Two types of stringency of selection in education

Maarten Buis

February 15, 2005

Abstract

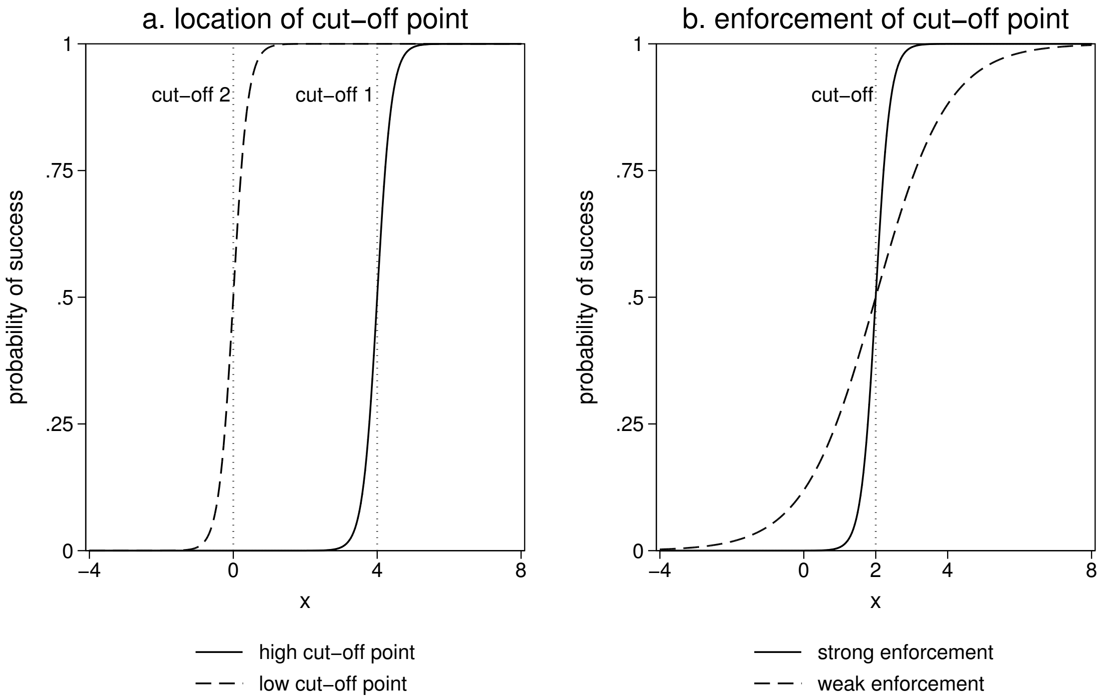
One of the tasks of education is to determine which student gets which diploma. In order to do so, schools select students on characteristics like intelligence, motivation, or father's status. The stringency of this selection contains consists of two components: first the minimal level required to pass a level of education (location of the cut-off value), and second the strength with which this level is enforced (enforcement of the cut-off value). The stringency of selection has often been studied using the Discrete Time Logistic model, but only the results regarding the enforcement were interpreted. This paper shows how to get information regarding the location of the cut-off value from this model. It shows that, in the Discrete Time Logistic model, the functional form of trends in location and enforcement of the cut-off point are related, and illustrates this with data concerning selection on father's status from the Netherlands for birthcohorts 1880-1975.

1 introduction

Education has two tasks: it helps students obtain useful knowledge and skills, and and it gives different students different diplomas according to various characteristics of the students, like intelligence, motivation, or social status. This paper will focus on the latter task of education: the selection of students on these characteristics, and more specifically the historical trend in the stringency of this selection. There are two types of stringency of selection. First, selection entails that a cut-off value is chosen: a student passes if he scores better than the cut-off value and fails if he scores worse than that cut-off value. Selection is less stringent if a lower cut-off value is chosen. Second, this cut-off value has to be enforced. Selection is less stringent if a given cut-off value is less strictly enforced. Within sociology the selection on father's status has long been studied, but the distinction between location and enforcement has been overlooked in this literature.

The distinction between the location (argument 1) and the enforcement (argument 2) of the cut-off value is shown visually in figure 1. It shows the relationship between some variable on which the student is selected (x) and the child's probability of passing a level of education. Graph 1.a shows how, with a given strength of enforcement, the stringency of selection can be decreased by moving the cut-off point to the left. The probability of success for students with a x between 0 and 4 have greatly improved, by moving the cut-off value from 4 to 0. Graph 1.b shows how, with a given cut-off point (in this case 2), the stringency of selection can be decreased by making the enforcement more lax. With weak enforcement, students with a x of less than 2 still have some probability of passing while students with a x of more than 2 still have some probability of failing. With strong enforcement, both the probability that low x children pass and the probability that high x children fail are a lot smaller.

