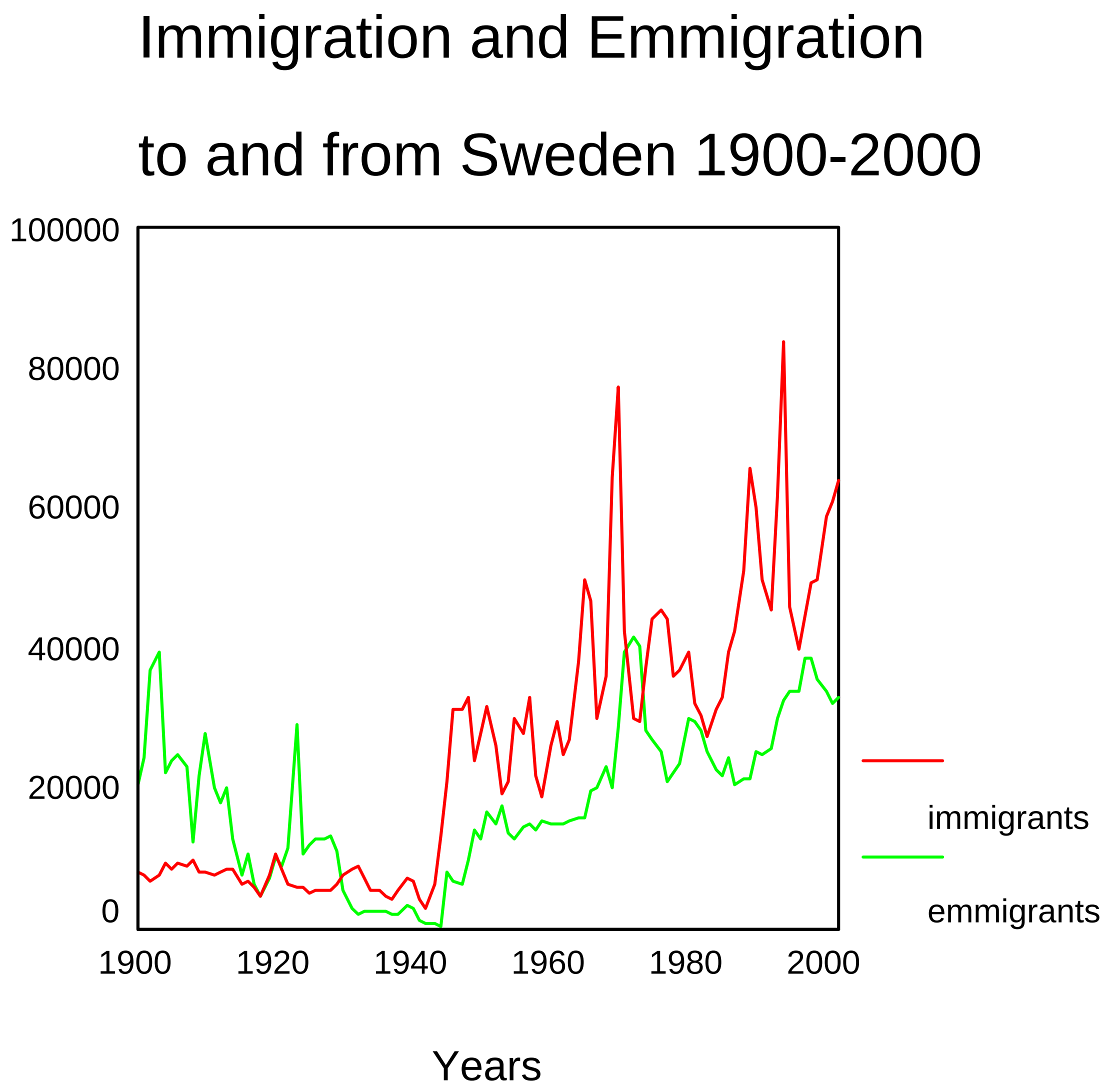
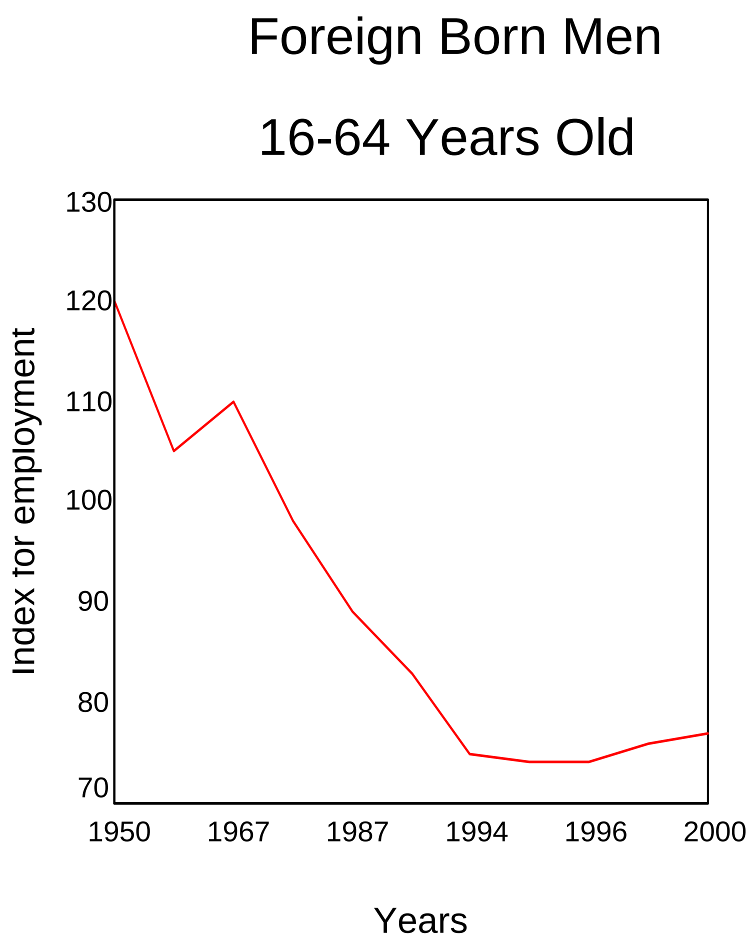


Figure 1

Source: Statistics Sweden, Statistical yearbook and Population statistics.



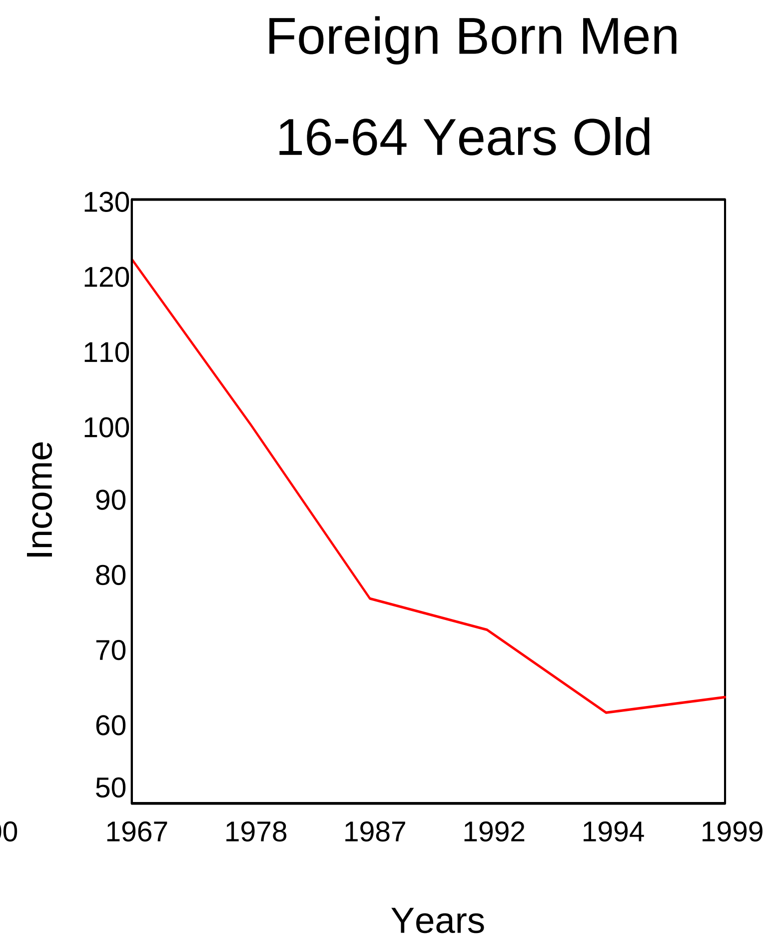
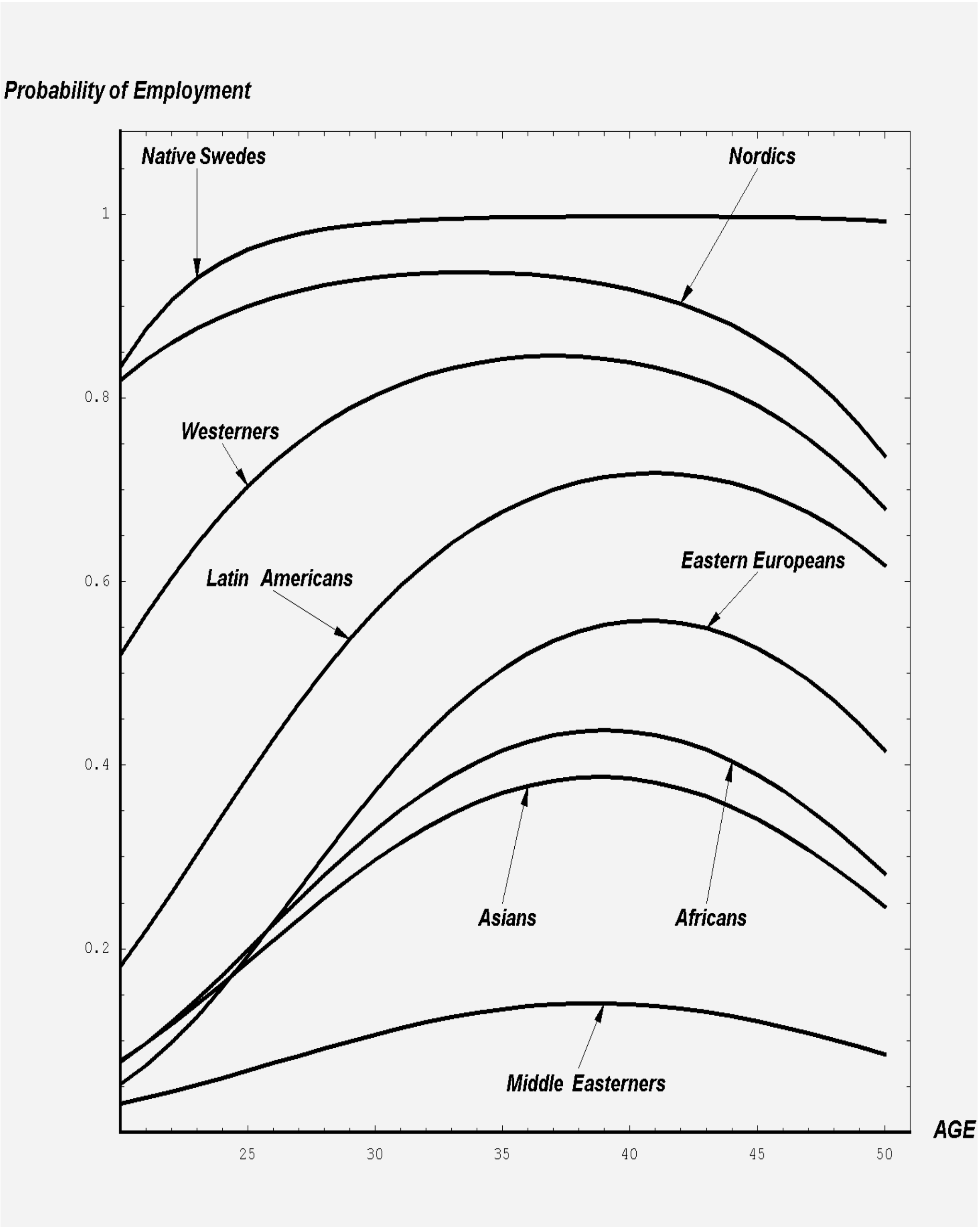


Figure 2a and Figure 2b

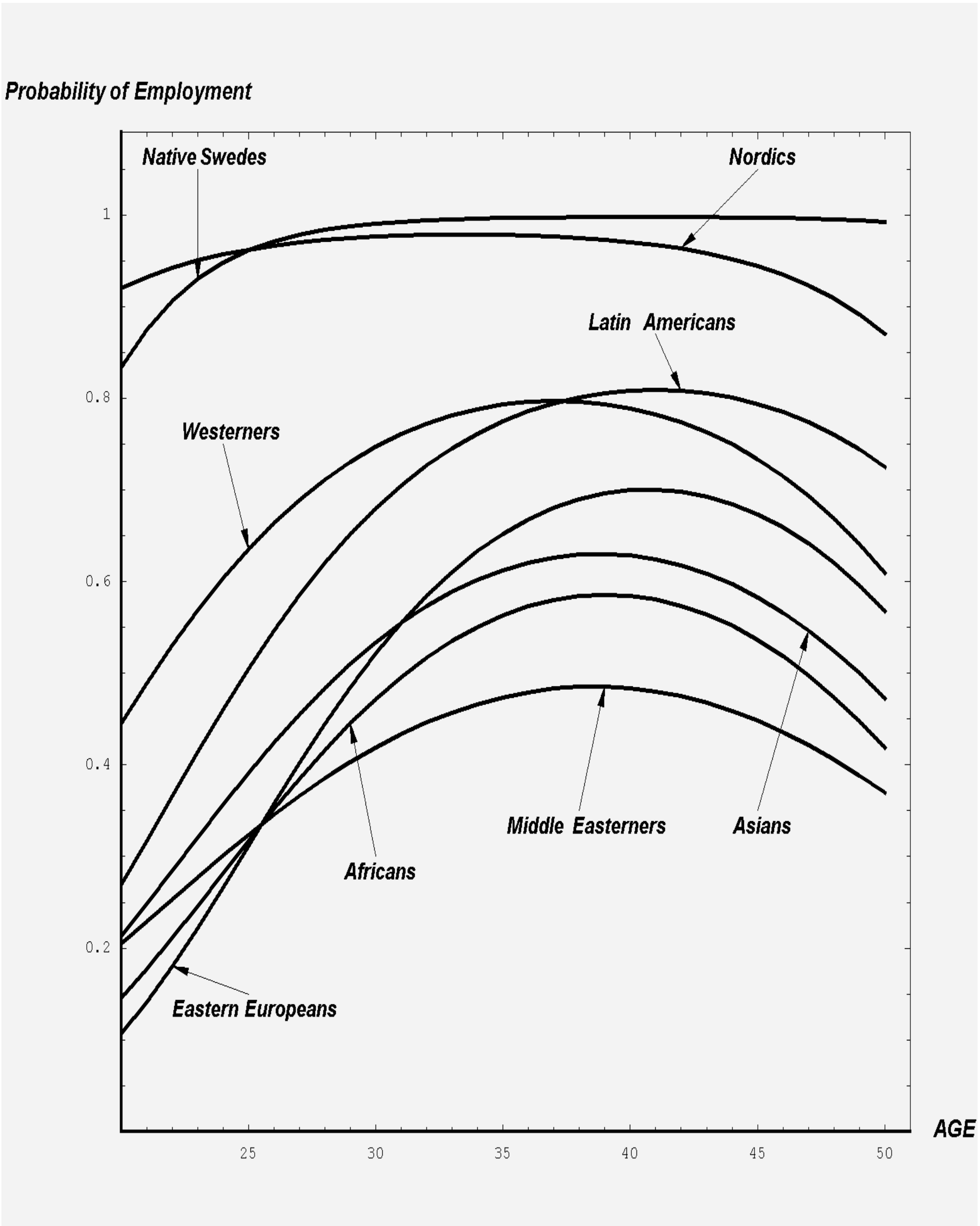
See Ekberg and Hammarstedt (2002, pp. 345) for the creation of the indexes in these figures.



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, arrival time unemployment rate=mean value

Figure 3

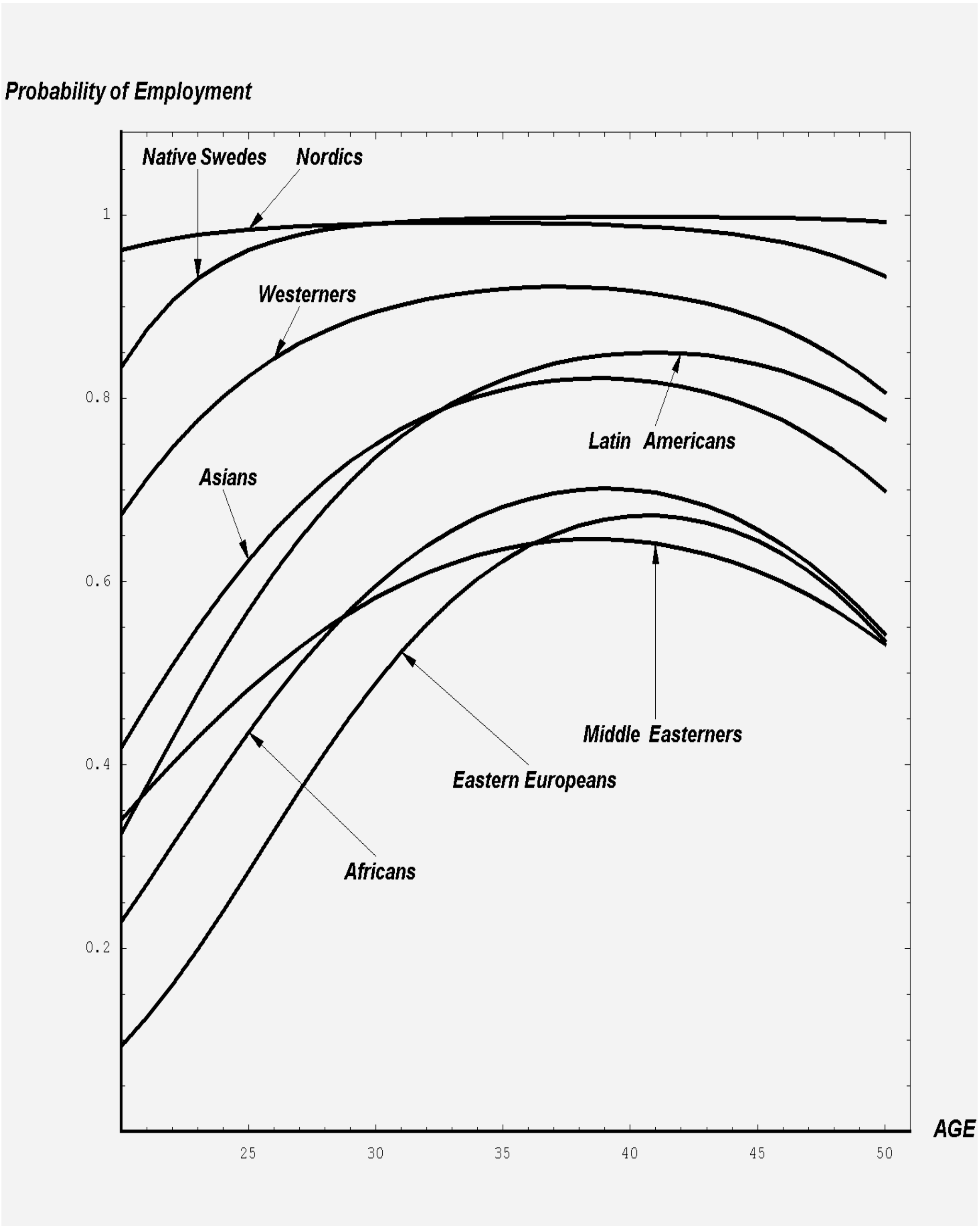
Probabilities of employment by age and geographic origin for arrival cohort 1995-2000



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, arrival time unemployment rate=mean value

Figure 4

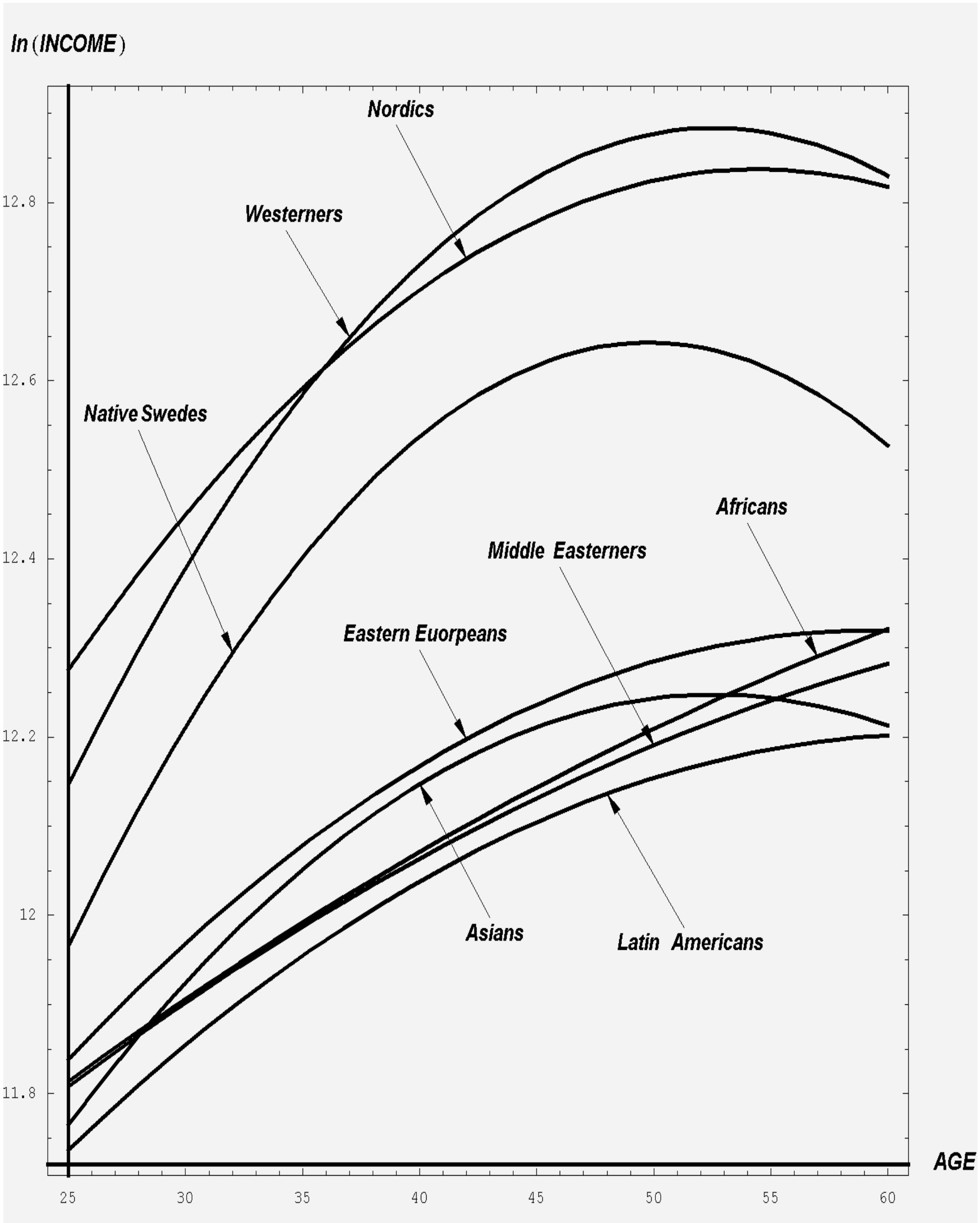
Probabilities of employment by age and geographic origin for arrival cohort 1985-1989



