Linking process to outcome

Inequality of educational opportunities and inequality of educational outcomes

Maarten L. Buis

Institut für Soziologie
Universität Tübingen

Two types of educational inequality

The difference between high and low status children in

Two types of educational inequality

The difference between high and low status children in

- probabilities of passing transitions between levels of education;

Two types of educational inequality

The difference between high and low status children in

- probabilities of passing transitions between levels of education; Inequality of Educational Opportunity (IEOpp), or

Two types of educational inequality

The difference between high and low status children in

- probabilities of passing transitions between levels of education; Inequality of Educational Opportunity (IEOpp), or
- highest achieved level of education;

Two types of educational inequality

The difference between high and low status children in

- probabilities of passing transitions between levels of education; Inequality of Educational Opportunity (IEOpp), or
- highest achieved level of education; Inequality of Educational Outcome (IEOut).

Two types of educational inequality

The difference between high and low status children in

- probabilities of passing transitions between levels of education; Inequality of Educational Opportunity (IEOpp), or
- highest achieved level of education; Inequality of Educational Outcome (IEOut).
- The aims of this presentation are to:
  - relate IEOut to the IEOpps,

Two types of educational inequality

The difference between high and low status children in

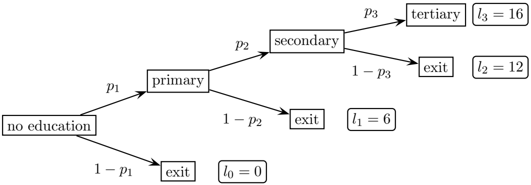
- probabilities of passing transitions between levels of education; Inequality of Educational Opportunity (IEOpp), or
- highest achieved level of education; Inequality of Educational Outcome (IEOut).
- The aims of this presentation are to:
  - relate IEOut to the IEOpps,
  - relate educational expansion to IEOut.

Outline

Outline

Example

Figure: Hypothetical educational system



The dominant model: the sequential logit or Mare model

- A series of logistic regressions on the transition probabilities

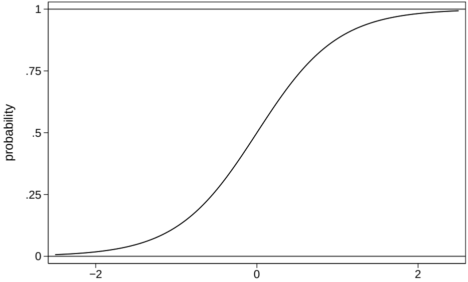
The dominant model: the sequential logit or Mare model

- A series of logistic regressions on the transition probabilities
- The aim is to explain the probability of passing a transition.

The dominant model: the sequential logit or Mare model

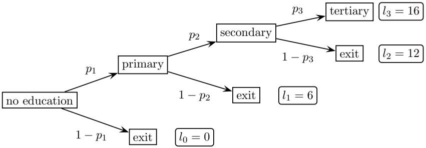
- A series of logistic regressions on the transition probabilities
- The aim is to explain the probability of passing a transition.

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



Modeling transition probabilities and the expected level of education

$$E(ed) = (1 - \hat{p}_{1i})l_0 + \hat{p}_{1i}(1 - \hat{p}_{2i})l_1 + \hat{p}_{1i}\hat{p}_{2i}(1 - \hat{p}_{3i})l_2 + \hat{p}_{1i}\hat{p}_{2i}\hat{p}_{3i}l_3$$



IEOpps and IEOut

- IEOut is the increase in expected highest achieved level of education for a unit increase in SES, i.e. a first derivative:

IEOpps and IEOut

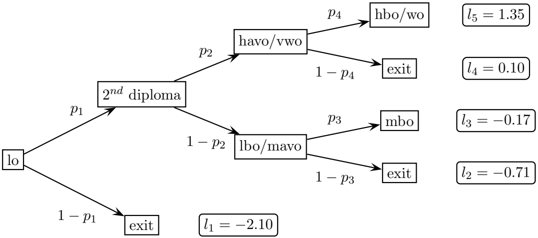
- IEOut is the increase in expected highest achieved level of education for a unit increase in SES, i.e. a first derivative:
- IEOut = weighted sum of IEOpps