Figure 1: Location and enforcement of the cut-off point

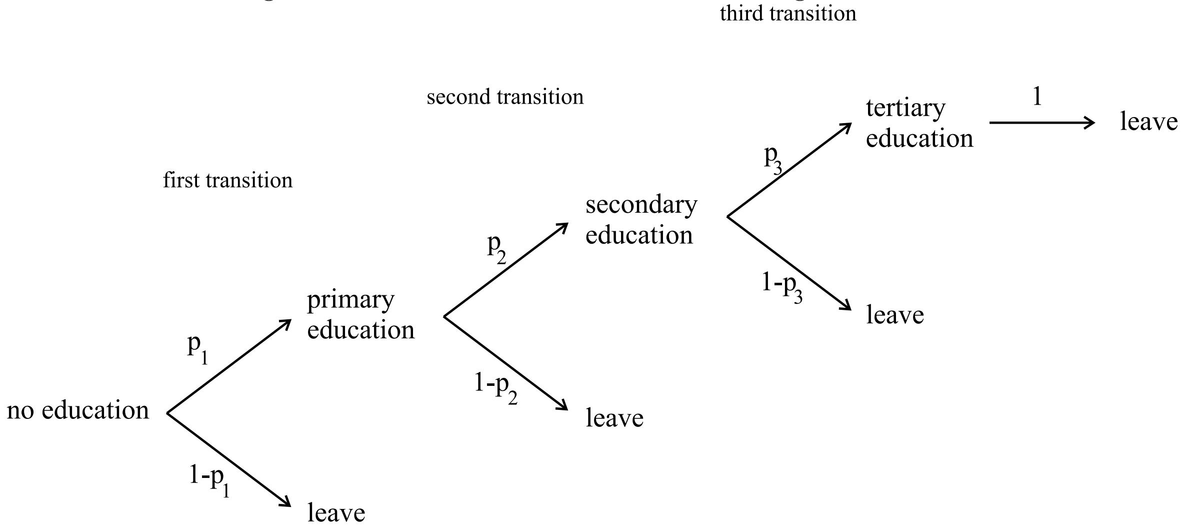


This paper will deal with the following questions: First, how can the location, and the enforcement of the cut-off point and a trend in these values be measured? Second, what are the trends in enforcement and location of the cut-off in the Netherlands for the birth-cohorts 1880-1975? This paper will consist of two parts: First, estimation of the location and the enforcement of the cut-off point will be discussed. Second, this method will be illustrated by estimating the model on data from the Netherlands.

2 Applying and interpreting the Discrete Time Logistic Model for educational transitions

In order to estimate the two forms of stringency of selection — the enforcement and the location of the cut-off point — one must model the relationship between the xs and the child's educational attainment. A frequently used model to estimate this relationship is the Discrete Time Logistic Model (e.g. Mare, 1980, 1981; Blossfeld and Shavit, 1993; Raftery and Hout, 1993). This model treats education as a hierarchically ordered series of levels. A student who achieved a certain level of education is assumed to have passed all 'lesser' levels of education. The Discrete Time Logistic Model estimates the effect of the explanatory variables on the probability of going from one level of education to the next level of education. The usual interpretation of the parameters only relate to enforcement of the cut-off value, but the parameters of the Discrete Time Logistic Model also contain information about the location of the cut-off value. This section will consist of two parts: the first part will describe the

Figure 2: Transitions in a Discrete Time Logistic Model



Discrete Time Logistic Model and the second part will describe how the estimated parameters can be interpreted in terms of the location and the enforcement of the cut-off value.

To estimate a Discrete Time Logistic Model one needs data on the highest achieved level of education, and the values of the explanatory variables. To estimate the effect of the explanatory variables on each transition one makes an assumption about which transitions a person must have passed in order to get to his highest achieved level of education. Figure 2 illustrates this assumption for four levels of education: no education, primary education, secondary education, and tertiary education. Each person starts with no education and can either obtain a diploma in primary education or leave the educational system. A person who obtains a diploma in primary education is said to have passed the first transition, and a person who leaves the educational system before obtaining a diploma in primary education is said to have failed the first transition. The probability of passing the first transition is p_1 and the probability of failing is $1 - p_1$. After obtaining a diploma in primary education, the student can either obtain a diploma in secondary education or be content with his diploma in primary education and leave. In other words: only students who have passed the first transition are 'at risk' of passing the second transition, only they can either pass or fail the second transition. Similarly, only students who have passed the second transition are at risk of passing the third transition. Students who have obtained a diploma in tertiary education have, in this example, a probability of one of leaving the educational system, since tertiary education is the highest possible level education. So a person who reports that his highest achieved level of education is secondary education is assumed to have passed the first and the second transition and has failed the third transition.

The Discrete Time Logistic Model can be used to estimate the effects of variables, like father's socio-economic status, on the probabilities of passing the first, second and third transition. Each transition has only two possible outcomes: passing or failing the transition. For each transition a dummy is created, which is zero for those students that have failed that transition, one for those students that have passed the transition, and missing if those students have failed one of the previous transitions. These dummies can be modelled using the standard techniques for a binary dependent variable. Of these techniques, logistic regression is most often used, both in the general application of this model in discrete time survival analysis (Cox, 1972; Hosmer and Lemeshow, 1999), and in the application of this model to educational transitions (Mare, 1980; Shavit and Blossfeld, 1993). Each transition is thus represented by a logistic regression in order to estimate the effect of the explanatory variables on the probability of passing that transition. By assigning 'missing values' to those students that have failed one of the previous transitions, one can ensure that these logistic regression are only estimated on students that have passed all previous transitions.

This means that the probability of passing a transition is represented by equation (1), where D_k is the dummy for transition k, x an explanatory variable, and β_{0k} and β_{1k} are the parameters for transition k that are to be estimated. Equation (1) basically fits a curve like figure 1 for each transition to the data. The estimated parameters determine the shape and location of curves.

$$\Pr(D_k = 1|D_{k-1} = 1, x) = \frac{e^{\beta_{0k}+\beta_{1k}x}}{1 + e^{\beta_{0k}+\beta_{1k}x}} \quad (1)$$

The parameters can be used to say something about the two aspects of selection — the location of the cut-off point, and the enforcement of the cut-off point. Consider a single transition, where the probability of successfully passing that transition is determined by x through the logistic regression function (1). The value of β_1 influence the steepness of the curve, and enforcement is stronger if the curve is steeper. The exponential of the parameter is easier to interpret. This is an odds ratio, i.e. the factor by which the odds change as a result of a unit change in x. Traditionally, only the odds ratio has been interpreted, even though it represents only the strength of the enforcement of the cut-off value.