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, arrival time unemployment rate=mean value

Figure 5

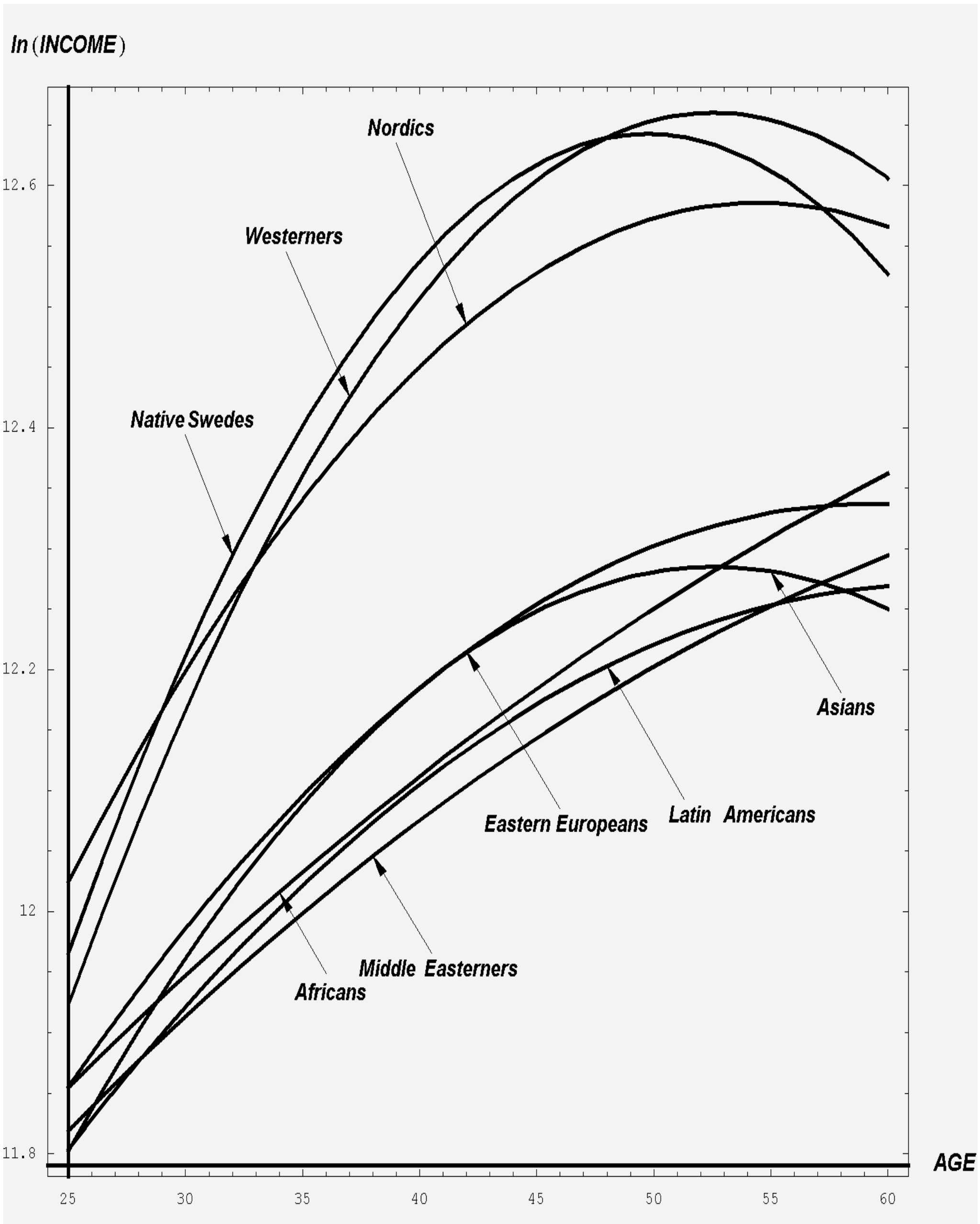
Probabilities of employment by age and geographic origin for arrival cohort 1975-1979



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value

Figure 6

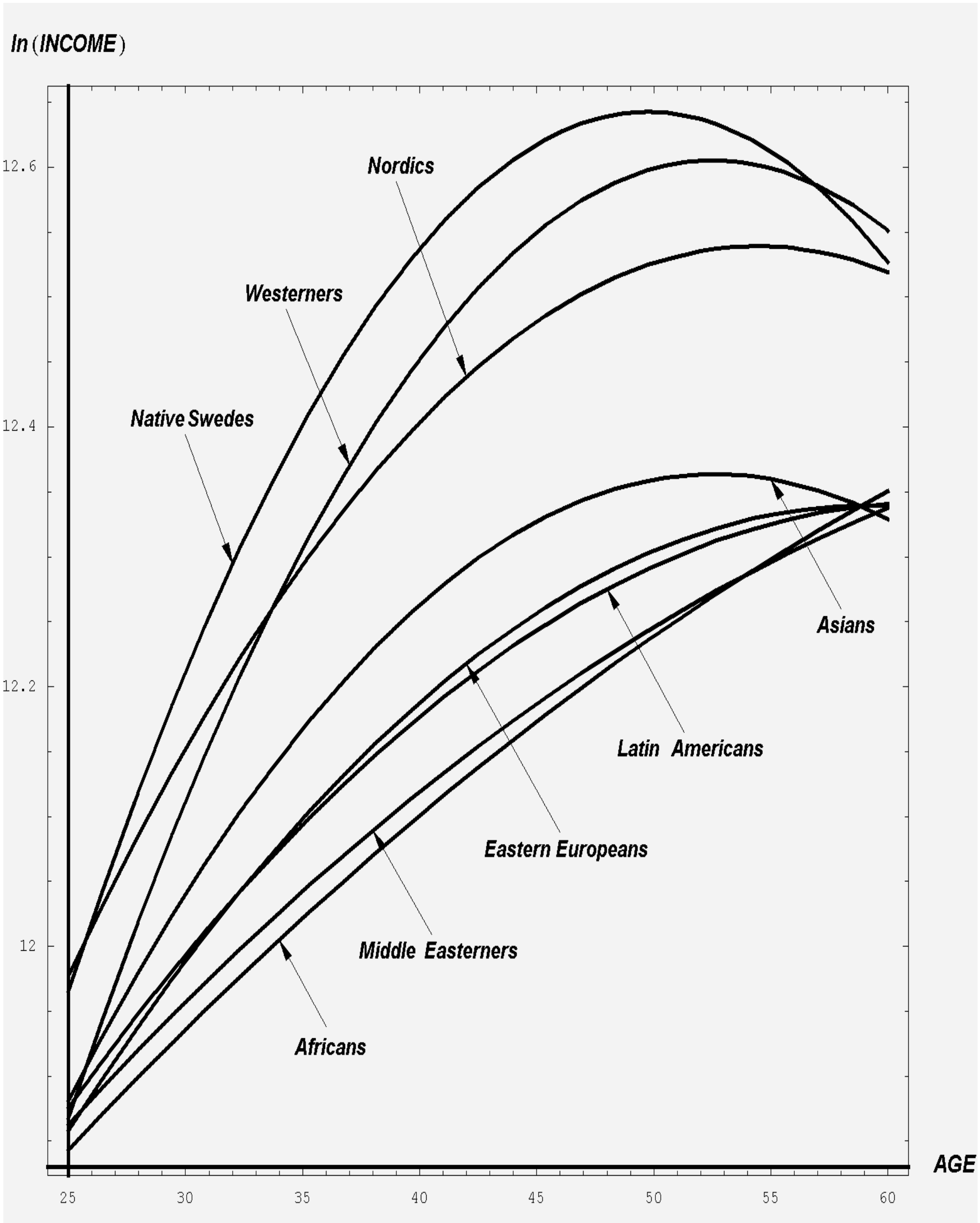
Log. of earnings by age and geographic origin for arrival cohort 1995-2000



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value

Figure 7

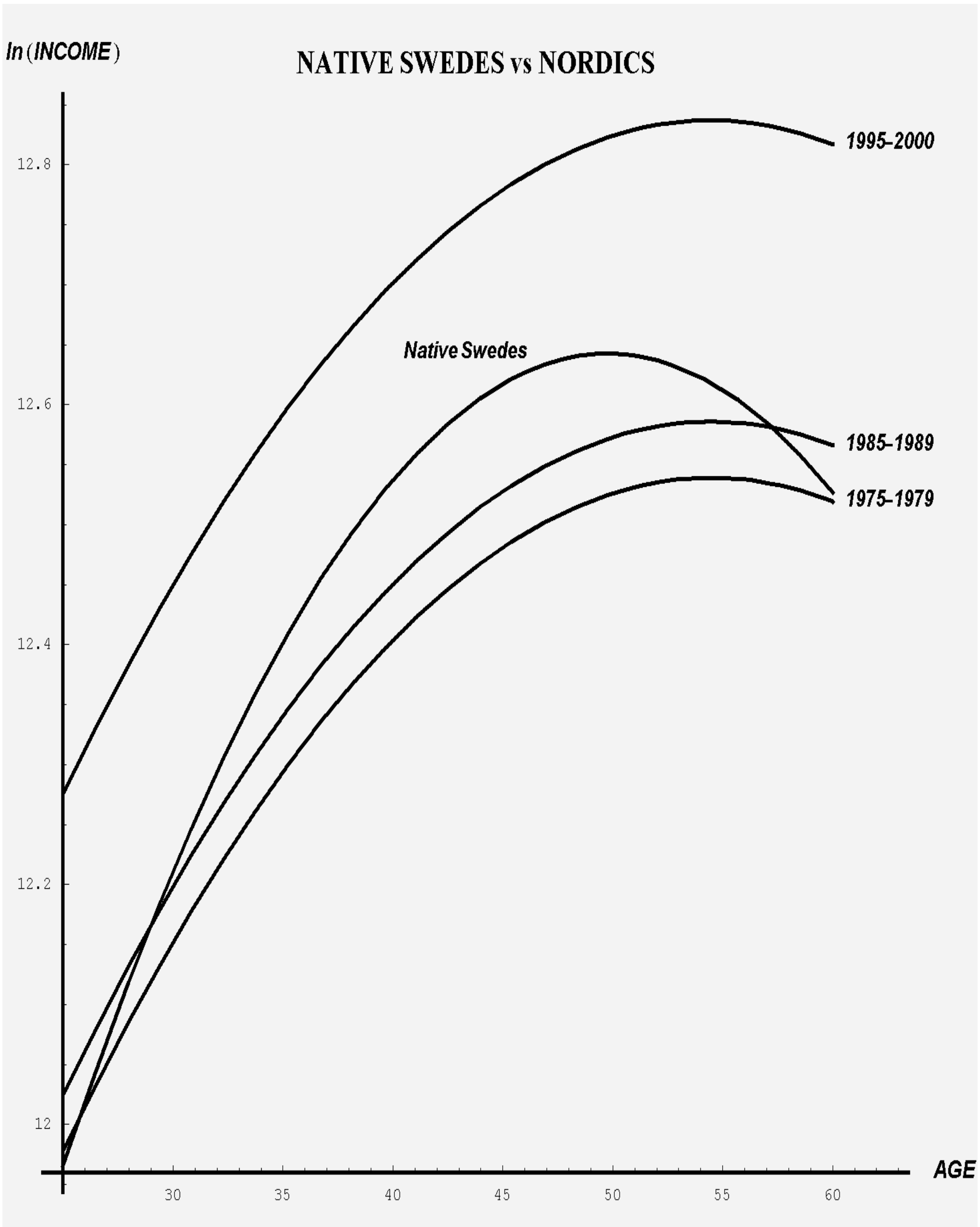
Log. of earnings by age and geographic origin for arrival cohort 1985-1989



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value

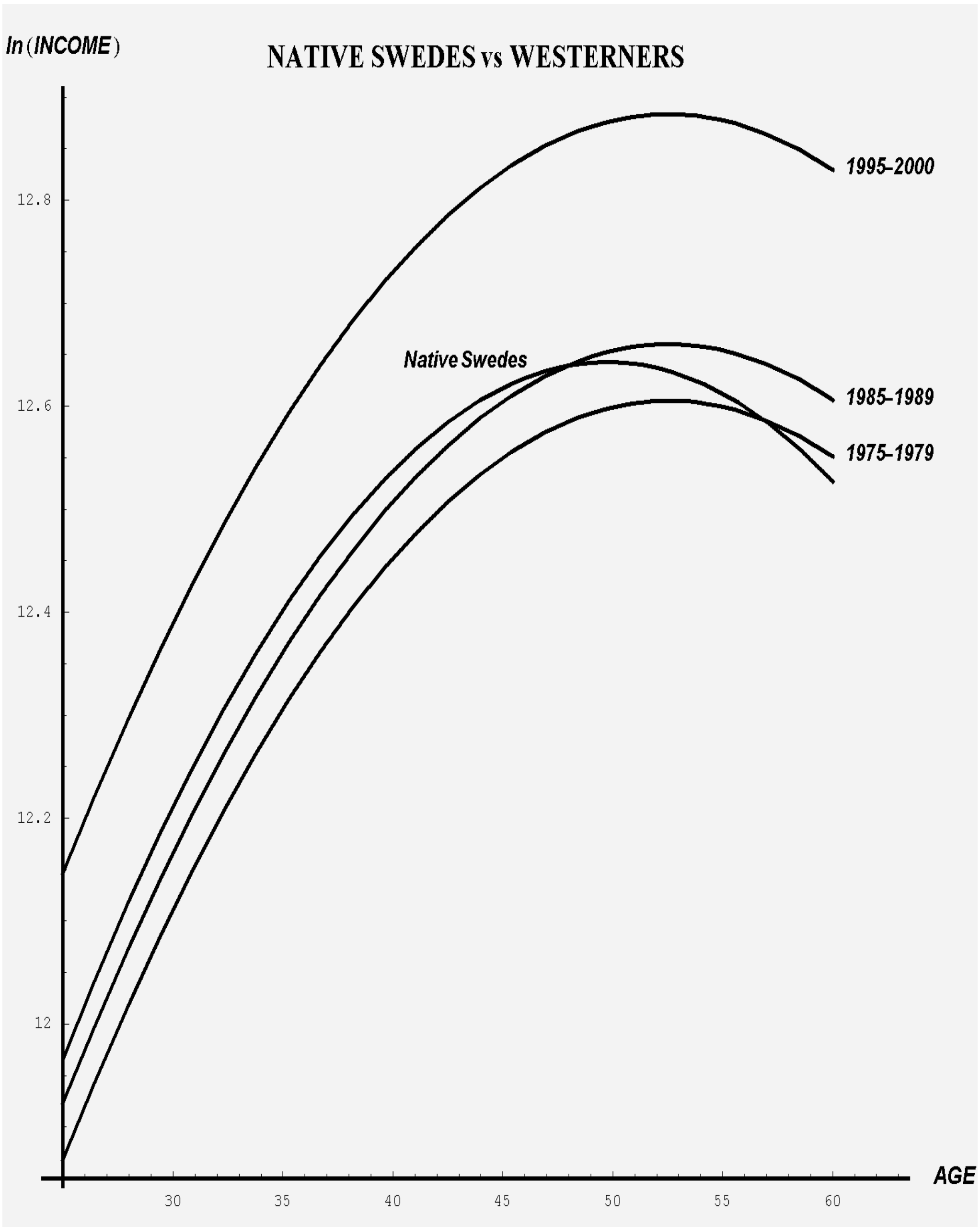
Figure 8

Log. of earnings by age and geographic origin for arrival cohort 1975-1979



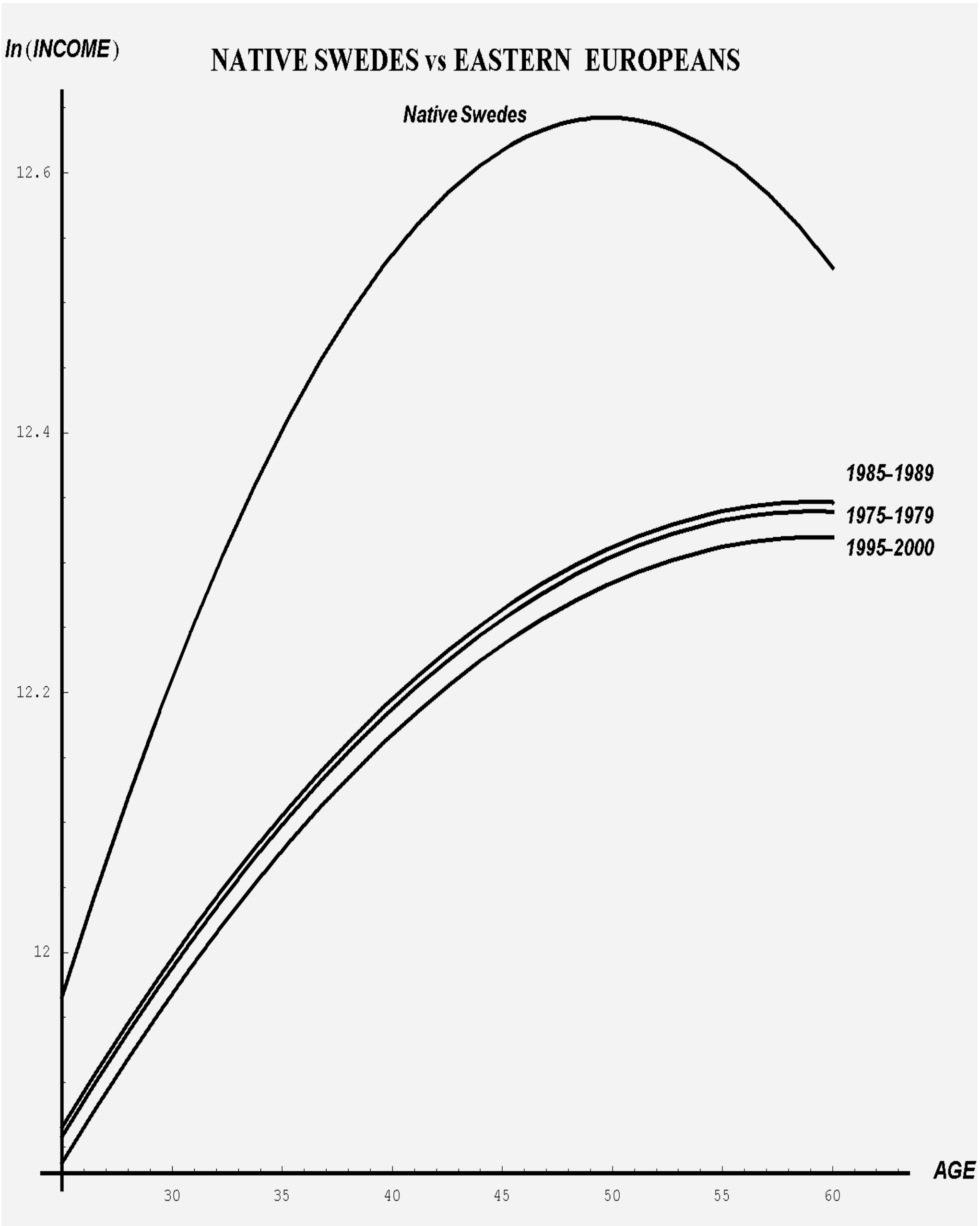
Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value

Figure 9



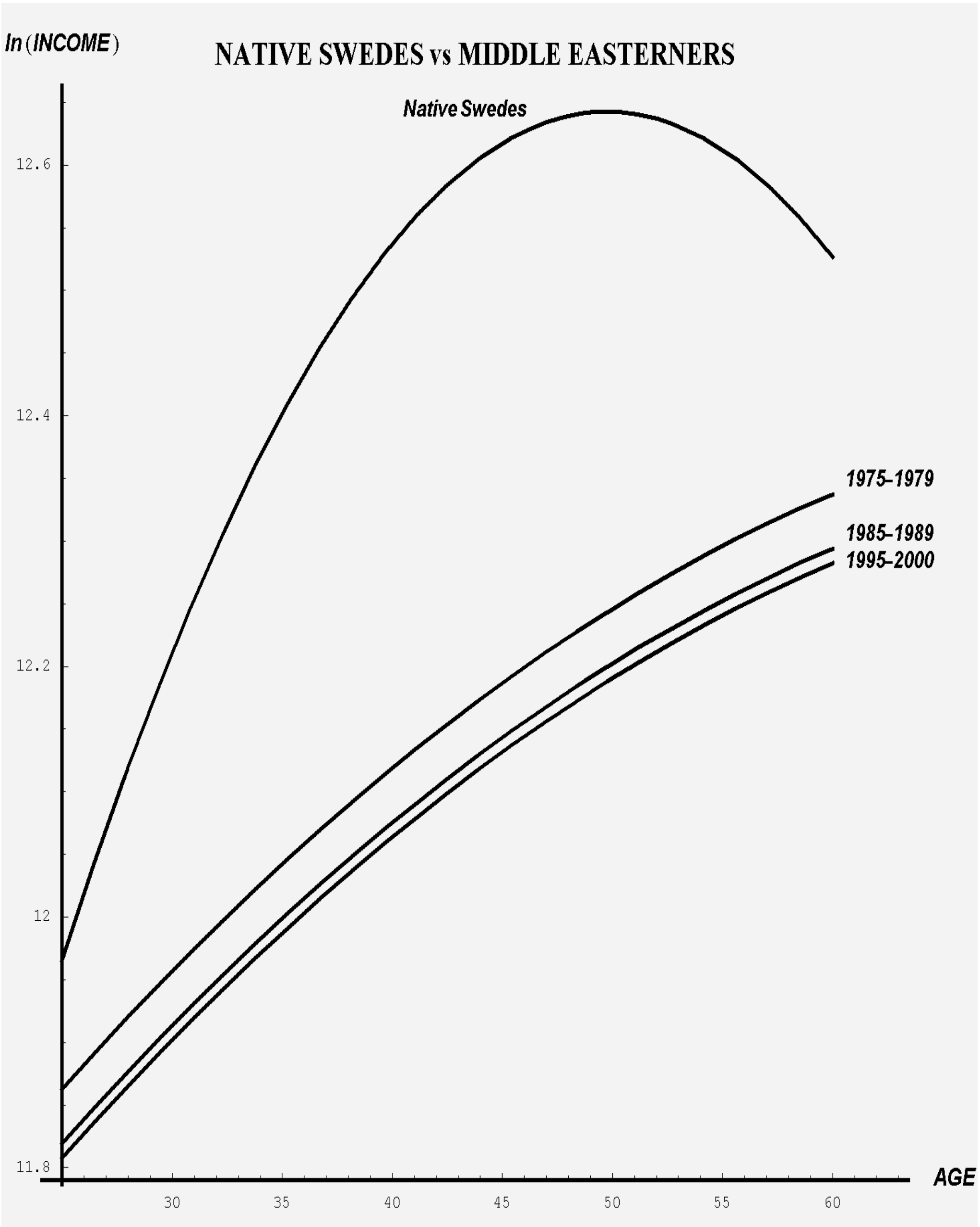
Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value

