The Consequences of Unobserved Heterogeneity in a Sequential Logit Model

Maarten L. Buis

Institut für Soziologie
Eberhard Karls Universität Tübingen
maarten.buis@ifsoz.uni-tuebingen.de

Introduction

- This talk will discuss the sequential logit model

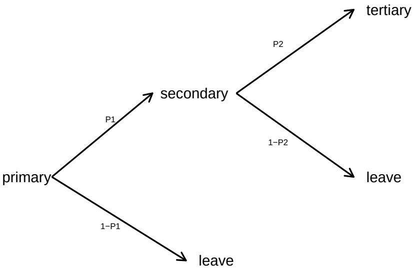
Introduction

- This talk will discuss the sequential logit model
- This model has been re-invented many times and is known under many names:
  - sequential logit model (Tutz 1991)
  - sequential response model (maddala 1983),
  - continuation ratio logit (Agresti 2002),
  - model for nested dichotomies (fox 1997), and
  - the Mare model (after Mare 1981)

Introduction

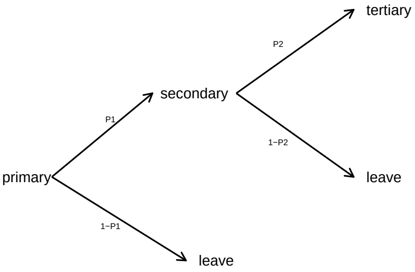
- This talk will discuss the sequential logit model
- This model has been re-invented many times and is known under many names:
  - sequential logit model (Tutz 1991)
  - sequential response model (maddala 1983),
  - continuation ratio logit (Agresti 2002),
  - model for nested dichotomies (fox 1997), and
  - the Mare model (after Mare 1981)
- The aim of this talk is assess how sensitive the conclusions from this model are to unobserved heterogeneity.

The sequential logit model



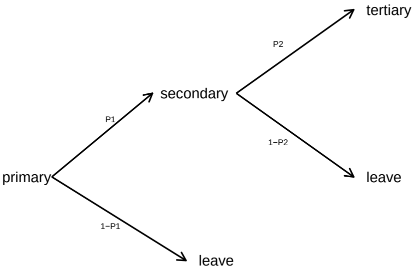
- A sequential logit model models a sequence of transitions.

The sequential logit model



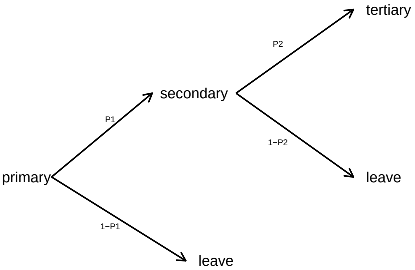
- A sequential logit model models a sequence of transitions.
- Each transition is modeled as a (multinomial) logistic regression using the sample which is 'at risk'.

The sequential logit model



- A sequential logit model models a sequence of transitions.
- Each transition is modeled as a (multinomial) logistic regression using the sample which is 'at risk'.
- $p_1 = \Lambda(\beta_{01} + \beta_{11}x + \beta_{21}z)$
- $p_2 = \Lambda(\beta_{02} + \beta_{12}x + \beta_{22}z)$

The sequential logit model



- A sequential logit model models a sequence of transitions.
- Each transition is modeled as a (multinomial) logistic regression using the sample which is 'at risk'.
- $p_1 = \Lambda(\beta_{01} + \beta_{11}x + \beta_{21}z)$
- $p_2 = \Lambda(\beta_{02} + \beta_{12}x + \beta_{22}z)$
- $\Lambda(u) = \frac{\exp(u)}{1+\exp(u)}$

Outline

The problem with unobserved variables

- The unobserved variable(s) could be confounding variables.