The estimated parameters can also be used to get an estimate of the location of the cut-off point. The cut-off point is defined as the minimum status needed to pass if the cut-off point would be perfectly enforced. When enforcement becomes stricter the function gravitates towards the value of x where the probability of success is .5. This is a direct result of the fact that logistic curves are symmetric around this point. This means that the cut-off point is the status where the probability of success is 50%. If the probability of success is 50% than the odds ratio of success is 1 and the log odds ratio is zero. Finding the location of the cut-off is simply a case of solving $\beta_0 + \beta_1 x = 0$, since equation (1) can be rewritten as a linear function of the log odds ratio with the same βs. This means equation (2) represents the cut-off value. Note that this estimate does not exist when β_1 is zero, and that this estimate will become unstable when β_1 is close to zero. However, this makes substantive sense: it is very difficult to see what the cut-off value is when it is not or very weakly enforced. Similarly a linear trend in the *location* of the cut-off value is estimated by including a variable representing the year of birth of the respondent (t). Solving $\beta_0 + \beta_1 x + \beta_2 t = 0$ results in a trend according to equation (3). This implies that a linear trend in the *location* of the cut-off value exist only when there is no trend in the *enforcement* of the cut-off value. The strength of the enforcement of the cut-off point remains unchanged over time, and is still represented by β_1. A linear trend in the *enforcement* of the cut-off value can be estimated by adding interaction term between x and t. The location of the cut-off can now be found by solving $\beta_0 + \beta_1 x + \beta_2 t + \beta_3 xt = 0$, which is shown in equation (4). A consequence of a linear trend in the enforcement of the cut-off value is a non-linear trend in the *location* of the cut-off value. This leads to the following conclusion: estimating a linear trend in the cut-off value implies estimating no trend in the enforcement, and estimating a linear trend in the enforcement implies estimating a non-linear trend in the location.

$$x_{\text{cut}-\text{off}} = -\frac{\beta_0}{\beta_1} \quad (2)$$

$$x_{\text{cut}-\text{off}} = -\frac{\beta_0 + \beta_2 t}{\beta_1} \quad (3)$$

$$x_{\text{cut}-\text{off}} = -\frac{\beta_0 + \beta_2 t}{\beta_1 + \beta_3 t} \quad (4)$$

however, the direction and size of the trends in both location and enforcement are largely independent, even though the functional form of the trend in location or enforcement fixes the functional form for the trend in the other parameter.[1] Consider for example a linear trend in enforcement, which forces the trend in location to be equation (4). This trend is either increasing or decreasing, and has a horizontal and a vertical asymptote. With a given trend in enforcement, the parameters β_0 and β_2 determine whether the trend in location is increasing or decreasing, how fast it is increasing or decreasing, and the horizontal asymptote. The trend in enforcement determines the vertical asymptote of the trend in location, i.e. the location can not be determined when the enforcement is zero. The independence of the trend in location and enforcement of the cut-off point is even clearer when both are estimated with a discrete trend. This is done by dividing the respondents into different cohorts and, with the use of dummy variables, estimate the cut-off values and the enforcement of the cut-off values for each individual cohort. This amounts to estimating for each cohort an enforcement (β_1) and a location ($-\frac{\beta_0}{\beta_1}$), so given an enforcement any location is possible. Any discrete trend in both the location and the enforcement of the cut-off value are now possible, and the trend in one does not depend on the trend in the other.

So, trends in location and enforcement can be estimated with the conventional Discrete Time Logistic Model. The only difference is the interpretation of the parameters. Traditionally one has only considered the odds ratio, which is a measure of the enforcement of the cut-off point, and ignored the location of the cut-off point. This section has shown how the location of the cut-off point can also be extracted from the parameters. It has also shown that the functional forms of the trends in the location and enforcement of the cut-off point are connected. One assumes that there is no trend in the enforcement of the cut-off point whenever one estimates a linear trend in the location of the cut-off point, and one assumes that there is a nonlinear trend in the location of the cut-off point whenever one estimates a linear trend in the enforcement of the cut-off point. The sign and size of the two trends are however largely independent.

3 Data and results from the Netherlands

The non-linearity of the estimated trend in the location of the cut-off point due to linearity of the estimated trend in the enforcement of the cut-off point can best be illustrated with data that spans a long period of time. Selection on father's status in the Netherlands is in this respect a good example since there is data spanning almost one hundred years. Empirical research on selection on father's status in education and a possible trend in selection has a long history (see: (Ganzeboom et al., 1991) and (Treiman and Ganzeboom, 2000) for a review). It has also sparked a fair amount of theories predicting a decline in selection on father's status,

[1] An obvious exception is a linear trend in location, which forces the trend in enforcement to be zero.

e.g. modernization theory (Blossfeld and Shavit, 1993; Rijken, 1999), the age effect and educational expansion (Mare, 1981), no decline in this type of selection, e.g. Relative Risk Aversion (Breen and Goldthorpe, 1997; Goldthorpe, 1996), or a decline in this type of selection if certain conditions are met, e.g. Human Capital Theory (Becker and Tomes, 1979, 1986; Becker, 1989), and Maximally Maintained Inequality (Raftery and Hout, 1993). Selection on father's status is interpreted as inequality of educational opportunity. However, both the empirical and the theoretical literature have ignored the distinction between the location and the enforcement of the cut-off value, while these two aspects of selection have very different interpretations. Lowering the enforcement of the cut-off value reduces inequality since those students who do not meet the 'required level of father's status' are less disadvantaged, while those students that meet the required level are less advantaged, so father's status has become increasingly irrelevant. Lowering the location of the cut-off value reduces the consequences of inequality since more students meet the 'required level of father's status', so less students suffer from the selection process.