Figure 10



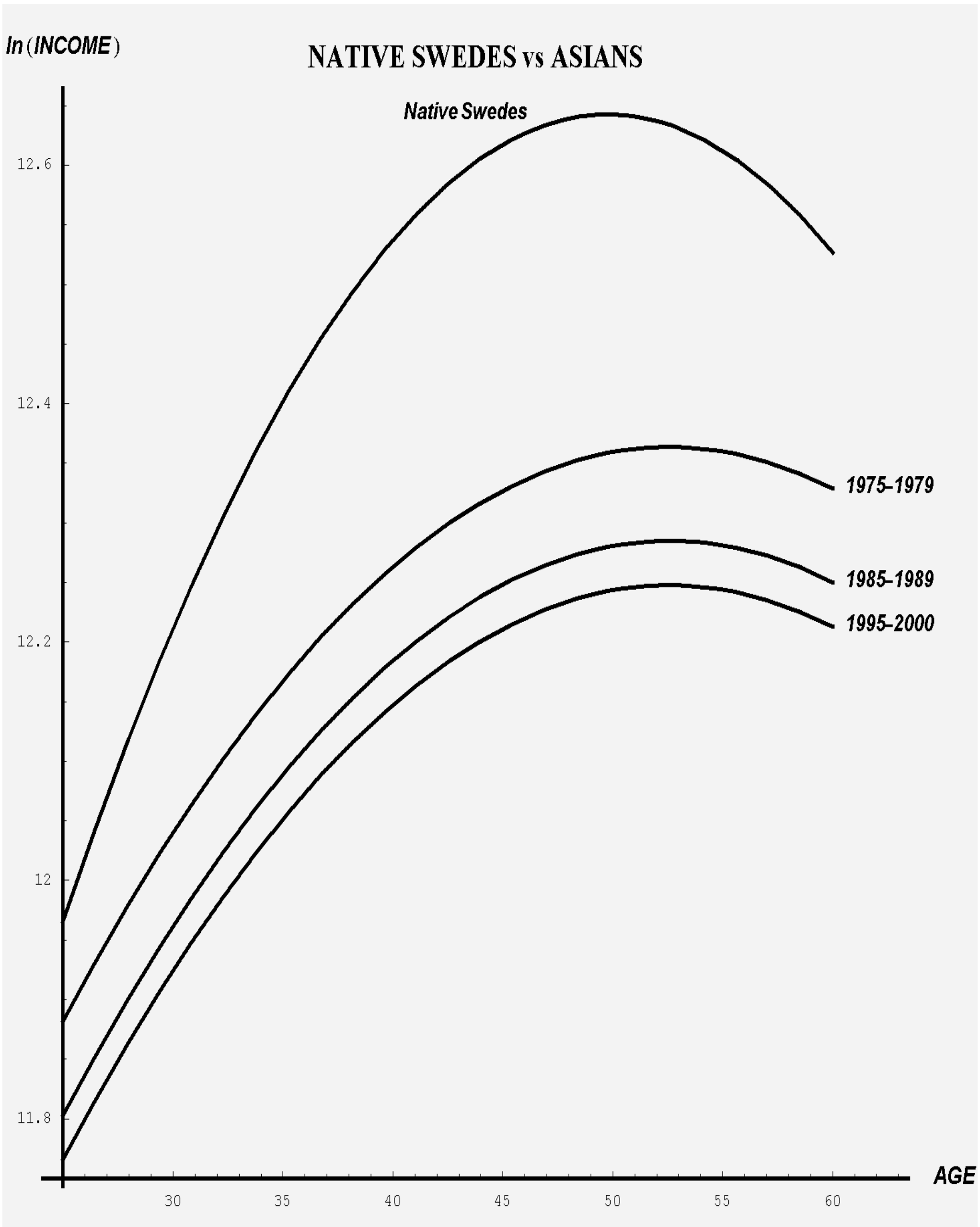
Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value

Figure 11



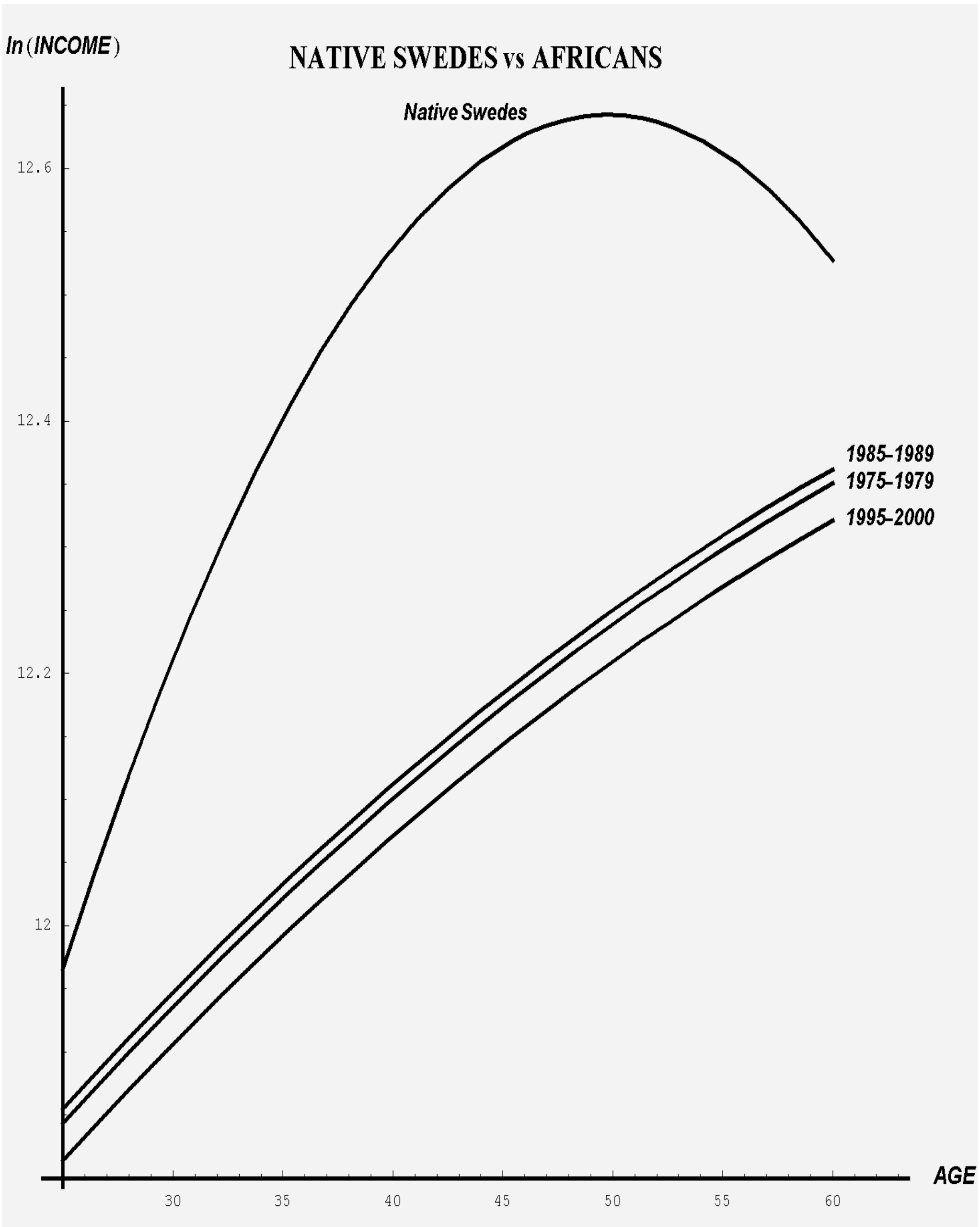
Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value

Figure 12



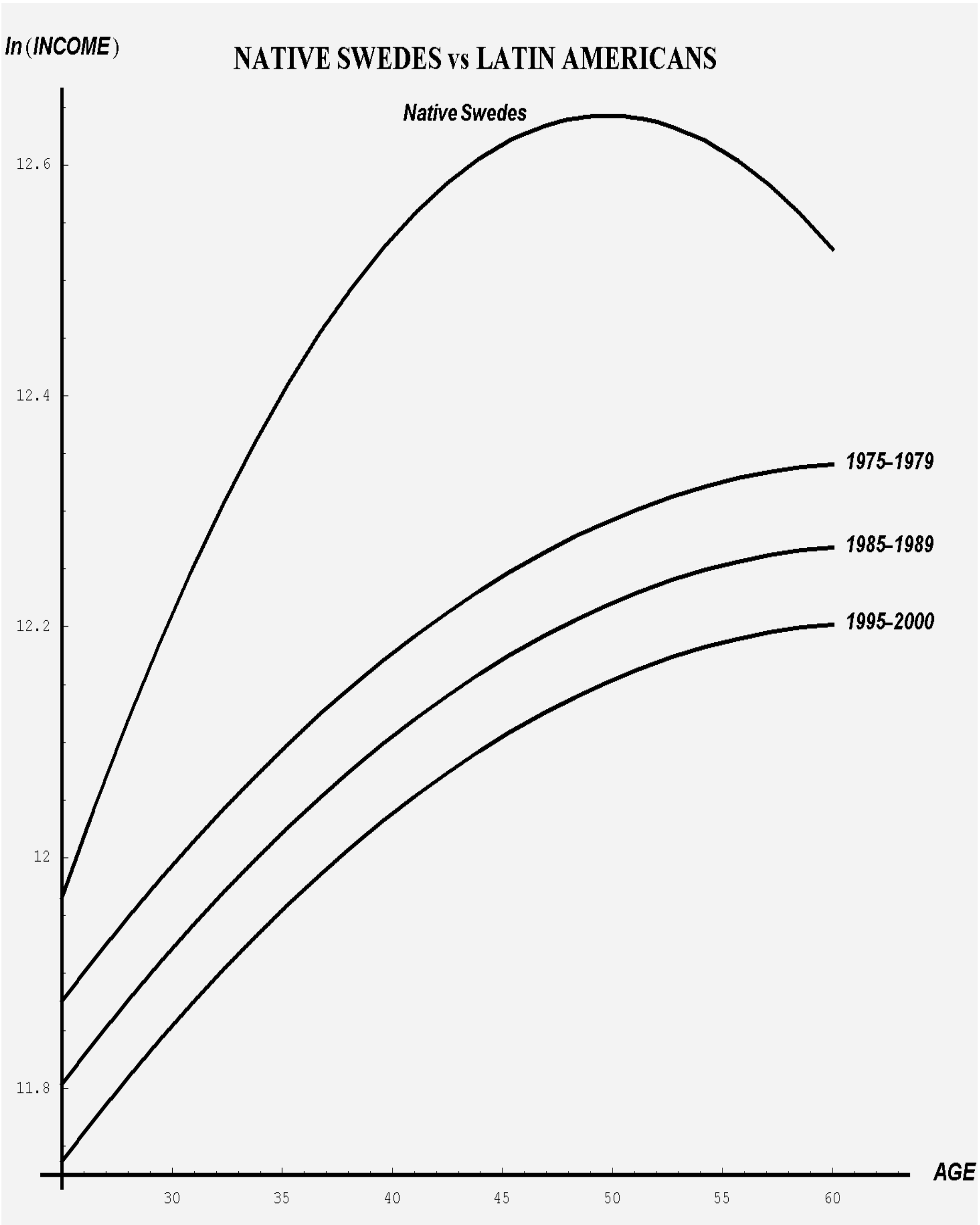
Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value

Figure 13



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value

[Figure 14]



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value

[Figure 15]

Arrival-Cohort Effects on the Incomes of Immigrant Women in Sweden

Kerem Tezic
Department of Economics, Göteborg University

March 2004

Abstract

Employment probabilities and earnings of first-generation immigrant women were estimated simultaneously in a random-effects model in order to control for individual effects and panel-selectivity due to missing earnings-information. Arrival-cohorts, family structure, husbands' working-status and earnings were all controlled for . The possible endogeneity-problem of the husbands' earnings was tackled by using predicted earnings of the husbands instead of observed earnings. Chiswick's immigrant catching-up hypothesis was not much supported by the data. Most of the immigrant groups had or have had problems establishing themselves in the Swedish labour-market. On the other hand, the last arrival-cohort of highly-educated Western immigrant women has performed slightly better than native Swedish women.

Keywords: Immigrants, labour-market integration, unbalanced panel, sample-selection, random-effects.

J.E.L Classification: C33, J15, J61.

Contents

1 Introduction

After not much immigration in earlier decades of the 20th century, Sweden welcomed many labour-market immigrants during the post-war period. At first it was largely high-skilled Nordic immigrants, followed by unskilled immigration from Southern-Europe to fill labour-needs in many industrial sectors. But in the mid-1970s there was a switch to refugee-immigration, and the labour-market experiences of the immigrants got worse, regardless of the economic-cycle (see Lundh and Ohlsson , 1994) and Ekberg and Gustafsson, 1995).

In recent decades, numerous analyzed the economic integration of immigrants, e.g., Chiswick (1978), Borjas (1985,1987), Lalonde and Topel (1991, 1992) for the U.S.; Baker and Benjamin (1994) and Grant (1999) for Canada; Husted *et al.* (2000) for Denmark; Longva and Raaum (2003) for Norway; and Wadensjö (1975), Aguilar and Gustafsson (1994), Ekberg (1994), and Hammarstedt (2001) for Sweden. But most of these were concentrated on the performance of immigrant men, or did not distinguish between men and women, despite the fact that women constitute a large proportion of immigrants and their labour-market experiences are not necessarily the same as men's.

The majority of immigrants living in Sweden are in fact women. According to Statistics Sweden the number of immigrant-women outnumbered immigrant-men by about 20,000-40,000 between the years 1985-2000.

In the few studies relevant to immigrant-women's labour-market experiences, either cross-sectional data was used, or the joint employment-decisions of women and their husband (if married) were not handled properly, or selection-issues were not handled properly when the earnings of immigrant-women were analyzed. Participation issues are especially important in direct earnings-comparisons between immigrants and native women (Dustmann and Schmidt, 2000).

The aim of this study therefore was to make an extensive analysis of immigrant women's economic experiences in Sweden. Eleven waves of detailed panel-data were used allowing unobserved heterogeneity to be accounted for. Possible selection problems were taken into account by estimating employment and earnings-equations jointly. If the women were married, her husband's employment status and earnings were not only included but the possible endogeneity problem was also handled by using husband's predicted-earnings as an instrument in woman's employment and earnings-equations, because the same unobservable factors that influenced the woman's choices might also have influenced his. Such an approach does not appear to have been used before in studies of this type.

There was no catch-up for observed African women no matter how long they had been in Sweden, and the integration process was slow for Middle Eastern, Latin American, Asian or Eastern European women.

The next section describes the previous studies more fully. Section 3 then discusses some hypotheses about the labour-market experiences of immigrant women while Section 4 presents the data. Section 5 develops the model used and discusses econometric issues. Section 6 gives the estimation results, and Section 7 summarizes and draws conclusions.

2 Previous studies

As mentioned above, there are few studies which report results on female immigrant earnings differentials. Long (1980), using 1970 US-census data —the same data and variables used by Chiswick (1978)— found that newly-arrived immigrant women earned about 13 percent more than native-born counterparts, but this advantage declined over time. These results were unexpected and puzzling, but as Borjas (1987) has shown, the use of cross-sectional data can conflate aging and arrival-cohort effects, leading the researcher to erroneous results. Besides, selectivity-bias was not addressed in Long's study.

On the other hand, Chiswick (1980), confirming his (1978) finding about immigrant men, found that even controlling for age and work-experience, females' hourly-earnings increased with time in the host country, and these findings were supported by Blau (1980), in her analysis of the earnings-assimilation of migrants. Blau (p.22) hypothesized that the relative earnings of immigrant-women increase "as they become acclimized to their new surroundings and seek out the best opportunities to utilize their skills and abilities".

Reimers (1985) also pointed out the extent to which cultural differences regarding language, family size, age, and education accounted for differences in the labour-force participation of different ethnic groups, and also found that the labour-force participation of immigrant-women adjusted to that of natives over time, due to cultural assimilation.

Field-Hendry and Balkan (1991), addressing the question of selectivity-bias but using two unmatched cross-sections, namely 1970 and 1980 U.S: census data, also found an initial disadvantage for newly arrived immigrant-women again, but that earnings differentials declined with time. They also found a selectivity-variable to be statistically significant and negative in all regressions, similar to Reimers' (1983) finding for male migrants. Their results indicated the importance of selectivity-issues when analyzing immigrant earnings.

Duleep and Sanders (1993), using a 5 percent sample of the 1980 Census and focusing on 25-65 year-old immigrant women married to men of the same nativity and race, compared the labour-force participation Asians with Europeans and Canadians. Their main interest was whether differences in U.S. specific skills explained differences in labour-force participation and to what extent variations were explained by differences in family circumstances. Using a probit-model, they showed that years since arrival, number of relatives at home, and proficiency in English were positively related with labour-force participation. They also found that whether husbands invested in human capital specific to the U.S. labour-market, and the extent of this investment, affected women's labour-force participation rates.

Beach and Worswick (1993) examined the possibility of a "double-negative" effect on the earnings of immigrant-women in Canada due to both gender and birthplace. Their 1973 Job Survey Mobility cross-sectional data permitted them to use a novel "home-time" variable, which proved to be highly significant and negative determinant of earnings, but with no significant difference between immigrant and natives. The double-negative effect did not appear for all immigrant women, particularly for the highly educated. Years since arrival was also not statistically significant, different from U.S. findings.

Dustmann and Schmidt (2000), using 12 waves of the German Socio-Economic Panel, containing 3841 native and 1073 immigrant-women who were married or cohabiting found that immigrants received lower wages in the same labour-market segment, which they attributed mainly to their lower average education.

Grossman (1984) examined the observed differences in the occupational distribution of native and immigrant-women in Sweden, using Level of Living surveys for the years 1978, 1979, and 1980. She came to the conclusion that socioeconomic variables explained very little.

3 Hypotheses

Researchers have identified several factors which can be relevant to the labour-force participation and earnings of immigrant-women, one of which is skill transferability (Chiswick 1978, 1980). According to this hypothesis, the underlying cause of initial earnings differentials was differences in country-specific labour-market skills. With time spent in the new country, those differences decrease and immigrants start to catch-up. The coefficients of arrival-cohorts and the steepness of the age-income profile yield evidence about this catch-up hypothesis. In addition, immigrant groups with more-readily transferable skills should have labour-market participation and initial earnings, and this must apply to women from countries culturally and economically similar to Sweden.

Another factor is the husband's role and family-circumstances. Seeking to explain his (1980) finding that recently-arrived female immigrants in the U.S. earned more than their native counterparts, but that their earnings decreased over time, Long put forward the Family Investment Hypothesis, whereby newly-arrived immigrant women have to work to finance their husbands' initial investments in country-specific human capital, but later, as their husbands' earnings rise they reallocate their time from market to nonmarket activities. In Sweden, the transfer system, along with government provided study-loans make it possible for immigrants as well as natives to finance investments in human capital, so this hypothesis is not so useful.

However another potentially important family-circumstance —if the woman is married or cohabiting— is whether the husband is working or not and his earnings along with the number of children in the household. If an immigrant woman is married or cohabiting with someone of the same race and national origin and has children, then their partner's earnings cannot be considered as exogenous, because, as mentioned earlier unobserved factors which enter into the error-term in the woman's employment-and earnings-equations may be correlated with the partner's earnings. But if the woman was not married or cohabiting, or not with someone of the same race and nationality, or had no children, then it should be reasonable to take the partner's earnings as observed. For this purpose, Swedes, Nordics, and Westerners might be culturally similar enough to induce endogeneity and perhaps should be considered as the same. Such an approach can be seen as the decision to work as a family investment strategy. It is also assumed that the effect of the number of children on women's earnings and her decision to work is not the same for married/cohabiting and those living alone with children. This hypothesis is tested by interacting number of children with the civil status of the woman.[1]

4 The data

The study was based on the 1990-2000 panel of the register-based Longitudinal Individual Data-set (LINDA), which contains a representative sample of approximately 3 percent of the Swedish population. The panel-data is updated with current household information each year. The information in LINDA is derived from the Population and Housing Censuses and official Income Registers, as well as a higher-education register. The income-register is based on filed tax returns, which makes the information on income contingent on the tax rules for that year. LINDA contains a panel of about 20 percent of the foreign-born population. For more details see Edin and Frederiksson (2001).

In order to avoid selection problems due to retirement at age 65, the 30,407 immigrant-women in LINDA aged 18-55 in 1990 were selected for the study, as well as an equal-sized control group of randomly-selected native Swedish women, matched for age and county (län) of residence. For each new year 1991-2000, approximately 3000 similar newly included (to LINDA) immigrant-women were added to the study, as well as an equal number of randomly-selected but matched Swedes. By 2000, the unbalanced panel consisted of 62,957 immigrant women (generating 505,362 annual observations and nearly the same number of Swedes.

