FAMILY BACKGROUND AND DANISH SECONDARY EDUCATIONAL CHOICES: RESULTS FROM FIVE COHORTS

James McIntosh[*,†], **Martin D. Munk**[†]

[*]Economics Department
Concordia University
1455 De Maisonneuve Blvd. W.
Montreal Quebec, H3G 1M8, Canada
and
[†]Danish National Institute of Social Research
Herluf Trolles Gade 11
DK-1052 Copenhagen K, Denmark

June 21, 2006

E-mail Addresses and Telephone Numbers. jamesm@vax2.concordia.ca, +001 514 848 2424 Ex.3910 and mdm@sfi.dk, +45 33 69 77 10.

Abstract

The purpose of this research is to determine whether the high levels of intergenerational immobility found by McIntosh and Munk (2006) prior to 1980 have declined for younger Danes. We examine the participation in secondary education of five cohorts of Danish males and females who were aged twenty starting in 1982 and ending in 2002 by estimating non-linear probability regression models. We find that the large expansion of secondary education in this period was characterized by an expansion of gymnasiums, the most advanced category of secondary education. Most of the increase in gymnasium attendance was generated by the inclusion of individuals whose parents had low levels of educational attainment and low status occupations. Not only did the educational opportunities for individuals with disadvantaged backgrounds improve absolutely, but their relative position also improved making Denmark both a much better educated society but also a fairer one as well.

1 Introduction

The purpose of this research is to determine whether the high levels of intergenerational educational immobility found by us prior to 1980 have declined for younger Danes. We examine participation in secondary education of five cohorts of Danish males and females who were aged twenty starting in 1982 and ending in 2002. This is done by estimating non-linear probability regression models on all five cohorts for both males and females.

In a recent paper, McIntosh and Munk (2006), we found, using sample survey data from a cohort of individuals born in 1954-55, that there was a significant amount of dependence of respondent's completed educational attainments on their social and economic origins. In particular, we discovered that parent's education and occupation along with an indicator of ability which was represented by a set of intelligence test scores explained a modest but significant portion of the variation in the respondents' educational success. In terms of explaining the variation in the probabilities of attaining a particular level of education our results showed that family background variables rather than ability as measured by educational test scores were the more important. This was especially true for women who were shown to respond differently to their environments than men.

In this project we examine whether there have been any changes in this dependence using register data on five cohorts of twenty year olds. The first cohort was born in 1962 and the last in 1982. The most recent register available is for the year 2002. Danish registers are very comprehensive and contain almost all the relevant information on every individual. Everyone who was aged twenty and was born in Denmark was included in the sample. These registers contain the central population register numbers for the parents of each individual. Hence, for each cohort it is possible to assemble a data set which contains all of the required information on the individual as well as a set of variables relating to his or her family background. This was done for all five cohorts. Register data for individuals born prior to 1962 is not as comprehensive and is characterized by large numbers of missing values for parental information so 1982 is the earliest cohort.[1] The variables used in the analysis are described in section 3.

Our previous project looked at completed attainments. This was possible using the survey because respondents were asked what education they had received when they were 37, long after most respondents had completed their schooling. In order to examine the dependence of educational attainments on family background for more recent cohorts we have to look at secondary education rather than tertiary education. Danes are notoriously slow at completing their educations. It is pointless to consider completed attainments for ages less that thirty and we want to say something about individuals who enrolled in educational institutions in the 1990's. Our focus is, therefore, on secondary education and in particular our criterion is which secondary educational programme the individual had started by age twenty.

Looking at secondary education is not a substitute for analyzing final educational

[1] In what follows we refer to a cohort by the year when the respondent was aged twenty and not the year of birth.

attainments but it is an important in its own right and the choices made at the secondary level largely determine what educational opportunities are available at the tertiary level. In Denmark all students are compelled to complete grade 9; about 60% go on to grade 10. After grade 9 or 10 there are two choices: the individual can elect to enroll in a vocational programme. Welding, carpentry, or being an electrician are typical options. Vocational programmes can take quite a long time to complete and sometimes involve apprenticeships. Gymnasiums offer a broad range of academic programmes in the humanities, arts, physical and social sciences. After grade nine or ten students can enroll in these programmes which last for about three years and provide qualifications which are required for entrance to university.[2] Many programmes involving short or intermediate tertiary programmes also require a completed gymnasium certificate for entry.

To summarize our results we find, as before, that household background variables explain a significant amount of the variation in secondary educational choices for both males and females. The educational attainment of the respondent's mother as household income in which the respondent resided at age twenty turned out to be the most important variables. However, father's education, the occupations of both parents, the number of siblings that the respondent had and whether the respondent's father was unemployed or had a single mother were also significant explanatory variables. Consequently, intergenerational dependence of educational choices continues to be a prominent feature of Danish society. But the degree of dependence has lessened over the twenty years under consideration. Summary statistics which measure overall dependence on family background variables have declined by about 11% for males and 19.6% for females. Although not particularly large these are highly significant. While family background variables have become less important there have been some dramatic increases in the intercept terms of our statistical models. These capture the effects of changes in Danish educational system and social policy as well as possible changes in preferences for educational streams. They also generate higher probabilities of being enrolled in a gymnasium for respondents coming from socially disadvantaged backgrounds. Our models fit the data very well so that our findings of a decline in the importance in family background variables is consistent with the observed expansion of the gymnasium system to include a much larger proportion of respondents whose parents were poorly educated or had low status occupations.

The paper is organized in the following way. The next section reviews the relevant literature on changes educational mobility. Section 3 outlines the data used in the study. Section 4 describes a new estimation procedure which is based on a generalization of the linear probability model. The results appear in section 5 and are discussed in section 6. Section 7 ends the paper with some concluding observations.

[2]Gymnasiums correspond to grammar schools or academic streams ending in sixth form colleges granting A level qualifications in the United Kingdom or to high schools in Canada or the United States.

2 LITERATURE REVIEW

There are a large number of studies that attempt to relate individual performance as measured by educational attainment, earnings, or occupation to the characteristics of the household in which the individual grew up. These are seen as crucial in determining children's outcomes as adults, both in the educational system and in the labour market. This literature is reviewed in considerable detail in the research papers arising out of our previous project so there is no need for any further discussion other than to point out that the results that we obtained on Danish sample survey data are similar to what many other researchers had found for the countries they examined.