The problem with unobserved variables

- The unobserved variable(s) could be confounding variables.
- Even if the unobserved variable(s) are not confounding variables, they will still influence the results through 2 mechanisms:

The problem with unobserved variables

- The unobserved variable(s) could be confounding variables.
- Even if the unobserved variable(s) are not confounding variables, they will still influence the results through 2 mechanisms:
  - The Averaging Mechanism
  - The Selection Mechanism

The Averaging Mechanism

- leaving a variable out means averaging the probability over this variable

The Averaging Mechanism

- leaving a variable out means averaging the probability over this variable
- So, if we think that the true model is:

$$Pr(pass) = \Lambda(\beta_0 + \beta_1 x + \beta_2 z)$$

The Averaging Mechanism

- leaving a variable out means averaging the probability over this variable
- So, if we think that the true model is:

$$Pr(pass) = \Lambda(\beta_0 + \beta_1 x + \beta_2 z)$$

- and we cannot observe z, then we should use:

$$E_z[Pr(pass)] = E_z[\Lambda(\beta_0 + \beta_1 x + \beta_2 z)]$$

The Averaging Mechanism

- leaving a variable out means averaging the probability over this variable
- So, if we think that the true model is:

$$Pr(pass) = \Lambda(\beta_0 + \beta_1 x + \beta_2 z)$$

- and we cannot observe z, then we should use:

$$E_z[Pr(pass)] = E_z[\Lambda(\beta_0 + \beta_1 x + \beta_2 z)]$$

- Because $\Lambda()$ is a non-linear transformation, this is not the same as a simple logistic regression excluding z:

$$E_z[Pr(pass)] \neq \Lambda(\beta_0 + \beta_1 x + E_z[\beta_2 z])$$

The Averaging Mechanism

- leaving a variable out means averaging the probability over this variable
- So, if we think that the true model is:

$$Pr(pass) = \Lambda(\beta_0 + \beta_1 x + \beta_2 z)$$

- and we cannot observe z, then we should use:

$$E_z[Pr(pass)] = E_z[\Lambda(\beta_0 + \beta_1 x + \beta_2 z)]$$

- Because $\Lambda()$ is a non-linear transformation, this is not the same as a simple logistic regression excluding z:

$$E_z[Pr(pass)] \neq \Lambda(\underbrace{\beta_0^*}_{=\beta_0 + E_z[\beta_2 z]} + \beta_1 x)$$

The Selection Mechanism

- At higher transitions the sample at risk is a selected sample

The Selection Mechanism

- At higher transitions the sample at risk is a selected sample
- This selection is likely to produce a negative correlation between the observed and unobserved variables

The Selection Mechanism

- At higher transitions the sample at risk is a selected sample
- This selection is likely to produce a negative correlation between the observed and unobserved variables
- This means that the unobserved variable is likely to become a confounding variable at higher transitions, even if it was not one at the first transition

Outline

Sensitivity analysis

- The aim of a sensitivity analysis is to show what would happen to the estimates of our model under a range of scenarios concerning unobserved heterogeneity.

Sensitivity analysis

- The aim of a sensitivity analysis is to show what would happen to the estimates of our model under a range of scenarios concerning unobserved heterogeneity.
- These scenarios are not intended to be true, but together they are meant to show what unobserved heterogeneity could do to the estimates.

Sensitivity analysis

- The aim of a sensitivity analysis is to show what would happen to the estimates of our model under a range of scenarios concerning unobserved heterogeneity.
- These scenarios are not intended to be true, but together they are meant to show what unobserved heterogeneity could do to the estimates.
- A sensitivity analysis is intended to show which conclusions are robust and which are not.

Scenarios concerning unobserved heterogeneity

- There is not one unobserved variable but many.

Scenarios concerning unobserved heterogeneity

- There is not one unobserved variable but many.
- So, if we have one observed variable x, the true model is for transition k is:
  $p_k = \Lambda(\beta_{0k} + \beta_{1k}x + \underbrace{\gamma_{1k}z_1 + \gamma_{2k}z_2 + \cdots + \gamma_{lk}z_l}_{\varepsilon})$

Scenarios concerning unobserved heterogeneity

- There is not one unobserved variable but many.
- So, if we have one observed variable x, the true model is for transition k is:
  $$p_k = \Lambda(\beta_{0k} + \beta_{1k}x + \underbrace{\gamma_{1k}z_1 + \gamma_{2k}z_2 + \cdots + \gamma_{lk}z_l}_{\varepsilon})$$
- ε is likely to be normally distributed at the first transition, because it is a sum of variables.