IEOpps and IEOut

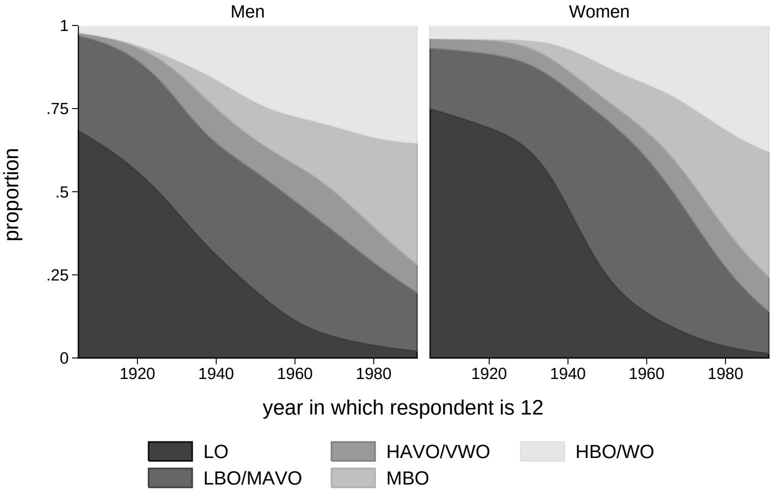
- IEOut is the increase in expected highest achieved level of education for a unit increase in SES, i.e. a first derivative:
- IEOut = weighted sum of IEOpps
- weights = at risk $\times$ variance $\times$ gain

Outline

Simplified model of Dutch educational system



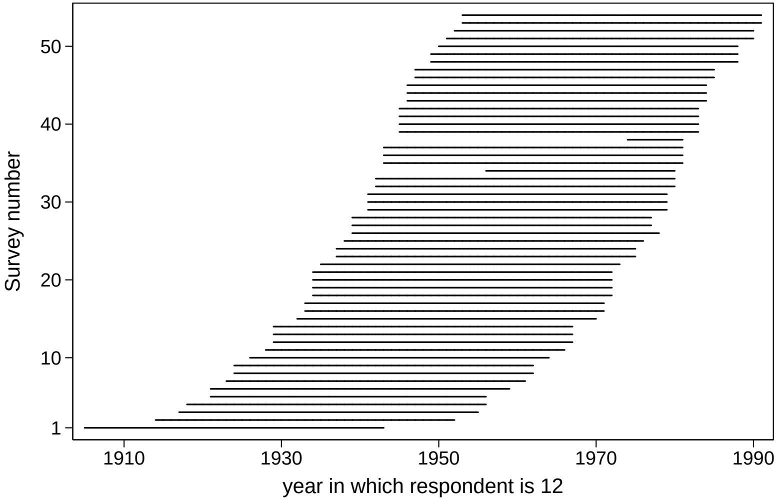
Distribution of highest achieved level of education



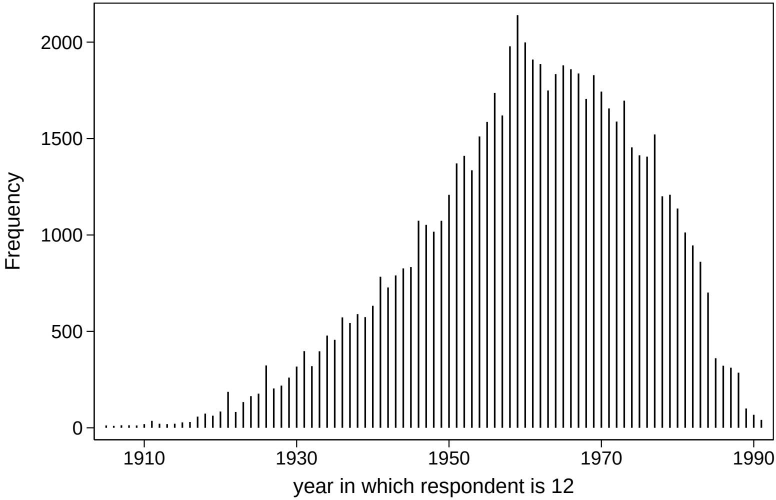
Data

- International Stratification and Mobility File (ISMF) on the Netherlands.
- 54 surveys held between 1958 and 2006 with information on cohorts 1905-1991.
- 67,000 respondents aged between 27 and 65 with complete information.

Surveys and cohorts



Cohorts and number of observations



Variables

- Father's occupational status is a standardized score using the mean and standard deviation from the cohort 1950.

Variables

- Father's occupational status is a standardized score using the mean and standard deviation from the cohort 1950.
- Level of education is standardized using the mean and standard deviation from the cohort 1950.

Variables

- Father's occupational status is a standardized score using the mean and standard deviation from the cohort 1950.
- Level of education is standardized using the mean and standard deviation from the cohort 1950.
- the main effect of cohort is measured by a restricted cubic spline with boundary knots at 1910 and 1970 and an interior knot in 1940.