However, the literature *on changes* in economic, educational, and social mobility over time and across cohorts is not well developed. In addition to the relatively small number of papers and monographs which deal with the topic, much of the research suffers from serious statistical flaws. In 1998, new standards were imposed on the analysis of educational stratification. In that year Cameron and Heckman published the first of their two classic papers on educational attainment. The main contribution of their work was to expose the flaws in the estimation of the Mare (1980) transition model which has been the preferred vehicle for analyzing educational success. They showed that (1), applying the Mare model to education transitions would lead to biased parameter estimates if there was no correction for unobservables (2), it was not a valid exercise to compare coefficients across transitions or across cohorts (3), the decline in coefficients with higher transitions was an artifact of the logit specification of the transition probabilities and (4), the model was unlikely to be non-parametrically identified.

These results raise serious questions about the validity of much of the research carried out on European educational mobility the most prominent of which is Shavit and Blossfeld (1993).This is an important study and involved the analysis of thirteen countries using a common methodology. Their main result for the transition model was

> "....while the effects of students' origins decline across transitions there is little change across cohorts. There are only two exceptions to this pattern: Sweden and the Netherlands, in which the effects of father's occupation and education on the low and intermediate transitions declined."

Wanner (1999) also applied Blossfeld-Shavit methodology to a series of Canadian cohorts and reported similar findings. However, McIntosh (2006) applying the Cameron-Heckman methodology to the 2001 Canadian General Social Survey found quite different results. He discovered a considerable reduction in the importance of social background variables on educational attainment for Canadians who were born on average around 1965 compared to those born twenty-five years earlier. The period under consideration was one where there was a very large expansion of educational facilities for the younger cohorts so it appears that the expansion of the Canadian educational system made Canada a fairer society as well as a better educated one. Until these new methods developed by Cameron and Heckman are applied to the original sample of countries that were covered in the Blossfeld-Shavit study it would, perhaps, be unwise to draw any conclusions about changes in their educational mobility.

The closest study to our work is a recent paper by Dustmann (2004) who uses the German Socio-Economic Panel data base to examine the secondary school outcomes of a sequence of cohorts the first of which was born in 1925 and the last in 1965. He finds using ordered probability models that the probability of completing high school for respondents with 'working class' parents increases moderately over the ten cohorts and is higher for males than females. This is much lower than for respondents with 'academic' parents whose probabilities also increase with females overtaking males by about 1960. The large gaps between these two probabilities leads him to conclude that considerable educational immobility still exists in Germany.

For the Netherlands de Graaf and Kalmijn (2001) examined trends in occupational mobility. Their measures of occupational mobility included both an economic and what they called a cultural dimension. Average income of the occupation was used to represent the economic dimension and the cultural dimension was represented by the number of years of schooling required to acquire the occupation's credentials which they call 'occupational education'. They refer to these as economic and cultural status scales in the tradition of Blau and Duncan (1967) and found that the intergenerational dependence of the sons status scales on that of the father had virtually disappeared by 1984. This is similar to the results that de Graaf and Ganzeboom (1993) found in their contribution to the Blossfeld-Shavit volume without all of its statistical shortcomings. In fact, the methodology that de Graaf and Kalmijn used has much to recommend it. Status scale comparisons can be properly handled by ordinary least squares and these measures of occupation are much more stable over time so that problems that arise in finding good measures of permanent income that occur in income comparisons do not arise here. There is one potential flaw in their procedure and this arises because of an omitted variable which could be individual ability which contributes to an individual's success through the educational system as well as in the labour market. However, it is both unobserved and as our results, McIntosh and Munk (2006), and the results of others have shown, ability, as measured by scholastic performance, is also correlated with family background variables. Consequently, their estimated coefficients may be biased.

Hauser (1998) using the same methodology as de Graaf and Kalmijn (2001) on American data concluded that "there is no global trend in the intergenerational persistence of occupational income or education from the 1960's to the 1990's". Cameron and Heckman (1998) report the effects of several family background variables on educational attainment. However, these are mixed with the effect of household income showing a slight decline in importance towards the end of their sample period. On the other hand, their parental education variables retain their importance. For France, Vallet (2004) reports a decline over thirteen cohorts over the period 1908-1972 using log-linear models to examine changes in associations between social origin and educational destination. He also notes that "The decline in origin-education association in France therefore seems largely independent of major secondary school reforms introduced to promote equality of educational opportunity (p. 31)". For Sweden, in many ways close to Denmark, Jonsson (1993), Erikson and Jonsson (1996), Jonsson and Erikson (2000) tried to show a decline in the social inheritance effect on educational attainment, including low and intermediate transitions. Erikson and Goldthorpe (1992) found that intergenerational occupational mobility increased among the youngest cohorts in Sweden (see also Esping-Andersen 2004). Bynner and Joshi (2002) examined sample survey data from the 1958 and 1970 cohorts in Britain. They found no change in the response of the probability of leaving school at age sixteen to family or social origin variables. Marks and McMillan (2003) found a decline in the dependence of educational attainment on social background variables for Australia for cohorts born during the period 1961-1985.

Blanden and Gregg (2004) found an increased dependence of tertiary educational attainments on household income over thew period 1958 to 1970 using the British National Child Development and British Cohort Surveys. Individuals were aged 33 and 30, respectively. We also find an increase in income dependence but it is dominated by other factors.

In summary, results differ by country and sometimes by type of procedure employed. Perhaps the most striking feature of the research on the evolution of intergenerational mobility is the almost uniform neglect of unobservable characteristics. The costs of this omission have emphasized by Cameron and Heckman for the Mare transition model but the simple fact remains that all probability and log-linear modelling procedures produce biased parameter estimates if there is no procedure which deals with unobservable effects.