The immigrants were categorized as being from the Nordic countries not including Sweden; Western countries consisting of the EU, USA, Canada, Australia, and New Zealand; Eastern Europe; the Middle-East; Asia; Africa; or Latin America.

Based on working-indicators in the data, an employment dummy was defined as 1 if the individual was employed, 0 if not.

The earnings-variable used in the study was calculated from the Tax Registers. The earnings were measured in thousands of SEK per year, adjusted using the consumer price-index to 2000 prices.

[1] Unfortunately the data only distinguished between children older or younger than 16, so we had no way to distinguish those less than school-age.

Tables 1 and 2 in the Appendix show descriptive statistics for both immigrants and native Swedes for their first year in LINDA.

Overall, only 39.4 percent of the immigrant women were employed (Table 1), compared to 72.1 percent of the native Swedes, and the Swedes averaged considerably more earnings, 128,927 SEK per year vs. 107,474 SEK for the immigrants, and the immigrants had slightly more children. While equal percentages of both groups (22 percent) had university degrees, another 52 percent of the Swedes had completed upper-secondary (high-school) education whereas only 38 percent of the immigrants had done so. Slightly more immigrants lived (38 percent vs. 35 percent) in big cities of Stockholm, Göteborg or Malmö. Far more Swedes (73 percent vs. 59 percent) had working partners (husbands or "sambos"), and they earned about 40,000 SEK more earnings per year on average. While 16 percent of the immigrants had arrived before 1970, about 11 percent arrived in each of the next 5-year cohorts, before the numbers went up dramatically during the 1990s, during the wars in former Yugoslavia, Somalia, etc. Nordic (almost 27 percent) and Eastern European (26 percent) were by far the largest geographic groups of immigrants, followed by 17 percent from the Middle East, 9 percent from the Western countries and from Asia, and 5-6 percent from Africa and from Latin America.

About 95 percent of the married or cohabiting native Swedish women were living with Swedish men (Table 2), an in-group phenomenon matched only by the Middle Easterners (nearly 97 percent). Africans (77 percent), Eastern Europeans (73 percent), and Latin Americans (65 percent) were also predominantly living with "their own", probably with someone from their home-country. At the other extreme, (but supporting the hypothesis of ethnic and cultural similarity), 55 percent of married or cohabiting Western women were living with Swedish men, as were 47 percent of Nordic women. Of the Asian women who were married or cohabiting, almost as many (41 percent vs. 48 percent) were living with Swedish as with Asian men. Latin Americans (29 percent), Eastern Europeans (21 percent), and Africans (15 percent) were also well-represented in this regard.

As we have just seen, the immigrant population was not at all homogenous, and there were considerable differences with respect to other socioeconomic characteristics as well. Tables 3 and 4 in the Appendix show those characteristics by geographic origin and by arrival cohort respectively.

During their first observation year in LINDA, Nordic and Western immigrants had much higher employment rates (68 percent and 47 percent), followed by Latin Americans (38 percent), Eastern Europeans (30 percent), Asians (29 percent), Africans (20 percent), and Middle Easterners (17 percent). The differences in average earnings for those who were employed were also substantial; Nordics averaged about 3000 SEK more earnings than did native Swedes, while all other groups had less: Westerners (4,600 less); Eastern Europeans (10,300) less etc.

The average Asian or African was about 28 , the average Latin American or Middle Easterner about 30, compared to 35 for the Eastern Europeans, Westerners (and Swedes), and 38 for the Nordics. The highest percentage married or cohabiting were the Middle Easterners (58 percent), the lowest percentages were the Asians (43 percent) and the Latin Americans (40 percent). The Middle Easterners also had the most children on average, the Asians least. The most Westerners (36 percent) had university degrees, followed by Eastern Europeans (24 percent); Africans had the least (15 percent), the rest all 19-21 percent. Similarly, the least Westerners (30 percent) had not finished upper-secondary education, compared to 47-51 percent for Middle Easterners, Africans and Asians. The most Africans (49 percent) followed by Westerners and Middle Easterners (44 percent) were living in big cities; the fewest Nordics (29 percent) were there. Far more of the Nordics (38 percent) had arrived before 1970, followed by 19 percent of the Westerners, and only 0.9 percent of the Africans. By 1979, 73 percent of the Nordics had arrived. Eastern Europeans and Westerners immigrated at steady rates during the 1970s and 1980s, with bulges throughout the 1990s for the Eastern Europeans. Latin American immigration has been steady since about 1975, whereas Asian immigration built up steadily throughout until the 1990s. Middle Eastern immigration picked up somewhat in the late 1970s and early 1980s, then held steady at a high rate for the rest of the century, a pattern also followed by Africans, with even more emphasis on the 1990s. The unemployment rate during the arrival-year was by far lowest for the Nordics (2.4 percent), highest for the Africans (3.9 percent), Middle Easterners (4.1 procent), Eastern Europeans (4.7 procent), and Asians (4.8 procent). All this is in accordance with earlier studies.

Table 4 (in the Appendix) shows the same statistics by arrival-cohort.

Although 2-3 years older than the previous cohort, those who arrived in 1995-2000 were considerably less likely to be married or cohabiting (43 percent vs. 51-58 percent for all others). Neverthless they had considerably more children than the cohort before; otherwise the number of children (living at home) follows a fairly predictable pattern associated with parental age. The educational pattern was rather steady from pre-1970 through 1989 (about 37-46 percent hadn't finished upper-secondary, 16-17 percent had a university degree); but after 1989 the low-education group fell to 36 percent and then 28 percent, while the university-educated rose to 28 percent and then 42 percent. Again, the cohort that arrived in the early-1990s recession had the lowest percentage of employed partners in their first year (21 percent), while the 1995-2000 cohort had 47 percent. Generally, those who had been here longest had partners earning the most , as one would expect. Immigrants who arrived before 1975 were 68-70 percent employed , compared to 46 percent of those who arrived in 1985-89. Only 10 percent of those who arrived during a major recession in 1990-94 and 14 percent of those who arrived in 1995-2000, were employed during their first years in Sweden. Earnings follow the same general pattern, partly due to age (those who arrived earlier are generally older), although the 1995-2000 cohort was both older and earned more than the previous cohorts.

Table 5 (in the Appendix) shows mean employment-percentages and log-of-earnings by years from 1990 to 2000 for everyone in the data set, not just in their first year in LINDA.

For Swedes, employment fell from 81.9 percent in 1990 to 74.1 percent in 1994 before recovering somewhat to 76.2 percent in 2000. Immigrants started much lower (60.7 percent) and fell much further (to 42.6 percent in 1994) before recovering somewhat to 52.8 percent in 2000. While immigrants' earnings were also consistently somewhat below Swedes' earnings (for those employed) didn't fall during the period and in fact gradually rose (for both Swedes and immigrants).

5 Econometric specifications

The main object of analysis was to compare the earning of natives and immigrants and to find out if there had been 'catch-up' as has been hypothesized in the literature. But since the data were in the form of an unbalanced panel and the immigrants' employment-rate was very low, possible selectivity-bias was an important issue. An econometric model was adopted which would both exploit the panel-aspect of data and at the same time correct for sample-selection bias.

It is well known that, with selection-bias, something other than random-sampling from the underlying population has taken place, distorting the representation of the true population and any inferences based on the observed data using standard methods. This can occur because of self-selection[2] by the individuals under investigation, or due to sample-selection decisions by the data analysts. Sample-selection models are frequently estimated in applied micro-econometric work using cross-sectional data, but less frequently when panel data are available. It has been common in many applied economic analyses of panel-data to study only the balanced sub-panel or the unbalanced panel, without correcting for selectivity-bias. One kind of selection-problem occurs when individuals do not disappear from the panel but certain variables are not observed for at least some time-periods. This is an incidental-truncation problem. Examples are the estimation of earnings-equations or hours of work equations. Some variables may be observed for everyone in each time-period but due to the fact that some individuals do not work in some of the periods, their earnings may not be observable. This in fact was the situation here.

Another main concern in empirical work is the presence of unobserved heterogeneity in the equation of interest. Heterogeneity across individuals may arise as a result of differences in preferences, individual characteristics, or endowments. Failure to account for such individual-specific effects may result in biased and inconsistent estimates of the structural parameters. In many applications, as here, both sample-selection and unobserved-heterogeneity problems can occur simultaneously. For more discussion of these topics see for instance Matyas and Sevestre (1995, ch.18), Kyriazidou (1997), or Vella and Verbeek (1999).

Due to the possibility of both sample-selectivity and unobserved heterogeneity, it was desirable to consider them simultaneously. This can be done in various ways. In our case, the random effects model suggested by Jensen *et al.* (2002) is specified and the model can be formulated as follows:

[2] A simple sample-selection bias test, suggested by Nijman and Verbeek (1992) was also performed by adding the lagged selection indicator $r_{i,t-1}$, to the equation, estimating the model by fixed effects on the unbalanced panel and doing a t test for the significance of $r_{i,t-1}$. For all the groups, $r_{i,t-1}$ was significant.

$$y_{it}^* = \beta' x_{it} + u_i + \epsilon_{it} \quad (1)$$

$$r_{it}^* = \gamma' z_{it} + v_i + \omega_{it} \quad (2)$$

$$r_{it} = \begin{array}{l} 1 \text{ if } r_{it}^* > 0, \\ 0 \text{ otherwise} \end{array}$$

$$y_{it} = y_{it}^* * r_{it}$$

where i denotes the individual; t denotes the time period; y^* denotes earnings; x_{it} and z_{it} are vector of exogenous variables; β and γ are column vectors of unknown parameters of interest; u_i and v_i are unobserved individual-specific effects; and ϵ_{it} and ω_{it} are idiosyncratic error terms. The observations for y_{it}^* may only be available if an unobserved latent variable r_{it}^*, measuring the extra benefits of being employed over not being employed, is non-negative. The following statistical assumptions are made for the idiosyncratic error terms:

$$\epsilon_{it} \sim N(0, \sigma_\varepsilon^2) \qquad i = 1, ...N \qquad t = 1, ...T \quad (3)$$

$$\omega_{it} \sim N(0, 1) \qquad i = 1, ...N \qquad t = 1, ...T \quad (4)$$

$$Corr(\epsilon_{it}, \omega_{it}) = \rho \quad (5)$$

The error-terms ϵ_{it} and ω_{it} are assumed to be non-autocorrelated. The likelihood of a single observation, conditional on the random effects, is

$$L_{it}(\gamma, \beta, \sigma_\varepsilon^2, \rho \mid u_i, v_i) = f(\epsilon_{it}, \omega_{it} \mid u_i, v_i) \quad (6)$$

$$
\begin{aligned}
&= \left[\int_{-\infty}^{-v_i-\gamma' z_{it}} \int_{-\infty}^{\infty} \phi_{\epsilon\omega}(\varepsilon,\omega) d\epsilon d\omega \right]^{1-r_{it}} \\
&\quad \times \left[\int_{-v_i-\gamma' z_{it}}^{\infty} \phi_{\epsilon\omega}(y_{it}-u_i-\beta' x_{it},\omega) d\omega \right]^{r_{it}} \\
&= \left[\int_{-\infty}^{-v_i-\gamma' z_{it}} \phi_{\omega}(\omega) d\omega \right]^{1-r_{it}} \\
&\quad \times \left[\int_{-v_i-\gamma' z_{it}}^{\infty} \phi_{\omega|\epsilon}(\omega \mid y_{it}-u_i-\beta' x_{it}) \cdot \phi_{\epsilon}(y_{it}-u_i-\beta' x_{it}) d\omega \right]^{r_{it}} \\
&= [\Phi_{\omega}(-v_i-\gamma' z_{it})]^{1-r_{it}} \\
&\quad \times \left[(1-\Phi_{\omega|\epsilon}(-v_i-\gamma' z_{it} \mid y_{it}-u_i-\beta' x_{it})) \cdot \phi_{\epsilon}(y_{it}-u_i-\beta' x_{it})\right]^{r_{it}} \qquad (7)
\end{aligned}
$$

where the conditional distribution of $\omega \mid \epsilon \sim N(\frac{\rho\epsilon}{\sigma_\epsilon}, (1-\rho^2)$. In this study, for the random effects and for their interactions with the idiosyncratic error terms, following specifications are made:

$$
\begin{aligned}
v_i &\sim N(0,\sigma_v^2) && (8)\\
u_i &\sim N(0,\sigma_u^2) && (9)\\
\epsilon_{it},\omega_{it} &\perp u_i, v_i \qquad u_i \perp v_i && (10)
\end{aligned}
$$

Thus, the individual-specific components in the selection equation and the equation of interest are assumed to be uncorrelated so that selectivity is assumed to show-up through the correlation of error terms ϵ_{it} and ω_{it}.

The contribution of the ith individual to the log likelihood conditional on random effects is:

$$
\begin{aligned}
LogL_{it} &= \sum_{r_{it}=0} \log \Phi(-v_i - \gamma' z_{it}) + \\
& \sum_{r_{it}=1} \frac{-\log 2\pi}{2} - \log \sigma_\varepsilon - \frac{(y_{it} - u_i - \beta' x_{it})^2}{2\sigma_\varepsilon^2} + \\
& \log \Phi \left[\frac{(v_i + \gamma' z_{it}) + (\rho/\sigma_\varepsilon)(y_{it} - u_i - \beta' x_{it})}{\sqrt{1-\rho^2}} \right]
\end{aligned} \tag{11}
$$

The parameters of interest are estimated by maximum likelihood method.

The conditional mean function for the sample selection model is the same as in the case of cross-sectional data and is not changed by the presence of random effects:

$$
E[y_{it} \mid x_{it}, r_{it=1}] = \beta' x_{it} + \rho \sigma_\varepsilon \sigma_\omega \frac{\phi(\gamma' z_{it})}{\Phi(\gamma' z_{it})} \tag{12}
$$

where $\sigma_\omega = 1$ due to the normalization restriction.

Further, a possible endogeneity-problem when using husbands'-earnings as a control variable was also corrected for by predicting husbands' earnings and using it as an instrument in woman's employment and earnings-equations. Husbands' earnings were predicted for each cross-sectional year correcting for selectivity-bias.

6 Results

The coefficient estimates and marginal effects of the joint estimation of employment and earnings-equations are shown on Tables 6, 7, 8 and 9 (in the Appendix).

Tables 6 and 8 shows the coefficients and marginal effects of the employment-equation. The binary-dependent variable denotes whether or not the woman was employed. For nearly all variables, except the arrival-cohorts, the signs were the same for all groups of immigrants. As one would expect, the probability of being employed increased with age, most for the Eastern Europeans, Westerners and native Swedes, and least for Middle Easterners and Africans. Those who were married or cohabiting were less likely to be employed, 10-22 percent less likely for Middle Easterners, Asians, Africans, Latin Americans, and Eastern Europeans; only 3-5 percent less likely for Westerners, native Swedes and Nordics. As Reimers (1985, p. 251) points out, "differences among ethnic subcultures may affect the labour-supply of wives more than they influence many other types of economic behaviour. Ethnic groups are distinguished by, among other things, views about male and female roles in the family..."

On the other hand, the effect of each additional child living at home was about the same for all groups, reducing the probability of working by 5-6 percent for Westerners, Africans, Middle Easterners, Eastern Europeans and Latin Americans, by 8 percent for Nordics and native Swedes. The interaction-term between married/cohabiting and number of children allowed for possibility that the effect of additional children was not the same for single women, and could vary across groups. The coefficients were statistically significant (except for native Swedes, Asians and Latin Americans) and negative for all groups (except Middle Easterners), with small marginal effects accounting for the difference in employment probabilities between single and married/cohabiting women with each additional child.

As expected, both upper-secondary education and a university degree increased the probability of being employed. Upper-secondary increased the employment probability by 11-15 percent for Africans, Middle Easterners and Latin Americans, with the others mostly in the 18-23 percent range. Relative to not having finished upper-secondary, native Swedes, Westerners and Nordics (Eastern Europeans being closed to native Swedes) with a university degree were 30-38 percent more likely to be employed but for Africans, Asians, Middle Easterners, and Latin Americans, a university degree did not substantially increase the employment probabilities over those who had finished upper-secondary. For these groups, adding a university degree made virtually no difference.

Living in a big city had a uniformly negative effect on the probability of wage-employment especially for Africans, Latin Americans, Middle Easterners and Eastern Europeans.

The coefficient for partner-employed was not statistically significant for Africans (negative), Asians, Westerners or Swedes, whereas for the other groups (except for Middle Easterners) it seemed to increase the probability of wage-employment. The coefficient of partners-earnings was also not statistically significant for the Swedes, whereas it was for all the immigrant groups positive though small. Apparently partner's earnings did not reduce the probability of women seeking wage-employment as one might have thought.

Considering arrival-cohort effects across immigrant-groups, those who came before 1970 were perhaps more likely to be retired, so that those who came in later cohorts were more likely to be employed. But the pattern changed by the mid-80s, with those coming later generally progressively less likely to be employed. Of course this is not unexpected as it takes time establish oneself in the labour-market, and it can thus even indicate “catching up”.

Figures 1,2 and 3 in the Appendix plot predicted employment-probabilities by age for a “standard” women in three different arrival-cohorts. Figure 1 shows the employment probabilities of those who arrived in Sweden during 1995-2000, compared to those of native Swedes. Of the immigrants, Nordics and Westerners were most likely to be employed, Africans and Middle Easterners least. But if we compare with Figures 2 (arrived 1985-89) and 3 (arrived 1975-79), we can see that cohorts that have been here longer have much better employment probabilities. Nevertheless, even among immigrants who arrived in Sweden in 1975-79 (Figure 3), Africans, Asians, Middle Easterners and even Eastern Europeans and Latin Americans are far from catching up, whereas Nordics are actually more likely to be employed than are native Swedes at most ages. As we have seen, the unemployment rate on arrival (Table 8) also matters, and it seems to have mattered most for Africans and Eastern Europeans.

Table 7 (in the Appendix) shows the results from the jointly-estimated earnings-equation and Table 9 displays the marginal effects. Earnings increased with age for all groups as one might expect, most for Westerners and Swedes, least for Middle Easterners and Africans, and at a decreasing rate for all groups. Being married or cohabiting reduced earnings by 15-20 percent for most groups, though only by 9-10 percent for Nordics, native Swedes and Westerners. Each additional child living at home reduced earnings by 6-10 percent for all groups. The interaction term for married/cohabiting and number of children was again not statistically significant for many groups and very small for others except for Africans and Middle Easterners where the effect was substantial. This is puzzling since the effect of a possible cultural preference for staying home with the kids should have been picked up in the employment-equation, but it may reflect less hours worked, although employed, since it is total not hourly earnings at issue here.

Having completed upper-secondary education increased earnings only 9-11 percent for all groups. Having a university degree had more effect least for Middle Easterners and Africans (about 20 percent), and most for Native Swedes and Nordics (about 28 percent).

Living in a big city increased earnings most for Swedes, Westerners, Nordics and Eastern Europeans, while for others groups the effect was either very small or negative. Having a partner employed reduced earnings by 3-9 percent for most groups. Presumably these women chose to spend more time at home. The effect of husband's earnings was positive, however but very small.

The arrival-cohort effects don't show much pattern, either across groups or over time. The most recent cohort of Westerners and Nordics had a substantial positive earnings differential compared to those who arrived before 1970. Latin American cohorts after the second one (1970-74) earned progressively less, which may be evidence of "catching up". Any patterns may be easier to see in the plots in Figures 4-13.

Figure 4 (in the Appendix) plots predicted log of income[3] by age and geographic origin for 1995-2000 arrival cohorts and Swedes. Nordics earned almost as much as Swedes, and Westerners more, but these facts may well depend upon the particular qualifications of those specific cohort.[4] The other groups were far behind, especially the Africans. Figures 5 (arrived 1985-89) and 6 (arrived 1975-79) show that earlier cohorts of Nordics and Westerners weren't earning quite as much as Swedes, but the other groups seem to have moved up a bit over time. Still the Africans were far from catching up. If one looked at Nordics and Westerners (see also figures 7 and 8) one might agree with Long's (1980) conclusion that recently-arrived immigrants initially earned more than their native counterparts, but based on all the evidence here that would seem to be spurious conclusion in general, most likely did not apply to the earlier Nordic and Western cohorts when they arrived either. Figures 9-13 for the other groups show very mixed patterns of possible small improvement over time but no general support either for Long's family-investment hypothesis or Chiswick's catching-up hypothesis.

[3]Those are plots of predicted earnings for those who work, based on the expected values of earnings formula at the end of the section 'econometric specifications'.

[4]About 60 percent of this last cohort of Western women had a university education. About 50 percent of this cohort was German, British, French or American.

7 Summary and conclusions

This has attempted to fill an important gap in the economic-integration of immigrants literature by analyzing arrival-cohort effects on the incomes of first-generation immigrant-women in Sweden. The register-based Longitudinal Individual Data set (LINDA) covering the period 1990-2000 was used. The study differed substantially from earlier ones, not only in using a very large 11-wave panel which allowed construction of a matched native control-group. In addition, a major problem in many earlier analyses of economic integration has been potential selectivity-bias due to high frequency of missing earnings-information, mostly because of non-participation. The econometric model used here handled this problem by estimating employment and earnings-equations simultaneously while allowing random-effects in both. Unobserved heterogeneity was controlled for using panel-estimation techniques which also allowed aging and cohort-effects —due to changes in the composition of immigrant groups over time— to be distinguished. A possible endogeneity-problem when using husbands'-earnings as a control variable was also corrected for by predicting husbands' earnings and using it as an instrument in woman's employment and earnings-equations. This variable was ignored in most earlier studies, and if used, the problem of endogeneity was not taken care off. The joint estimation of employment and earnings-equations was performed separately for each immigrant group defined with respect to geographical origin.

The results show that there has been little catch-up for most immigrant groups. Thus Chiswick's skill-transferability hypothesis, according to which the underlying cause of initial earnings-differentials is differences in country-specific labour-market skills, seems not to apply. With time spent in the adopted country, those differences should decrease and immigrants should start to catch up. But except for Nordics and Westerners, immigrant groups have improved slightly, despite in some cases, 30+ years in Sweden. A further study might illuminate whether the observed differences are due to measured socioeconomic traits or due to some unobserved characteristics.

References

[1] Aguilar,R., and Gustafsson, B. (1994), "*Immigrants in Sweden's Labour-Market during the 1980's,*" Scandinavian Journal of Social Welfare, No. 3, pp. 139-47.