A Discrete Time Logistic Model will be estimated using data from the Netherlands from respondents that were born between 1880 and 1975. The data come from the International Stratification and Mobility File (ISMF) (Ganzeboom, 2004). The data for the Netherlands consists of 41 surveys, held between 1958 and 2000. These different surveys were merged to increases the number of respondents and the time-span covered, and to diminish the effect of idiosyncracies of individual surveys. The job of the father is measured 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 range from 1 to 9. The average of the recoded variable has increased over time from about 3.5 to 4.5. Time has been measured by the year of birth for the respondents, which ranges from 1880 to 1975. This variable has been recoded to range from -0.2 to 0.75, so a unit change in the variable birthyear is equivalent to one hundred years. The level of education is measured by the highest achieved level of education. Figure 3 shows how the proportions of respondents for each level of education has developed over time. The graph is more erratic in the early years since they are based on a lot less observations (10-20 compared to a maximum of 2,267 for cohort 1942). It shows that lower education (LO) has declined sharply as highest achieved level of education, while HAVO, VWO, MBO (primarily MBO) and HBO and WO have become more important. In order to estimate a Discrete Time Logistic Model, assumptions have to be made about the transition structure. These assumptions are presented in figure 4. The resulting proportions of successes and the number of respondents at risk at different cohorts are reported in table 1. It shows the proportions of successes during the first two transitions have increased substantially. Interestingly, the proportion of successes have remained approximately constant for the third transition. However, the number of students at risk has greatly increased due to the increased probability of passing the first two transitions. So, the number of students going to higher tertiary education has still risen sharply even though the proportion of students continuing to higher tertiary education has remained constant. Table 1 also shows that the cohort 1880-1895 should be treated with care because it contains few respondents. Furthermore, the proportion of failures in the first transition has become very small in recent cohorts.

A Discrete Time Logistic Model with a linear trend in the enforcement of the cut-off value, and a Discrete Time Logistic Model with dummies for the different cohorts has been estimated using this data. The estimated parameters are shown in tables 2 and 3 in appendix A. These parameters are transformed to the parameters of interest — trends in the enforcement and location of the cut-off value. The results are shown in figure 5. For instance, applying

Table 1: proportion of successes and number of children at risk

	transition 1		transition 2		transition 3	
cohort	proportion	n[a]	proportion	n[a]	proportion	n[a]
1880-1885	0.15	70.4	0.72	10.8	0.74	7.8
1885-1890	0.25	116.3	0.28	28.9	0.90	8.2
1890-1895	0.22	240.3	0.32	53.4	0.61	17.3
1895-1900	0.27	465.3	0.41	127.5	0.41	52.3
1900-1905	0.35	754.8	0.43	263.9	0.45	113.9
1905-1910	0.41	1196.1	0.38	489.1	0.49	187.7
1910-1915	0.43	1666.7	0.40	718.6	0.42	291.0
1915-1920	0.53	2199.1	0.43	1176.0	0.43	501.0
1920-1925	0.61	3074.9	0.45	1877.1	0.39	852.9
1925-1930	0.64	3792.1	0.46	2417.6	0.43	1104.9
1930-1935	0.70	4447.3	0.46	3128.7	0.45	1454.3
1935-1940	0.79	4734.6	0.49	3733.7	0.47	1813.5
1940-1945	0.84	5826.0	0.50	4902.9	0.47	2434.6
1945-1950	0.88	7314.9	0.53	6449.7	0.46	3411.4
1950-1955	0.91	6463.7	0.60	5883.9	0.47	3554.7
1955-1960	0.95	6090.6	0.66	5762.8	0.45	3822.6
1960-1965	0.96	5147.2	0.70	4957.2	0.42	3490.2
1965-1970	0.97	3456.5	0.76	3346.3	0.42	2543.4
1970-1975	0.98	1060.1	0.78	1037.3	0.44	813.8
total	0.80	58116.8	0.57	46365.5	0.45	26475.5

[a] fractions are the result of sampling weights

Figure 3: Trend in the distribution of highest achieved level of education

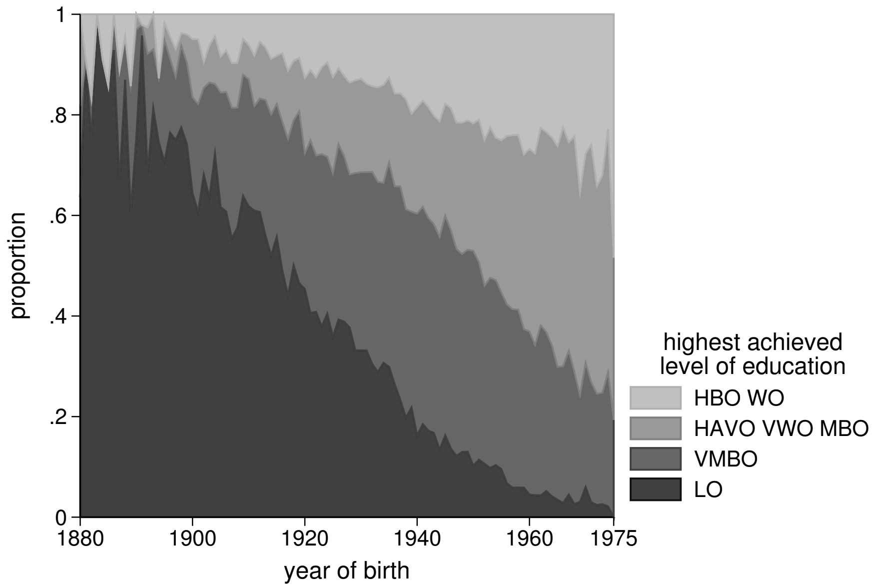
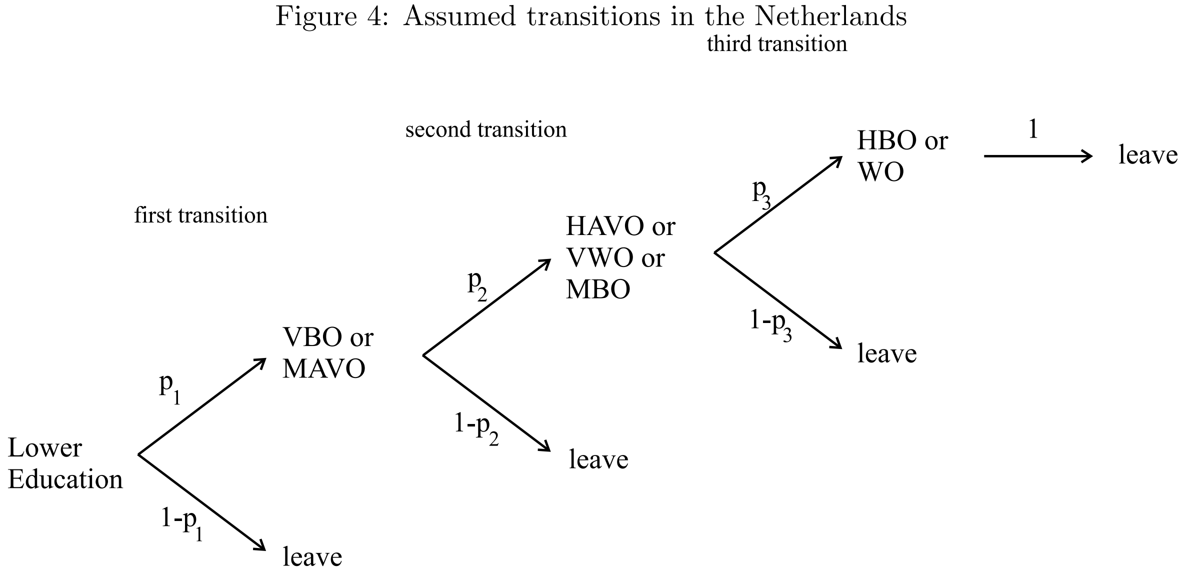