Scenarios concerning unobserved heterogeneity

- There is not one unobserved variable but many.
- So, if we have one observed variable x, the true model is for transition k is:
  $p_k = \Lambda(\beta_{0k} + \beta_{1k}x + \underbrace{\gamma_{1k}z_1 + \gamma_{2k}z_2 + \cdots + \gamma_{lk}z_l}_{\varepsilon})$
- ε is likely to be normally distributed at the first transition, because it is a sum of variables.
- The scenarios can differ with respect to the initial standard deviation of ε and the initial correlation between x and ε

Scenarios concerning unobserved heterogeneity

- There is not one unobserved variable but many.
- So, if we have one observed variable x, the true model is for transition k is:
  $p_k = \Lambda(\beta_{0k} + \beta_{1k}x + \underbrace{\gamma_{1k}z_1 + \gamma_{2k}z_2 + \cdots + \gamma_{lk}z_l}_{\varepsilon})$
- ε is likely to be normally distributed at the first transition, because it is a sum of variables.
- The scenarios can differ with respect to the initial standard deviation of ε and the initial correlation between x and ε
- This standard deviation can be seen as the effect of a standardized version of ε.

Scenarios concerning unobserved heterogeneity

- There is not one unobserved variable but many.
- So, if we have one observed variable x, the true model is for transition k is:
  $p_k = \Lambda(\beta_{0k} + \beta_{1k}x + \underbrace{\gamma_{1k}z_1 + \gamma_{2k}z_2 + \cdots + \gamma_{lk}z_l}_{\varepsilon})$
- ε is likely to be normally distributed at the first transition, because it is a sum of variables.
- The scenarios can differ with respect to the initial standard deviation of ε and the initial correlation between x and ε
- This standard deviation can be seen as the effect of a standardized version of ε.
- The effect of x given a scenario can be computed in Stata using the `seqlogit` package.

Outline

Sequential logit model for the Netherlands (Males)

- Replication of De Graaf and Ganzeboom (1993) and Buis (2009).

Sequential logit model for the Netherlands (Males)

- Replication of De Graaf and Ganzeboom (1993) and Buis (2009).
- Data comes from the International Stratification and Mobility File.

Sequential logit model for the Netherlands (Males)

- Replication of De Graaf and Ganzeboom (1993) and Buis (2009).
- Data comes from the International Stratification and Mobility File.
- Covers the birth cohorts 1890–1980.

Sequential logit model for the Netherlands (Males)

- Replication of De Graaf and Ganzeboom (1993) and Buis (2009).
- Data comes from the International Stratification and Mobility File.
- Covers the birth cohorts 1890–1980.
- father's education measured in categories primary, lower secondary, higher secondary, tertiary.

Sequential logit model for the Netherlands (Males)

- Replication of De Graaf and Ganzeboom (1993) and Buis (2009).
- Data comes from the International Stratification and Mobility File.
- Covers the birth cohorts 1890–1980.
- father's education measured in categories primary, lower secondary, higher secondary, tertiary.
- father's occupation measured in ISEI scores.

Sequential logit model for the Netherlands (Males)

- Replication of De Graaf and Ganzeboom (1993) and Buis (2009).
- Data comes from the International Stratification and Mobility File.
- Covers the birth cohorts 1890–1980.
- father's education measured in categories primary, lower secondary, higher secondary, tertiary.
- father's occupation measured in ISEI scores.
- both are standardized.

Sequential logit model for the Netherlands (Males)

	primary versus lower secondary	lower secondary versus higher secondary	higher secondary versus tertiary
father's education	1.442	0.713	0.447
	(11.51)	(11.79)	(7.13)
father's eduction X cohort	-0.117	-0.026	-0.031
	(-4.82)	(-2.34)	(-2.80)
father's occupation	0.814	0.455	0.226
	(13.06)	(8.15)	(3.37)
father's occupation x cohort	-0.087	-0.016	0.015
	(-6.32)	(-1.53)	(1.28)
N	42271		

log odds ratios, *z* statistics in parentheses

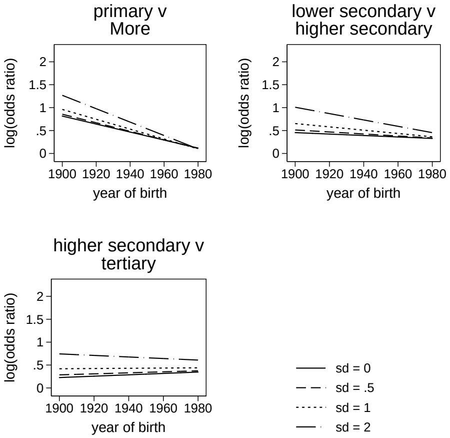
Scenarios

- The standard deviation of ε ranges from .5 to 5.