[2] Antelius, J., and Björklund, A.(2000), "*How Reliable are Register Data for Studies of the Return on Schooling? An Examination of Swedish Data,*" Scandinavian Journal of Educational Research, Vol. 44, No. 4, pp. 341-55.

[3] Baker, M., and Benjamin, D.(1994), "*The Performance of Immigrants in the Canadian labour-market,*" Journal of labour Economics, Vol. 12, No. 3, pp. 369-405.

[4] Beach, C.M., and Worswick, C. (1993), "*Is There a Double Negative Effect on the Earnings of Immigrant Women?*" Canadian Public Policy, Vol. 19, No.1, pp. 36-53.

[5] Blau, F. (1980), "*Immigration and Labour Earnings in Early Twentieth Century America,*" Research in Population Economics, Vol. 2, ed. Julian L. Simon and Julie Da Vanzo (Greenwich, Conn. :Jai Press, 1980), pp. 21-41.

[6] Borjas, G.J. (1985), "*Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants,*" Journal of Labour Economics, Vol. 3, No. 4, pp. 463-89.

[7] Borjas, G.J. (1987), "*Self-Selection and the Earnings of Immigrants,*" American Economic Review, Vol. 77, No. 4, pp. 531-53.

[8] Chiswick, B.R. (1978), "*The Effect of Americanisation on the Earnings of Foreign-born Men,*" The Journal of Political Economy, Vol. 86, No. 5, pp. 897-921.

[9] Chiswick, B.R. (1980), "*An Analysis of the Economic Progress and Impact of Immigrants,*" National Technical Information Service, No. PB80-200454, Prepared for the Employment and Training Administration, U.S. Department of Labour Monograph.

[10] Duleep, H.O., and Sanders, S. (1993), "*The Decision to Work by Married Immigrant Women,*" Industrial and Labour Relations Review, Vol. 46, No. 4, pp. 677-90.

[11] Duleep, H.O., and Sanders, S. (1994), "*Empirical Regularities across Cultures: The Effect of Children on Woman's Work,*" The Journal of Human Resources, Vol. 29, No. 2, pp. 328-47.

[12] Dustmann, C., and Schmidt, C. (2000), "*The Wage Performance of Immigrant Women: Full-Time Jobs, Part-Time Jobs, and the Role of Selection,*" Discussion Paper No. 233, The Institute for the Study of Labour IZA, Bonn.

[13] Edin, P.-A., and Frederiksson, P. (2001), "*LINDA-Longitudinal Individual Data for Sweden,*" Working Paper 2001:6, Department of Economics, Uppsala University.

[14] Ekberg, J. (1994), "*Economic Progress Among Immigrants in Sweden,*" Scandinavian Journal of Social Welfare, Vol. 3, pp. 148-57.

[15] Ekberg, J., and Gustafsson, B. (1995), "*Invandrare på Arbetsmarknaden (Immigrants in the Labour-market),*" SNS Förlag.

[16] Field-Hendry, E., and Balkan, E. (1991), "*Earnings and Assimilation of Female Immigrants,*" Applied Economics, Vol. 23, pp. 1665-672.

[17] Grant, M.L. (1999), "*Evidence of New Immigrant Assimilation in Canada,*" The Canadian Journal of Economics, Vol. 32, No. 4, pp. 930-55.

[18] Grossman, J.B. (1984), "*The Occupational Attainment of Immigrant Women in Sweden,*" Scandinavian Journal of Economics, Vol. 86, No. 3, pp. 337-51.

[19] Hammarstedt, M. (2001), "*Disposable Income Differences between Immigrants and Natives in Sweden,*" International Journal of Social Welfare, Vol.10, No. 2, pp. 117-26.

[20] Husted, L., Nielsen, H.S., Rosholm, M., and Smith, N. (2000), "*Employment and Wage Assimilation of Male First Generation Immigrants in Denmark,*" Discussion Paper No. 101, The Institute for the Study of labour IZA, Bonn.

[21] Lalonde R.J., and Topel R.H. (1991), "*Immigrants in the American Labourmarket: Quality, Assimilation, and Distributional Effects,*" American Economic Review, Vol. 81, No. 2, pp. 297-302.

[22] Lalonde R.J., and Topel R.H. (1992), "*The Assimilation of Immigrants in the U.S. Labour-market,*" Working Paper No. 3573, NBER Working Paper Series§, pp. 67-92.

[23] Long, J.E. (1980), "*The Effect of Americanization on Earnings: Some Evidence for Women,*" The Journal of Political Economy, Vol. 88, No. 3, pp. 620-29.

[24] Longva, P., and Raaum, O. (2003), "*Earnings Assimilation of Immigrants in Norway: A Reappraisal,*" Journal of Population Economics, Vol. 16, pp. 177-93.

[25] Lundh, C., and Ohlsson, R. (1994), "*Från Arbetskraftsimport till Flyktning Invandring*", SNS Förlag.

[26] Nijman, T., and Verbeek, M. (1992), "*Nonresponse in Panel Data: The Impact on Estimates of a Life Cycle Consumption Function,*" Journal of Applied Econometrics, Vol. 7, pp.243-257.

[27] Reimers, C.W. (1983), "*Labour-market Discrimination Against Hispanic and Black Men,*" The Review of Economics and Statistics, Vol. 65, No. 4, pp. 570-79.

[28] Reimers, C.W. (1985), "*Cultural Differences in Labour Force Participation Among Married Women,*" The American Economic Review, Vol. 75, No. 2, Papers and Proceedings of the Ninety-Seventh Annual Meeting of the American Economic Association, pp. 251-55.

[29] Österberg, T. (2000), "*Economic Perspectives on Immigrants and Intergenerational Transmissions,*" Economic Studies 102. Thesis. Göteborg University. Sweden.

Appendix: Tables and Figures

Table 1: Mean characteristics of immigrants and native Swedes, first year in the dataset (LINDA) n=62,957 (std.dev. in parantheses)

	Immigrants		Native Swedes	
	Mean	Std. dev.	Mean	Std. dev
Employed[1]	0.394	(0.488)	0.721	(0.448)
Log of earnings	11.585	(0.534)	11.767	(0.502)
Age	33.741	(11.189)	35.045	(12.013)
Age-squared	1264.820	(823.366)	1399.641	(979.475)
Married /cohabiting	0.522	(0.499)	0.450	(0.497)
Number of children at home	0.714	(1.116)	0.663	(1.005)
Education (highest level):				
Lower-secondary	0.394	(0.486)	0.256	(0.436)
Upper-secondary	0.382	(0.483)	0.520	(0.499)
University degree	0.223	(0.413)	0.223	(0.416)
Big city (>250,000)	0.380	(0.484)	0.351	(0.444)
Husband/partner employed	0.538	(0.485)	0.732	(0.445)
Log of husband's/partner's earnings[2]	12.084	(0.567)	12.293	(0.466)
Log of husband's/partner's predicted-earnings[3]	12.186	(0.274)	12.365	(0.184)
Arrival Cohort:				
before 1970	0.162	(0.341)		
1970-74 (5 years)	0.110	(0.288)		
1975-79 (5 years)	0.117	(0.296)		
1980-84 (5 years)	0.094	(0.269)		
1985-89 (5 years)	0.130	(0.311)		
1990-94 (5 years)	0.194	(0.367)		
1995-2000 (6 years)	0.193	(0.366)		
Unemployment rate in arrival year	3.722	(2.527)		
Birthplace:				
Nordic countries	0.268	(0.443)		
Western countries	0.091	(0.287)		
Eastern Europe	0.262	(0.443)		
Middle East	0.172	(0.376)		
Asia	0.089	(0.285)		
Africa	0.052	(0.224)		
Latin America	0.063	(0.243)		

[1] "Employed" means wage-employment; those attending school or otherwise occupied as well as those unemployed, were considered non-employed.

[2] This is the log of the mean-earnings of the partner of the women in our sample who are married or cohabiting.

[3] The partners's predicted-earnings were estimated for each year, taking selection into account, and used in both employment and earnings equations.

Table 2: Percent of married/cohabiting Swedish and immigrant women, by geographic origin of their partners by geographic origin of their partners

Men from:	Women from: Sweden	Nordic countries	Western countries	Eastern Europe	Middle East	Asia	Africa	Latin America
Sweden	94.89	45.86	55.46	20.54	2.26	41.23	14.56	28.62
Nordic countries	2.17	48.54	2.24	1.66	0.10	2.96	1.04	1.51
Western countries	1.63	2.13	35.48	1.68	0.31	3.32	1.47	2.01
Eastern Europe	0.74	1.03	2.45	72.66	0.33	0.73	0.87	0.93
Middle-East	0.21	0.87	1.74	2.37	96.68	2.13	3.55	1.15
Asia	0.11	0.41	1.12	0.12	0.06	48.50	1.04	0.14
Africa	0.14	0.68	0.62	0.33	0.18	0.99	77.30	0.43
Latin America	0.13	0.48	0.87	0.64	0.08	0.16	0.17	65.21

Table 3: Mean characteristics of immigrants, first year in the data-set LINDA, by geographical origin

	Nordic n=16945	Western n=5707	Eastern- Eur. n=16531	Middle East n=10799	Asia n=5648	Africa n=3336	Latin America n=3991
Employed	0.680	0.470	0.303	0.166	0.288	0.202	0.375
	(0.466)	(0.499)	(0.459)	(0.372)	(0.453)	(0.402)	(0.484)
Log of earnings	11.790	11.730	11.683	11.376	11.452	11.455	11.516
	(0.494)	(0.576)	(0.541)	(0.504)	(0.503)	(0.503)	(0.525)
Age	38.136	35.091	35.168	30.270	27.751	27.927	30.288
	(11.134)	(11.703)	(10.939)	(9.892)	(9.002)	(8.347)	(10.126)
Age-squared	1578.34	1368.36	1356.46	1014.13	851.14	849.59	1019.93
	(869.73)	(893.83)	(818.76)	(690.82)	(578.93)	(559.65)	(675.44)
Married/cohabiting	0.493	0.505	0.520	0.580	0.429	0.455	0.396
	(0.499)	(0.500)	(0.493)	(0.499)	(0.495)	(0.498)	(0.489)
Number of children at home	0.671	0.603	0.726	0.947	0.539	0.736	0.753
	(1.059)	(0.986)	(1.038)	(1.406)	(0.874)	(1.175)	(1.070)
Education (highest level):							
Lower-secondary	0.368	0.301	0.334	0.474	0.510	0.503	0.431
	(0.482)	(0.457)	(0.469)	(0.497)	(0.499)	(0.499)	(0.494)
Upper-secondary	0.431	0.342	0.423	0.322	0.300	0.345	0.364
	(0.495)	(0.472)	(0.492)	(0.461)	(0.452)	(0.469)	(0.479)
University degree	0.200	0.357	0.243	0.205	0.190	0.152	0.205
	(0.399)	(0.473)	(0.426)	(0.396)	(0.385)	(0.399)	(0.401)
Big city (>250,000)	0.289	0.444	0.362	0.440	0.361	0.485	0.396
	(0.453)	(0.496)	(0.481)	(0.469)	(0.480)	(0.499)	(0.489)
Husband/partner employed	0.766	0.704	0.423	0.319	0.599	0.413	0.664
	(0.485)	(0.470)	(0.435)	(0.393)	(0.437)	(0.406)	(0.440)
Log of husband's/partner's earnings	12.235	12.286	12.057	11.667	12.018	11.803	11.969
	(0.486)	(0.602)	(0.549)	(0.569)	(0.543)	(0.577)	(0.541)
Log of husband's/partner's predicted earnings	12.304	12.324	12.193	11.875	12.127	11.880	12.085
	(0.167)	(0.218)	(0.246)	(0.278)	(0.276)	(0.334)	(0.254)
Arrival Cohort:							
before 1970-	0.378	0.188	0.117	0.012	0.026	0.009	0.012
	(0.473)	(0.371)	(0.285)	(0.100)	(0.147)	(0.144)	(0.097)
1970-74 (5 years)	0.190	0.114	0.116	0.025	0.088	0.033	0.032
	(0.376)	(0.300)	(0.283)	(0.141)	(0.261)	(0.220)	(0.155)
1975-79 (5 years)	0.161	0.114	0.080	0.082	0.144	0.042	0.189
	(0.351)	(0.300)	(0.239)	(0.251)	(0.325)	(0.183)	(0.353)
1980-84 (5 years)	0.077	0.098	0.093	0.073	0.150	0.071	0.179
	(0.255)	(0.281)	(0.257)	(0.238)	(0.330)	(0.270)	(0.345)
1985-89 (5 years)	0.067	0.110	0.107	0.241	0.127	0.156	0.228
	(0.238)	(0.295)	(0.274)	(0.399)	(0.307)	(0.334)	(0.381)
1990-94 (5 years)	0.054	0.111	0.268	0.305	0.228	0.368	0.178
	(0.215)	(0.297)	(0.404)	(0.434)	(0.451)	(0.459)	(0.344)
1995-2000 (6years)	0.071	0.264	0.218	0.263	0.237	0.320	0.181
	(0.244)	(0.422)	(0.372)	(0.413)	(0.322)	(0.440)	(0.347)
Unemployment rate in arrival year	2.433	3.586	4.712	4.065	4.770	3.863	3.277
	(1.587)	(2.361)	(2.817)	(2.390)	(2.618)	(2.473)	(2.213)

Table 4: Mean characteristics of immigrants, first year in the data-set LINDA, by arrival-cohort

(Std. dev. in parantheses)

	before 1970	1970-74	1975-79	1980-84	1985-89	1990-94	1995-99
Employed	0.677	0.696	0.576	0.565	0.458	0.100	0.141
	(0.467)	(0.459)	(0.494)	(0.495)	(0.498)	(0.291)	(0.348)
Log of earnings	11.906	11.748	11.689	11.642	11.488	11.296	11.639
	(0.481)	(0.489)	(0.509)	(0.506)	(0.506)	(0.516)	(0.668)
Age	45.969	36.768	32.825	32.171	30.472	29.479	32.137
	(10.416)	(10.118)	(10.469)	(9.592)	(9.245)	(9.298)	(10.089)
Age-squared	2221.72	1454.27	1187.10	1126.97	1014.05	955.51	1134.60
	(916.15)	(718.89)	(691.24)	(642.89)	(623.81)	(623.30)	(756.38)
Married/cohabiting	0.583	0.543	0.506	0.505	0.543	0.541	0.427
	(0.492)	(0.498)	(0.499)	(0.500)	(0.498)	(0.498)	(0.494)
Number of children at home	0.571	0.926	1.032	1.073	0.961	0.296	0.511
	(0.928)	(1.092)	(1.245)	(1.231)	(1.167)	(0.795)	(1.032)
Education (highest level):							
Lower-secondary	0.368	0.403	0.430	0.441	0.462	0.360	0.277
	(0.482)	(0.490)	(0.495)	(0.496)	(0.498)	(0.480)	(0.447)
Upper-secondary	0.457	0.437	0.424	0.391	0.338	0.361	0.303
	(0.498)	(0.496)	(0.491)	(0.488)	(0.472)	(0.480)	(0.459)
University degree	0.174	0.156	0.163	0.158	0.173	0.279	0.420
	(0.378)	(0.363)	(0.368)	(0.364)	(0.378)	(0.437)	(0.493)
Big city (>250,000)	0.316	0.357	0.374	0.442	0.433	0.334	0.453
	(0.465)	(0.479)	(0.483)	(0.496)	(0.495)	(0.471)	(0.497)
Husband/partner employed	0.679	0.747	0.750	0.736	0.646	0.212	0.468
	(0.489)	(0.491)	(0.479)	(0.483)	(0.477)	(0.319)	(0.399)
Log of husband's/partner's earnings	12.292	12.168	12.149	12.074	11.917	11.834	11.639
	(0.482)	(0.479)	(0.508)	(0.510)	(0.560)	(0.593)	(0.668)
Log of husband's/partner's predicted earnings	12.305	12.301	12.272	12.215	12.121	11.972	12.064
	(0.181)	(0.187)	(0.210)	(0.230)	(0.224)	(0.333)	(0.434)
unemployment rate in arrival year	1.775	2.120	1.821	2.718	2.217	4.812	4.124
	(0.345)	(0.524)	(0.276)	(0.587)	(0.427)	(2.473)	(0.977)

Table 5: Mean employment and log of earnings (in 1000 SEK), immigrants and native Swedes, 1990-2000 (Std. dev. in parantheses)

		Immigrants		**Native Swedes**	
		Employed	Log of earnings	Employed	Log of earnings
		Mean		Mean	
1990	n=30407	0.607	11.660	0.819	11.749
		(0.488)	(0.503)	(0.384)	(0.469)
1991	n=33136	0.566	11.667	0.803	11.748
		(0.495)	(0.502)	(0.397)	(0.509)
1992	n=35579	0.523	11.725	0.781	11.786
		(0.499)	(0.505)	(0.413)	(0.522)
1993	n=39168	0.463	11.718	0.748	11.779
		(0.498)	(0.509)	(0.433)	(0.472)
1994	n=43673	0.426	11.721	0.741	11.790
		(0.494)	(0.528)	(0.438)	(0.482)
1995	n=46183	0.428	11.728	0.751	11.799
		(0.496)	(0.626)	(0.432)	(0.482)
1996	n=48560	0.429	11.791	0.753	11.867
		(0.494)	(0.541)	(0.431)	(0.491)
1997	n=53884	0.425	11.815	0.743	11.894
		(0.494)	(0.554)	(0.436)	(0.499)
1998	n=57446	0.480	11.852	0.758	11.945
		(0.499)	(0.558)	(0.427)	(0.509)
1999	n=58915	0.508	11.878	0.764	11.976
		(0.499)	(0.560)	(0.424)	(0.518)
2000	n=58411	0.528	11.906	0.762	12.007
		(0.499)	(0.565)	(0.425)	(0.525

Standard deviations in parantheses

Table 6: **Jointly estimated parameters for employment equation from the random-effects model of employment and earnings**[1]

	Native Swedes n=505080	Nordics n=143748	Westerners n=42222	Eastern Europeans n=131063	Middle Easterners n=83092	Asians n=45010	Africans n=25399	Latin Americans n=34828
constant	-4.449	-4.061	-4.734	-6.665	-5.192	- 5.255	-5.626	-4.959
	(.872e-02)	(.103e-01)	(.153e-01)	(.445e-02)	(.661e-02)	(.700e-02)	(.136e-01)	(.896e-02)
Age	0.273	0.226	0.257	0.259	0.243	0.289	0.298	0.267
	(.896e-03)	(.813e-03)	(.140e-2)	(.103e-2)	(.903e-03)	(.108e-2)	(.170e-02)	(.116e-02)
Age-squared	-0.0033	-0.0038	-0.0032	-0.0044	-0.0031	-0.0035	-0.0036	-0.0033
	(.140e-01)	(.122e-01)	(.207e-01)	(.118e-01)	(.145e-1)	(.177e-01)	(.291e-01)	(.182e-01)
Married/cohabiting	-0.099	-0.144	-0.083	-0.574	-0.396	-0.438	-0.514	-0.341
	(.140e-02)	(.896e-02)	(.151e-01)	(.762e-02)	(.989e-02)	(.125e-01)	(.179e-01)	(.150e-01)
Number of children at home	-0.210	-0.201	-0.125	-0.161	-0.181	-0.207	-0.205	-0.181
	(.675e-02)	(.414e-02)	(.840e-02)	(.458e-02)	(.462e-02)	(.671e-02)	(.606e-02)	(.615e-02)
Married/coh.*no. of children	0.985e-03$^{n.s}$	-0.024	-0.083	-0.024	0.026	-0.008$^{n.s}$	-0.514	0.001$^{n.s}$
	(.824e-02)	(.527e-02)	(.101e-01)	(.551-e02)	.558e-02	(.125e-01)	(.179e-01)	(.846e-02)
Education (highest level):								
Upper-secondary	0.470	0.482	0.634	0.545	0.403	0.482	0.457	0.416
	(.117e-02)	(.715e-02)	(.136e-01)	(.749e-02)	(.850e-02)	(.103e-01)	(.142-01)	(.116e-01)
University degree	0.777	0.940	0.947	0.771	0.429	0.505	0.458	0.557
	(.135e-01)	(.903e-02)	(.146e-01)	(.851e-02)	(.102e-01)	(.116e-01)	(.193e-01)	(.143e-01)
Big city (>250,000)	-0.031	-0.033	0.003$^{n.s}$	-0.149	-0.116	.523e-03$^{n.s}$	-0.120	-0.060
	(.102e-01)	(.699e-02)	(.110e-01)	(.637e-02)	(.735e-02)	(.907e-02)	(.130e-01)	(.104e-01)
Husband/partner employed	0.071$^{n.s}$	0.059$^{n.s}$	0.054$^{n.s}$	0.308	-0.270	0.062$^{n.s}$	-0.002$^{n.s}$	0.326
	(.596e-01)	(.351e-01)	(.563e-01)	(.261e-01)	(.269e-01)	(.441e-01)	(.649e-01)	(.529e-01)
husband's/partner's predicted earnings	0.128$^{n.s}$	0.033	0.042	0.063	0.029	0.034	0.050	0.018
	(.100e-00)	(.315e-02)	(.504e-02)	(.236e-02)	(.247e-02)	(.402e-02)	(.594e-02)	(.485e-02)

"continued"

Table 6 : (cont.) **Jointly estimated parameters for employment equation from the random-effects model of employment and earnings**[1]

	Native Swedes	Nordics	Westerners	Eastern Europeans	Middle Easterners	Asians	Africans	Latin Americans
Arrival Cohort :								
1970-74 (5 years)		0.049	0.161	0.280	0.677	0.536	0.713	0.401
		(.906e-02)	(.193e-01)	(.115e-01)	(.223e-01)	(.171e-01)	(.363e-01)	(.335e-01)
1975-79 (5 years)		-0.077	0.074	0.312	0.697	0.099	0.624	0.299
		(.965e-02)	(.195e-01)	(.130e-01)	(.144e-01)	(.161e-01)	(.360e-01)	(.167e-01)
1980-84 (5 years)		-0.194	0.054	0.510	0.646	0.293	0.646	0.298
		(.123e-01)	(.202e-01)	(.120e-01)	(.149e-01)	(.156e-01)	(.279e-01)	(.169e-01)
1985-89 (5 years)		-0.646	-0.451	0.238	0.230	-0.063	-0.406	0.058
		(.124e-01)	(.195e-01)	(.111e-01)	(.110e-01)	(.159e-01)	(.225e-01)	(.154e-01)
1990-94 (5 years)		-0.676	-0.745	-0.219	-0.408	-0.452	-0.303	-0.444
		(.147e-01)	(.206e-01)	(.894e-02)	(.121e-01)	(.148e-01)	(.197e-01)	(.181e-01)
1995-2000 (6 years)		-0.645	-0.698	-0.555	-0.911	-0.723	-0.327	-0.397
		(.269e-01)	(.294e-01)	(.144e-01)	(.218e-01)	(.234e-01)	(.303e-01)	(.323e-01)
Unemployment rate in arrival year		-0.108	-0.115	-0.153	-0.069	-0.100	-0.137	-0.104
		(.335e-02)	(.421e-02)	(.154e-02)	(.248e-02)	(.297e-02)	(.351e-02)	(.411e-02)
σ_v	1.253	1.729	1.677	1.402	1.062	1.096	1.135	1.110
	(.649e-02)	(.641e-02)	(.107e-01)	(.446e-02)	(.454e-02)	(.529e-02)	(.880e-02)	(.641e-02)

[1] Due to the size of the sample almost all parameters are significant, at the 1 percent level except for those marked $^{n.s}$ or for those marked on the power of coefficients which indicate p-values in percents, i.e power 6 means p-value between 0.05 and 0.06...so on.
Note: The reference variables are: single, lower-secondary education, small city, husband/partner not employed, arrival cohort before 1970.

