Not all transitions are equal: The effect of educational expansion and the decreased disadvantaged position of women in education on educational inequality in the Netherlands between 1906 and 1990 *

Maarten L. Buis

m.buis@fsw.vu.nl

May 1, 2007

Abstract

Mare (1981) showed that the effect of family socioeconomic status on the highest achieved level of education (overall inequality of educational opportunity or overall IEO) depends in part on the distribution of education. However this relationship has been primarily used as an argument for a model that studies the the effect of family socioeconomic status on the probabilities of passing the transitions between levels of education (partial IEOs), because estimates of inequality of educational opportunity in this model do not depend on the distribution of education. The effect of the distribution of education on overall IEO is treated as a black box. The aim of this paper is to break this black box open. It will show that overall IEO is a weighted sum of the partial IEOs, and that a partial IEO receives more weight if more people are at risk, if passing or failing the transition is not universal, i.e. if the proportion of students who pass is closer to 50%, and if the difference in expected level of education of people who pass and not pass is bigger. An application of this result to the Netherlands between 1906 and 1990 will show how this result can be used to describe the impact of educational expansion and the disadvantaged position of women on IEO.

*DRAFT: Please request permission before citing. I am grateful for comments and suggestions made by Harry B.G. Ganzeboom, Adriaan Hoogendoorn, Ineke Maas, Robert D. Mare, and Paul Nieuwbeerta.

1 Introduction

The association between family socioeconomic status (SES) and the child's educational attainment has long been studied in social stratification and social mobility research. A strong positive association means that people with a higher SES background are more likely to achieve higher levels of education than people with a lower SES background. The strength of the association is, for this reason, often called inequality of educational opportunity (IEO). Different types of IEO exist: One can look at the association between family SES and highest achieved level of education or at the association between family SES and the probabilities of passing the transitions between the levels of education that make up the educational system. In this paper IEO in terms of highest achieved level of education will be called overall IEO and IEOs in terms of probabilities of passing transitions will be called partial IEOs. Overall IEO focusses more on the end result of the educational career. It is primarily of interest to those who also study the effects of education and family background on other live chances like success on the labor market, success on the marriage market, or health (Blau and Duncan, 1967; Hout and DiPrete, 2006; Treiman and Ganzeboom, 2000). The partial IEOs focuss more on the educational career itself. The partial IEOs are primarily of interest to those who study IEO as an interesting phenomenon in it's own right (Boudon, 1974; Mare, 1981; Shavit and Blossfeld, 1993).

Mare (1981) showed that differences in overall IEOs between cohorts are in part due to differences in the distribution of education. These effects can be considerable since the distribution of education has changed substantially over cohorts. In almost all countries people born in later cohorts obtain much more education, a process called educational expansion (Hout and DiPrete, 2006). Furthermore, Mare (1981) showed that differences across cohorts in partial IEOs are not affected by educational expansion. As a result, Mare (1981) argues that the partial IEOs is the more pure measure of IEO.

The relationship between overall IEO, the partial IEOs, and the distribution of education has since been treated as a black box. Those studies that did investigate this relationship (Mare, 1981; Smith and Cheung, 1986; Nieuwbeerta and Rijken, 1996) compared the observed overall IEO with simulated results of two counterfactual scenarios: the distribution of education remained unchanged and the partial IEOs changed as observed, and the distribution of education changed as observed, but the partial IEOs remained unchanged. Simulations like these can tell us how much overall IEO is affected by changes in the distribution of education, but does not give any insight into why. The aim of this paper is to break this black box open. The main question this paper will answer is:

> What is the substantive interpretation of the relationship between overall IEO, the partial IEOs and the distribution of education found by Mare (1981)?

Mare (1981) already showed that the overall IEO is a weighted sum of the partial IEOs and that these weights capture the effect of the distribution of education, whereby the distribution of education is represented by the set of probabilities of passing each transitions. I will show that the effect of the distribution of education has a substantive interpretation: A partial IEO receives more weight if 1) the proportion of people at risk of passing that transition increases, 2) the proportion passing that transition is closer to 50%, i.e. passing or failing that transition is not close to universal, and 3) the difference in expected level of education between those who pass and those who fail the transition increases, i.e. the gain from passing increases. I will show that the weight is actually the product of these three elements.

This result is interesting in its own right, as it sheds light on the relationship between inequality during the process of obtaining education (the partial IEOs) and inequality in the end result (overall IEO). It also shows that overall IEO is a reasonable summary measure of IEO. Finally, studying the influence of differences in the distribution of education across cohorts is substantively interesting as these differences represent educational expansion, a process associated with the modernization of society. So changes across cohorts in the weights attached to each partial IEO describe how a educational expansion influences overall IEO. The usefulness of this point is is not limited to comparisons across cohorts. A similar argument holds for comparisons across for instance gender or ethnic groups, except that differences in the distribution of education now represents the disadvantaged position of women or certain ethnic groups.

This paper will start with a description of the model proposed by Mare (1981), and the derivation of the relationship between partial IEO, overall IEO, and the distribution of education. This section will also discuss an influential criticism of this model by Cameron and Heckman (1998). The next section will illustrate the way that this result can enrich the description of IEO by applying it to differences in IEO between men and women and across cohorts between 1906 and 1990 in the Netherlands. This section will also generalize the relation between overall IEO, partial IEOs and the distribution of education to models that are suitable for a tracked system.

2 Model of education

2.1 partial and overall IEO

Partial IEOs are the effects of family background on the probabilities of passing the different transitions between levels of education. These are typically estimated using a sequence of logistic regression. This model is known under a variety of names: sequential response model (Maddala, 1983), continuation ratio logit (Agresti, 2002), model for nested dichotomies (Fox, 1997), and the Mare model (Shavit and Blossfeld, 1993). In this paper I will use the term sequential response model or, to emphasize that the probabilities of passing sequences are modeled using logistic regression, sequential logit model.

Consider for instance the hypothetical educational system in figure 1. In this system a person can face three transitions: he can go to primary education or settle for no education at all, if he chose to go to primary education he can choose to leave afterwards or go to secondary education, and if he chose to go to secondary education he can choose to leave afterwards or go to tertiary education. Furthermore the values assigned to each level are 0, 6, 12, and 16, which are quite typical values if the levels are scaled according the years of education. This is by no means the only or best way to scale education, but it is common. However, another scale will be used in the empirical application.

[Figure 1 about here.]

The model assumes that one has to be at risk of passing a transition, i.e. have passed all lower transitions, in order to make a decision at that transition on whether to continue or to leave school. Otherwise, these decisions are assumed to be completely independent. As a consequence one can estimate the partial IEOs by running separate logistic regression for each transition on the appropriate sub-sample. This model is shown in equation (1).

$$p_{ki} = \frac{\exp(\alpha_k + \lambda_k SES_i)}{1 + \exp(\alpha_k + \lambda_k SES_i)} \quad \text{if} \quad y_{k-1\,i} = 1 \tag{1}$$

The probability that person i passes transition k is p_{ki}. The partial IEO belonging to transition k is λ_k. Whether or not individual i has passed the previous transition is indicated by the indicator variable $y_{k-1\,i}$ with the understanding that everybody is at risk of passing the first transition, so $y_{0i} = 1$. The differences in partial IEO between men, women, and cohorts are obtained by adding the appropriate interaction terms. The constant for transition k is α_k. Once these models are estimated it is straightforward to compute predicted probabilities for passing each transition. These predicted probabilities can than be used to compute the expected highest achieved level of education ($E(ed)$), as is done in equation (2). The expected highest achieved level of education is the sum of the value of each level of education (l_k) times the probability of attaining that level. The probabilities and levels can be derived from figure 1. One way of thinking about equation 2 is that it is a regression equation showing a non-linear relationship between family's SES and highest achieved level of education. [1]

$$E(ed) = (1 - p_{1i})l_0 + p_{1i}(1 - p_{2i})l_1 + p_{1i}p_{2i}(1 - p_{3i})l_2 + p_{1i}p_{2i}p_{3i}l_3 \tag{2}$$

Remember that the overall status IEO is the effect of family's SES on child's highest achieved level of education, or in other words, by how much the expected highest achieved level of education changes if family's SES changes. This shows that overall IEO is the first derivative of equation (2) with respect to family SES. Notice that SES is part of equation (2) through the p_{ki}s as described in equation (1). This derivative is shown in equation (3).