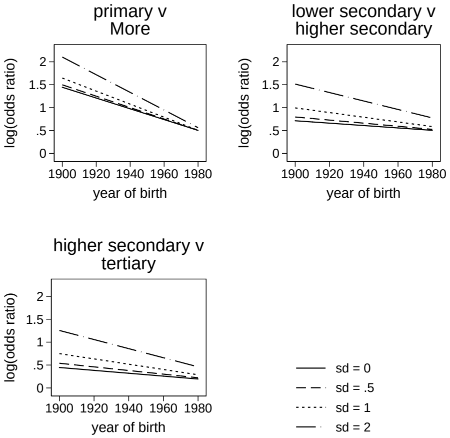
Scenarios

- The standard deviation of ε ranges from .5 to 5.
- The correlation between ε and father's education or occupation ranges from -.60 to .60.

Effect of father's occupation in various scenarios



Effect of father's education in various scenarios



Two conclusions

decline of the effect of father's education and occupation over cohorts

p-values, H0: trend ≥ 0

	father's education	father's occupation
transition 1	.000	.000
transition 2	.010	.063
transition 3	.003	.900

Two conclusions

decline of the effect of father's education and occupation over cohorts

p-values, H0: trend ≥ 0

	father's education	father's occupation
transition 1	.000	.000
transition 2	.010	.063
transition 3	.003	.900

decline of the effect of father's education and occupation over transitions

p-value, H0: father's resource is equal across transitions

father's education	father's occupation
.000	.000

H0: effect father's education equal across transitions

p-values

		sd									
		.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
	-.60	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003
	-.50	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003
	-.40	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.003	0.004
	-.30	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003	0.004
	-.25	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003	0.004
	-.20	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003	0.004
	-.15	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003	0.005
	-.10	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.002	0.004	0.005
	-.05	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.003	0.004	0.005
corr	0	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.003	0.004	0.005
	.05	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.003	0.004	0.005
	.10	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.002	0.004	0.005
	.15	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003	0.005
	.20	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003	0.004
	.25	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003	0.004
	.30	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003	0.004
	.40	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.003	0.004
	.50	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003
	.60	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003

H0: effect father's occupation equal across transitions

p-values

		sd									
		.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
	-.60	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.003	0.005	0.006
	-.50	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.004	0.006	0.008
	-.40	0.000	0.000	0.000	0.000	0.001	0.002	0.003	0.005	0.007	0.009
	-.30	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
	-.25	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
	-.20	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
	-.15	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
	-.10	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
	-.05	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
corr	0	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
	.05	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
	.10	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
	.15	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
	.20	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
	.25	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
	.30	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.006	0.008	0.010
	.40	0.000	0.000	0.000	0.000	0.001	0.002	0.003	0.005	0.007	0.009
	.50	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.004	0.006	0.008
	.60	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.003	0.005	0.006

H0: Trend effect father's education in transition 1 $\geq$ 0

p-values

		sd									
		.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
	-.60	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.50	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.40	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.30	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.25	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.20	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
corr	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.20	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.25	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.30	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.40	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.50	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.60	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

H0: Trend effect father's education in transition 2 $\geq$ 0

p-values

		sd									
		.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
	-.60	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.50	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.40	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.30	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.25	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.20	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.15	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.10	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.05	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
corr	0	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.05	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.10	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.15	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.20	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.25	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.30	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.40	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.50	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.60	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

H0: Trend effect father's education in transition 3 $\geq$ 0

p-values

		sd									
		.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
	-.60	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.50	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.40	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.30	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.25	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.20	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
corr	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.20	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.25	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.30	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.40	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.50	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.60	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