Variables

- Father's occupational status is a standardized score using the mean and standard deviation from the cohort 1950.
- Level of education is standardized using the mean and standard deviation from the cohort 1950.
- the main effect of cohort is measured by a restricted cubic spline with boundary knots at 1910 and 1970 and an interior knot in 1940.
- The IEOpps are allowed to change linearly over cohorts.

sequential response model for men

	LO v more	LBO/MAVO v HAVO/VWO	LBO/MAVO v MBO	HAVO/VWO v HBO/WO
father's status	0.910	0.692	0.268	0.447
	(15.25)	(14.15)	(3.51)	(5.92)
father's status X cohort	-0.068	-0.014	-0.005	-0.034
	(-5.06)	(-1.58)	(-0.37)	(-2.36)
cohort	0.566	0.315	0.460	0.458
	(17.53)	(9.14)	(9.45)	(7.89)
cohort_1	-0.000	0.013	0.002	0.014
	(-0.03)	(7.07)	(0.96)	(4.79)
constant	-0.588	-1.469	-2.891	-0.797
	(-6.34)	(-13.12)	(-18.00)	(-4.19)
N	43768			
log likelihood	-50030.862			

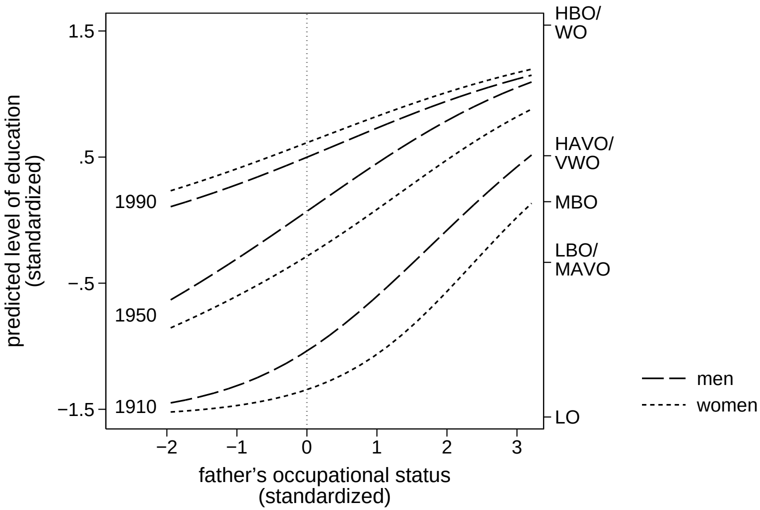
z statistics in parentheses

sequential response model for women

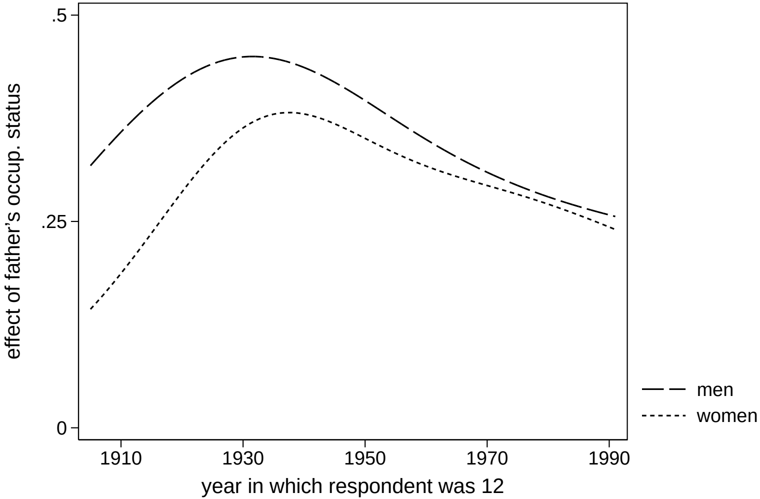
	LO v more	LBO/MAVO v HAVO/VWO	LBO/MAVO v MBO	HAVO/VWO v HBO/WO
father's status	0.876	1.025	0.406	0.080
	(15.36)	(17.29)	(5.13)	(0.89)
father's status X cohort	-0.068	-0.064	-0.020	0.029
	(-5.37)	(-6.06)	(-1.44)	(1.81)
cohort	0.744	0.104	0.130	0.349
	(21.28)	(2.30)	(2.34)	(4.72)
$cohort_1$	-0.000	-0.008	-0.022	0.009
	(-0.22)	(-3.55)	(-8.26)	(2.40)
constant	-1.730	-1.699	-2.432	-0.777
	(-17.07)	(-10.91)	(-12.87)	(-3.05)
N	43677			
log likelihood	-45829.805			

z statistics in parentheses

Predicted level of education



Change in IEOut over cohorts



Decomposition of IEOut

- IEOut is a weighted sum of IEOpps:

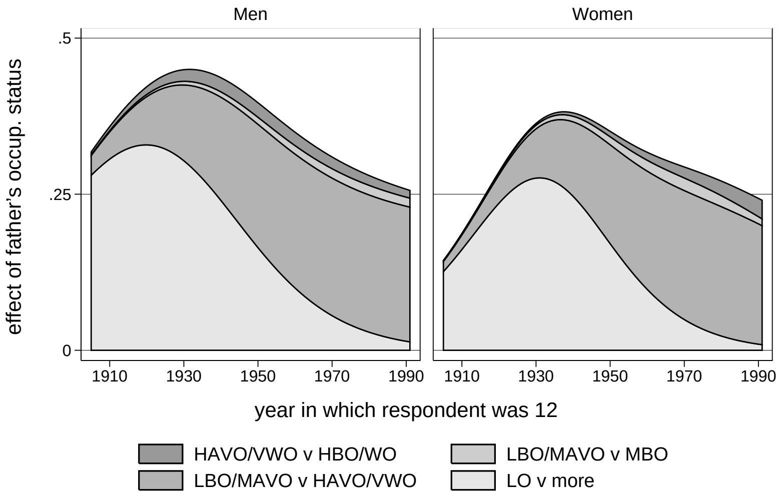
Decomposition of IEOut

- IEOut is a weighted sum of IEOpps:
  IEOut = w_1 $IEOpp_1$ + w_2 $IEOpp_2$ + w_3 $IEOpp_3$ + w_4 $IEOpp_4$

Decomposition of IEOut

- IEOut is a weighted sum of IEOpps:
  IEOut = w_1 $IEOpp_1$ + w_2 $IEOpp_2$ + w_3 $IEOpp_3$ + w_4 $IEOpp_4$
- The contribution of the first transition is: w_1 $IEOpp_1$

Change in IEOut over cohorts



Decomposition of IEOut

- IEOut is a weighted sum of IEOpps:
  IEOut = w_1 $IEOpp_1$ + w_2 $IEOpp_2$ + w_3 $IEOpp_3$ + w_4 $IEOpp_4$
- The contribution of the first transition is: w_1 $IEOpp_1$

Decomposition of IEOut

- IEOut is a weighted sum of IEOpps:
  $\text{IEOut} = w_1 \text{ IEOpp}_1 + w_2 \text{ IEOpp}_2 + w_3 \text{ IEOpp}_3 + w_4 \text{ IEOpp}_4$
- The contribution of the first transition is: $w_1 \text{ IEOpp}_1$
- This can be visualized as the area of a rectangle

Decomposition of IEOut

- IEOut is a weighted sum of IEOpps:
  IEOut = w_1 $IEOpp_1$ + w_2 $IEOpp_2$ + w_3 $IEOpp_3$ + w_4 $IEOpp_4$
- The contribution of the first transition is: w_1 $IEOpp_1$
- This can be visualized as the area of a rectangle with width w_1

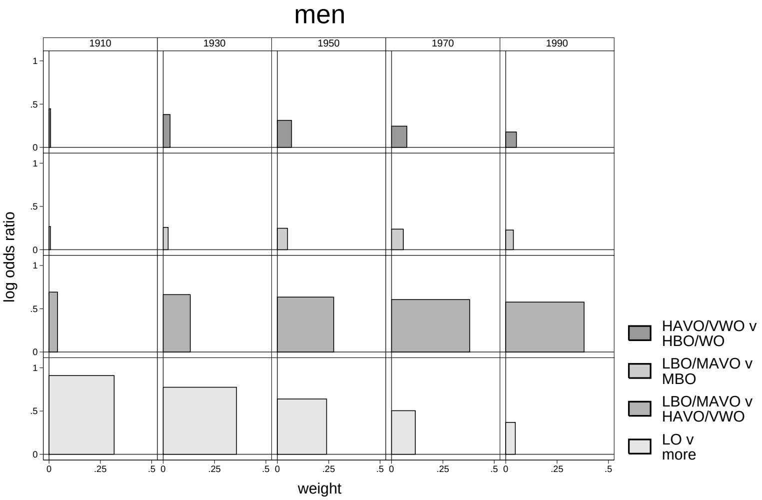
Decomposition of IEOut

- IEOut is a weighted sum of IEOpps:
  IEOut = w_1 $IEOpp_1$ + w_2 $IEOpp_2$ + w_3 $IEOpp_3$ + w_4 $IEOpp_4$
- The contribution of the first transition is: w_1 $IEOpp_1$
- This can be visualized as the area of a rectangle with width w_1 and height $IEOpp_1$.