Figure 4: Assumed transitions in the Netherlands



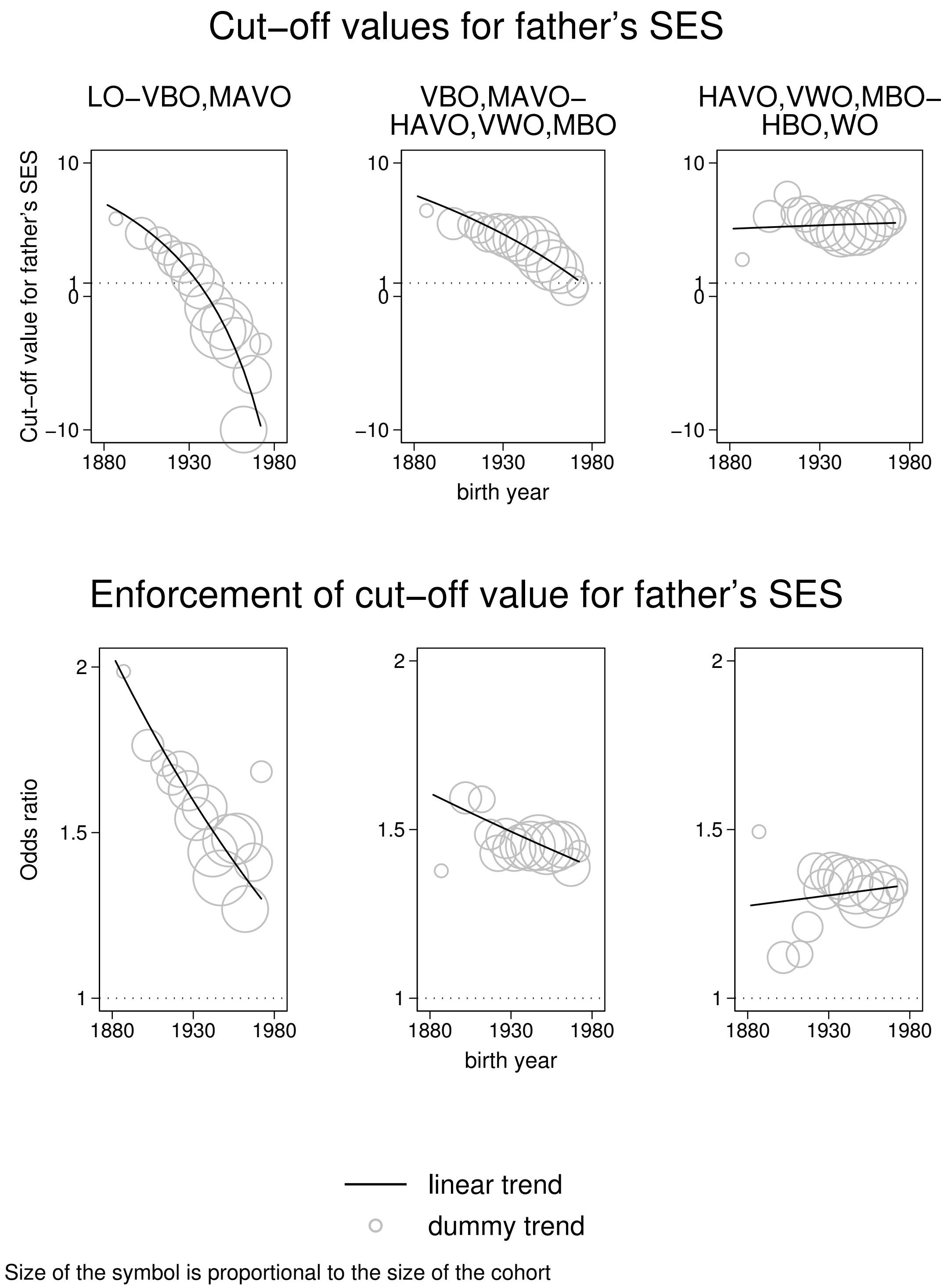
equation (4) to the estimates for the first transition in table 2 means that the cut-off value for father's status (fisei) is $-\frac{-3.341+8.176\times\text{birthyear}}{0.614-0.489\times\text{birthyear}}$, which is the solid line in the upper left graph of figure 5. The solid line in the lower left graph, representing the enforcement of the cut-off point, is $e^{0.614-0.489\times\text{birthyear}}$. The first transition shows a clear and strong downward trend in both the location and the enforcement of the cut-off value. The cut-off value has even left the range of father's status (1 till 9). The reason for this is that almost everyone passes the first transition. The cut-off value is not only the minimum value of fisei needed to pass the level if this value was perfectly enforced, but also the value at which 50% of the children pass if enforcement is not perfect. The downward trend in both the location and the enforcement of the cut-off value is less extreme for the second transition, although the trend in cut-off value is still very strong. The cut-off value has moved from almost the maximum value to almost the minimum value. The trend during the third transition is for both the cut-off value and the enforcement of the cut-off value slightly positive. Results from earlier studies assessing the enforcement of the cut-off value in the Discrete Time Logistic have also shown a large effect of father's status during the first transition, and ever smaller effects for subsequent transitions (Mare, 1980, 1981; Shavit and Blossfeld, 1993).