H0: Trend effect father's occupation in transition 1 $\geq$ 0

p-values

		sd									
		.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
	-.60	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.50	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.40	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.30	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.25	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.20	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
corr	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.20	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.25	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.30	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.40	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.50	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.60	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

H0: Trend effect father's occupation in transition 2 $\geq$ 0

p-values

		sd									
		.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
	-.60	0.031	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.50	0.028	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.40	0.025	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.30	0.023	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.25	0.023	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.20	0.022	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.15	0.022	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.10	0.021	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	-.05	0.021	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
corr	0	0.021	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.05	0.021	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.10	0.021	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.15	0.022	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.20	0.022	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.25	0.023	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.30	0.023	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.40	0.025	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.50	0.028	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	.60	0.031	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

H0: Trend effect father's occupation in transition 3 $\geq$ 0

p-values

		sd									
		.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
	-.60	0.846	0.681	0.464	0.286	0.171	0.104	0.066	0.044	0.031	0.023
	-.50	0.836	0.645	0.414	0.241	0.139	0.083	0.052	0.035	0.025	0.019
	-.40	0.828	0.618	0.378	0.212	0.119	0.070	0.044	0.030	0.021	0.016
	-.30	0.822	0.597	0.353	0.192	0.106	0.062	0.039	0.026	0.019	0.015
	-.25	0.820	0.589	0.344	0.185	0.102	0.059	0.037	0.025	0.018	0.014
	-.20	0.818	0.583	0.336	0.180	0.098	0.057	0.036	0.024	0.018	0.014
	-.15	0.816	0.578	0.331	0.176	0.096	0.056	0.035	0.024	0.018	0.014
	-.10	0.815	0.574	0.327	0.173	0.094	0.055	0.034	0.023	0.017	0.013
	-.05	0.815	0.572	0.325	0.172	0.093	0.054	0.034	0.023	0.017	0.013
corr	0	0.814	0.572	0.324	0.171	0.093	0.054	0.034	0.023	0.017	0.013
	.05	0.815	0.572	0.325	0.172	0.093	0.054	0.034	0.023	0.017	0.013
	.10	0.815	0.574	0.327	0.173	0.094	0.055	0.034	0.023	0.017	0.013
	.15	0.816	0.578	0.331	0.176	0.096	0.056	0.035	0.024	0.018	0.014
	.20	0.818	0.583	0.336	0.180	0.098	0.057	0.036	0.024	0.018	0.014
	.25	0.820	0.589	0.344	0.185	0.102	0.059	0.037	0.025	0.018	0.014
	.30	0.822	0.597	0.353	0.192	0.106	0.062	0.039	0.026	0.019	0.015
	.40	0.828	0.618	0.378	0.212	0.119	0.070	0.044	0.030	0.021	0.016
	.50	0.836	0.645	0.414	0.241	0.139	0.083	0.052	0.035	0.025	0.019
	.60	0.846	0.681	0.464	0.286	0.171	0.104	0.066	0.044	0.031	0.023

Conclusion

- There are three ways in which unobserved heterogeneity can influence results in a sequential logit model:

Conclusion

- There are three ways in which unobserved heterogeneity can influence results in a sequential logit model:
  - Through possible confounding variables

Conclusion

- There are three ways in which unobserved heterogeneity can influence results in a sequential logit model:
  - Through possible confounding variables
  - Through the averaging mechanism

Conclusion

- There are three ways in which unobserved heterogeneity can influence results in a sequential logit model:
  - Through possible confounding variables
  - Through the averaging mechanism
  - Through the selection mechanism

Conclusion

- There are three ways in which unobserved heterogeneity can influence results in a sequential logit model:
  - Through possible confounding variables
  - Through the averaging mechanism
  - Through the selection mechanism
- The potential influence of unobserved heterogeneity on your results can be investigated with a sensitivity analysis.