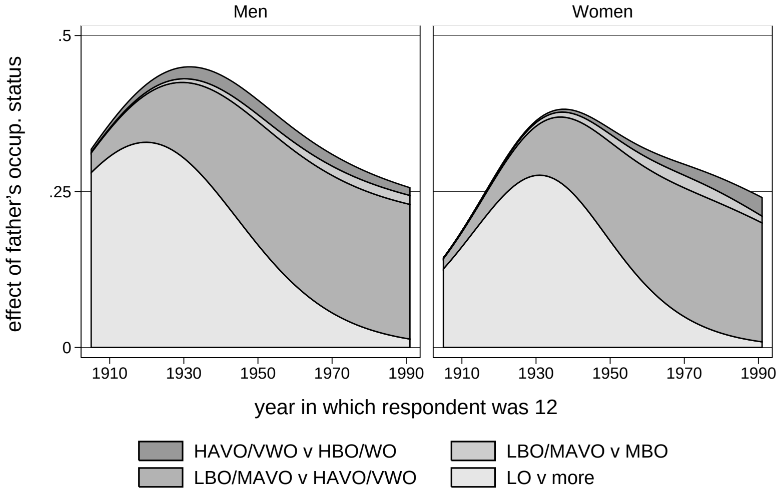
Decomposition of IEOut

- IEOut is a weighted sum of IEOpps:
  IEOut = w_1 $IEOpp_1$ + w_2 $IEOpp_2$ + w_3 $IEOpp_3$ + w_4 $IEOpp_4$
- The contribution of the first transition is: w_1 $IEOpp_1$
- This can be visualized as the area of a rectangle with width w_1 and height $IEOpp_1$.
- IEOut is the sum of the areas of these rectangles

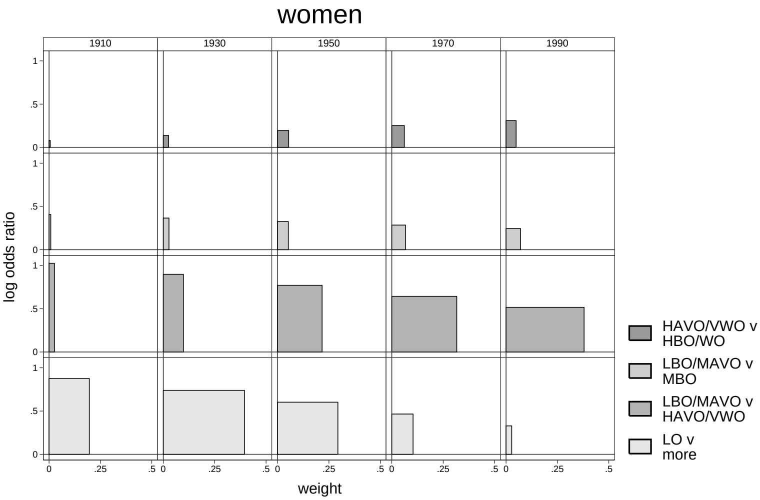
Decomposition of IEOut for men



Change in IEOut over cohorts



Decomposition of IEOut for women



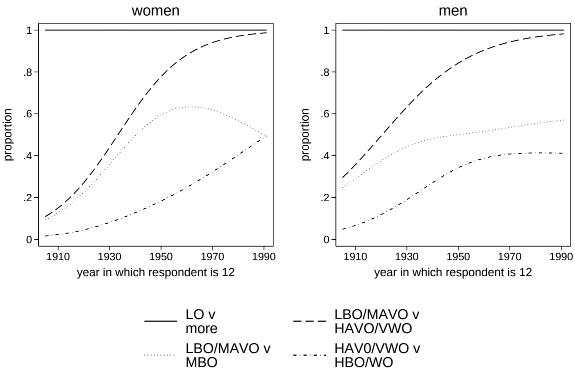
Decomposition of weights

- The weights are:
  at risk $\times$ variance $\times$ gain

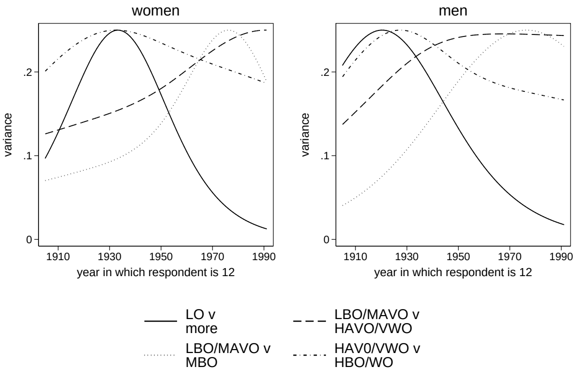
Decomposition of weights

- The weights are:
  at risk $\times$ variance $\times$ gain
- These three elements are all a function of the proportions that pass the transitions