The striking point is that trends in the location and the enforcement of the cut-off point show a remarkable similar pattern. The similarity of trends in enforcement and location of the cut-off value looks suspicious. In the previous section it was analytically shown that especially the two estimated discrete trends are independent. This means that the similarity in trend is an empirical finding and not an artefact of the model. Several simulations were run to see if this similarity in trends also occurs in simulated datasets with dissimilar 'true' trends, to further test this assertion. The results presented in Appendix B show that different trends in location and enforcement of the cut-off point can be estimated. This is even the case if the model is somewhat misspecified. This gives additional support for the conclusion that the similarity in trend is an empirical finding and not an artefact.

4 Conclusion

This paper promised to answer two questions. 1) How can the location and the enforcement of the cut-off point and a trend in these values be measured? 2) What are the trends in enforcement and location of the cut-off in the Netherlands for the birth-cohorts 1880-1975? The answer to the first question is that the both enforcement and the location of the cut-off values can be estimated using conventional Discrete Time Logistic Model, using simple transformations of the parameters. The enforcement can be measured with the usual odds ratio, and the cut-off value can be measured with the transformation in equation. These transformations have also led to the conclusion that the functional form of the trend in the location and the enforcement of the cut-off value are connected. There can only be a linear trend in the location if there is no trend in the enforcement, and there is a nonlinear trend in the location if there is a linear trend in the enforcement. The answer to the second question is that there are strong negative trends in both the location and the enforcement of the cut-off value during the first transition, the trends during the second transition are still negative but less strong than during the first transition, and the trends are slightly positive during the third transition. So educational opportunities have become more equal during the first two transitions and have become slightly less equal in during the last transition. Another conclusion is that the trends in the location and the enforcement of the cut-off point are surprisingly similar. Analytical results and simulations show that the Discrete Time Logistic Model is capable of estimating different trends in location and enforcement, so the similarity in the trends seems to be a new empirical finding.

Figure 5: Trends in cut-off point and enforcement of the cut-off point



A Parameter estimates

Table 2: linear trend in the enforcement of cut-off point

	transition 1		transition 2		transition 3	
	b	z	b	z	b	z
fisei	0.614	32.00	0.446	22.35	0.252	9.80
birthyear	8.176	40.13	3.472	18.39	-0.382	-1.51
fisei*birthyear	-0.489	-9.73	-0.147	-3.56	0.048	0.97
Constant	-3.341	-43.03	-2.919	-31.73	-1.303	-10.03
log likelihood	-26439		-34929		-21116	
df	3		3		3	

Table 3: dummy trend

	transition 1		transition 2		transition 3	
	b	z	b	z	b	z
fisei(1880-1895)	0.686	8.48	0.320	2.55	0.401	2.00
fisei(1895-1910)	0.567	18.86	0.466	10.32	0.114	1.82
fisei(1910-1915)	0.537	14.87	0.464	9.58	0.123	1.83
fisei(1915-1920)	0.507	15.52	0.396	10.00	0.192	3.47
fisei(1920-1925)	0.526	18.19	0.358	11.87	0.320	7.59
fisei(1925-1930)	0.486	18.46	0.388	14.34	0.280	7.69
fisei(1930-1935)	0.433	17.40	0.365	15.93	0.314	9.97
fisei(1935-1940)	0.456	16.41	0.370	17.38	0.307	10.71
fisei(1940-1945)	0.366	13.81	0.372	19.90	0.297	12.05
fisei(1945-1950)	0.310	12.07	0.379	23.06	0.287	13.71
fisei(1950-1955)	0.389	11.90	0.368	20.54	0.252	12.31
fisei(1955-1960)	0.393	9.19	0.372	19.91	0.289	14.71
fisei(1960-1965)	0.239	4.69	0.375	18.23	0.267	13.15
fisei(1965-1970)	0.344	5.06	0.328	12.24	0.290	11.73
fisei(1970-1975)	0.521	3.10	0.362	7.40	0.280	6.64
cohort 1895-1910	1.316	3.46	-0.483	-0.71	0.423	0.39
cohort 1910-1915	1.737	4.49	-0.430	-0.63	0.170	0.16
cohort 1915-1920	2.232	5.85	0.027	0.04	-0.091	-0.09
cohort 1920-1925	2.523	6.72	0.387	0.58	-0.867	-0.83
cohort 1925-1930	2.767	7.42	0.235	0.36	-0.431	-0.42
cohort 1930-1935	3.300	8.89	0.401	0.61	-0.541	-0.52
cohort 1935-1940	3.659	9.79	0.453	0.69	-0.454	-0.44
cohort 1940-1945	4.301	11.53	0.517	0.79	-0.322	-0.31
cohort 1945-1950	4.799	12.84	0.581	0.89	-0.364	-0.35
cohort 1950-1955	4.805	12.62	0.955	1.47	-0.161	-0.16
cohort 1955-1960	5.365	13.56	1.190	1.82	-0.446	-0.44
cohort 1960-1965	6.376	15.23	1.324	2.03	-0.458	-0.45
cohort 1965-1970	6.009	13.21	1.812	2.75	-0.607	-0.59
cohort 1970-1975	5.852	7.88	1.807	2.64	-0.525	-0.50
Constant	-3.995	-11.11	-2.061	-3.18	-1.109	-1.09
log likelihood	-26419		-34802		-21057	
df	29		29		29	