3 DATA AND VARIABLES

For the dependent variable our choice is the three category variable: choice of secondary education at age twenty. The three categories are nothing past grade nine or ten, a vocational educational programme, or gymnasium. An individual is in a particular category if he or she had ever been enrolled in the programme or had completed it. The numbers and percentage allocations for all five cohorts are displayed in Table 1. There are number of variables for parental characteristics. Parents education is a five category variable where the first category is no education past grade nine or ten. The second category is a vocational qualification and there are three categories of tertiary education which in Denmark are characterized by their durations: short, medium and long. Examples for the three types are police training, primary school teacher training and university, respectively. The residual category is no education past grade nine or ten. There are eight parental occupations; the first three are white collar occupations starting with high level managerial, low level managerial and ordinary employee. Occupations four, five, and six are self-employed and skilled and unskilled blue collar workers and occupation seven is the occupation missing category. For the first cohort there are many parents whose occupations are not known and it does not seem appropriate to combine them with the unemployed so they are represented a separate category for all of the cohorts although there is very little missing parental information for the last cohort. The residual category consists of those who are unemployed or not in the labour force.

The data set also contains the number of siblings, whether the father was unemployed, whether the respondent's mother was a single mother, and the standardized (mean zero unit variance) household income, all collected when the respondent was twenty.

4 Estimation methods

A natural way to estimate these choices is to estimate an ordered or a mixed ordered probability model. This reflects the theoretical model outlined in our first paper and these choices while not being sequential can certainly be ordered in terms of their academic difficulty. We applied these procedures to the data but the results were unsatisfactory in several respects.

Ordered probability models were first estimated and these were compared to mixed models using the Heckman-Singer (1984) procedure. At most two mixtures could ever be estimated and the probability of the second type was always very small indicating that unobserved heterogeneity played a very small role in the determination of secondary educational choices. The more general latent class models that were employed in our first paper always failed to converge. However, these models were dominated in terms of the Akaike criterion by multinomial logit models. These suggest that the respondents choose different outcomes because they respond differently to their family backgrounds and not because of differences in unobservable factors. A plausible outcome, perhaps, but the conclusion that unobservables are unimportant is hard to accept especially since there are at least two different types within the gymnasium category: those who will not obtain any further education and those who will go on to some form of tertiary education. The problematic nature of these models is confirmed by their inability to fit the data. Both the ordered probability and multinomial logit models have low psuedo-$R^{2's}$ and are simply unable to predict the actual categorical proportions.

Clearly, a different approach is needed. The superiority of the logit models over the ordered models suggests that single index models are inadequate to explain the data and multiple index models are required[3]. We also want to combine this aspect of the model with the idea that the choices are ordered in terms of their difficulty. There are three possible outcomes that the respondent can choose: no education past grade nine or ten, a vocational programme, or the gymnasium programme. We represent these choices by three dummy variables (y_1, y_2, y_3) each of which takes on the values 1 and 0 so that $y_3 = 1$ if the respondent decides to go the gymnasium, for example. We specify the probabilities of these events, which are the expectations of these dummy variables, as

$$\begin{aligned} E(y_1) &= F_I(-X\varphi\ ) && (1) \\ E(y_2) &= [1 - F_I(-X\varphi\ )]F_{II}(-X\gamma\ ) && (2) \\ E(y_3) &= [1 - F_I(-X\varphi\ )][1 - F_{II}(-X\gamma\ )] && (3) \end{aligned}$$

where $F_i() \in [0,1] \quad i = I, II$.

These probabilities, which by construction sum to unity, are generated by two latent attribute variables A_I^* and A_{II}^*. These attributes are assumed to depend on family background variables but measure two qualitatively different individual characteristics and are defined as

$$\begin{aligned} A_I^* &= X\varphi\ + v_I && (4) \\ A_{II}^* &= X\gamma\ + v_{II} && (5) \end{aligned}$$

where X is a vector of family background variables and (v_I, v_{II}) are independent identically distributed zero mean random disturbance terms[4]. We interpret A_I^* as a variable representing the motivation or enthusiasm of the respondent and v_I is its unobserved component. A_{II}^* represents some general measure of ability with v_{II} as its unobserved component. If $A_I^* \leq 0$ then the respondent will choose not to be involved in any further education. The probability that this occurs is $F_I(-X\varphi\ )$ where $F_I()$ is the cumulative

[3]By an index for individual i we mean $X_i\varphi = \sum_{j=0}^{J} x_{ij}\varphi_j$, for example. The ordered probit model is a single index model whereas the multinomial logit model for three unordered categories has two indexes.

[4]It is possible to allow (v_I, v_{II}) to be jointly distributed. In this case the probability of attending gymnasium is $\int_{-X\varphi}^{\infty}\int_{-X\gamma}^{\infty} h(v_I, v_{II})dv_I dv_{II}$ where $h(v_I, v_{II})$ is the joint distribution of (v_I, v_{II}). We have not explored this alternative yet.

distribution function of the random variable v_I. If $A_I^* > 0$ then the respondent will enter a vocational or an academic programme. This decision is determined by the second attribute variable. If $A_{II}^* \leq 0$ the respondent selects a vocational education; otherwise an academic stream is selected.

Once these probabilities have been specified a probability model can be estimated. Alternatively, if additional unobservable effects need to be accommodated, a non-linear regression model of the form

$$y_1 = F_I(-X\varphi) + u_1 \tag{6}$$

$$y_2 = [1 - F_I(-X\varphi)]F_{II}(-X\gamma) + u_2 \tag{7}$$

$$y_3 = [1 - F_I(-X\varphi)][1 - F_{II}(-X\gamma)] + u_3 \tag{8}$$

can be estimated. (u_1, u_2, u_3) represent unobservable characteristics of the respondent's family background or possible unobservable external effects. Our preference is for the non-linear regression specification since it permits a more general random effects specification. It, unlike the probability models or mixed probability models, gives very good predictions of the category proportions.

This procedure is a generalization of the linear probability model in two directions. In the binary case the linear probability model is a regression of a dichotomous variable y on the vector X. The first generalization is to use a cumulative distribution function, $F(X\varphi)$, for the right hand side which makes all of the fitted values lie in the unit interval. The second generalization is to extend the model to more than two alternatives. This is done by making sure that the probabilities of the three possibilities are non-negative and sum to unity. The estimation of the model is carried out by estimating any two of the three equations. The jacobian of the model is singular if all three equations are estimated.