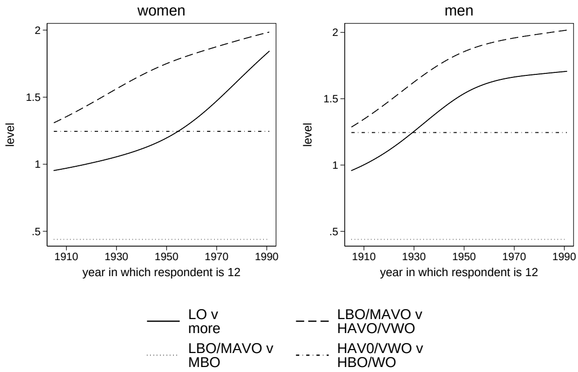
Decomposition of the weights: Proportion at risk



Decomposition of the weights: Variance



Decomposition of the weights: Gain



Outline

Conclusion (1)

- The dominant model for IEOpp — The Mare model — implies a substantively interesting relationship between IEOpp and IEOut:

Conclusion (1)

- The dominant model for IEOpp — The Mare model — implies a substantively interesting relationship between IEOpp and IEOut:
  - IEOut is a weighted sum of IEOpps

Conclusion (1)

- The dominant model for IEOpp — The Mare model — implies a substantively interesting relationship between IEOpp and IEOut:
  - IEOut is a weighted sum of IEOpps
  - The weights are at risk $\times$ variance $\times$ gain

Conclusion (1)

- The dominant model for IEOpp — The Mare model — implies a substantively interesting relationship between IEOpp and IEOut:
  - IEOut is a weighted sum of IEOpps
  - The weights are at risk $\times$ variance $\times$ gain
- This also implies a substantively interesting relationship between IEOut and educational expansion

Conclusion (2)

The trend in IEOut in the Netherlands is the result of a shift as the dominant source of IEOut between

Conclusion (2)

The trend in IEOut in the Netherlands is the result of a shift as the dominant source of IEOut between the transition between wether or not to continue after primary

Conclusion (2)

The trend in IEOut in the Netherlands is the result of a shift as the dominant source of IEOut between the transition between wether or not to continue after primary to the transition between entering the high or the low track:

Conclusion (2)

The trend in IEOut in the Netherlands is the result of a shift as the dominant source of IEOut between the transition between wether or not to continue after primary to the transition between entering the high or the low track:

- The first transition lost his dominant position as passing it became near universal

Conclusion (2)

The trend in IEOut in the Netherlands is the result of a shift as the dominant source of IEOut between the transition between wether or not to continue after primary to the transition between entering the high or the low track:

- The first transition lost his dominant position as passing it became near universal
- The second transition gained in prominence because ever more students became at risk and the probability of entering the high track moved to close to 50%.

Conclusion (2)

The trend in IEOut in the Netherlands is the result of a shift as the dominant source of IEOut between the transition between wether or not to continue after primary to the transition between entering the high or the low track:

- The first transition lost his dominant position as passing it became near universal
- The second transition gained in prominence because ever more students became at risk and the probability of entering the high track moved to close to 50%.
- The subsequent transitions contributed very little to IEOut, as the IEOpps where relatively small, fewer students were at risk, and the expected gain from passing these transitions is relatively small compared to the lower transitions.

Discussion: controlling for educational expansion? (1)

- The argument by Mare (1981) that the measure of IEOut discussed in presentation is influenced by educational expansion has led to use of various alternative measures.

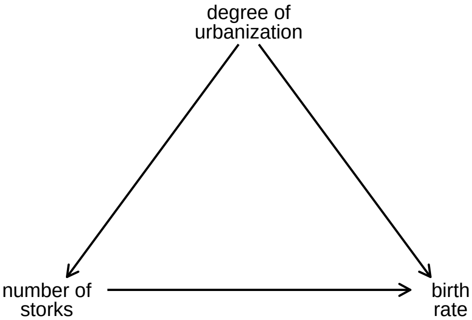
Discussion: controlling for educational expansion? (1)

- The argument by Mare (1981) that the measure of IEOut discussed in presentation is influenced by educational expansion has led to use of various alternative measures.
- For example coefficients from ordered logit or scaled association models, which are odds ratios and thus control for educational expansion.