Table 7: **Jointly estimated parameters for earnings equation from the random-effects model of employment and earnings**[1]

	Native Swedes	Nordics	Westerners	Eastern Europeans	Middle Easterners	Asians	Africans	Latin Americans
constant	10.289	10.230	9.882	10.629	11.206	10.910	11.028	10.834
	(.886e-02)	(.130e-01)	(.269e-01)	(.177e-01)	(.144e-01)	(.139e-01)	(.288e-01)	(.170e-01)
Age	0.071	0.064	0.088	0.058	0.032	0.046	0.027	0.045
	(.353e-03)	(.652e-03)	(.113e-02)	(.325e-03)	(.479e-03)	(.526e-03)	(.923e-03)	(.515e-03)
Age-squared	-0.0007	-0.0006	-0.0009	-0.0004	-0.0003	-0.0005	-0.0002	-0.0004
	(.365e-02)	(.798e-02)	(.536e-02)	(.300e-02)	(.597e-02)	(.649e-02)	(.899e-02)	(.552e-02)
Married/cohabiting	-0.108	-0.082	-0.089	-0.100	-0.035	-0.064	-0.026	-0.076
	(.335e-02)	(.197e-02)	(.477e-02)	(.254e-02)	(.561e-02)	(.612e-02)	(.947e-02)	(.653e-02)
Number of children at home	-0.073	-0.063	-0.063	-0.039	-0.030	-0.038	-0.044	-0.046
	(.180e-02)	(.104e-02)	(.262e-02)	(.176e-02)	(.256e-02)	(.320e-02)	(.327e-02)	(.267e-02)
Married/coh.*no. of children	0.008	.552e-04$^{n.s}$	-0.016	0.002$^{n.s}$	-0.100	0.003$^{n.s}$	0.004$^{n.s}$	0.013
	(.218e-02)	(.125e-02)	(.318e-02)	(.209e-02)	(.318e-02)	(.412e-02)	(.475e-02)	(.364e-02)
Education (highest level):								
Upper-secondary	0.104	0.062	0.052	0.025	0.027	0.011^{2}	0.016	0.017
	(.328e-02)	(.166e-02)	(.489e-02)	(.286e-02)	(.457e-02)	(.474e-02)	(.690e-02)	(.501e-02)
University degree	0.272	0.248	0.276	0.169	0.092	0.133	0.129	0.115
	(.339e-02)	(.185e-02)	(.462e-02)	(.292e-02)	(.502e-02)	(.512e-02)	(.606e-02)	(.557e-02)
Big city (>250,000)	0.034	0.024	0.028	0.022	0.014	0.004$^{n.s}$	0.007	-0.003$^{n.s}$
	(.246e-02)	(.148e-02)	(.334e-02)	(.214e-02)	(.382e-02)	(.399e-02)	(.599e-02)	(.416e-02)
Husband/partner employed	-0.097	-0.026	-0.102	-0.078	-0.071	-0.057	-0.032$^{n.s}$	-0.035$^{n.s}$
	(.142e-01)	(.771e-02)	(.188e-01)	(926e-02)	(.145e-02)	(.211e-01)	(.327e-01)	(.231e-01)
husband's/partner's predicted earnings	0.007	0.006	0.010	0.009	0.003	0.004^{3}	0.002$^{n.s}$	0.003$^{n.s}$
	(.127e-02)	(.690e-03)	(.167e-02)	(837e-03)	(.133e-02)	(.191e-02)	(.296e-02)	(.209e-02)

"continued"

Table 7: (cont.) **Jointly estimated parameters for earnings equation from the random-effects model of employment and earnings[1]**

	Native Swedes	Nordics	Westerns	East-Europeans	Middle.-Easterns	Asians	Africans	Latin Americans
Arrival Cohort :								
1970-74 (5 years)		-0.003^{8}	-0.022	0.042	-0.031	0.041	0.080	0.125
		(.187e-02)	(.536e-02)	(.338e-02)	(.991e-02)	(.713e-02)	(.134e-01)	(.118e-01)
1975-79 (5 years)		0.016	-0.050	$0.004^{n.s}$	-0.058	-0.013^{5}	0.030^{3}	0.016^{1}
		(.200e-02)	(.564e-02)	(.386e-02)	(.685e-02)	(.693e-02)	(.137e-01)	(.637e-02)
1980-84 (5 years)		0.035	$-0.004^{n.s}$	0.025	-0.036	-0.041	$0.009^{n.s}$	0.015^{2}
		(.263e-02)	(.578e-02)	(.349e-02)	(.734e-02)	(.658e-02)	(.112e-01)	(.640e-02)
1985-89 (5 years)		0.018	-0.082	0.020	-0.045	-0.015^{2}	$0.005^{n.s.}$	-0.017
		(.291e-02)	(.590e-02)	(.353e-02)	(.559e-02)	(.676e-02)	(.992e-02)	(.619e-02)
1990-94 (5 years)		-0.008^{1}	-0.030	-0.050	-0.030	$-0.003^{n.s}$	0.021^{3}	$-0.002^{n.s}$
		(.365e-02)	(.695e-02)	(.341e-02)	(.653e-02)	(.683e-02)	(.948e-02)	(.777e-02)
1995-2000 (6 years)		0.089	0.092	-0.046	-0.048	$-0.016^{n.s.}$	$-0.005^{n.s}$	-0.059
		(.672e-02)	(.101e-02)	(.604e-02)	(.132e-01)	(.114e-01)	(.154e-01)	(.138e-01)
σ_u	0.321	0.352	0.414	0.354	0.249	0.252	0.245	0.253
	(.100e-02)	(.701e-03)	(.163e-02)	(.971e-03)	(.163e-02)	(.166e-02)	(.276e-02)	(.181e-02)
σ_ε	0.362	0.335	0.393	0.403	0.526	0.500	0.484	0.461
	(.327e-03)	(.437e-02)	(.110e-02)	(.453e-03)	(.764e-03)	(.655e-03)	(.129e-02)	(.742e-03)
ρ	-0.303	-0.422	-0.517	-0.701	-0.875	-0.903	-0.778	-0.854
	(.249e-02)	(.471e-02)	(.639e-02)	(.237e-02)	(.321e-02)	(.152e-02)	(.436e-02)	(.241e-02)
Log-likelihood	-328562.7	-100190.8	-32086.7	-87279.9	-46290.5	-34973.3	-16539.9	-27324.8

Note: The reference variables are: single, lower-secondary education, small city, husband/partner not employed, arrival cohort before 1970.

Table 8: **Marginal effects in employment equation (reference Table 6)**

		Native Swedes n=505080	Nordics n=143748	Westerners n=42222	Eastern Europeans n=131063	Middle Easterners n=83092	Asians n=45010	Africans n=25399	Latin Americans n=34828
Age	coefficient	0.273	0.220	0.257	0.259	0.243	0.289	0.298	0.267
	slope	0.105	0.087	0.110	0.139	0.069	0.114	0.071	0.102
Married/cohabiting	coefficient	-0.099	-0.144	-0.083	-0.574	-0.396	-0.438	-0.514	-0.341
	slope	-0.038	-0.058	-0.032	-0.222	-0.116	-0.174	-0.116	-0.132
Number of children at home	coefficient	-0.210	-0.201	-0.125	-0.161	-0.181	-0.207	-0.205	-0.181
	slope	-0.080	-0.081	-0.048	-0.063	-0.052	-0.081	-0.048	-0.069
Married/coh.*no. of children	coefficient	0.985e-03$^{n.s}$	-0.024	-0.083	-0.024	0.026	-0.008$^{n.s}$	-0.514	0.001$^{n.s}$
	slope	0.0003	-0.009	-0.032	-0.009	0.007	-0.003	-0.122	0.0003
Education (highest level):									
Upper-secondary	coefficient	0.470	0.482	0.634	0.545	0.403	0.482	0.457	0.416
	slope	0.180	0.195	0.236	0.211	0.122	0.186	0.114	0.158
University degree	coefficient	0.777	0.940	0.947	0.771	0.429	0.505	0.458	0.557
	slope	0.298	0.380	0.342	0.286	0.130	0.192	0.127	0.202
Big city (>250,000)	coefficient	-0.031	-0.033	0.003$^{n.s}$	-0.149	-0.116	.523e-03$^{n.s}$	-0.120	-0.060
	slope	-0.011	-0.013	0.001	-0.058	-0.033	.207e-03	-0.029	-0.023
Husband/partner employed	coefficient	0.071$^{n.s}$	0.059$^{n.s}$	0.054$^{n.s}$	0.308	-0.270	0.062$^{n.s}$	-0.002$^{n.s}$	0.326
	slope	0.027	0.023	0.021	0.120	-0.082	0.024	-0.0004	0.122
husband's/partner's predicted earnings	coefficient	0.128$^{n.s}$	0.033	0.042	0.063	0.029	0.034	0.050	0.018
	slope	0.049	0.013	0.016	0.025	0.008	0.013	0.012	0.007

"continued"

Table 8 : (cont.) **Marginal effects in employment equation** **(reference Table 6)**

		Native Swedes	Nordics	Westerners	Eastern Europeans	Middle Easterners	Asians	Africans	Latin Americans
Arrival Cohort :									
1970-74 (5 years)	coefficient		0.049	0.161	0.280	0.677	0.536	0.713	0.401
	slope		0.019	0.061	0.108	0.236	0.199	0.222	0.143
1975-79 (5 years)	coefficient		-0.077	0.074	0.312	0.697	0.099	0.624	0.299
	slope		-0.031	0.028	0.119	0.239	0.038	0.189	0.111
1980-84 (5 years)	coefficient		-0.194	0.054	0.510	0.646	0.293	-0.646	0.298
	slope		-0.078	0.021	0.189	0.221	0.113	-0.195	0.111
1985-89 (5 years)	coefficient		-0.646	-0.451	0.238	0.230	-0.063	-0.406	0.058
	slope		-0.261	-0.177	0.092	0.069	-0.025	-0.123	0.022
1990-94 (5 years)	coefficient		-0.676	-0.745	-0.219	-0.408	-0.452	-0.303	-0.444
	slope		-0.273	-0.290	-0.086	-0.108	-0.179	-0.091	-0.174
1995-2000 (6 years)	coefficient		-0.645	-0.698	-0.555	-0.911	-0.723	-0.327	-0.397
	slope		-0.261	-0.273	-0.217	-0.184	-0.278	-0.099	-0.156
Unemployment rate	coefficient		-0.108	-0.115	-0.153	-0.069	-0.100	-0.137	-0.104
in arrival year	slope		-0.043	-0.044	-0.060	-0.019	-0.039	-0.057	-0.041

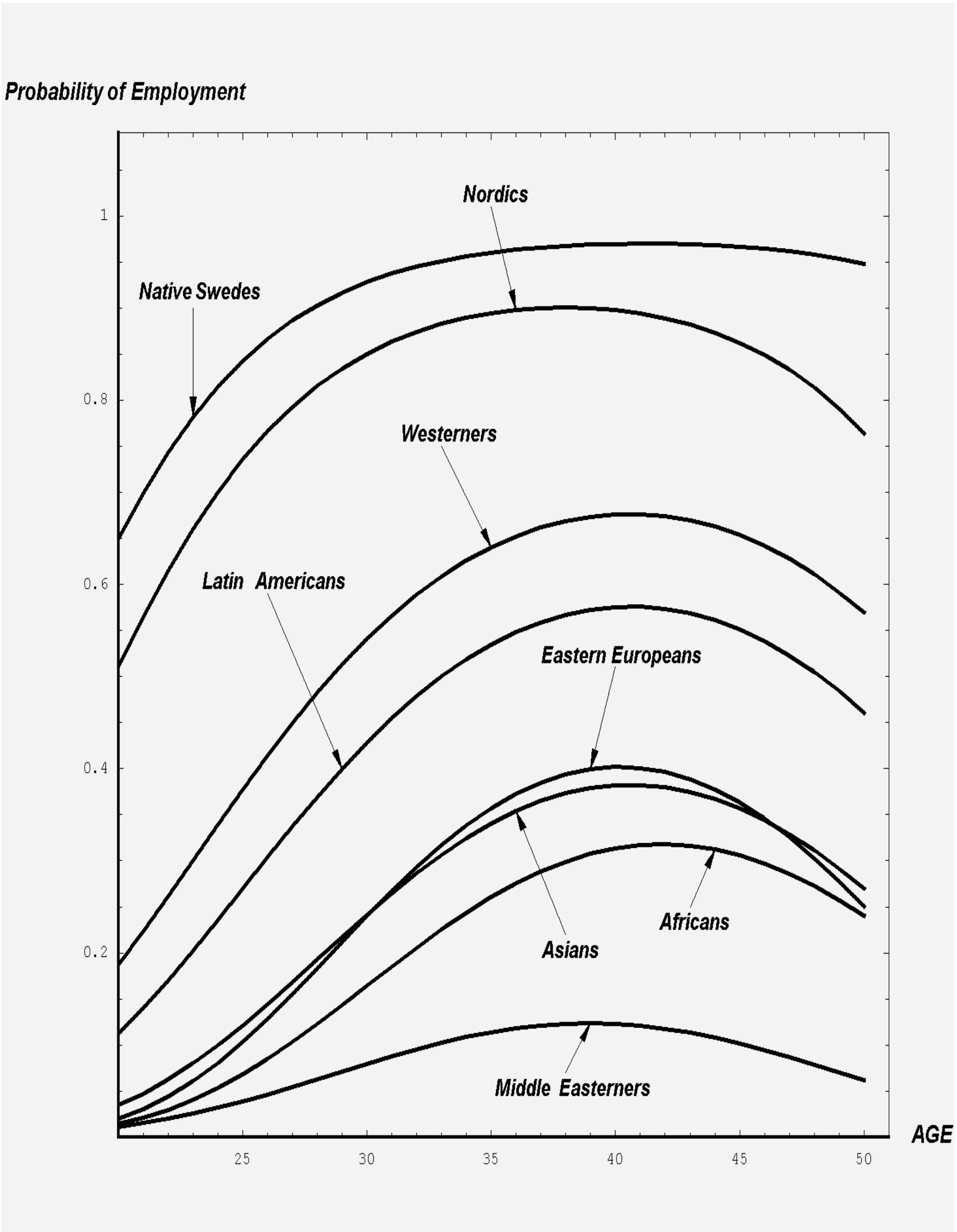
Table 9: **Marginal effects in earnings equation** **(reference Table 7)**

		Native Swedes	Nordics	Westerners	Eastern Europeans	Middle Easterners	Asians	Africans	Latin Americans
Age	direct effect	0.028	0.026	0.030	0.022	0.011	0.013	0.014	0.015
	indirect effect	0.004	0.002	0.003	0.003	0.007	0.012	0.006	0.007
	total effect	0.032	0.028	0.033	0.025	0.018	0.025	0.020	0.023
Married/cohabiting	direct effect	-0.108	-0.082	-0.089	-0.100	-0.035	-0.064	-0.026	-0.076
	indirect effect	-0.001	-0.004	-0.006	-0.095	-0.140	-0.121	-0.156	0.078
	total effect	-0.109	-0.086	-0.095	-0.195	-0.175	-0.185	0.182	-0.154
Number of children at home	direct effect	-0.073	-0.063	-0.063	-0.039	-0.030	-0.038	-0.044	-0.046
	indirect effect	-0.003	-0.006	-0.010	-0.027	-0.064	-0.056	-0.061	-0.040
	total effect	-0.076	-0.069	-0.073	-0.066	-0.094	-0.094	-0.105	-0.086
Married/coh.*no. of children	direct effect	0.008	.552e-04$^{n.s}$	-0.016	0.002$^{n.s}$	-0.100	0.003$^{n.s}$	0.004$^{n.s}$	0.013
	indirect effect	.144e-04	.825e-03	-0.006	-0.004	0.009	-0.002	-0.155	0.0002
	total effect	0.008	.770e-03	-0.022	-0.002	-0.091	0.001	-0.151	0.013
Education (highest level):									
Upper-secondary	direct effect	0.104	0.062	0.052	0.025	0.027	0.011[2]	0.016	0.017
	indirect effect	0.007	0.016	0.049	0.071	0.079	0.083	0.106	0.092
	total effect	0.111	0.078	0.102	0.096	0.106	0.094	0.112	0.109
University degree	direct effect	0.272	0.248	0.176	0.169	0.092	0.133	0.129	0.115
	indirect effect	0.015	0.034	0.074	0.083	0.105	0.103	0.084	0.117
	total effect	0.287	0.282	0.250	0.252	0.197	0.236	0.213	0.232
Big city (>250,000)	direct effect	0.034	0.024	0.028	0.022	0.014	0.004$^{n.s}$	0.007	-0.003$^{n.s}$
	indirect effect	-0.001	-0.001	0.0002	-0.025	-0.041	0.0001	-0.036	-0.013
	total effect	0.033	0.023	0.028	-0.003	-0.027	0.004	-0.029	-0.016
Husband/partner employed	direct effect	-0.097	-0.026	-0.102	-0.078	-0.071	-0.057	-0.032$^{n.s}$	-0.035$^{n.s}$
	indirect effect	0.001	-0.002	0.004	0.051	0.033	0.016	-0.0006	0.011
	total effect	-0.096	-0.028	-0.098	-0.026	-0.038	-0.041	-0.032	-0.024
husband's/partner's predicted earnings	direct effect	0.007	0.006	0.010	0.009	0.003	0.004[3]	0.002$^{n.s}$	0.003$^{n.s}$
	indirect effect	0.001	0.001	0.003	0.010	0.011	0.009	0.008	0.004
	total effect	0.008	0.007	0.013	0.019	0.014	0.013	0.010	0.007

"continued"

Table 9: (cont.) **Marginal effects in earnings equation**

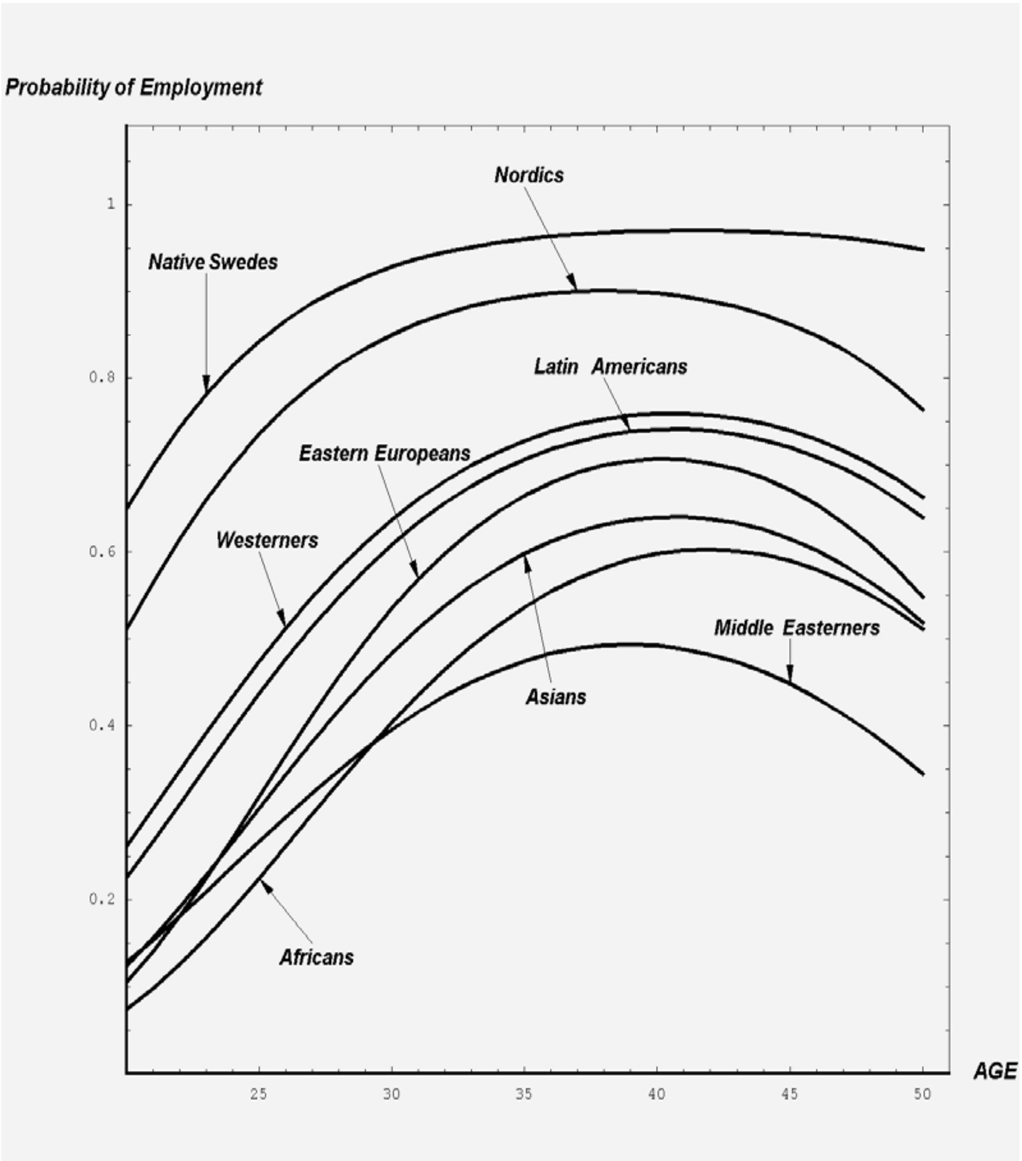
		Native Swedes	Nordics	Westerns	East-Europeans	Middle.-Easterns	Asians	Africans	Latin Americans
Arrival Cohort :									
1970-74 (5 years)	direct effect		-0.003[8]	-0.022	0.042	-0.031	0.041	0.080	0.125
	indirect effect		0.002	0.012	0.046	0.226	0.134	0.203	0.083
	total effect		-0.001	-0.010	0.088	-0.195	0.175	0.283	0.208
1975-79 (5 years)	direct effect		0.016	-0.050	$0.004^{n.s}$	-0.058	-0.013[5]	0.030	0.016[1]
	indirect effect		-0.002	0.006	0.050	0.235	0.026	0.179	0.064
	total effect		0.014	-0.044	0.054	0.176	0.013	0.209	0.080
1980-84 (5 years)	direct effect		0.035	$-0.004^{n.s}$	0.025	-0.036	-0.041	0.009	0.015[2]
	indirect effect		-0.007	0.0044	0.080	0.218	0.077	-0.202	0.064
	total effect		0.028	0.004	0.105	0.182	0.036	-0.193	0.079
1985-89 (5 years)	direct effect		0.018	-0.082	0.020	-0.045	-0.015[2]	0.005	-0.017
	indirect effect		-0.029	-0.039	0.039	0.081	-0.017	-0.125	0.013
	total effect		-0.011	-0.121	0.059	0.036	-0.032	-0.120	-0.004
1990-94 (5 years)	direct effect		-0.008[1]	-0.030	-0.050	-0.030	$-0.003^{n.s}$	0.021	$-0.002^{n.s}$
	indirect effect		-0.031	-0.067	-0.037	-0.147	-0.129	-0.092	-0.106
	total effect		-0.039	-0.097	-0.087	-0.177	-0.132	-0.071	-0.108
1995-2000 (6 years)	direct effect		0.089	0.092	-0.046	-0.048	$-0.016^{n.s}$	-0.005	-0.059
	indirect effect		-0.030	-0.063	0.102	-0.342	-0.218	-0.100	-0.096
	total effect		0.058	0.028	-0.148	-0.390	-0.234	-0.105	-0.155