[1] Normal linear regressions of highest achieved level of education on family SES, as used by for instance Blau and Duncan (1967) or Shavit and Blossfeld (1993), can be seen as linear approximations of this non-linear relationship. How well this linear approximation works is an empirical question.

$$
\begin{array}{llllllll}
\frac{\partial E(ed)}{\partial SES} & = & & & & & & \\
\{1 & \times & p_{1i}(1-p_{1i}) & \times & [(l_1-l_0)+p_{2i}(l_2-l_1)+p_{2i}p_{3i}(l_3-l_2)] & \} & \lambda_1 & + \\
\{p_{1i} & \times & p_{2i}(1-p_{2i}) & \times & [(l_2-l_1)+p_{3i}(l_3-l_2)] & \} & \lambda_2 & + \\
\{p_{1i}p_{2i} & \times & p_{3i}(1-p_{3i}) & \times & [(l_3-l_2)] & \} & \lambda_3 &
\end{array}
\tag{3}
$$

It shows that overall IEO is a weighted sum of the partial IEOs, the λ_ks. The weights, the parts between curly brackets, consist of three parts that are all related to the distribution of education: 1) The proportion of people at risk. For the first transition this proportion is one, for the second it is the proportion who went to primary education, p_{1i}, and for the third transition it is the proportion who went to secondary education, $p_{1i}p_{2i}$ 2) A function of the proportion passing that transition that is low if virtually everybody passes or failsand is highest when the probability of passing is .5, $p_{ki}(1-p_{ki})$. Substantively this makes sense because if only a few people pass or fail a transition then any inequality at this transition will affect only a few people. 3) The differences in expected level of education between those who pass the transitions and those who do not. These are the parts in the square brackets. They are the immediate gain of passing that transition plus the gain of passing any subsequent transition times the probability of passing that one.

When applying these results one should take into account that the weights assigned to partial IEOs depend on the values of all the explanatory variables. The weight depend on the probabilities of passing the various transitions, and these probabilities in turn depend on all the explanatory variables through equation (1). A simple and reasonable choice would be to evaluate overall IEO at the mean values of the explanatory variables.

Figure 2 summarizes how educational expansion or gender IEO could influence overall status IEO. Arrows a represent the partial status IEOs; the odds ratios of passing the different transitions for a unit change in socioeconomic status. Arrows b show that these partial IEOs can be combined to one effect of socioeconomic status on highest achieved level of education, but that not every transitions receives equal weight (arrow c). These weights depend of the transition probabilities (arrow d). Cohort or gender can influences status IEO both directly partial IEOs (arrow e) and through affecting the transition probabilities (arrow f). This latter arrow is educational expansion or gender IEO, and its influence on social IEO is represented by arrows c and d. This argument can easily be extended to comparisons among other groups, like countries or ethnic groups.

[Figure 2 about here.]

2.2 Critique on Mare (1981)

The model discussed so far is an extension of the work done by Mare (1981). However, this model has been subject to an influential critique by Cameron and Heckman (1998). They have two main objections against the sequential response model. Their first point of criticism is that differences in the estimated odds ratios between transitions or between groups (e.g. men and women or cohorts) do not represent causal effects. This is correct and Mare has confirmed this (e.g. Mare, 2006). However, both Cameron and Heckman (1998) and Mare (2006) tend to be too negative, because this does not mean that the parameter estimates of a sequential response model cannot be compared between groups or transition equations. The parameter estimates still represent the ratio of the odds of passing versus not passing between higher status and lower status pupils, and as such are comparable. The sequential logit model is what is sometimes called a population average model (e.g. Agresti, 2002), i.e. it describes differences between groups, but it does not give causal effects. Differences between the parameters do not represent causal effects since these odds ratios do not control for all variables that are not in the model, a phenomenon referred to as unobserved heterogeneity. There are two ways in which unobserved heterogeneity makes causal inference difficult: First, in models with a non-linear link function the effect found with conventional models is influenced by the variance of the sum of the effects of the unobserved variables (Allison, 1999; Neuhaus, Kalbfleisch and Hauck, 1991; Neuhaus and Jewell, 1993). Differences in estimated parameters between groups (e.g. cohort, or gender) can be the result of a causal effect of group membership, or of differences in the variance of the unobserved effects. This variance can differ across groups because the variance of the unobserved variable differs across groups and because the effects of the unobserved variables differ across groups. The second reason why unobserved heterogeneity makes causal inference difficult is that the unobserved variables are likely to become correlated with the observed variables at higher transitions, even if they were uncorrelated during the first transition (Mare, 1981). For instance, consider a situation where a) the transition probabilities are influenced by family SES and intelligence, b) SES and IQ are initially uncorrelated, c) a logistic regression of both variables without their interaction is the appropriate model, and d) that intelligence is not observed by the researcher. After the first transition SES and IQ will most likely be negatively correlated, and the effect of SES will in that case be underestimated. The odds of passing versus failing may be a factor x higher for intelligent students for both high and low status children, but when the odds are translated to proportions this factor will be lower for high status children then for low status children. Thus the proportion less intelligent children will be higher among high status children than among low status children, even if the odds ratio of passing versus failing for intelligent versus less intelligent children is the same for high and low status children. This means that after the first transition low status children will tend to be on average smarter then high status children, thus leading to an underestimation of the effect of SES[2]. Moreover, the relative advantage of less intelligent high status children tends to decrease when both low and high status children are more likely to pass (educational expansion). Consequently over time one can expect the underestimation to decrease and thus the observed odds ratios to increase.

The second point of criticism by Cameron and Heckman (1998) is that they considered the underlying behavioral model to be unrealistic. They show that the sequential response model implies that the decisions at each transition is taken separately without taking the consequences for future transitions into account. The main problem with this criticism is that there is not really an alternative. Cameron and Heckman (1998) propose a model in which the final level of education is chosen at the beginning of the educational career and the decisions at each transition are just a function of this initial decision. The first important decision in the Dutch educational system is made at about age 12. The assessment of the child's abilities and the child's ambition is still likely to change as the child gets older (and

[2]Cameron and Heckman (1998) showed that this does not always hold, in particular if the selection on the unobserved variable is so strong that only intelligent children survive regardless of family background. However, such strong selection is unlikely to exist in practice.

passes through puberty). Consequently the sequential response model seems less problematic then the model proposed by Cameron and Heckman.

In conclusion, Cameron and Heckman (1998) show some limitations of the sequential response model. However, these limitations are not serious enough to make the model unusable. The differences in odds ratios do not represent causal mechanisms, but they can still be interpreted as a description of the differences across groups. Showing that the sequential response model implies a particular behavioral model is a valuable contribution to the understanding of the model, but it is not a major concern if the parameters are interpreted as descriptive, and alternative models make even more unrealistic assumptions.

3 Empirical application

This section will illustrate how the relation between the distribution of education, the partial IEOs and overall IEO can be used to get a more complete picture of IEO. In particular this section will describe the influence of educational expansion and gender IEO on status IEO in the Netherlands between 1906 and 1990. Gender IEO and educational expansion influence overall status IEO because they imply that the probabilities of passing the different transitions differ between men and women and between cohorts. In turn these different transition probabilities cause differences in the impact of the partial status IEO of each transition on overall status IEO. This secion will also compare the estimates of overall IEO based on the sequential response model with the estimates of overall IEO based on a simple linear regression of highest achieved level of education on father's occupational status. Linear regression provides a simple and easy to interpret alternative to estimating overall IEO, if one is only interested in overall IEO (for instance because is it part of a larger status attainment model like in (Blau and Duncan, 1967)).

This section will start with a description of the Dutch educational system, and the data. Next, results for the sequential logit model discussed in the previous section will be generalized to be applicable to a tracked educational system. Finally, the model will be estimated on the Dutch data and the results will be presented.

3.1 The Dutch educational system

The Dutch educational system is stratified. The educational system as it was formalized in 1968 is shown in figure 3. A child finishes primary education when he or she is about 12 years old and can then choose between 4 levels: junior vocational, junior general secondary, senior general secondary, and pre-university education. Junior vocational and junior general secondary both give excess to senior secondary vocational. Senior general secondary gives access to higher professional education, and pre-university gives access to university. To make things more complicated, they can also choose to move up within their current column (junior vocational to junior secondary general, senior secondary vocational to higher professional, etc.), or move down in the next column (senior general secondary to senior secondary vocational, and pre-university to higher professional).

[Figure 3 about here.]