Conclusion

- There are three ways in which unobserved heterogeneity can influence results in a sequential logit model:
  - Through possible confounding variables
  - Through the averaging mechanism
  - Through the selection mechanism
- The potential influence of unobserved heterogeneity on your results can be investigated with a sensitivity analysis.
- This has been done for a replication of De Graaf and Ganzeboom (1993) and Buis (2009)

Conclusion

- There are three ways in which unobserved heterogeneity can influence results in a sequential logit model:
  - Through possible confounding variables
  - Through the averaging mechanism
  - Through the selection mechanism
- The potential influence of unobserved heterogeneity on your results can be investigated with a sensitivity analysis.
- This has been done for a replication of De Graaf and Ganzeboom (1993) and Buis (2009)
  - both the effect and the trend are likely to be underestimated

Conclusion

- There are three ways in which unobserved heterogeneity can influence results in a sequential logit model:
  - Through possible confounding variables
  - Through the averaging mechanism
  - Through the selection mechanism
- The potential influence of unobserved heterogeneity on your results can be investigated with a sensitivity analysis.
- This has been done for a replication of De Graaf and Ganzeboom (1993) and Buis (2009)
  - both the effect and the trend are likely to be underestimated
  - most conclusions are robust

Conclusion

- There are three ways in which unobserved heterogeneity can influence results in a sequential logit model:
  - Through possible confounding variables
  - Through the averaging mechanism
  - Through the selection mechanism
- The potential influence of unobserved heterogeneity on your results can be investigated with a sensitivity analysis.
- This has been done for a replication of De Graaf and Ganzeboom (1993) and Buis (2009)
  - both the effect and the trend are likely to be underestimated
  - most conclusions are robust
  - the exception is the non-significant trend in the effect of father's occupation in the second transition.

References

Agresti, A. (2002).
Categorical Data Analysis, 2nd edition.
Hoboken, NJ: Wiley-Interscience.

Buis, M.L. (2009).
The Consequences of Unobserved Heterogeneity in a Sequential Logit Model
http://home.fsw.vu.nl/m.buis/wp/unobserved_het.pdf

Fox, J. (1997).
Applied Regression Analysis, Linear Models, and Related Methods.
Thousand Oaks: Sage.

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

Mare, R.D. (1981).
Change and Stability in educational Stratification.
American Sociological Review, 46(1):72–87.

Tutz, G. (1991)
Sequential Models in Categorical Regression.
Computational Statistics & Data Analysis, 11(3):275–295.

Observe IQ

transition	IQ	class	y fail	y pass	N	Pr(pass)	odds(pass)	log odds ratio
1	low	low	300	100	400			
		high	200	200	400			
	high	low	200	200	400			
		high	100	300	400			
2	low	low	75	25	100			
		high	100	100	200			
	high	low	100	100	200			
		high	75	225	300			

Not observe IQ

transition	x	y fail	y pass	N	Pr	odds	log odds ratio
1	low	500	300	800			
	high	300	500	800			
2	low	175	125	300			
	high	175	325	500			

Observe IQ

transition	IQ	class	y fail	y pass	N	Pr(pass)	odds(pass)	log odds ratio
1	low	low	300	100	400			
		high	200	200	400			
	high	low	200	200	400			
		high	100	300	400			
2	low	low	75	25	100			
		high	100	100	200			
	high	low	100	100	200			
		high	75	225	300			

Not observe IQ

transition	x	y fail	y pass	N	Pr	odds	log odds ratio
1	low	500	300	800			
	high	300	500	800			
2	low	175	125	300			
	high	175	325	500			

Observe IQ

transition	IQ	class	y fail	y pass	N	Pr(pass)	odds(pass)	log odds ratio
1	low	low	300	100	400			
		high	200	200	400			
	high	low	200	200	400			
		high	100	300	400			
2	low	low	75	25	100			
		high	100	100	200			
	high	low	100	100	200			
		high	75	225	300			

Not observe IQ

transition	x	y fail	y pass	N	Pr	odds	log odds ratio
1	low	500	300	800			
	high	300	500	800			
2	low	175	125	300			
	high	175	325	500			

Observe IQ

transition	IQ	class	y: fail	y: pass	N	Pr(pass)	odds(pass)	log odds ratio
1	low	low	300	100	400			
		high	200	200	400			
	high	low	200	200	400			
		high	100	300	400			
2	low	low	75	25	100			
		high	100	100	200			
	high	low	100	100	200			
		high	75	225	300			