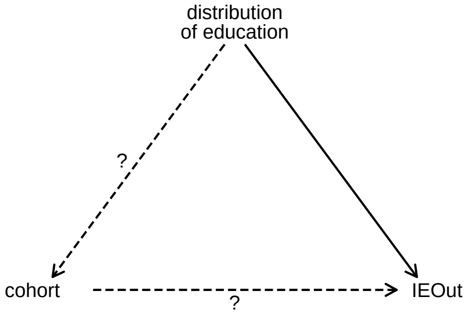
Discussion: controlling for educational expansion? (1)

- The argument by Mare (1981) that the measure of IEOut discussed in presentation is influenced by educational expansion has led to use of various alternative measures.
- For example coefficients from ordered logit or scaled association models, which are odds ratios and thus control for educational expansion.
- Is this controlling for educational expansion a desirable characteristic?

Intermezzo: Why control?



Intermezzo: Why control?



Discussion: controlling for educational expansion? (2)

- There is no direct causal relationship between cohort and IEOpp

Discussion: controlling for educational expansion? (2)

- There is no direct causal relationship between cohort and IEOpp
- Instead IEOpp differs across cohorts because the educational system and society differs across cohorts

Discussion: controlling for educational expansion? (2)

- There is no direct causal relationship between cohort and IEOpp
- Instead IEOpp differs across cohorts because the educational system and society differs across cohorts
- The aim should be to see the impact of these changes, rather than control for them.

IEOpps and IEOut

IEOut is the increase in expected highest achieved level of education for a unit increase in SES, i.e. a first derivative:

IEOpps and IEOut

IEOut is the increase in expected highest achieved level of education for a unit increase in SES, i.e. a first derivative:

$$
\begin{aligned}
&\frac{\partial E(ed)}{\partial SES} = \\
&\{1 \times \hat{p}_{1i}(1 - \hat{p}_{1i}) \times [(1 - \hat{p}_2)l_1 + \hat{p}_2(1 - \hat{p}_3)l_2 + \hat{p}_2\hat{p}_3 l_3 - l_0]\}\lambda_1 + \\
&\{\hat{p}_{1i} \times \hat{p}_{2i}(1 - \hat{p}_{2i}) \times [(1 - \hat{p}_3)l_2 + \hat{p}_3 l_3 - l_1]\}\lambda_2 + \\
&\{\hat{p}_{1i}\hat{p}_{2i} \times \hat{p}_{3i}(1 - \hat{p}_{3i}) \times [l_3 - l_2]\}\lambda_3
\end{aligned}
$$

IEOpps and IEOut

$$\frac{\partial E(ed)}{\partial SES} =$$
$$\{1 \times \hat{p}_{1i}(1-\hat{p}_{1i}) \times [(1-\hat{p}_2)I_1 + \hat{p}_2(1-\hat{p}_3)I_2 + \hat{p}_2\hat{p}_3I_3 - I_0]\}\lambda_1 +$$
$$\{\hat{p}_{1i} \times \hat{p}_{2i}(1-\hat{p}_{2i}) \times [(1-\hat{p}_3)I_2 + \hat{p}_3I_3 - I_1]\}\lambda_2 +$$
$$\{\hat{p}_{1i}\hat{p}_{2i} \times \hat{p}_{3i}(1-\hat{p}_{3i}) \times [I_3 - I_2]\}\lambda_3$$

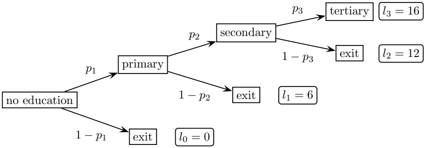
IEOpps and IEOut

$$\frac{\partial E(ed)}{\partial SES} =$$
$$\{1 \times \hat{p}_{1i}(1-\hat{p}_{1i}) \times [(1-\hat{p}_2)I_1 + \hat{p}_2(1-\hat{p}_3)I_2 + \hat{p}_2\hat{p}_3 I_3 - I_0]\}\lambda_1 +$$
$$\{\hat{p}_{1i} \times \hat{p}_{2i}(1-\hat{p}_{2i}) \times [(1-\hat{p}_3)I_2 + \hat{p}_3 I_3 - I_1]\}\lambda_2 +$$
$$\{\hat{p}_{1i}\hat{p}_{2i} \times \hat{p}_{3i}(1-\hat{p}_{3i}) \times [I_3 - I_2]\}\lambda_3$$