The Dutch educational system has changed over the period under study (1906-1990). An important change was that senior secondary vocational training only came into existence during this period, and was initially not very popular. It was long believed that the best way for young people to acquire vocational skill was on the job (Boekholt and de Booy, 1987). This also had an impact on the nature of junior general secondary education: its primary purpose gradually changed from preparing for the labor market to preparing for senior secondary vocational education. Other differences are that senior general secondary education was prior to 1968 reserved for girls and that pre-university education was before 1968 much more diversified then after 1968. Moreover, general access to pre-university education was at the beginning of the study period a recent event: prior to 1906 women had to ask the minister of education permission to enter pre-university education (van Essen, 1990). Table 1 shows the levels of education before and after 1968, the number of years of education they represent, (Boekholt and de Booy, 1987) the English names for these levels, and their ISCED classification (UNESCO, 1997).

[Table 1 about here.]

3.2 data

The data come from the International Stratification and Mobility File (ISMF) (Ganzeboom, 2005). For the Netherlands, the ISMF consists of 51 surveys. These surveys were held between 1958 and 2004, and were merged to increase the time-span covered, the number of respondents, and to diminish the effect of idiosyncrasies of individual surveys. The purpose of this paper is to compare the effect of family SES on highest achieved level of education between men and women and across cohorts. Time is measured by annual birth-cohorts and represented by the year in which the members of the cohort were 12 years old (approximately the age at which the most important educational transition is made in the Netherlands). Information is available for the cohorts 1906-1990. Overall IEO is measured by the strength of the association between child's highest achieved level of education and parental socioeconomic status in a linear regression model. The socioeconomic status of the father is measured by father's occupational status in ISEI scores (Ganzeboom and Treiman, 2003). The original ISEI score is a continuous variable ranging from 10 to 90, but it has been recoded to a range from 0 to 1. A unit increase in ISEI is thus equivalent to moving from the lowest status to the highest status. Where available survey weights have been used. The weighted number of respondents is approximately 77,000. There remain approximately 67.000 weighted respondents, after removing respondents with missing observations on any of the variables. The number of respondents is unequally distributed over the cohorts, ranging from 47 observations in 1912, to 2,476 in 1959, to 205 in 1988. Finally, a scale for the level of education is needed in order to estimate the relationship between partial and overall IEO using equation (3). The scaling of education used in this paper was estimated in such a way that it maximizes the direct effect of education on income controlling for father's occupational status. Details are shown in the appendix. The advantage of using this scaling rather then years of education is that it defines levels of education in terms of how much a student will gain from achieving a level of education, instead of defining it in terms of how much was invested to achieve that level. The former is more relevant for measuring inequality of educational opportunity. This scale is for interpretability coded in such a way that the mean is zero and the variance is one for the cohort that was twelve in 1970.

3.3 Generalizing the sequential response model to a tracked system

Thus far the sequential response model has been discussed in terms of a simple educational system represented in figure 1. The Dutch educational system is shown to be much more complicated, however for estimation purposes this system will be simplified according to figure 4. This involves two simplifications: 1) Some levels are merged, and 2) each level of education is assumed to be reachable through only one sequence of transitions. All children are assumed to have finished primary education, and then face a choice between leaving the schooling system or get a second diploma. If they choose to get a second diploma they will have to choose whether to get it in a high (havo/vwo and hbo/wo) or a low track (lbo/mavo and mbo). Once they have finished their second diploma in either the high or low track they have to choose whether or not to get a third diploma: mbo if they are in the low track and hbo/wo if they are in the high track. The assumption that everybody has finished at least primary education is justified by the very small proportions of people without primary education and the fact that during the entire period under study there was compulsory education of at least six years. This simplification also implies that everybody stays in his or her track. With this assumption one only needs the highest achieved level of education in order to derive which transitions a person made. However, this assumption may not hold, since in the Netherlands it is possible to switch between tracks. Three of the fifty-one surveys that will be used in this study contain information about the actual transitions respondents experienced. These surveys show that only a small proportion of all moves between levels of education that could involve crossing the line between the low and the high track actually do cross that line. The percentages are 5.7%, 7.6%, 8.6%, and 8.8% for cohorts that were 12 between 1939-1944, 1945-1959, 1960-1974, and 1974-1984 respectively. Hence the assumption that people stay within their track seems justified.

[Figure 4 about here.]

As before logistic regressions are used to model the probabilities of passing the different transitions. However, whereas a conventional sequential response model consists of sequence of decisions to either continue or to stop, this model also contains a choice between tracks, a branching point. In that sense it is akin to models proposed by Lucas (2001) and by Breen and Jonsson (2000). Again the partial IEO and the predicted probabilities belonging to transition k are represented by λ_k and p_{ki} respectively. the predicted level of education is now represented by equation 4.

$$E(ed) = (1-p_{1i})l_1 + p_{1i}(1-p_{2i})(1-p_{3i})l_2 + p_{1i}(1-p_{2i})p_{3i}l_3 + p_{1i}p_{2i}(1-p_{4i})l_4 + p_{1i}p_{2i}p_{4i}l_5 \quad (4)$$

Remember that the overall IEO is first derivative of equation (4) with respect to family SES. This derivative is shown in equation (5).

$$\begin{array}{llllllll}
\frac{\partial E(ed)}{\partial SES} & = & & & & & & \\
\{1 & \times & p_{1i}(1-p_{1i}) & \times & [(1-p_{2i})(l_2-l_1)+p_{2i}(l_4-l_1)+ & & & \\
 & & & & (1-p_{2i})p_{3i}(l_3-l_2)+p_{2i}p_{4i}(l_5-l_4)] & \} & \lambda_1 & + \\
\{p_{1i} & \times & p_{2i}(1-p_{2i}) & \times & [(l_4-l_1)+p_{4i}(l_5-l_4)- & & & \\
 & & & & (l_2-l_1)-p_{3i}(l_3-l_2)] & \} & \lambda_2 & + \\
\{p_{1i}(1-p_{2i}) & \times & p_{3i}(1-p_{3i}) & \times & [(l_3-l_2)] & \} & \lambda_3 & + \\
\{p_{1i}p_{2i} & \times & p_{4i}(1-p_{4i}) & \times & [(l_5-l_4)] & \} & \lambda_4 &
\end{array} \quad (5)$$

Again, overall IEO is a weighted sum of the partial IEOs, the λ_ks. The weights, the parts between curly brackets, consist of the same three parts: 1) The proportion of people at risk, only this is defined slightly different for the third and fourth transition: For the third transition it is the proportion who continued after primary education and then went to the lower track, $p_{1i}(1-p_{2i})$, and for the fourth transition it is the proportion who continued after primary education and went to the high track, $p_{1i}p_{2i}$. 2) A part $(p_{ki}(1-p_{ki}))$ that is small if virtually everybody passes or fails that transition and is largest when the probability of passing is .5. 3) The differences in expected level of education between those who pass the transitions and those who do not. These are the parts in the square brackets. The expected gain of passing the second transition is the expected increase for those who pass *minus* the expected increase of those who fail. The latter is usually zero, since failing usually means leaving the schooling system, which means no gain in level of education. However the second transition is between the low and the high track, and those who fail, go to the low track, and still get an increase in level of education.

This illustrates that the relation between overall IEO and partial IEOs can be extended to tracked educational systems. Moreover, using the same logic this result could be extended to more complicated systems, e.g. multiple branching points or branching points with more then two destinations. In the latter case the partial IEOs would be estimated with a multinomial logit.

3.4 results

The analysis will consist of three parts. First, a descriptive analysis is performed on the differences in transitions probabilities between men and women and between cohorts. Second, the generalized sequential response model described in the previous section is estimated. The results from this model will be used to compute the partial IEOs, the weights and the overall IEO. Together they will provide a detailed picture of status IEO and how it is influenced by educational expansion and gender IEO. Third, the relation between the transition probabilities and the weights will be investigated in more detail by looking at the three components of the weights; the proportion at risk, how close the transition probability is to 50%, and the expected increase in level of education when passing a transition.

Changes in transition proportions for males and females are shown in figure 5. As in most other countries, the Netherlands experienced a period of educational expansion during the previous century. The proportion of pupils at risk who passed a transition has increased over cohorts in all transitions. It also shows that senior secondary vocational education is a relatively recent level. In the earliest cohort only 2% (women) or 3% (men) of those at risk actually finished this level of education. But this has rapidly grown to about 50%. The difference between men and women was primarily located in the first two transitions. Both the proportion continuing after primary education and the proportion entering the high track increased slower for women then for men. However, the difference between men and women disappeared in the subsequent transitions. Gender IEO in the first transition gradually disappeared since the post World War II cohort, while gender IEO in the second transition decreased much slower and disappeared only in the most recent cohort.

[Figure 5 about here.]