All that is left is the choice of functional form for $F_I()$and $F_{II}()$ and the estimation procedure. Several distribution functions were tried all with the same qualitative results. The results in the next section are based on the cumulative normal distribution function. The estimation procedure is the non-linear seemingly unrelated regression model which makes no assumptions about the correlation structure of (u_1, u_2, u_3). Robust standard errors are used to correct for possible heteroscedasticity due to the binary nature of the dependent variables.

5 Results

The results for the model outlined in section 3 are shown in Tables 2M and 2F for males and females, respectively for the 1982, 1987, 1992, 1997, and 2002 cohorts. The first point to note is that $||\varphi||$, $||\gamma||$, and $||(\varphi, \gamma)||$ are significantly different from zero indicating the expected dependence of educational choices on family background variables. These are the norms of the parameter vectors[5]. Using the norms of the parameter vectors in the probability functions is a simple way of summarizing the overall effect of family background on educational decisions. There are five cohorts for each gender and sixty parameters to be estimated for each model so some summary measures of the results are required. Using the norm of the parameter vector is a simple and appealing measure since all variables have equal weight and the sign of the parameter does not matter.

The significance of these norms means that family background variables have a significant impact on the educational choices of twenty year old Danish men and women. While the total effect is important individual variables are also important. To capture the effects of education we calculate the averages of the φ and γ parameters for the educational and occupational dummies for both fathers and mothers. These are shown as $(\overline{\varphi}_{fe}, \overline{\varphi}_{me})$ and $(\overline{\gamma}_{fe}, \overline{\gamma}_{me})$ in Tables 2M and 2F. Averages of father's and mother's occupations are also included in these tables[6].

In addition to the sets of dummy variables relating to education and occupation there are some individual variables which are important and have captured the attention of other researchers. These are the number of the respondent's siblings, the standardized income of the parental household, whether the respondent's mother was a single parent, and whether the father was unemployed, all of which were collected when the respondent was twenty years old.

The major result of our new study is the significant decline in the importance of family background variables as determinants of secondary educational choices for both males and females. The declines in $||(\varphi, \gamma)||$ of 11% for males and 19.6% for females, respectively for the period 1982-2002 (when all respondents were twenty), although not particularly large, are significant at the 1% level. There are significant declines in the individual norms, $||\varphi||$ and $||\gamma||$, of 30.5% and 7.0% for males and 20.9% and 21.2% for females, respectively. There were also significant declines in $||\varphi||$ for the 2002 cohort relative to cohorts born earlier.

Although there is a significant decline in the aggregate measures of dependence some of the variables which define the family background of the respondent actually increase in importance. The pattern varies by both gender and the type of parameter. For males the parameters which increase significantly are $\varphi_{hi}$and$\overline{\gamma}_{mo}$; those significantly decreasing

[5] For example, $||(\varphi, \gamma)|| = \sqrt{\sum \varphi_i^2 + \sum \gamma_i^2}$. The intercept is not included in these norms.

[6] For males in 1982 the φ coefficients and their standard errors associated with father's education are 0.274** (0.019), 0.200 (0.146), 0.359** (0.061) 0.359** (0.040), and 0.313** (0.057). The corresponding γ coefficients are 0.080** (0.025), 0.721** (0.152), 0.231** (0.060), 0.472** (0.040), and 0.920* (0.064). These are typical of the father's education parameter estimates for other cohorts.

are $\overline{\varphi}_{fe}, \overline{\varphi}_{fo}, \overline{\varphi}_{me}, \varphi_{fu}, \varphi_{sib}, \varphi_{sm}, \overline{\gamma}_{fe}$, and γ_{sib}. For females the parameters which increase significantly are φ_{hi} , φ_{sm} and γ_{hi}; those significantly decreasing are $\overline{\varphi}_{fo}, \overline{\varphi}_{me}$, and $\overline{\gamma}_{me}$. In terms of the relative importance of the household background variables the educational attainment of the respondent's mother ranks as the most important variable. The estimated parameter values are larger for this variable than any other. This variable exhibits the largest change as a determinant of both decisions for both genders. The second most important variable is household income. Like mother's education it affects both decisions for both genders and becomes more important over the five cohorts.

Like our previous project there are large differences between the two genders. There are some significant differences in the parameter estimates. An example is the set of parameters associated with mother's education. These are more important for females than for males. In contrast to father's occupation which was the most important variable in our first study, parental education and household income are the most important variables when register data for the whole population is used.

According to our model these declines in the family background parameters should contribute to a reduction in the probabilities of enrolling in either of the two educational alternatives and given that an educational alternative was chosen it should reduce the probability of attending a gymnasium programme. While these estimated coefficients for the family background variables have declined the estimated intercept terms have increased significantly over the period. For females especially, the change in γ_0 between the 1982 and 2002 cohorts is three times larger than the changes in any of the other parameters. Because the changes in the family background parameters are dominated by the changes in the intercept terms our model predicts that enrollment probabilities in gymnasium should increase for all respondents but the largest increases should come respondents with the most disadvantaged family backgrounds. The reason for this is because respondents with parents with low levels of education or low status occupations have relatively low levels of $X\varphi$ or $X\gamma$ since their (φ, γ) coefficients are much smaller than those with well educated or high status occupation parents. Since the probably function is highly non-linear increases in $X\varphi$ or $X\gamma$ from low levels will have a much larger impact on the probability of attending gymnasium than would be the case for those respondents with parents with high levels of education or high status occupations.

This is in part what actually happens and can be seen by looking at figures 1 and 2. The averages of the actual values of the probabilities of being enrolled in a gymnasium programme as a function of parental educational achievement are shown in Table 3 for both males and females. For both males and females mother's education is used since this is the most important family background variable. It also declines significantly in importance over the period. In figures 1 and 2 these probabilities[7] are plotted against

[7] These come from columns 3 and 6 of Table 3.

the parental educational category. In each part of Table 3 there are two entries: the first is the actual proportion and the second is the proportion predicted by the model. The averages of the probabilities in each category are always very close to the proportion of the gymnasium attendees in the category indicating that the model does a very good job of fitting the data.