IEOpps and IEOut

proportion at risk

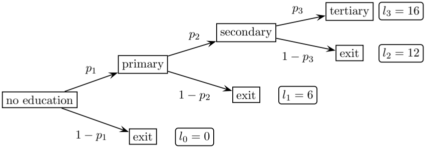
$$\frac{\partial E(ed)}{\partial SES} =$$
$$\{1 \times \hat{p}_{1i}(1-\hat{p}_{1i}) \times [(1-\hat{p}_2)l_1 + \hat{p}_2(1-\hat{p}_3)l_2 + \hat{p}_2\hat{p}_3 l_3 - l_0]\}\lambda_1 +$$
$$\{\hat{p}_{1i} \times \hat{p}_{2i}(1-\hat{p}_{2i}) \times [(1-\hat{p}_3)l_2 + \hat{p}_3 l_3 - l_1]\}\lambda_2 +$$
$$\{\hat{p}_{1i}\hat{p}_{2i} \times \hat{p}_{3i}(1-\hat{p}_{3i}) \times [l_3 - l_2]\}\lambda_3$$



IEOpps and IEOut

variance of the variable indicating whether one passes or not

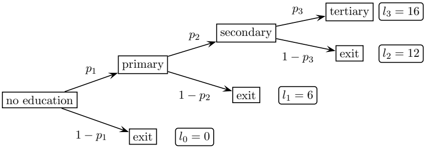
$$\frac{\partial E(ed)}{\partial SES} =$$
$$\{1 \times \hat{p}_{1i}(1-\hat{p}_{1i}) \times [(1-\hat{p}_2)l_1 + \hat{p}_2(1-\hat{p}_3)l_2 + \hat{p}_2\hat{p}_3 l_3 - l_0]\}\lambda_1 +$$
$$\{\hat{p}_{1i} \times \hat{p}_{2i}(1-\hat{p}_{2i}) \times [(1-\hat{p}_3)l_2 + \hat{p}_3 l_3 - l_1]\}\lambda_2 +$$
$$\{\hat{p}_{1i}\hat{p}_{2i} \times \hat{p}_{3i}(1-\hat{p}_{3i}) \times [l_3 - l_2]\}\lambda_3$$



IEOpps and IEOut

expected increase in the level of education after passing

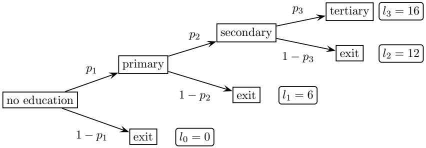
$$\frac{\partial E(ed)}{\partial SES} =$$
$$\{1 \times \hat{p}_{1i}(1-\hat{p}_{1i}) \times [(1-\hat{p}_2)l_1 + \hat{p}_2(1-\hat{p}_3)l_2 + \hat{p}_2\hat{p}_3 l_3 - l_0]\}\lambda_1 +$$
$$\{\hat{p}_{1i} \times \hat{p}_{2i}(1-\hat{p}_{2i}) \times [(1-\hat{p}_3)l_2 + \hat{p}_3 l_3 - l_1]\}\lambda_2 +$$
$$\{\hat{p}_{1i}\hat{p}_{2i} \times \hat{p}_{3i}(1-\hat{p}_{3i}) \times [l_3 - l_2]\}\lambda_3$$



IEOpps and IEOut

expected level of education for those that pass

$$\frac{\partial E(ed)}{\partial SES} =$$
$$\{1 \times \hat{p}_{1i}(1-\hat{p}_{1i}) \times [(1-\hat{p}_2)l_1 + \hat{p}_2(1-\hat{p}_3)l_2 + \hat{p}_2\hat{p}_3 l_3 - l_0]\}\lambda_1 +$$
$$\{\hat{p}_{1i} \times \hat{p}_{2i}(1-\hat{p}_{2i}) \times [(1-\hat{p}_3)l_2 + \hat{p}_3 l_3 - l_1]\}\lambda_2 +$$
$$\{\hat{p}_{1i}\hat{p}_{2i} \times \hat{p}_{3i}(1-\hat{p}_{3i}) \times [l_3 - l_2]\}\lambda_3$$



IEOpps and IEOut

minus the expected level of education for those that fail

$$\frac{\partial E(ed)}{\partial SES} =$$
$$\{1 \times \hat{p}_{1i}(1-\hat{p}_{1i}) \times [(1-\hat{p}_2)l_1 + \hat{p}_2(1-\hat{p}_3)l_2 + \hat{p}_2\hat{p}_3 l_3 - l_0]\}\lambda_1 +$$
$$\{\hat{p}_{1i} \times \hat{p}_{2i}(1-\hat{p}_{2i}) \times [(1-\hat{p}_3)l_2 + \hat{p}_3 l_3 - l_1]\}\lambda_2 +$$
$$\{\hat{p}_{1i}\hat{p}_{2i} \times \hat{p}_{3i}(1-\hat{p}_{3i}) \times [l_3 - l_2]\}\lambda_3$$