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, arrival time unemployment rate=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

[Figure 1]

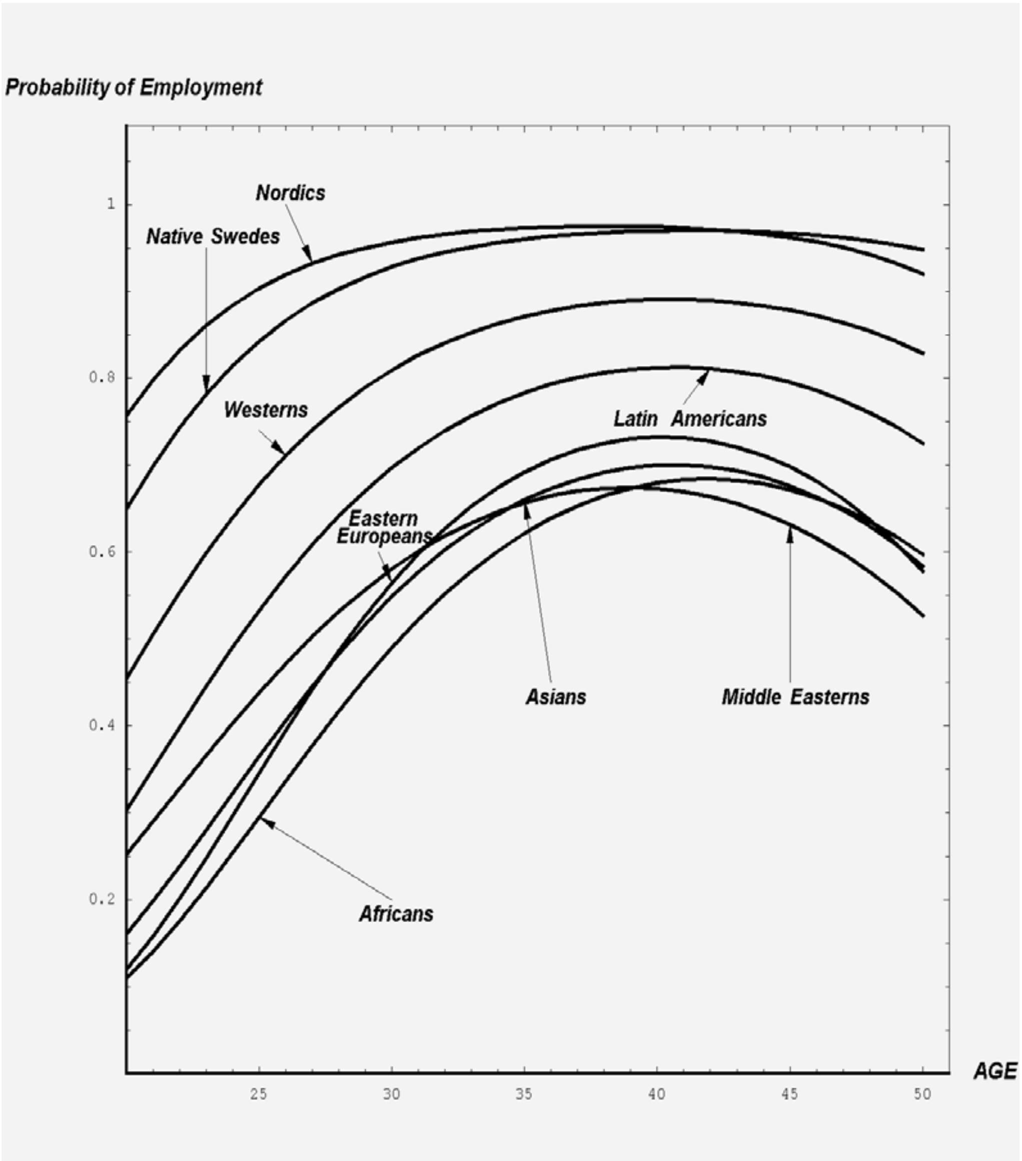
Probabilities of employment by age and and geographic origin for arrival cohort 1995-2000



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, arrival time unemployment rate=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

[Figure 2]

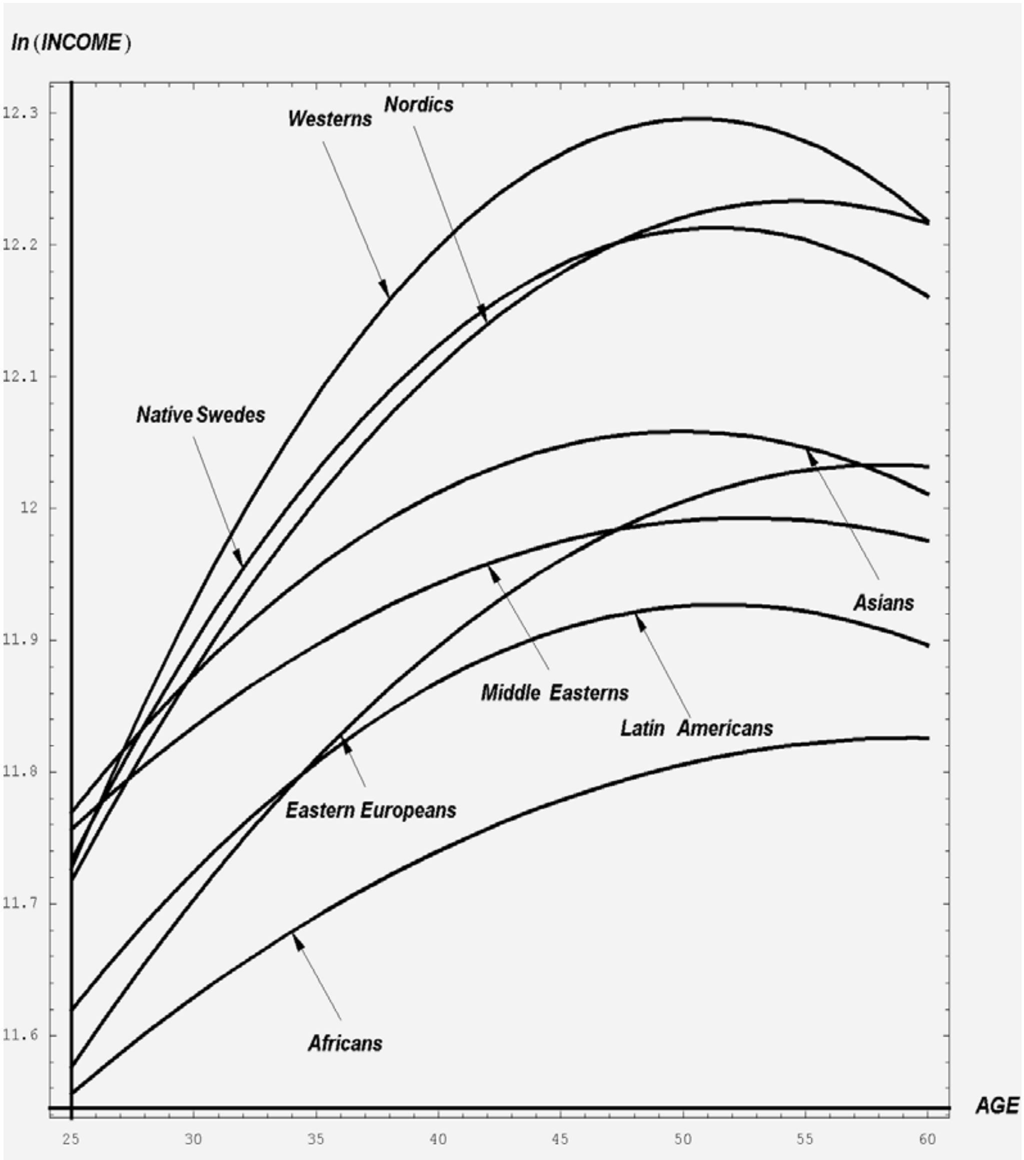
Probabilities of employment by age and and geographic origin for arrival cohort 1985-1989



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, arrival time unemployment rate=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

[Figure 3]

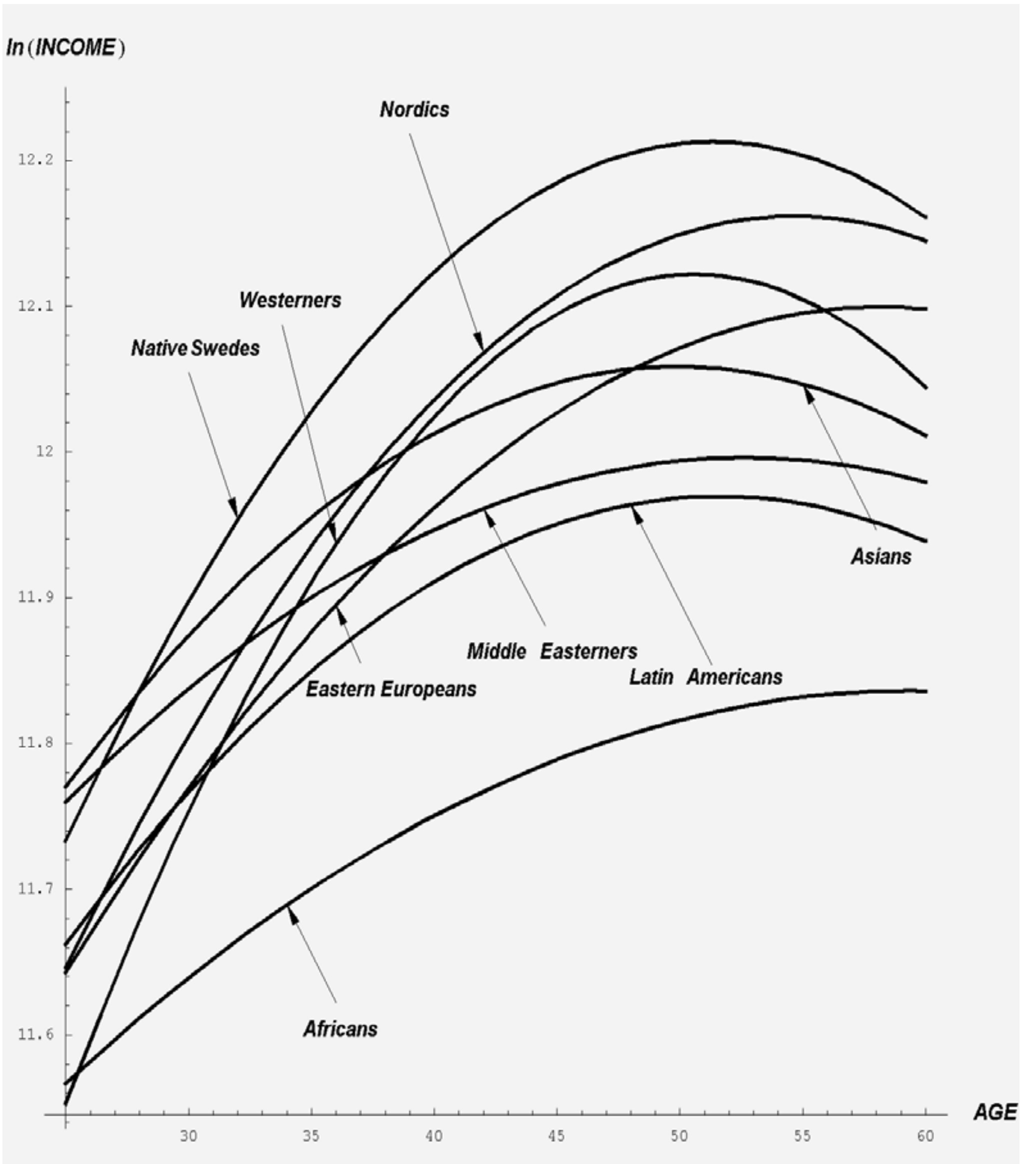
Probabilities of employment by age and and geographic origin for arrival cohort 1975-1979



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

[Figure 4]

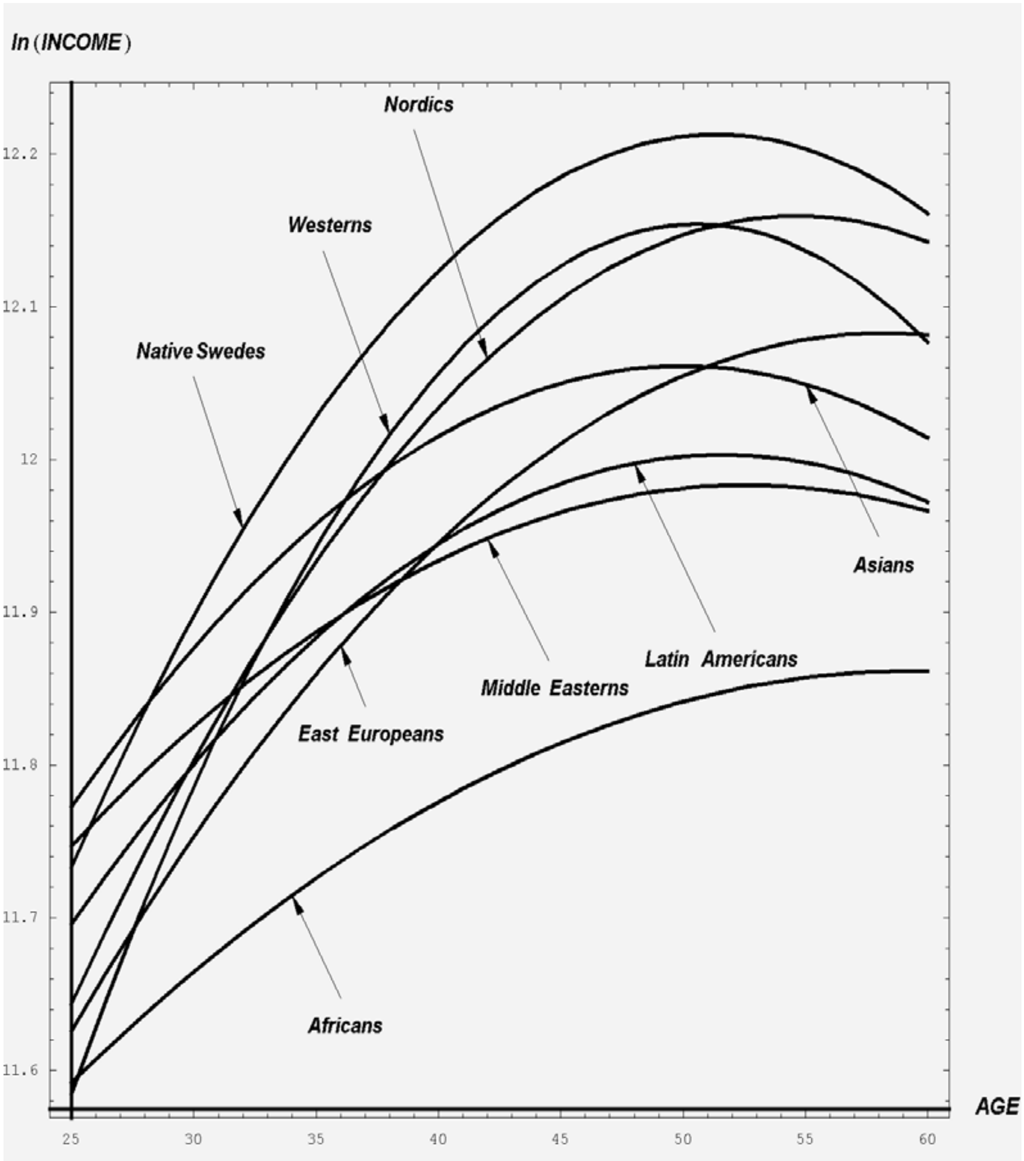
Log. of income by age and and geographic origin for arrival cohort 1995-2000



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

[Figure 5]

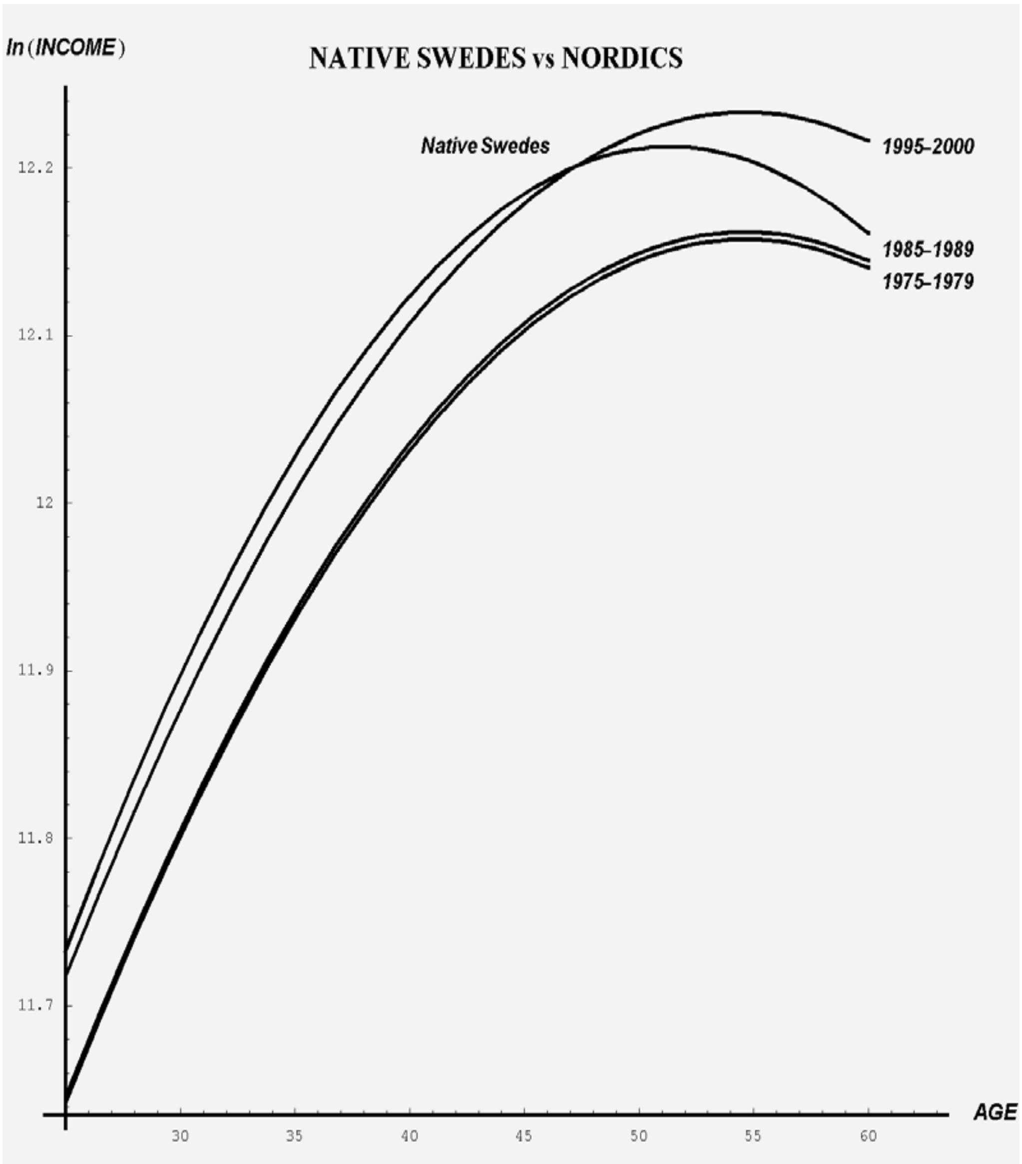
Log. of income by age and and geographic origin for arrival cohort 1985-1989