To investigate status IEO and how it is influenced by gender IEO and educational expansion generalized sequential logit models are estimated separately for both men and women, in effect adding interaction terms between gender and all other variables. The other variables are dummies for the cohorts 1930-1945, 1945-1960, 1960-1975, 1975-1990, excluding the reference category 1906-1930 and interaction terms with all cohort dummies and father occupational status. All interaction terms could be added because the main effect of father's occupational status is excluded. The interaction effects now measure the effect of father's occupational status in each cohort, the partial status IEOs. The effects are log odds ratios. The most striking features are the partial IEOs in higher transitions (in particular lbo/mavo versus mbo and havo/vwo versus hbo/wo) are smaller than the lower transitions. This pattern has also been found in many other studies using an sequential response model Mare (1980); Shavit and Blossfeld (1993) and is usually attributed to unobserved heterogeneity, as was discussed in section 2.2. In this section it was also shown that unobserved heterogeneity would also cause the observed effects in higher transitions to go up as educational expansion progresses. Tables 2 and 3 do not support this prediction (with the exception of the transition havo/vwo versus hbo/wo for females). In particular the partial IEO at the second transition shows a downward trend, suggesting that at least at that level the causal effect has gone down faster then could be compensated with changes due to unobserved heterogeneity.

[Table 2 about here.]

[Table 3 about here.]

From these results one can derive predicted levels of education for each father's occupational status, which together form a regression line, which will be non-linear. The slope of this regression line will tell how much the expected level of education changes when father's occupational status changes by one unit, and is thus the overall IEO. It is instructive to first look at these regression lines, both as an interesting description in its own right and as a means to assess how well a simple linear regression approximates the overall IEO estimated using the sequential response model. This is particularly relevant since trends in the sequential response model are created by comparing overall status IEOs for five rather coarse cohorts: Those respondents that were 12 in 1906–1930, 1931–1945, 1946–1960, 1961–1975, and 1976–1990. To check if the observed pattern is a artefact of the rather coarse cohorts one could compare these results with results for much more and finer cohorts. However, using much more very fine cohorts is not possible in a sequential response model, since the sequential response model estimates the effects of higher transitions on only those who are at risk of passing that transition, and these sub-samples can get very small. However, estimating overall IEOs for much finer cohorts would be possible if overall status IEO could be well approximated by a linear (OLS) regression of highest achieved level of education on father's occupational status. Figure 6 checks this approximation by comparing the expected levels of education according to the sequential response model and linear regression. It shows that the sequential response model implies a non-linear relationship between father's occupational status and highest achieved level of education. This non-linearity appears to be guided by minimum and maximum levels of education. Linear regression is shown not to respect these bounds, but otherwise seem to fit remarkably well.

[Figure 6 about here.]

Figure 7 shows the overall IEOs according to the sequential response model using five cohorts, and according to an OLS regression model using 75 (annual where the number of observations allow) cohorts. Generally, the pattern found in the sequential response model seems to fit the pattern found using the finer cohorts. The only exception seems to be that the initial strong increase for men appears to be the result of a cluster of five early cohorts with very small IEOs. However, these outliers are caused by the relatively small sample size (50–120) in each cohort combined with the small proportion who went to university or higher professional education. Of these men who went to higher tertiary education a relatively small proportion had a father with a medium to low occupational status. As a result not every cohort should find a low status university educated man. If such a man is found, he will be an influential data-point and pull the effect of father's occupational status on educational attainment down. The effect of father's occupational status on educational attainment will be underestimated in that cohort since the proportion low status university educated men in that cohort will be overestimated. However, deleting these outliers is not the right solution either, since it would brush away the fact that some men from low status background did make it to university. Moreover, in those cohorts where no low status university educated men are found the status IEO is likely to be overestimated, since in these case the likelihood of a low status male getting a university degree is underestimated. A solution is to increase the sample size by combining cohorts, and this is exactly what the coarse cohorts used in the sequential response model do. This pattern can not be seen for women since low status women had virtually no opportunity of entering the high track, and continue to university. As a result not a single low status university educated woman was found in this era.

Education is scaled in such a way that the mean for the cohort 1970 is 0 and the standard deviation is 1, and father's occupational status is scaled so that 0 is the lowest value and 1 the highest. So the overall IEO measures the difference between a child from the lowest and the highest background in standard deviations education. Overall IEO shows two striking features: The first feature is the trend in overall IEO, it first increases and than decreases. As will be shown later, both the initial increase and the later decrease can be explained with educational expansion. The second feature is the difference in shape of the trend between men and women. The initial increase and subsequent decrease is much more pronounced for men then for women. Moreover the decline in overall IEO started later for women then for men. This will be shown to be the result of gender IEO.

[Figure 7 about here.]

This pattern in overall IEO can be explained by looking at the differences in transition probabilities between men and women and cohorts, as these probabilities determine how much each transition contributes to overall IEO. Remember that each transitions' contribution is the weight times the partial IEO. So each transitions' contribution could be visualized by the area of a rectangle with a height equal to the partial IEO and a length equal to the weight. This is shown for men and women in figures 8 and 9. The horizontal axis shows the weights and the vertical axis the partial IEOs. The columns represent the cohorts and the rows represent the transitions. The overall IEO for a cohort (as shown in figure 7) is the sum of the areas of the rectangles within one column. This way the relative contribution of the different transitions to overall IEO can be compared, and how these differences are caused by differences in partial IEOs and their weights. This can be compared across groups.

[Figure 8 about here.]

[Figure 9 about here.]

Figures 8 and 9 show that the main sources of IEO are the first two transitions. The main change over time is the loss of importance of the first transition (continue after lo) and its replacement by the second transition (lbo/mavo versus havo/vwo). However, the increase in importance of the second transition happened faster then the decrease of the first transition, thus causing the initial increase in overall IEO. Later, overall IEO decreases as the first transition starts contributing less and less to overall IEO. For men these two processes happen consecutively, while for women these two processes overlapped, thus canceling each other out.

These changes in weights can be related to changes in three aspects of the distribution of education: the number of people at risk at each transition, the proportion of people passing each transition, and the expected increase in level of education if someone passes a transition. This is shown in table 4. It shows for each transition the predicted probability of passing that transition for someone with a father with average occupational status (.403), the three elements that make up the weight, and the weight itself. The proportions can be calculated using the estimates of the sequential response model in tables 2 and 3. The three components are labeled at risk, variance, and gain for respectively the proportion at risk, how close the proportion passing is to 50% ($p_k(1 - p_k$ is the variance of the dummy indicating whether or not someone passes transition k), and the expected gain in level of education from passing.

These can be calculated using the predicted proportions, the scaling of education shown in figure 4, and equation (3). The proportions at risk for the first, second, third and fourth transition are respectively: 1, p_1, $p_1(1-p_2)$, and p_1p_2. The second component of the weight for transition k is $p_k(1-p_k)$. The gain for the first, second, third, and forth transition is respectively: $[(1-p_2)(l_2-l_1)+p_2(l_4-l_1)+(1-p_2)p_3(l_3-l_2)+p_2p_4(l_5-l_4)]$, $[(l_4-l_1)+p_4(l_5-l_4)-(l_2-l_1)-p_3(l_3-l_2)]$, $[l_3-l_2]$, and $[l_5-l_4]$. The weight itself can be calculated by multiplying the three components with one another.

[Table 4 about here.]

The table shows that the decline in influence of the first transition is primarily due to the fact that passing has become virtually universal (educational expansion). This happened a bit later for women, and women for a long time gained less from passing (gender IEO). As a result the decline in the contribution of the first transition started later for women then for men. For the second transition table 4 shows the effect of the choice between the high track and the low track became more influential because the proportion of people at risk increased and the proportion choosing to go to the high track increased to about 50%. The increase in the proportion at risk and especially the increase in the proportion going to the high track increased slower for women then for men. These two transitions together explain the trend in overall IEO and the difference in trend between men and women. Table 4 shows that the last two transitions are relatively unimportant because relatively few people are at risk of passing these transitions and those who pass gain relatively little. Those who passed the first two transitions gained both the immediate increase in level of education and the probability of gaining an extra level of education (either mbo or hbo/wo) while in the third and fourth transition people only gain the immediate increase in level of education.

4 Conclusion

This paper started with the finding by Mare (1981) that differences in overall IEO across groups (e.g. cohorts) depends on both the differences in partial IEOs and differences in the distribution of education. However, this relationship has been treated as a black box and as an argument for controlling for differences in the distribution of education instead of studying its effects. This is a lost opportunity since the differences in distribution of education often represent substantively interesting phenomena. For instance the differences across cohorts represent educational expansion and the differences across the genders represent the degree to which women had a disadvantaged position in the educational system. This leads to the following question: What is the substantive interpretation of the relationship between overall IEO, the partial IEOs and the distribution of education found by Mare (1981)? The distribution of education, in the form of the set of transition probabilities, is shown to affect overall status IEO by altering the importance of the different transitions between levels of education. Specifically, overall IEO is shown to be a weighted sum of the partial IEOs, and the partial IEO of a transition receives more weight if more people are at risk, if passing or failing the transition is not universal, i.e. if the proportion of students who pass is closer to 50%, and if the difference in expected level of education of people who pass and not pass is bigger. Through these weights one can describe the influence of differences in the distribution of education.