It is easy to see that from Table 1 that over the period 1982 to 2002 much higher proportions of Danish adolescents enrolled in gymnasium programmes. However, the biggest percentage changes in enrollment rates came from respondents whose parents were poorly educated or had low status occupations. For example, the estimated probability of a male attending gymnasium given that his mother had no education past grade nine of ten rose from 0.167 in 1982 to 0.0.255 in 2002 which is a 52.8% increase. These percentage increases are shown by the thick graph in each figure. However, it is also important point to note that the majority of those attending gymnasium in both 1982 and 2002 came from households from the bottom of the maternal education distributions. The bar graphs in these two figures give the distribution of males and females by parental educational category. In 1982, for example, 38.6% of the twenty year old male population had mothers with no education past grade nine or ten.

While the expansion of gymnasium attendance rates has been facilitated by extending gymnasium education to respondents with poorly educated or low status occupation parents, probabilities of attending gymnasium have increased only marginally or actually decreased for both males with better educated mothers. Gymnasium participation rates for these parental categories were already quite high so the percentage increases here could not be very large. However, mothers of males in the 2002 cohort who had gymnasium only or lower tertiary education actually sent fewer of their children to gymnasium than the 1982 cohort.

Given that access to gymnasium has become much wider has Denmark become a fairer society in terms of access to secondary education? The answer to this question is emphatically yes! The predicted odds of going to gymnasium for a male respondent whose mother had no education at all compared to a male respondent whose father had a university education are $\frac{0.167}{0.809} = 0.206$ for the 1982 cohort. The same figure for the 2002 cohort is $\frac{0.255}{0.856} = 0.298$ 0.484 so that the odds for attending gymnasium for an individual whose father had no education at all have improved dramatically over the period. Similar results hold for females and for other educational category comparisons.

Results, like these based on Table 3, which is a mobility table of the type commonly used by sociologists, could be potentially misleading. The analysis is bivariate in nature and neglects the effects of other factors. So we follow Dustmann (2004: 220) and define a disadvantaged category which has parents with no education past grade nine or ten and being in the lowest occupational group. This is similar to what he had in mind in his definition of 'working class' parents. The predicted probabilities of going to gymnasium for males in this group increased from 0.088 in 1982 to 0.136 in 2002. For females the increase was much bigger: from 0.172 to 0.286.

6 Discussion and Conclusions

It is clear from the results in the previous section that the gymnasium educational stream has become the preferred option of most twenty year olds in Denmark. Of course, this could not have happened had it not have been the policy to expand the gymnasium system but making more gymnasium places available does not explain why these were occupied in such large numbers by students with socio-economic backgrounds that traditionally not attended such institutions.

As we mentioned in the previous section the increase in the intercept terms represented increased effects of economic and social policy or an increase in the 'taste' for more secondary education. The difficulty is determining which of these processes are at work. There are, however, some clues in Tables 2M and 2F which provide a profile of the changes of the parameter profiles over the five cohorts.

While there has been a decline in the norms of the parameter vectors associated with the respondent's family background variables, most of the decline is of recent origin: 1992-2002 for males and 1997-2002 for females. The 1990's was not a period of radical change in Danish social or educational policy although there were some changes that could have affected educational decisions. By 1975, tracking in primary education had come to an end so that allocating students to an academic or vocational track prior to age fifteen or sixteen was no longer being done[8]. However, changes to the gymnasium system that gave students more choice and introduced vocational options may have made the gymnasium choice more suitable for members of the cohort born in 1982. Social programmes including welfare support and unemployment insurance programmes had been established much earlier in the 1960's and the 1970's and but there were some changes during this period. Reduced entitlements to welfare programmes and tying benefits to schooling decisions made the costs of not getting more education much higher (Munk 1999).

These policies no doubt, had some impact but they do not represent a major change to the system. As a result, we are forced to conclude that it was probably more a change in attitudes or perceptions about the value of going to a gymnasium in terms of the options it gave to attendees for acquiring tertiary education. But this is not to say that the cumulative effects of social policy have not been important for making the

[8]Dustmann(2004: 226) sugests that German educational immobility may be due to "the relatively young age at which the secondary track decision has to be taken".

gymnasium alternative possible for socially disadvantaged groups. Looking at two such subgroups for males: those with household incomes one standard deviation below average and those with a father unemployed when the respondent was twenty, we find that the probabilities of attending gymnasium for both of these subgroups have doubled over the period 1982-2002.

One of the more important variables which matter in educational choices is household income. Like Blanden and Gregg (2004) we also find highly significant parameter estimated associated with this variable. This is consistent with our results using sample survey data. In Denmark, like Britain, the effect of household income on choices is very dramatic. The probability of attending gymnasium for males coming from households with incomes one standard deviation above average is more that four times as high compared to respondents coming from households with incomes one standard deviation below average. In our previous paper we interpreted high household incomes as proxies for parental competence rather than something which eases credit constraints since secondary education is free in Denmark. However, it is possible that higher income households are able to provide more of the things for their children that matter in the human capital accumulation process like access to personal computers, reading materials in the home, high quality day-care etc.

However, it should not be forgotten that in spite of this dramatic increase in the importance of household income; when all factors are considered dependence of secondary educational choices on family background variables has actually declined over the period. It would be interesting to see whether this is what has happened in Britain.

Cameron and Heckman (1998: 281) caution researchers about making comparisons of normal ordered probability models across cohorts. When normal probability models are being used the estimated coefficients are of the form β/σ so that changes in the β coefficients could be mistaken for changes in σ. We certainly use normal cumulative distribution functions to define our probabilities. As a check we also used the functions $\sin(X\phi)^2$ and $\sin(X\gamma)^2$. These generated results which were qualitatively similar to those reported in Tables 2M and 2F and there is no problem about comparing the (φ, γ) across cohorts. Moreover, there were no significant differences in the variances of (u_1, u_2, u_3) over the five cohorts. This suggest to us that it is unlikely that there would be significant differences in the variances of (v_I, v_{II}).

One of the more surprising outcomes of this study was the difference in the methods and the results they produced compared with those arising from the use of survey data. Because of the much larger sample sizes we had expected that mixtures of ordered probability models would adequately explain the data. This turned out not to be the case. The superiority of logit models over ordered probability models indicated that it was observables differences in responses to family background variables rather than unobservable effects that could be recovered by estimating mixtures of ordered models. Our probability regression models do support the presence of unobserved attributes or random effects, but these suggest that these unobservables are not well represented by mixtures of distributions with a finite number of support points.