B Simulation Results

One issue is whether the Discrete Time Logistic Model can distinguish trends in location and enforcement of the cut-off point. Simulation is one way to assess whether these different trends can be disentangled. A dataset was created with 10,000 'observations'[2] whereby each observation received a value of x as a draw from a standard normal distribution and a value of t as a draw from a uniform distribution between 0 and 1. These values of x and t where used to determine the probability that the individual passed a transition. The probability of successfully completing each transition conditional on having passed the previous transition is represented in equation (5). Note that there is no interaction term between t and x, so there is a linear trend in the location of the cut-off point but no trend in the enforcement. Whether or not an individual passed a transition was determined at random, given the calculated probability of success. After the data was generated in this way, the parameters were re-estimated using a Discrete Time Logistic Model with x, t, and an interaction term between x and t. The difference in the location and the enforcement when $t = 1$ and $t = 0$ is recorded. The data were created in such a way that the true values are 0.5 for the change in location and 0 for the change in enforcement. This process was repeated 1,000 times, and the results are presented in figure 6. This figure shows that the results are centered around their true value during each transition, and that the spread has increased over the transitions. This last finding is easily explained by the fact that each higher transition will contain less observations. So, this simulation shows that different trends in location and enforcement of the cut-off value can be disentangled if the Mare-model is correctly specified.

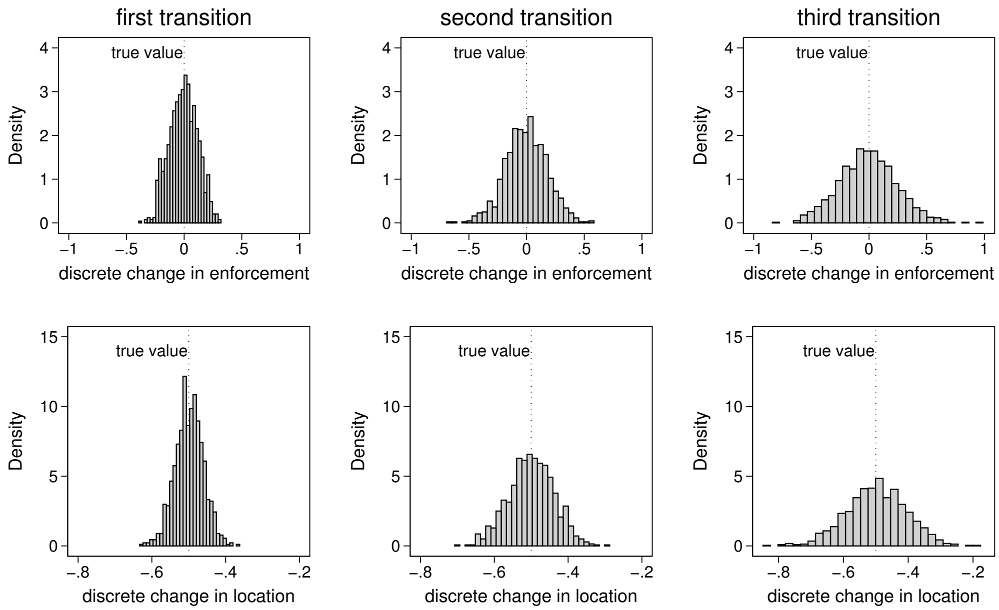
$$\Pr(D_k = 1 | D_{k-1} = 1) = \frac{e^{.5+x+.5t}}{1 + e^{.5+x+.5t}} \tag{5}$$

This begs the question whether this is also true if the model is somewhat incorrectly specified, for instance if a third variable is erroneously left out of the model. To answer this question datasets were created using the conditional probabilities of passing a transition as represented by equation (6). Equation (6) shows that a third variable (x_2) influences the probability of success, but this variable will be ignored during estimation. This creates unobserved heterogeneity. The same Discrete Time Logistic Model as before, which is now misspecified, was estimated on these datasets. The results are shown in graph 7. The trends in location and enforcement of the cut-off value remain remarkably accurate, even in the misspecified model. So, the Discrete Time Logistic Model can disentangle trends in location and enforcement of the cut-off value even if it is somewhat misspecified.

$$\Pr(D_k = 1 | D_{k-1} = 1) = \frac{e^{.5+x_1+.5t+x_2}}{1 + e^{.5+x_1+.5t+x_2}} \tag{6}$$