Not observe IQ

transition	x	y: fail	y: pass	N	Pr	odds	log odds ratio
1	low	500	300	800			
	high	300	500	800			
2	low	175	125	300			
	high	175	325	500			

Observe IQ

transition	IQ	class	y: fail	y: pass	N	Pr(pass)	odds(pass)	log odds ratio
1	low	low	300	100	400			
		high	200	200	400			
	high	low	200	200	400			
		high	100	300	400			
2	low	low	75	25	100			
		high	100	100	200			
	high	low	100	100	200			
		high	75	225	300			

Not observe IQ

transition	x	y: fail	y: pass	N	Pr	odds	log odds ratio
1	low	500	300	800			
	high	300	500	800			
2	low	175	125	300			
	high	175	325	500			

Observe IQ

transition	IQ	class	y: fail	pass	N	Pr(pass)	odds(pass)	log odds ratio
1	low	low	300	100	400			
		high	200	200	400			
	high	low	200	200	400			
		high	100	300	400			
2	low	low	75	25	100			
		high	100	100	200			
	high	low	100	100	200			
		high	75	225	300			

Not observe IQ

transition	x	y: fail	pass	N	Pr	odds	log odds ratio
1	low	500	300	800			
	high	300	500	800			
2	low	175	125	300			
	high	175	325	500			

Observe IQ

transition	IQ	class	y: fail	pass	N	Pr(pass)	odds(pass)	log odds ratio
1	low	low	300	100	400			
		high	200	200	400			
	high	low	200	200	400			
		high	100	300	400			
2	low	low	75	25	100			
		high	100	100	200			
	high	low	100	100	200			
		high	75	225	300			

Not observe IQ

transition	x	y: fail	pass	N	Pr	odds	log odds ratio
1	low	500	300	800			
	high	300	500	800			
2	low	175	125	300			
	high	175	325	500			

Observe IQ

transition	IQ	class	y fail	y pass	N	Pr(pass)	odds(pass)	log odds ratio
1	low	low	300	100	400	0.25		
		high	200	200	400	0.5		
	high	low	200	200	400	0.5		
		high	100	300	400	0.75		
2	low	low	75	25	100	0.25		
		high	100	100	200	0.5		
	high	low	100	100	200	0.5		
		high	75	225	300	0.75		

Not observe IQ

transition	x	y fail	y pass	N	Pr	odds	log odds ratio
1	low	500	300	800	0.375		
	high	300	500	800	0.625		
2	low	175	125	300	0.417		
	high	175	325	500	0.65		

Observe IQ

transition	IQ	class	y: fail	y: pass	N	Pr(pass)	odds(pass)	log odds ratio
1	low	low	300	100	400	0.25	0.333	
		high	200	200	400	0.5	1	
	high	low	200	200	400	0.5	1	
		high	100	300	400	0.75	3	
2	low	low	75	25	100	0.25	0.333	
		high	100	100	200	0.5	1	
	high	low	100	100	200	0.5	1	
		high	75	225	300	0.75	3	

Not observe IQ

transition	x	y: fail	y: pass	N	Pr	odds	log odds ratio
1	low	500	300	800	0.375	0.6	
	high	300	500	800	0.625	1.667	
2	low	175	125	300	0.417	0.714	
	high	175	325	500	0.65	1.857	

Observe IQ

transition	IQ	class	y fail	y pass	N	Pr(pass)	odds(pass)	log odds ratio
1	low	low	300	100	400	0.25	0.333	log(3)
		high	200	200	400	0.5	1	
	high	low	200	200	400	0.5	1	log(3)
		high	100	300	400	0.75	3	
2	low	low	75	25	100	0.25	0.333	log(3)
		high	100	100	200	0.5	1	
	high	low	100	100	200	0.5	1	log(3)
		high	75	225	300	0.75	3	

Not observe IQ

transition	x	y fail	y pass	N	Pr	odds	log odds ratio
1	low	500	300	800	0.375	0.6	log(2.778)
	high	300	500	800	0.625	1.667	
2	low	175	125	300	0.417	0.714	log(2.6)
	high	175	325	500	0.65	1.857	