That the rankings of the family background variables in terms of how important they were should depend on the type of data used is worrisome. The survey data was much richer in terms of providing detailed information on the individual respondent's characteristics while the registers provided much better information on the characteristics of the parents. Both types of data have their strengths and weaknesses so it may not be so surprising that different results are obtained depending on the data set used.

Finally, although we have commented on some of the differences across the two genders there is one amazing feature of these that needs to be mentioned. In 1982 25.3% of twenty year old males were attending gymnasium. For females, this figure is 37.9% (recall Table 1). By 2002 60.1% of twenty year old females were attending gymnasium but only 41.8% of twenty year old males were attending gymnasium. In other words the males had just barely been able to pass the levels that females had achieved twenty years earlier! More recent data shows that even more females are going into the gymnasium stream. If these trends continue then Denmark is likely to become Europe's first modern matriarchal society.

References

[1] Blanden, Jo and Paul Gregg (2004). "Family Income And Educational Attainment: A Review Of Approaches And Evidence For Britain" *Oxford Review of Economic Policy* 20 245-263

[2] Blau, P. and O. D. Duncan 1967. *The American Occupational Structure*, New York: John Wiley

[3] Bynner, J. and Joshi, H. 2002. Equality and Opportunity in Education: evidence from the 1958 and 1970 birth cohort studies. *Oxford Review of Education* 28: 405-427

[4] Cameron, S. V. and J. J. Heckman 1998. Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males. *Journal of Political Economy* 106: 262-333

[5] Dustmann, Christian. 2004 "Parental background, secondary school track choice, and wages" *Oxford Economic Papers* 56 209-239

[6] Erikson, R. and J.H. Goldthorpe 1992. *The Constant Flux*, Oxford: Clarendon Press

[7] Erikson, R. and J. O. Jonsson. 1996. *Can Education be Equalised? The Swedish Case in Comparative Perspective*, Boulder: WestviewPress

[8] Esping-Andersen, G. 2004. What Might Create More Equal Opportunity? Money, Cultural Capital, and Government, In M. Corak, *Intergenerational Mobility*, Cambridge: Cambridge University Press

[9] de Graaf, P. M. and H. B. G. Ganzeboom 1993. Family Background and Educational Attainment in the Netherlands for the 1891-1960 Cohorts, In Shavit and Blossfeld 1993, 75-99

[10] de Graaf, P. M. and M. Kalmijn 2001. Trends in the Intergenerational Transmission of Cultural and Economic Status, *Acta Sociologica*, 44, 51-66

[11] Hauser, R. M. 1998. Intergenerational Economic mobility in the United States: Measures Differentials and Trends. Working Paper, Department of Sociology, Center for Demography and Ecology, University of Wisconsin-Madison

[12] Heckman, J.J. and L. Lochner 2000. Rethinking education and training policy: Understanding the sources of skill formation in a modern society. In S. Danziger and J. Waldfogel, *Securing the Future. Investing in Children from Birth to College*. New York: Russell Sage Foundation: 47-83

[13] Jonsson, J.O. and R. Erikson 2000. Understanding educational inequality: The Swedish experience, *L'Année sociologique* 50: 342-382

[14] Lauer, C. 2003. Family Background, Cohort, and Education: A French-German Comparison Based on a Multivariate Ordered Probit Model of Educational Attainment. *Labour Economics* 10: 231-251

[15] Mare, R. D. 1980. Social Background and School Continuation Decisions. *Journal of the American Statistical Association* 75: 295-305

[16] Marks, G. and J. McMillan 2003. Declining inequality? The changing impact of socio-economic background and ability on education in Australia. *British Journal of Sociology* 54: 453-471

[17] McIntosh, J. and Munk, M. D. 2006. Scholastic Ability vs. Family Background in Educational Success: Evidence from Danish Sample Survey Data, Journal of Population Economics

[18] McIntosh, J. 2004. Educational mobility in Canada: Results from the 2001 General Social Survey, Working Paper, Concordia University and SFI

[19] Munk, Martin D. 1999. Social selektion i uddannelsessystemet, Report to Labour Office.

[20] Munk, M.D. 2001. The Same Old Story? Reconversions of Educational Capital in the Welfare State, *Jahrbuch für Bildung und Arbeit* 1999/2000, Deregulierung der Arbeit - Pluralisierung der Bildung? Eds. A. Bolder, W. R. Heinz, G. Kutscha, Opladen: Leske + Budrich Verlag: 87-98 + litt

[21] Munk, M. D. 2002. Institutionalised Legitimate Informational Capital in the Welfare State. Has Policy Failed? In C. Torres, *The International Handbook on the Sociology of Education: An Assessment of New Research Theory,* Boulder: Rowman & Littlefield Publisher: 285-302

[22] Munk, M. D. 2003a. *Social mobilitet. Social mobilitet i Danmark set i et internationalt perspektiv*, rapport til Undervisningsministeriet, Socialforskningsinstituttet

[23] Munk, M. D. 2003b. Livschancer og social mobilitet - forskellige fødselsårganges vilkår. *Dansk Sociologi* 14: 41-58.

[24] Shavit, Y. and Blossfeld, H.-P. 1993. *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Boulder Co:Westview Press.