The empirical application of this finding focused on the impact of differences in the distribution of education across cohorts (educational expansion) and gender (gender IEO) on status IEO in the Netherlands between 1906 and 1990. These two influences can explain two main features of the trend in overall IEO: First, the trend over cohorts showed an initial increase followed by a decrease. Second, this pattern was much less pronounced for women. In particular the initial increase was much smaller for women. The initial increasing trend could be explained by educational expansion, in particular the increase in the proportion of students that passed the second transition — the transition between entering the low track (lbo/mavo) or the high track (havo/vwo). Initially a small proportion of pupils at risk entered the high track, causing the contribution of this transition to overall IEO to be small. However, this proportion increased to about 50% greatly increasing this transitions' contribution, and as a result causing an increase in overall IEO. The subsequent decrease in the overall status IEO was also caused by educational expansion, in this case the increasing pass rate at the first transition — whether or not to continue after primary education. The contribution of this transition to overall IEO decreased as passing this transition became near universal. The difference in trend of overall IEO between men and women can be explained by gender IEO, in particular the fact that the pass rate for the second transition (entering the high track) for women remained low for a much longer period than for men. Consequently, for men these two processes occurred consecutively, while for women they overlapped. As a result for women the decrease in overall IEO due to the decrease in the contribution of the first transition partially canceled out the increase in overall IEO due to the increase in contribution of the second transition. As a result the trend in overall IEO for women was much less pronounced then the trend for men.

In conclusion, this paper showed how the study of IEO could be improved by also studying the impact of the distribution of education instead of just controlling for it. The main advantage is that this makes it possible to study the impact of phenomena like educational expansion or the disadvantaged position of women on status IEO.

Appendix: Scaling levels of education

To estimate the scale of education I use the fact that when estimating the effect of education on income one does not need to *a priori* fix the scale of education, one can just add education as a series of dummies. One way to interpret this model is that it simultaneously estimates the scaling and the effect of education. This feature will be used to estimate a scale for the levels of education. This would result in equation (6) if five diplomas are distinguished: *lo*, *lbo_mavo*, *havo_vwo*, *mbo*, and *hbo_wo*.

$$\ln(inc) = \beta_0 + \underbrace{\beta_1}_{0} lo + \beta_2 lbo\_mavo + \beta_3 havo\_vwo + \beta_4 mbo + \beta_5 hbo\_wo + \cdots \tag{6}$$

In order for this to be identified β_1 is constrained to be zero, in other words the dummy for primary education is left out. A scaling of education will measure the relative distances between diplomas. So if the value of primary education is fixed to 0 and that of hbo/wo to 1, then the scaling will assign positions to all other diplomas relative to these two diplomas.

These two constraints will fix the origin at primary education and the unit at the distance between primary and hbo/wo. One way to write this new variable is like equation (7):

$$ed = \underbrace{\alpha_1}_{0} lo + \alpha_2 lbo\_mavo + \alpha_3 havo\_vwo + \alpha_4 mbo + \underbrace{\alpha_5}_{1} hbo\_wo \tag{7}$$

A person with only primary education gets α_1, a person with lbo or mavo gets α_2, etc. In other words, the αs form the scale. If the effect of this scaled education is called γ_1, then the effect of education on income can be written like equation (8).

$$\begin{aligned} \ln(inc) &= \beta_0 + \gamma_1 ed + \cdots \\ &= \beta_0 + \gamma_1(\underbrace{\alpha_1}_{0} lo + \alpha_2 lbo\_mavo + \alpha_3 havo\_vwo + \alpha_4 mbo + \underbrace{\alpha_5}_{1} hbo\_wo) + \cdots \end{aligned} \tag{8}$$

All parameters in model (8) can be calculated from the parameters in model (6). The relationship between the parameters in the two models is given below:

$$\begin{aligned} \gamma_1 &= \beta_5 \\ \alpha_1 &= 0 \\ \alpha_2 &= \frac{\beta_2}{\beta_5} \\ \alpha_3 &= \frac{\beta_3}{\beta_5} \\ \alpha_4 &= \frac{\beta_4}{\beta_5} \\ \alpha_5 &= 1 \end{aligned}$$

Model (8) is thus just a reparameterization of model (6). This relationship becomes more complicated when interactions with education are added. An important interaction to add would be the interaction with time: over time one might expect to need more education to get the same income. If one assumes a scaling of education that is constant over time, then this implies a constraint that the relative distances between diplomas remain the same. Such a model was estimated for men only using data from the ISMF that contain information about the income at the time of the interview[3]. To take care of inflation and the change in currency (Dutch guilders to euros) these incomes were transformed to represent income in terms of euros from 2000. As is common with income data, the natural logarithm is taken to take care of the fact that income tends to be right skewed. Income is explained with the following variables:

- education (measured in a scale that is to be estimated)
- age and age squared (in 10s of years and zero when someone is 40),
- year (in 10s of years and zero in 1958, the earliest survey) added as three splines with knots in 1975 and 1990,
- father's occupational status (in ISEI points, re-scaled to range between zero and one),
- interaction of father's occupational status with year splines, and
- interaction of education with year splines.

The change of effect of education over time is shown in the γ panel in table 5. It shows that in 1958 someone with hbo/wo earned about 47% more than someone with primary education. This remained constant between 1958 and 1975, decreased by about 17 percentage points per 10 years between 1975 and 1990, and increased by about 19 percentage points per 10 years between 1990 and 2005. However, the scaling is constrained to remain constant (the α panel in table 5). This constraint was tested and rejected, but this is not surprising given the large sample size (31,253 respondents). However the difference in BIC scores is 26.09 points smaller for the constrained model, indicating 'very strong' (Raftery, 1995) or 'decisive' (Jeffreys, 1961) evidence in favor of the model with constant scaling of education over time. So the overall conclusion is that the relative distances between diplomas remained unchanged between 1958 and 2005.

[3]Surveys used to estimate the scaling of education are: net58, net67t, net70, net71, net71c, net74p, net76j, net77, net77e, net79p, net81e, net82e, net82n, net82u, net85o, net86e, net86l, net87i, net87j, net87s, net88o, net90, net90o, net91j, net92f, net92o, net92t, net94e, net94h, net94o, net95s, net96, net96c, net96o, net96y, net98, net98e, net98o, net99, net99i, nex00s, nex02e, nex03, nex04i. Codes refer to (Ganzeboom, 2005).

[Table 5 about here.]

An alternative method of scaling education would be to look at the official number of years needed to obtain a diploma. These two scalings are compared in figure 10, by transforming both to the same metric: the mean level of education for the cohort which is 12 in 1970 is fixed to zero and the standard deviation for that same cohort is fixed to one. The two main differences are that 1) mbo and havo/vwo have changed place in such a way that in the estimated scale mbo is less valuable then havo/vwo, and 2) the distance between hbo/wo and the rest is larger in the estimated scale. The estimated scale is preferred since education measured in this scale actually measures a resource — a greater ability to earn income — while education measured in the number of years officially needed to obtain the diploma represents the input rather than how much this input will pay.

[Figure 10 about here.]

References

Agresti, Alan. 2002. *Categorical Data Analysis.* 2nd ed. Hoboken, NJ: Wiley-Interscience.

Allison, Paul D. 1999. "Comparing Logit and Probit Coefficients Across Groups." *Sociological Methods & Research* 28(2):186–208.

Blau, Peter M. and Otis Dudley Duncan. 1967. *The American Occupational Structure.* New York: Wiley.

Boekholt, P. Th. F. M. and E. P. de Booy. 1987. *Geschiedenis van de school in Nederland vanaf de Middeleeuwen tot aan de huidige tijd.* Assen: Van Gorcum.

Boudon, Raymond. 1974. *Education, Opportunity and Social Inequality.* New York: Wiley.

Breen, Richard and Jan O. Jonsson. 2000. "Analyzing Educational Careers: A Multinomial Transition Model." *American Sociological Review* 65(5):754–772.

Cameron, Stephen V. and James J. Heckman. 1998. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males." *The Journal of Political Economy* 106(2):262–333.

Fox, John. 1997. *Applied Regression Analysis, Linear Models, and Related Methods.* Thousand Oaks: Sage.