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

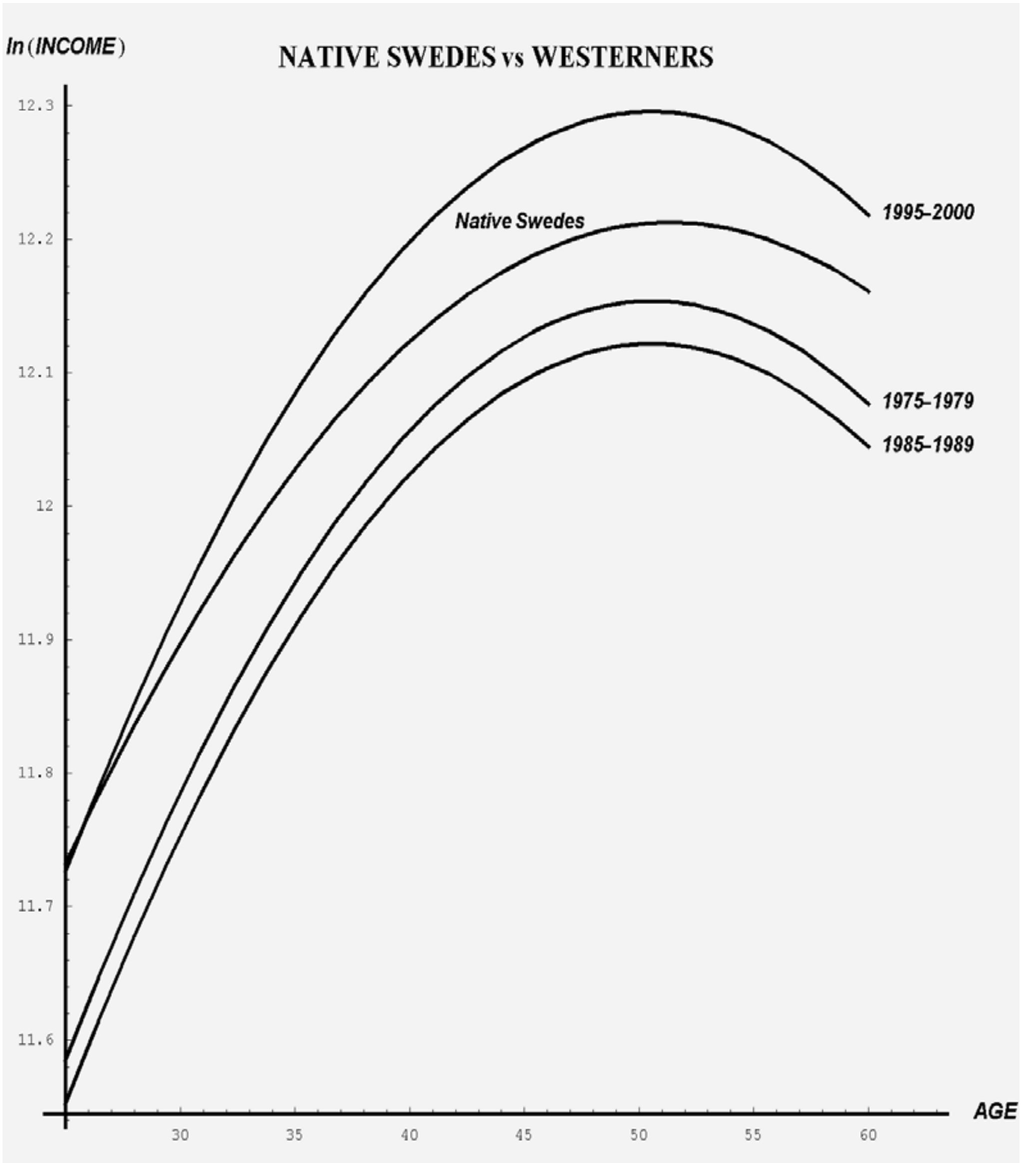
[Figure 6]

Log. of income by age and and geographic origin for arrival cohort 1975-1979



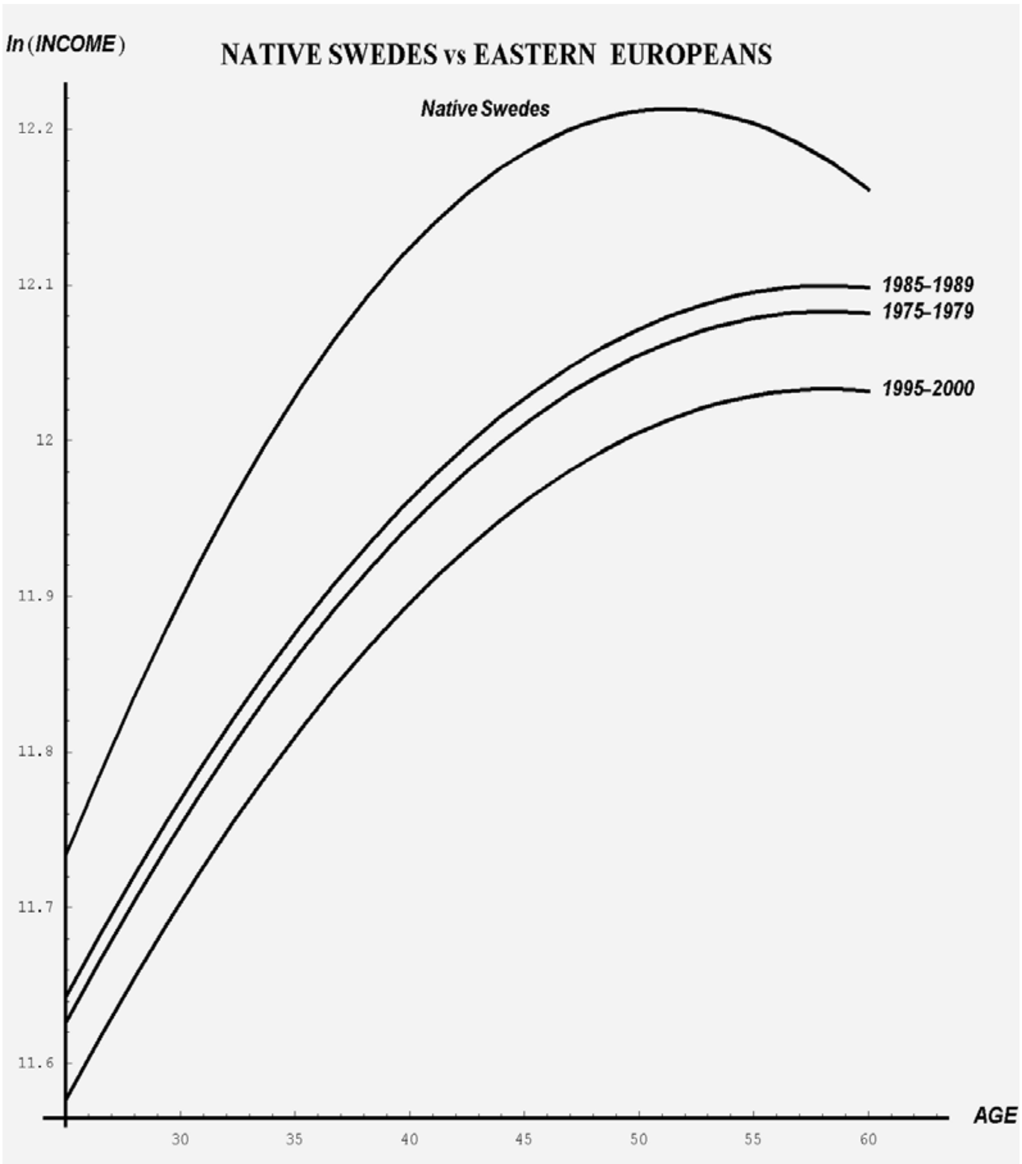
Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

[Figure 7]



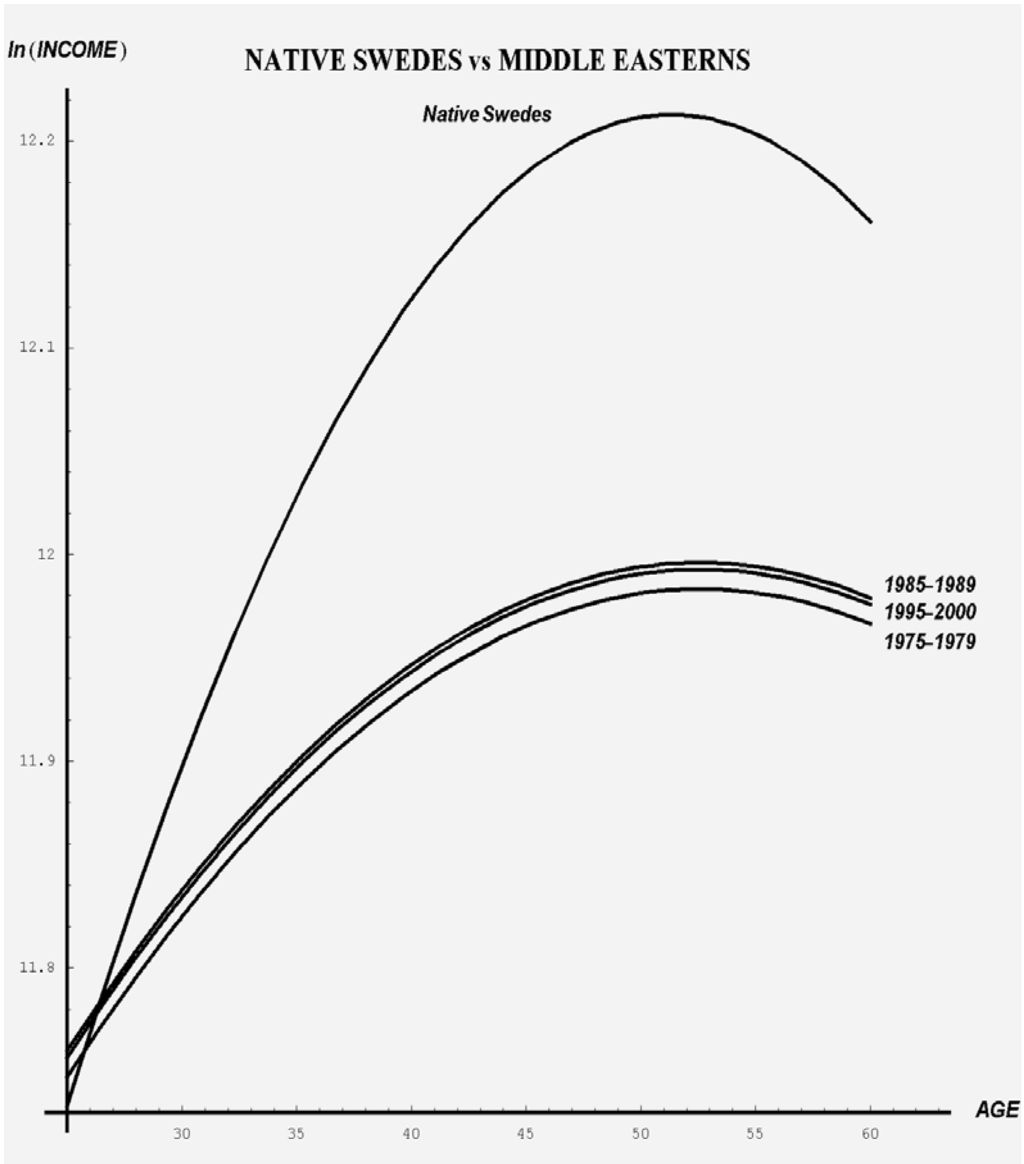
Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

[Figure 8]



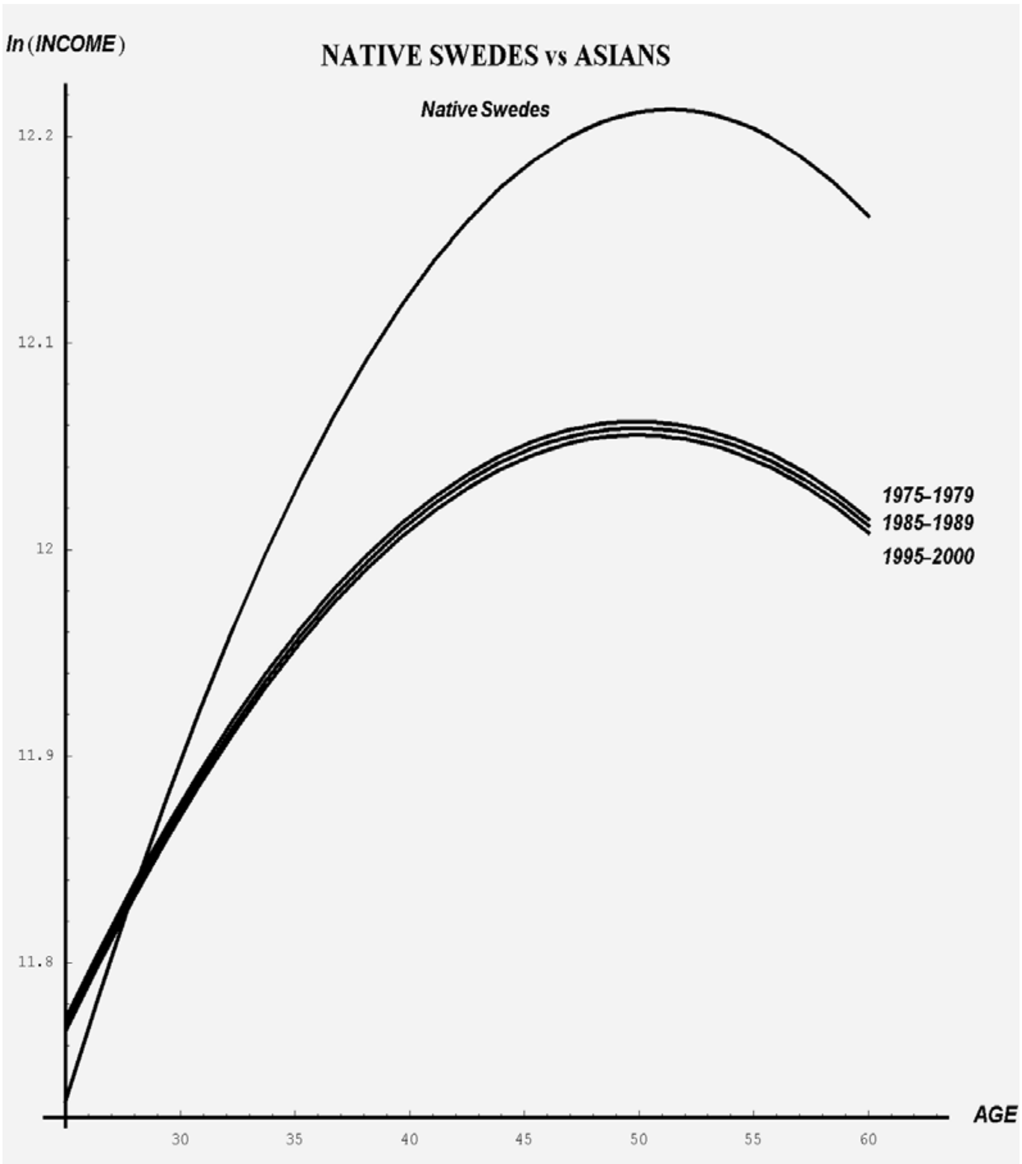
Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

[Figure 9]



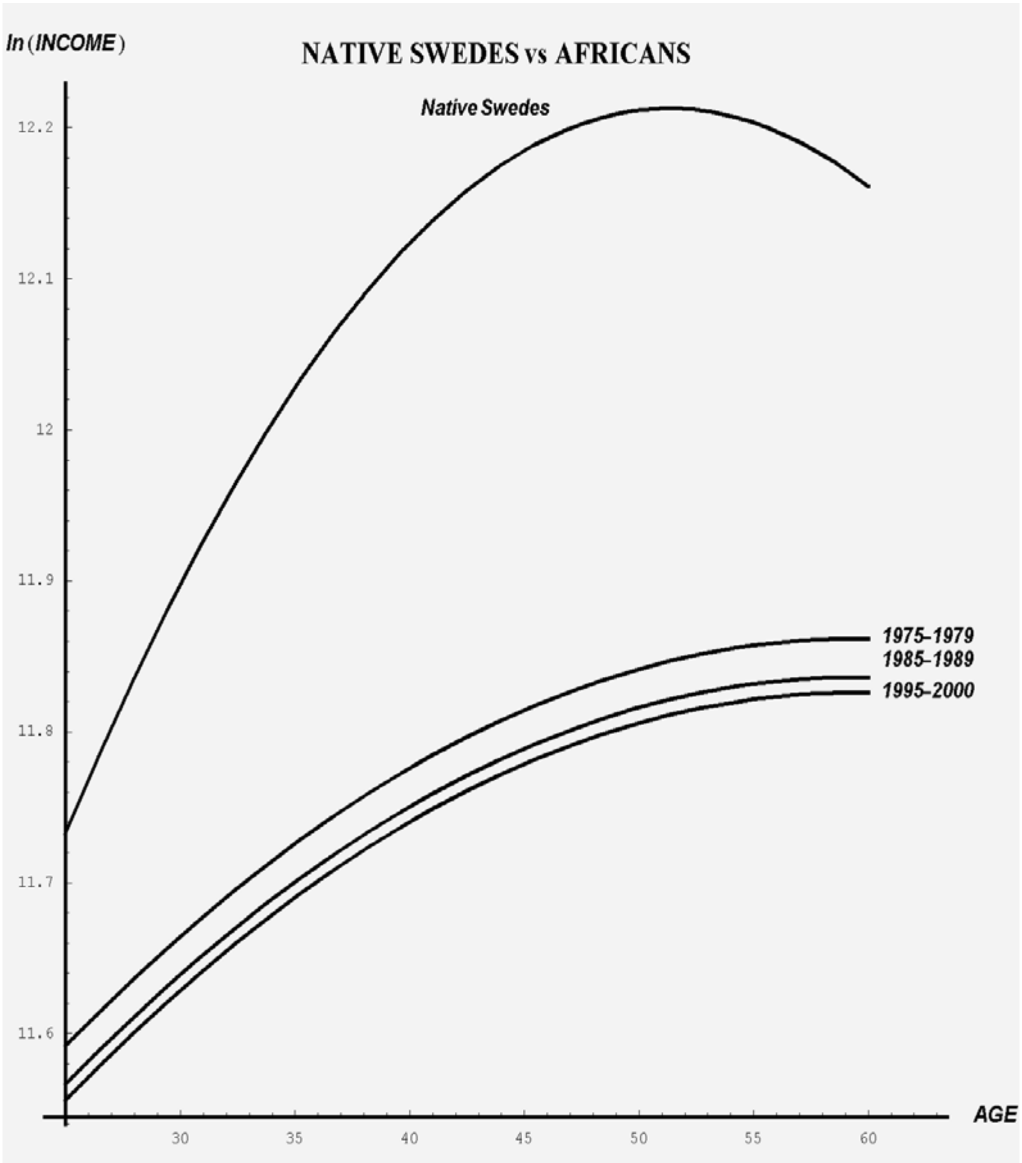
Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

[Figure 10]



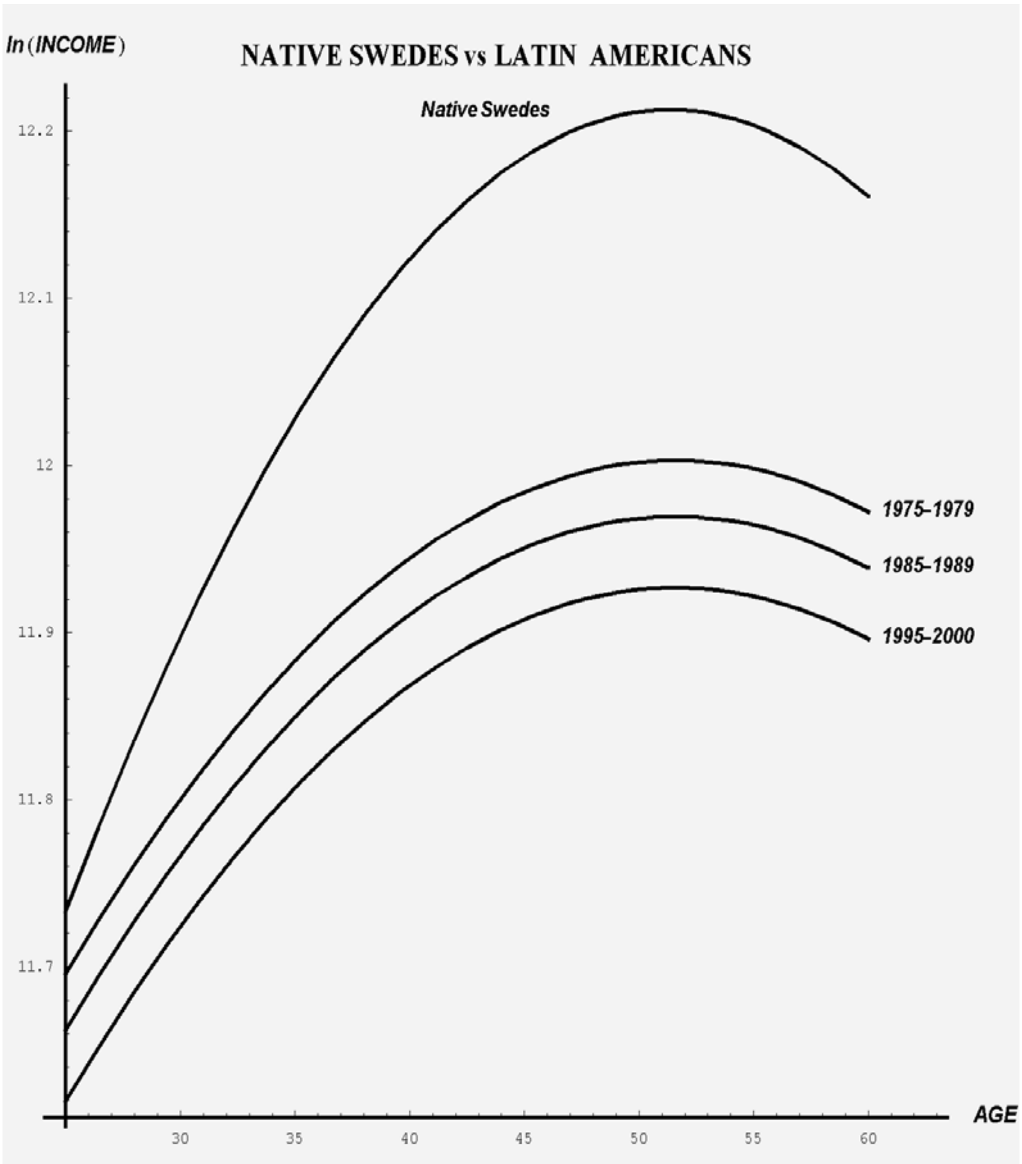
Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

[Figure 11]



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

[Figure 12]



Reference group: Married or cohabiting, having university degree, living in big cities, number of children=mean value, husband/partner working, husband's/partner's predicted earnings=mean value

[Figure 13]