[25] Vallet, L. A. 2004. The Dynamics of Inequality of Educational Opportunity in France. Change in the Association Between Social Background and Education in Thirteen Five-Year Birth Cohorts (1908-1972), Working paper, CREST-CNRS

[26] Wanner, Richard A. 1999 Expansion and Ascription: Trends in Educational Opportunity in Canada. *Canadian Journal of Sociology and Anthropology* 36: 409-44

TABLES

Table I

The Distribution of Secondary Educational Choices (Proportions) by Cohort

			Males		
Type of Education	**1982**	**1987**	**1992**	**1997**	**2002**
Gymnasium	9531 (0.253)	12347 (0.307)	14087 (0.374)	13271 (0.418)	11594 (0.418)
Vocational	15854 (0.422)	16881 (0.420)	13082 (0.348)	10823 (0.341)	8730 (0.315)
None	12228 (0.325)	10996 (0.273)	10476 (0.278)	7622 (0.240)	7430 (0.268)
Total	37613	40224	37645	31716	27554
			Females		
Gymnasium	13678 (0.379)	17501 (0.456)	19521 (0.544)	18038 (0.602)	15675 (0.601)
Vocational	9144 (0.254)	10134 (0.264)	7414 (0.207)	5360 (0.179)	4206 (0.161)
None	13225 (0.367)	10789 (0.281)	8963 (0.250)	6567 (0.219)	6222 (0.238)
Total	38371	38424	33762	29965	22037

TABLE 2M

Parameter Estimates For Males

Parameter (se)	**1982**	**1987**	**1992**	**1997**	**2002**
$\|\|\varphi\|\|$	1.422**(0.084)	1.529**(0.074)	1.472**(0.104)	1.364**(0.091)	1.090**(0.072)
$\|\|\gamma\|\|$	2.242**(0.104)	2.232**(0.086)	2.330**(0.096)	2.101**(0.071)	2.096**(0.076)
$\|\|(\varphi,\gamma)\|\|$	2.655**(0.093)	2.706**(0.080)	2.753**(0.096)	2.505**(0.077)	2.363**(0.075)
Father's Education					
$\overline{\varphi}_{fe}$	0.267**(0.032)	0.297**(0.029)	0.237**(0.026)	0.229**(0.026)	0.190**(0.025)
$\overline{\gamma}_{fe}$	0.458**(0.035)	0.458**(0.029)	0.523**(0.028)	0.502**(0.028)	0.471**(0.031)
Father's Occupation					
$\overline{\varphi}_{fo}$	0.226**(0.018)	0.244**(0.019)	0.188**(0.020)	0.189**(0.022)	0.146**(0.022)
$\overline{\gamma}o$	0.214**(0.025)	0.248**(0.024)	0.159**(0.025)	0.211**(0.024)	0.269**(0.026)
Mother's Education					
$\overline{\varphi}_{me}$	0.276**(0.039)	0.128**(0.033)	0.237**(0.026)	0.096**(0.028)	0.184**(0.026)
$\overline{\gamma}_{me}$	0.597**(0.038)	0.571**(0.036)	0.523**(0.028)	0.465**(0.031)	0.451**(0.031)
Mother's Occupation					
$\overline{\varphi}_{mo}$	0.044 (0.029)	0.125**(0.027)	0.116**(0.024)	0.177**(0.027)	0.087**(0.025)
$\overline{\gamma}_{m0}$	0.074**(0.033)	0.089**(0.030)	0.096**(0.028)	0.170**(0.028)	0.200**(0.029)
Number of Siblings					
φ_{sib}	-0.117**(0.007)	-0.131**(0.007)	-0.110**(0.007)	-0.088**(0.008)	-0.069**(0.008)
γ_{sib}	-0.082**(0.009)	-0.069**(0.010)	-0.091**(0.010)	-0.060**(0.0010)	-0.043**(0.01)
Household Income					
φ_{hi}	0.141**(0.018)	0.216**(0.021)	0.240**(0.022)	0.435**(0.045)	0.330**(0.025)
γ_{hi}	0.135**(0.015)	0.111**(0.013)	0.147**(0.018)	0.195**(0.026)	0.156**(0.021)
Father Unemployed					
φ_{fu}	-0.158**(0.026)	-0.186**(0.031)	-0.101**(0.027)	-0.107**(0.029)	-0.071**(0.029)
γ_{fu}	0.010 (0.041)	-0.008 (0.049)	-0.170**(0.036)	0.163**(0.036)	0.076*(0.038)
Single Mother					
φ_{sm}	-0.329**(0.024)	-0.318**(0.021)	-0.298**(0.020)	-0.286**(0.022)	-0.257**(0.022)
γ_{sm}	-0.092**(0.032)	-0.064**(0.025)	-0.088**(0.025)	-0.061**(0.025)	-0.054*(0.027)
Intercept terms					
φ_0	0.488**(0.026)	0.611 **(0.027)	0.534**(0.029)	0.612**(0.033)	0.494**(0.033)
γ_0	-0.653**(0.034)	-0.662**(0.035)	-0.410**(0.036)	-0.434**(0.037)	-0.444**(0.041)
Sample size	37613	41450	39283	32410	27754

†, *, and ** indicate significant at 10, 5, and 1 percent levels, respectively.