Ganzeboom, Harry B.G. 2005. "International Stratification and Mobility File." `http://home.scw.vu.nl/~ganzeboom/ismf/index.htm`.

Ganzeboom, Harry B.G. and Donald J. Treiman. 2003. Three Internationally Standardised Measures for Comparative Research on Occupational Status. In *Advances in Cross-National Comparison: A European Working Book for Demographic and Socio-Economic Variables*, ed. Jürgen H.P. Hoffmeyer-Zlotnik and Christof Wolf. Kluwer Academic Press chapter 9, pp. 159–193.

Hout, Michael and Thomas A. DiPrete. 2006. "What we have learned: RC 28's contributions to knowledge about social stratification." *Research in Social Stratification and Mobility* 24(1):1–20.

Jeffreys, Harold. 1961. *Theory of Probability.* 3rd ed. Oxford: Oxford University Press.

Lucas, Samuel R. 2001. "Effectively Maintainted Inequality: Education Transitions, Track Mobility, and Social Background Effects." *American Journal of Sociology* 106(6):1642–1690.

Maddala, G.S. 1983. *Limited Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press.

Mare, Robert D. 1980. "Social Background and School Continuation Decisions." *Journal of the American Statistical Association* 75(370):295–305.

Mare, Robert D. 1981. "Change and Stability in Educational Stratification." *American Sociological Review* 46(1):72–87.

Mare, Robert D. 2006. "Response: Statistical Models of Educational Stratification — Hauser and Andrew's Models for School Transitions." *Sociological Methodology* 11:27–37.

Neuhaus, John M., J.D. Kalbfleisch and W.W. Hauck. 1991. "A Comparison of Cluster-Specific and Population-Averaged Approaches for Analyzing Correlated Binary Data." *International Statistical Review* 59(1):25–35.

Neuhaus, John M. and Nicholas P. Jewell. 1993. "A Geometric Approach to Assess Bias Due to Omited Covariates in Generalized Linear Models." *Biometrika* 80(4):807–815.

Nieuwbeerta, Paul and Susanne Rijken. 1996. "Educational Expansion and Educational Reproduction in Eastern Europe, 1940–1979." *Czech Sociological Review* 4(2):187–210.

Raftery, Adrian E. 1995. "Bayesian Model Selection in Social Research." *Sociological Methodology* 25:111–163.

Shavit, Yossi and Hans-Peter Blossfeld. 1993. *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Boulder: Westview Press.

Smith, Herbert L. and Paul P.L. Cheung. 1986. "Trends in the Effects of Family Background on Educational Attainment in the Philippines." *The American Journal of Sociology* 91(6):1387–1408.

Treiman, Donald J. and Harry B.G. Ganzeboom. 2000. The Fourth Generaton of Comparative Stratification Research. In *The International Handbook of Sociology*, ed. Stellah Ouah and Arnaud Sales. Sage chapter 6, pp. 123–150.

UNESCO. 1997. *ISCED 1997, International Standard Classification of Education*. UNESCO.

van Essen, M. 1990. *Opvoeden met een dubbel doel: twee eeuwen meisjesonderwijs in nederland*. Amsterdam: SUA.