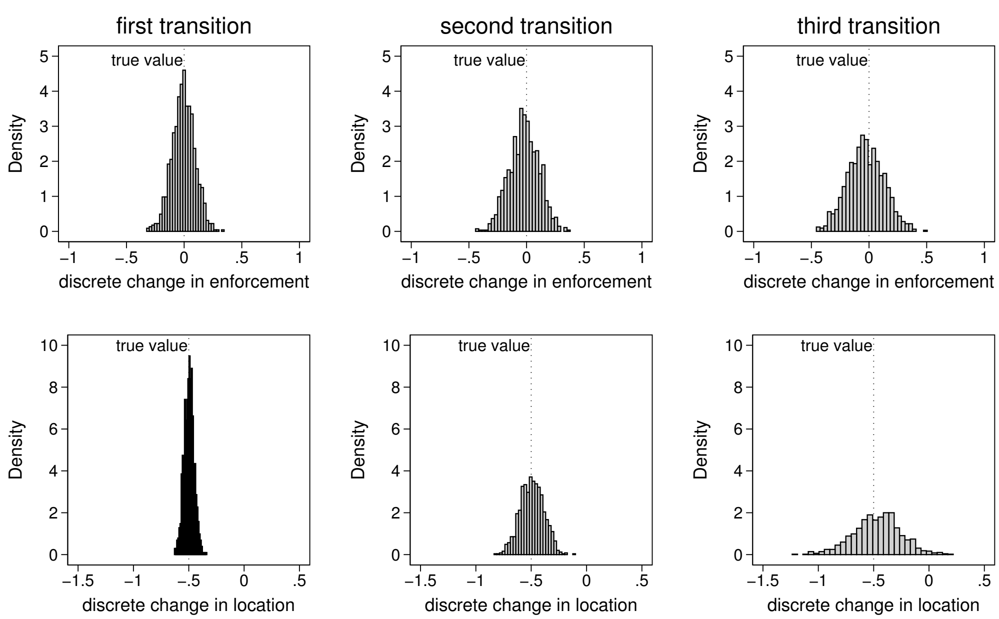
[2]The dataset which will used in the empirical example contains approximately 40,000 observations.

Figure 6: Simulation results when model is correctly specified



1,000 datasets containing 40,000 observations each

Figure 7: Simulation results with an omitted variable



1,000 datasets containing 40,000 observations each

C Studies part of ISMF

name of study[a]	men	women	Total
net58	662	807	1,469
net67t	245	302	547
net70	1,230	1,020	2,250
net71	1,202	995	2,198
net74p	622	622	1,244
net76j	717	136	853
net77	2,290	2,277	4,568
net77e	991	1,112	2,103
net79p	952	932	1,884
net81e	1,232	1,427	2,659
net82e	839	935	1,773
net82n	1,374	1,397	2,771
net82u	601	306	907
net85o	1,586	1,520	3,106
net86e	915	1,005	1,920
net86l	2,143	2,274	4,417
net87i	1,032	1,198	2,230
net87j	484	495	979
net87s	533	553	1,085
net88o	1,053	804	1,857
net90o	919	727	1,646
net90s	1,419	1,306	2,725
net91j	1,113	999	2,112
net92f	1,147	1,098	2,245
net92o	844	720	1,564
net94e	994	1,135	2,129
net94h	641	604	1,245
net94o	689	653	1,343
net95h	1,311	1,278	2,590
net95y	890	827	1,717
net96	461	466	927
net96c	984	1,290	2,275
net96o	981	925	1,906
net96y	256	350	606
net98	580	546	1,126
net98e	1,209	1,283	2,492
net98f	1,263	1,272	2,535
net98o	1,400	1,175	2,574
net99	1,798	1,336	3,134
Total	39,603	38,108	77,711

[a]A more detailed description of these studies can be found at (Ganzeboom, 2004)

References

Becker, Gary S., "On the Economics of the Family: Reply to a Skeptic," *The American Economic Review*, June 1989, *79* (3), 514–518.

_ **and Nigel Tomes**, "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *The Journal of Political Economy*, December 1979, *87* (6), 1153–1189.

_ **and** _ , "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, july 1986, *4* (3, Supplement), S1–S39.

Blossfeld, Hans-Peter and Yossi Shavit, "Persisting Barriers: Changes in Educational Opportunities in Thirteen Countries," in Yossi Shavit and Hans-Peter Blossfeld, eds., *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*, Boulder: Westview Press, 1993, chapter 1, pp. 1–23.

Breen, Richard and John H. Goldthorpe, "Explaining Educational Differentials, Towards a Formal Rational Action Theory," *Rationality and Society*, 1997, *9* (3), 275–305.

Cox, David R., "Regression Models and Life-Tables," *Journal of the Royal Statistical Society, Series B (Methodological)*, 1972, *34* (2), 187–220.

Ganzeboom, Harry B.G., "International Stratification and Mobility File," 2004. `http://home.scw.vu.nl/~ganzeboom/ismf/index.htm`.

_ **and Donald J. Treiman**, "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," Kluwer Academic Press, 2003, chapter 9, pp. 159–193.

_ **,** _ **, and Wout Ultee**, "Comparative Intergenerational Stratification Research: Three Generations and Beyond," *Annual Review of Sociology*, 1991, *17*, 277–302.

Goldthorpe, John H., "Class Analysis and the Reorientation of Class Theory: The Case of Persisting Differentials in Educational Attainment," *The British Journal of Sociology*, 1996, *47* (3), 481–505.

Hosmer, David W. and Stanley Lemeshow, *Applied Survaval Analysis: Regression Modeling of Time to Event Data* Wiley Series in Probabilty and Statistics, New York: John Wiley & Sons, Inc., 1999.

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

_ , "Change and Stability in Educational Stratification," *American Sociological Review*, February 1981, *46* (1), 72–87.

Raftery, Adrian E. and Michael Hout, "Maximally Maintained Inequality: Expansion, Reform, and Opportunity in Irish Education, 1921–75," *Sociology of Education*, January 1993, *66* (1), 41–62.

Rijken, Susanne, "Education Expansion and Status Attainment: A Cross-National and Over-Time Comparison." PhD dissertation, ICS, Utrecht University, Utrecht 1999.

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

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