TABLE 2F

Parameter Estimates (Standard Error) For Females

Parameter	1982	1987	1992	1997	2002
$\|\|\varphi\|\|$	1.581**(0.101)	1.552**(0.090)	1.548**(0.112)	1.656**(0.141)	1.191**(0.088)
$\|\|\gamma\|\|$	2.426**(0.157)	1.925**(0.100)	1.925**(0.130)	2.157.**(0.189)	2.001**(0.097)
$\|\|(\varphi,\gamma)\|\|$	2.896**(0.143)	2.473**(0.094)	2.471**(0.122)	2.720**(0.173)	2.329**(0.093)
Father's Education					
$\overline{\varphi}_{fe}$	0.156**(0.030)	0.225**(0.030)	0.179**(0.028)	0.182**(0.029)	0.201**(0.028)
$\overline{\gamma}_{fe}$	0.355**(0.040)	0.399**(0.032)	0.379**(0.032)	0.490**(0.053)	0.439**(0.035)
Father's Occupation					
$\overline{\varphi}_{fo}$	0.299**(0.018)	0.211**(0.019)	0.237**(0.021)	0.205**(0.023)	0.147**(0.023)
$\overline{\gamma}_{fo}$	0.219**(0.025)	0.141**(0.023)	0.129**(0.026)	0.187**(0.027)	0.232**(0.029)
Mother's Education					
$\overline{\varphi}_{me}$	0.355**(0.0043)	0.217**(0.036)	0.248**(0.031)	0.185**(0.032)	0.140**(0.028)
$\overline{\gamma}_{me}$	0.708**(0.051)	0.482**(0.042)	0.516**(0.039)	0.490**(0.053)	0.455**(0.038)
Mother's Occupation					
$\overline{\varphi}_{mo}$	0.113** (0.029)	0.166**(0.027)	0.162**(0.025)	0.205**(0.028)	0.124**(0.027)
$\overline{\gamma}_{m0}$	0.174**(0.033)	0.192**(0.030)	0.127**(0.030)	0.234**(0.031)	0.200**(0.034)
Number of Siblings					
φ_{sib}	-0.105**(0.007)	-0.124**(0.008)	-0.094**(0.008)	-0.086**(0.008)	-0.081**(0.009)
γ_{sib}	-0.025**(0.009)	-0.041**(0.009)	-0.044**(0.010)	-0.057**(0.010)	-0.025**(0.011)
Household Income					
φ_{hi}	0.166**(0.016)	0.235**(0.020)	0.373**(0.028)	0.329**(0.028)	0.336**(0.028)
γ_{hi}	0.070**(0.016)	0.067**(0.013)	0.119**(0.022)	0.115**(0.021)	0.188**(0.027)
Father Unemployed					
φ_{fu}	-0.120**(0.027)	-0.186**(0.032)	-0.086**(0.028)	-0.046 (0.031)	-0.091**(0.031)
γ_{fu}	0.001 (0.041)	-0.012 (0.047)	-0.040 (0.037)	-0.011 (0.037)	0.023 (0.042)
Single Mother					
φ_{sm}	-0.230**(0.024)	-0.282**(0.022)	-0.336**(0.021)	-0.227**(0.023)	-0.309**(0.023)
γ_{sm}	-0.018 (0.031)	-0.031 (0.024)	-0.059**(0.026)	-0.059*(0.028)	-0.046 (0.030)
Intercept terms					
φ_0	0.264**(0.025)	0.578**(0.028)	0.571**(0.032)	0.644**(0.034)	0.649**(0.035)
γ_0	-0.172**(0.033)	-0.041 (0.032)	0.186**(0.036)	0.295**(0.040)	0.821**(0.098)
Sample size	38371	38424	33762	29965	22037

$\dagger, *,$ and $**$ indicate significant at 10, 5, and 1 percent levels, respectively.

TABLE 3
The Distribution of Secondary Educational Attainments

By Mother's level of Education for 1982 and 2002. Actual/Predicted

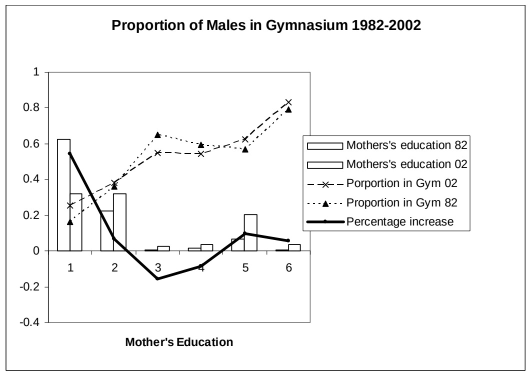
Males

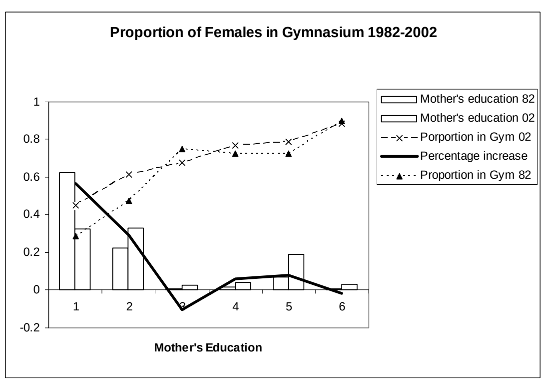
	1982			2002		
Father's Education	**None**	**Vocational**	**Gymnasium**	**None**	**Vocational**	**Gymnasium**
.	1	2	3	4	5	6
Category 1	0.384/0.383	0.452/0.450	0.165/0.167	0.362/0.361	0.384/0.384	0.253/0.255
Category 2	0.209/0.207	0.433/0.434	0.358/0.359	0.236/0.233	0.381/0.382	0.382/0.385
Category 3	0.177/0.182	0.173/0.166	0.649/0.653	0.271/0.274	0.180/0.176	0.548/0.550
Category 4	0.150/0.137	0.255/0.263	0.595/0.601	0.194/0.182	0.260/0.266	0.545/0.553
Category 5	0.176/0.175	0.228/0.254	0.567/0.571	0.172/0.170	0.185/0.183	0.643/0.647
Category 6	0.103/0.097	0.107/0.095	0.790/0.809	0.101/0.092	0.065/0.053	0.835/0.856

Females

	1982			2002		
Father's	**None**	**Vocational**	**Gymnasium**	**None**	**Vocational**	**Gymnasium**
Education	1	2	3	4	5	6
Category 1	0.430/0.428	0.282/0.282	0.288/0.290	0.322/0.321	0.226/0.227	0.452/0.451
Category 2	0.267/0.264	0.257/0.258	0.476/0.478	0.203/0.204	0.182/0.181	0.614/0.616
Category 3	0.163/0.253	0.082/0.075	0.755/0.771	0.232/0.225	0.093/0.093	0.675/0.682
Category 4	0.163/0.161	0.112/0.111	0.775/0.728	0.143/0.140	0.086/0.087	0.770/0.773
Category 5	0.167/0.160	0.106/0.106	0.727/0.734	0.153/0.146	0.061/0.061	0.786/0.793
Category 6	0.082/0.068	0.017/0.015	0.901/0.918	0.099/0.088	0.015/0.014	0.886/0.899

Notes: The elements in each row are the proportions of the respondents in each of the three educational categories. Each row represents a different category of father's education and hence the row entries sum to unity.


Proportion of Males in Gymnasium 1982-2002
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
1
2
3
4
5
6
Mother's Education
Mothers's education 82
Mothers's education 02
Porportion in Gym 02
Proportion in Gym 82
Percentage increase


Proportion of Females in Gymnasium 1982-2002
1
0.8
0.6
0.4
0.2
0
-0.2
1
2
3
4
5
6
Mother's Education
Mother's education 82
Mother's education 02
Porportion in Gym 02
Percentage increase
Proportion in Gym 82