Figure 1: Hypothetical educational system

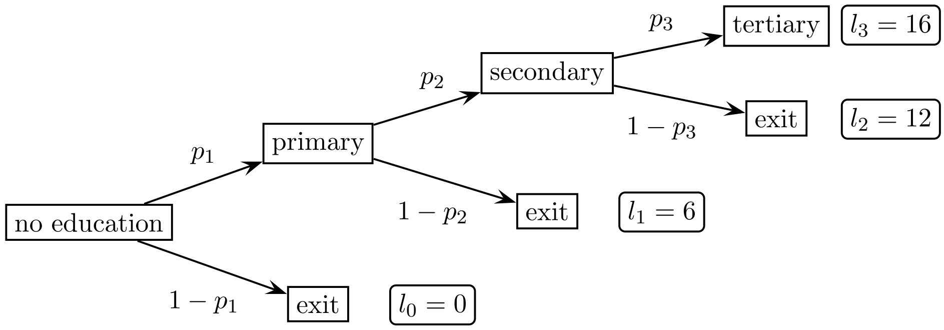

Figure 2: Direct and indirect effect of cohort on status IEO

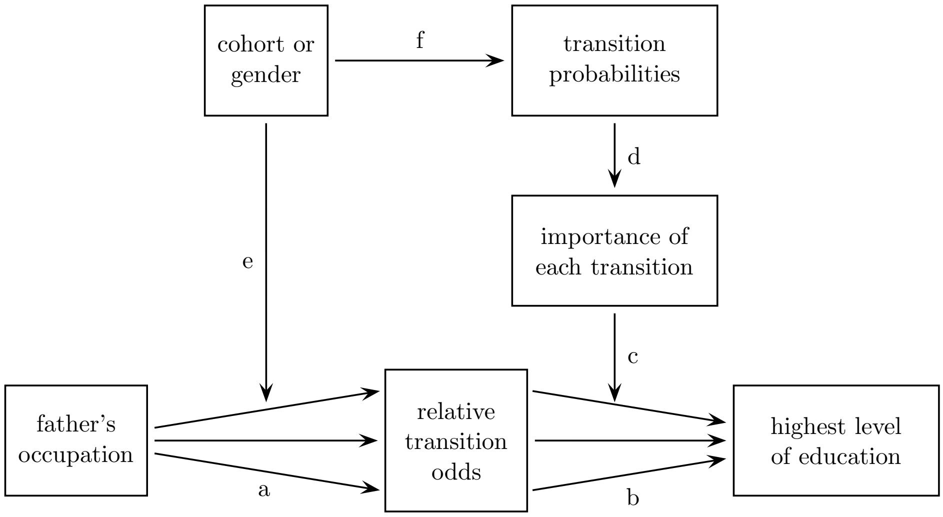

Figure 3: Dutch educational system

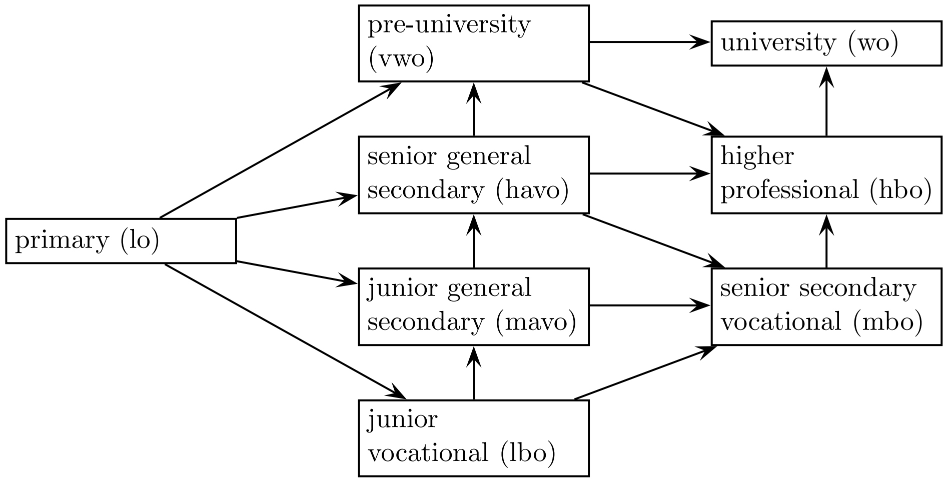

Figure 4: Simplified model of Dutch educational system

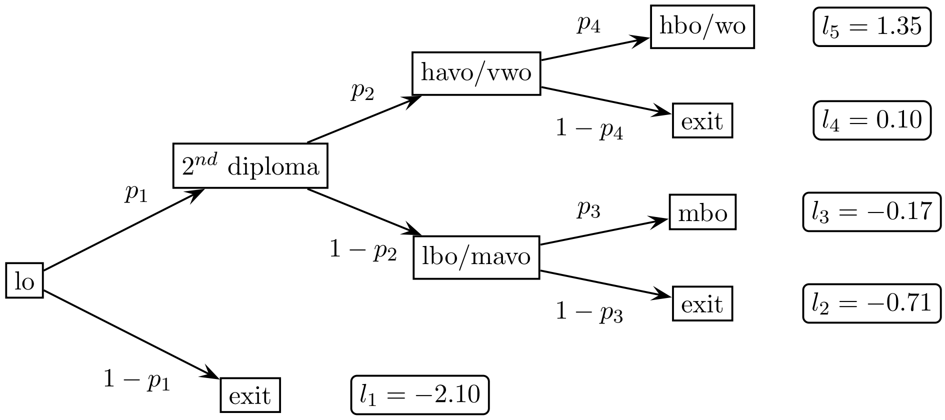

Figure 5: Proportions passing transitions

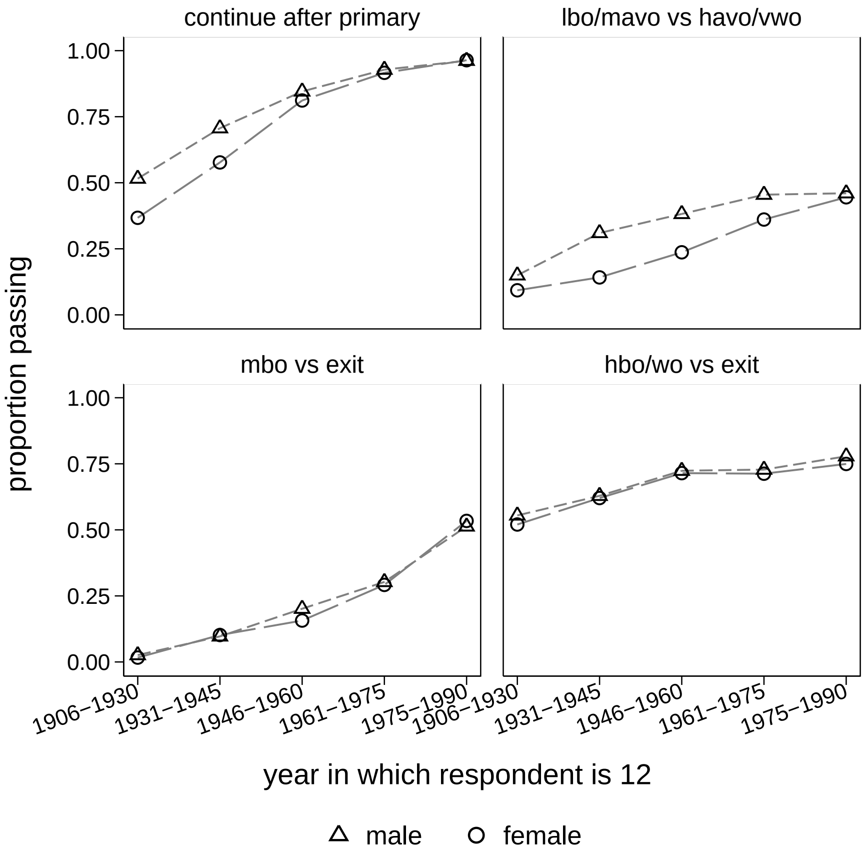

Figure 6: Average level of education according to the sequential response model (black) and OLS regression (gray)

Figure 7: Overall IEO according to the sequential response model and OLS regression

Figure 8: differences in partial IEOs across cohorts for men

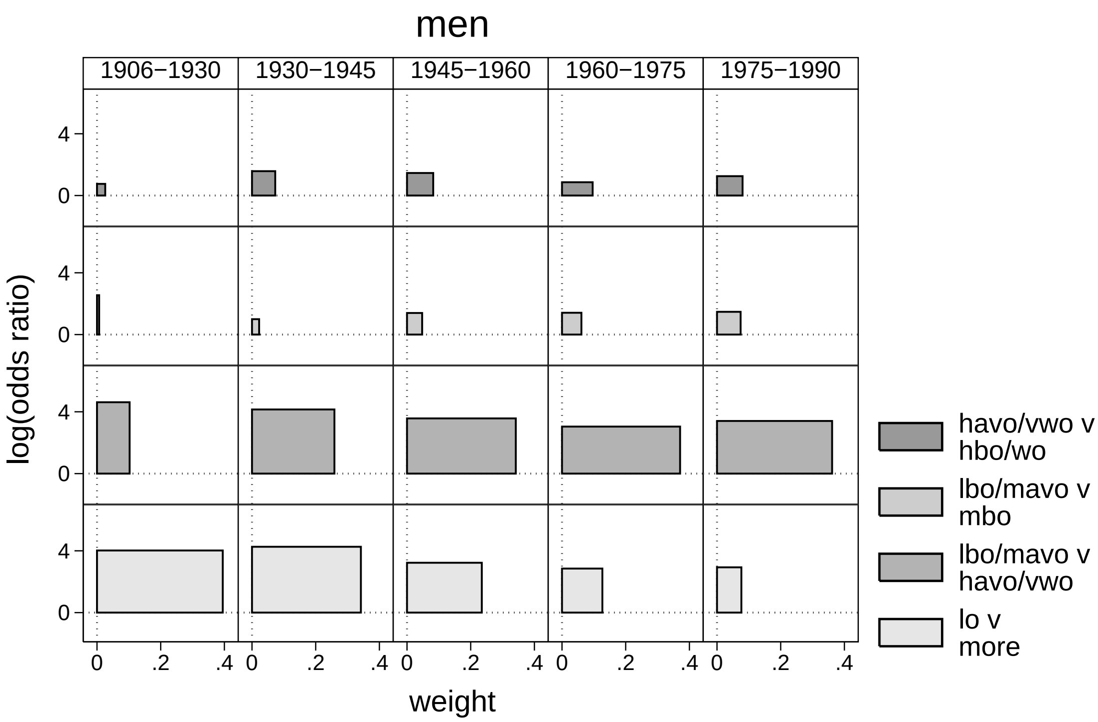

Figure 9: differences in partial IEOs across cohorts for women

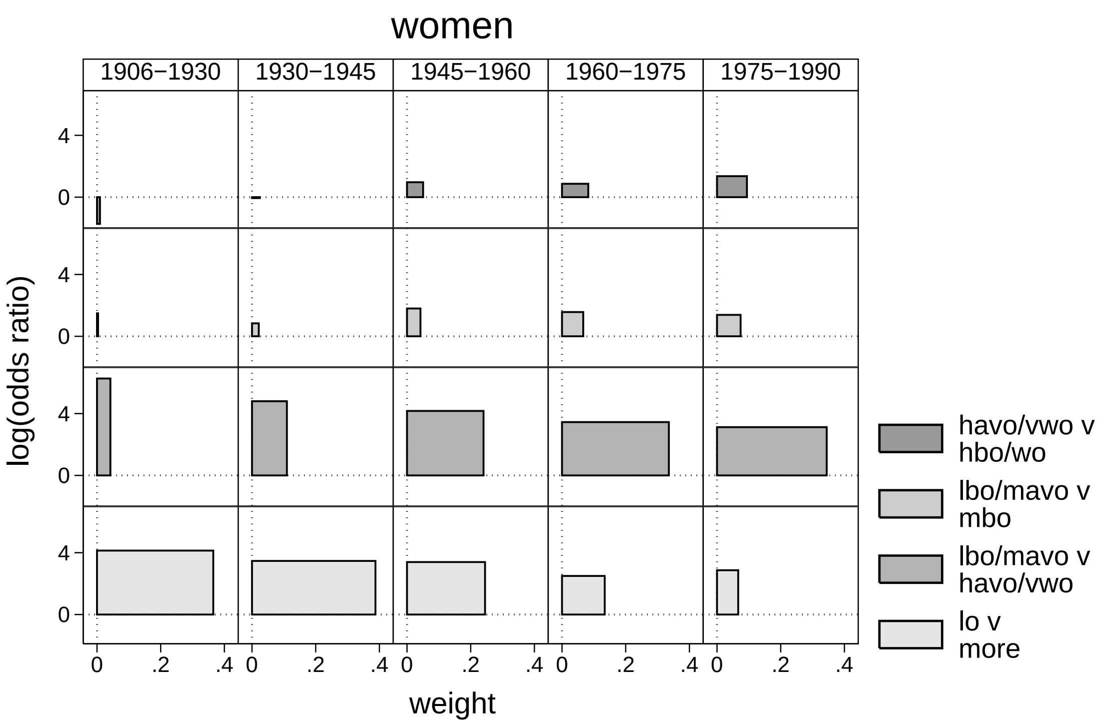

Figure 10: Estimated scale of education versus a scale based on years of education

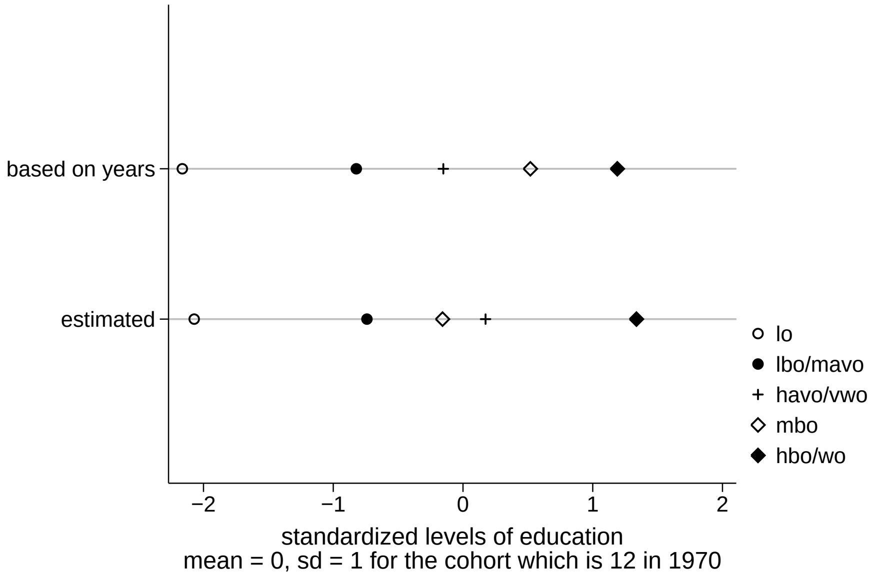

Table 1: Conversion of old educational levels into new educational levels and simplified educational levels

English name	before 1968	after 1968	years†	ISCED
primary	LO	lo	6	1
extended primary	VGLO	-	7	1
junior vocational	LTS / ambachtschool	lbo	10	2C
junior vocational	LHNO / huishoudschool	lbo	10	2C
junior general secondary	ULO / MULO	mavo	9 / 10	2B‡
senior secondary vocational	MTS	mbo	14	3C
senior general secondary	MMS	havo	11	3B‡
pre-university	HBS	vwo	12	3A‡
pre-university	lyceum	vwo	12	3A
pre-university	gymnasium	vwo	12	3A
higher professional	HTS	hbo	15	5B
university	universiteit	wo	16	5A

† Years refer to the situation after 1968 except VGLO.

‡ These levels were originally intended to be terminal levels of education for most students (so 2C or 3C) but evolved into levels that primarily grant access to subsequent levels of education.

Table 2: sequential response model for men

	lo v. continue b/se	lbo/mavo v. havo/vwo b/se	lbo/mavo v. mbo b/se	havo/vwo v. hbo/wo b/se
1930-1945	0.824	1.221	1.890	0.077
	(0.15)	(0.22)	(0.75)	(0.40)
1945-1960	1.981	1.784	2.771	0.715
	(0.14)	(0.21)	(0.72)	(0.38)
1960-1975	2.956	2.233	3.363	1.054
	(0.15)	(0.21)	(0.72)	(0.38)
1975-1990	3.496	1.969	4.187	1.143
	(0.23)	(0.22)	(0.73)	(0.40)
fiseiX1906-1930	4.022	4.620	2.549	0.756
	(0.35)	(0.43)	(1.48)	(0.72)
fiseiX1930-1945	4.262	4.152	0.998	1.578
	(0.23)	(0.19)	(0.49)	(0.29)
fiseiX1945-1960	3.227	3.574	1.398	1.458
	(0.17)	(0.11)	(0.20)	(0.18)
fiseiX1960-1975	2.852	3.042	1.410	0.861
	(0.23)	(0.10)	(0.16)	(0.16)
fiseiX1975-1990	2.933	3.405	1.474	1.255
	(0.47)	(0.16)	(0.23)	(0.27)
constant	-1.331	-3.637	-4.638	-0.219
	(0.13)	(0.21)	(0.72)	(0.37)
ll	-1.19e+04	-2.21e+04	-8932.048	-7897.968
N	36143	35925	16633	14729

Table 3: sequential response model for women

	lo v. continue b/se	lbo/mavo v. havo/vwo b/se	lbo/mavo v. mbo b/se	havo/vwo v. hbo/wo b/se
1930-1945	1.103	1.259	1.913	-0.459
	(0.16)	(0.34)	(1.11)	(0.62)
1945-1960	2.330	2.172	2.158	-0.479
	(0.16)	(0.32)	(1.10)	(0.59)
1960-1975	3.638	3.093	3.177	-0.303
	(0.17)	(0.32)	(1.10)	(0.58)
1975-1990	4.360	3.489	4.374	-0.355
	(0.24)	(0.33)	(1.10)	(0.59)
fiseiX1906-1930	4.135	6.269	1.486	-1.730
	(0.36)	(0.58)	(2.29)	(0.98)
fiseiX1930-1945	3.466	4.801	0.837	-0.057
	(0.19)	(0.24)	(0.47)	(0.40)
fiseiX1945-1960	3.394	4.171	1.799	0.968
	(0.16)	(0.13)	(0.20)	(0.22)
fiseiX1960-1975	2.496	3.449	1.569	0.868
	(0.21)	(0.10)	(0.15)	(0.17)
fiseiX1975-1990	2.866	3.123	1.385	1.359
	(0.47)	(0.15)	(0.22)	(0.24)
constant	-2.028	-5.059	-4.685	0.950
	(0.15)	(0.32)	(1.09)	(0.57)
ll	-1.22e+04	-1.86e+04	-9052.902	-6115.834
N	34616	34353	18168	10784

Table 4: Educational expansion, gender IEO and the weights of transitions

cohort	lo v more					lbo/mavo v havo/vwo				
		components of weight					components of weight			
	p	at risk	variance	gain	weight	p	at risk	variance	gain	weight
men										
1906-1930	0.57	1.00	0.24	1.61	0.39	0.14	0.57	0.12	1.45	0.10
1930-1945	0.77	1.00	0.18	1.93	0.34	0.32	0.77	0.22	1.54	0.26
1945-1960	0.88	1.00	0.11	2.15	0.23	0.40	0.88	0.24	1.63	0.34
1960-1975	0.94	1.00	0.06	2.29	0.13	0.46	0.94	0.25	1.59	0.37
1975-1990	0.97	1.00	0.03	2.33	0.08	0.43	0.97	0.24	1.53	0.36
women										
1906-1930	0.41	1.00	0.24	1.51	0.37	0.07	0.41	0.07	1.50	0.04
1930-1945	0.62	1.00	0.24	1.64	0.39	0.13	0.62	0.12	1.53	0.11
1945-1960	0.84	1.00	0.13	1.84	0.25	0.23	0.84	0.18	1.61	0.24
1960-1975	0.93	1.00	0.06	2.11	0.13	0.36	0.93	0.23	1.56	0.34
1975-1990	0.97	1.00	0.03	2.31	0.07	0.42	0.97	0.24	1.46	0.34

cohort	lbo/mavo v mbo					havo/vwo v hbo/wo				
		components of weight					components of weight			
	p	at risk	variance	gain	weight	p	at risk	variance	gain	weight
men										
1906-1930	0.03	0.49	0.03	0.54	0.01	0.52	0.08	0.25	1.25	0.03
1930-1945	0.09	0.52	0.08	0.54	0.02	0.62	0.25	0.24	1.25	0.07
1945-1960	0.21	0.53	0.17	0.54	0.05	0.75	0.35	0.19	1.25	0.08
1960-1975	0.33	0.51	0.22	0.54	0.06	0.77	0.43	0.18	1.25	0.10
1975-1990	0.54	0.55	0.25	0.54	0.07	0.81	0.41	0.16	1.25	0.08
women										
1906-1930	0.02	0.38	0.02	0.54	0.00	0.56	0.03	0.25	1.25	0.01
1930-1945	0.08	0.53	0.07	0.54	0.02	0.61	0.08	0.24	1.25	0.02
1945-1960	0.14	0.65	0.12	0.54	0.04	0.70	0.19	0.21	1.25	0.05
1960-1975	0.29	0.60	0.21	0.54	0.07	0.73	0.33	0.20	1.25	0.08
1975-1990	0.56	0.56	0.25	0.54	0.07	0.76	0.41	0.18	1.25	0.09

Table 5: Scaling of education

		b	se
α			
	lo	0	.
	lbo/mavo	.391	.017
	mbo	.562	.023
	havo/vwo	.659	.022
	hbo/wo	1	.
γ			
	1958-1975	.060	.050
	1975-1990	-.166	.025
	1990-2005	.192	.027
	constant	.474	.074
other			
	1958-1975	.865	.034
	1975-1990	.347	.019
	1990-2005	.161	.022
	fisei	.496	.125
	1958-1975Xfisei	-.077	.086
	1975-1990Xfisei	-.132	.044
	1990-2005Xfisei	.073	.042
	age	.115	.004
	age2	-.071	.003
	constant	4.88	.049