Working Paper Series
CSER WP No. 0002

Multiple paths in educational transitions: A multinomial transition model with unobserved heterogeneity

Af Kristian Bernt Karlson

August 20, 2010

Centre for Strategic Educational Research
DPU, Aarhus University

Multiple paths in educational transitions: A multinomial transition model with unobserved heterogeneity

Kristian Berndt Karlson*

This version: August 20, 2010

Running head: Multiple paths in educational transitions

Abstract:

In many countries educational branching points consist of more than two qualitatively different alternatives, and only some alternatives provide the opportunity of continuing into higher education. I develop a multinomial transition model for modeling the effects of family background characteristics and individual characteristics on these complex educational careers. The model controls for unobserved heterogeneity that may, if ignored, result in biased estimates at later transitions. Bias arises as a consequence of the non-random selection of individuals who "survive" to face later transitions in the educational system. Controlling for unobserved heterogeneity moreover relaxes the somewhat restrictive assumption of Independence from Irrelevant Alternatives on which the standard multinomial logit model is based. I apply the model to the Danish case and analyze data which covers the educational careers of a cohort of Danes born around 1954.

I find that the model brings forward nontrivial heterogeneity in the influence of family background and ability on qualitatively different choice alternatives both within and across transitions. I also find that not controlling for unobserved heterogeneity leads to marked underestimation of the family background effect on later transitions in the educational career.

Acknowledgements: I thank Mads Meier Jæger and Anders Holm for helpful comments and suggestions.

* The Danish School of Education, Aarhus University, Tuborgvej 164, DK-2400 Copenhagen NV, Denmark. Telephone: 0045 23369285. Fax: 0045 88889001. SFI – The Danish National Centre for Social Research, Herlufs Trolle Gade 11, DK-1052 Copenhagen K. Email: kbk@dpu.dk.

T7733 words (including all text, notes, and references) 5 tables, 1 figure
Keywords: multinomial transition model, unobserved heterogeneity, educational transitions, educational decisions, inequality of educational opportunity

Multiple paths in educational transitions: A multinomial transition model with unobserved heterogeneity

1. Introduction

Individuals obtain educational qualifications through various routes in the educational system. These routes are defined by the structure of the educational system and can be thought of as comprising a set of sequential branching points (Boudon, 1974; Gambetta, 1987). In many countries particularly in Western Europe these branching points involve more than two choice alternatives. In these situations, the standard sequential logit model (SLM) for educational transitions developed by Mare (1980, 1981) is inappropriate because it does not capture the multiple and unordered nature of choice alternatives. Consequently, using the SLM for diversified educational systems may ignore important heterogeneity in the ways family background influences educational decisions. In an influential paper Breen and Jonsson (2000) proposed a multinomial transition model (MTM) to accommodate the multiple and unordered choice alternatives of diversified educational systems.[1] Using large-scale Swedish administrative register data, Breen and Jonsson model the transition from primary to secondary education and the transition from secondary to tertiary education, with each transition comprising three choice alternatives. Their model allows for path dependence (i.e., that fact that previous tracks completed influence the likelihood of completing later tracks), and they test whether their results are robust to unobserved heterogeneity. Thus, the MTM suggested by Breen and Jonsson is a viable alternative to the SLM.

However, although Breen and Jonsson (2000) proposed the MTM more than a decade ago the model has so far not been adopted in mainstream stratification research.

[1] Other alternatives to the SLM have been suggested, in particular the ordered logit or probit model (see Cameron & Heckman, 1998; Lucas, 2001; Breen *et al.*, 2009; Ballarino and Shadee, 2010).

Rather, stratification researchers using multinomial logit models typically focus on either the transition from primary to secondary education or the transition from secondary to tertiary education. For the transition from primary to secondary education, Need and de Jong (2001) provide results for the Netherlands, Becker (2003) for Germany, Hansen (2007) for Norway, Kreidl (2004) for the Czech Republic, Jao and McKeever (2006) for Taiwan, Ayalon and Shavit (2004) for Israel, and Jæger (2009) for Denmark. For the transition from secondary to tertiary or higher education, Tieben and Wolbers (2010) and Tolsma *et al*. (2010) provide results for the Netherlands, Becker and Hecken (2009) for Germany, and Mastakaasa (2006) for Norway. The fact that joint modeling of two or more transitions with a MTM has not found its way into mainstream practice of stratification research is problematic for two reasons. First, the standard multinomial logit model is based on the often unrealistic assumption of Independence from Irrelevant Alternatives (IIA).[2] Second, applying multinomial logit models to later transitions (e.g., from secondary to tertiary education) ignores the fact that individuals who face these decisions represent a selective sample. Regression coefficients estimated on selective samples may be influenced by unobserved heterogeneity and may therefore suffer from selection bias (cf. Heckman, 1979).

In this paper I continue the work by Breen and Jonsson (2000) by estimating a MTM that incorporates unobserved heterogeneity. I name this model the multinomial transition model with unobserved heterogeneity (MTMU). The model is a flexible finite mixture model that accommodates both selection bias and violations of the IIA assumption. With the MTMU I jointly model the effects of family background and individual characteristics on the probability of making two transitions using data from the Danish Longitudinal Survey of Youth (DLSY). First, in the transition from primary to secondary education, individuals

[2] Similar to the model proposed in this paper, Jæger (2009) uses a finite mixture model to overcome this limitation. However, Jæger (2009) is an exception to the rule.

complete the academic track, complete the vocational track, or leave the educational system. Second, in the transition from secondary to tertiary education, individuals complete the university track, complete the short-cycle track, or leave the educational system. Thus, in contrast to the model by Breen and Jonsson (2000), my MTMU is simpler in terms of the number of branching points to be estimated and in terms the possible pathways to pursue. Moreover, in my analysis the academic track in secondary education is an absorbing state, which means that only individuals completing academic secondary education "survive" to make the transition into tertiary education. This property reflects the institutional structure of the educational system in Denmark in the 1960s and 1970s. I use Stata command *gllamm* to estimate my model, and sample data and code is available from the author. I proceed as follows. First, I present the multinomial transition model with unobserved heterogeneity. Second, I introduce the data from the DLSY. Third, I present the results. Fourth, I conclude with a discussion of the advantages of the MTMU.

2. A multinomial transition model with unobserved heterogeneity

In this section I present the MTMU. The MTMU is an extension of the SLM popularized by Mare (1980, 1981) which, first, allows for more than two choice alternatives at two or more branching points and, second, controls for the possible selection bias caused by unobserved heterogeneity. I capture these unobserved variables with a finite number of latent classes. This specification is highly flexible and makes the model a finite mixture model. Conceptually, the model may be thought of as two or more multinomial logit models with a common, unobserved variable affecting each choice alternative relative to a baseline alternative for each transition. In the next sections I first present the model formally, and then I provide an intuitive explanation of the model.

2.1 The multinomial logit model

I first consider a multinomial logit model presented in a latent variable framework (cf. McFadden, 1974; Powers & Xie, 2000:238-9). Let y^*_{ia} be a continuous latent propensity of individual *i* to choose the *a*th educational alternative, where $i = 1,..,N$ and $a = 1,..,A$. Let x_{ij} be the *j*th explanatory variable for individual *i*, where $j = 1,\ldots,J$. Let y^*_{ia} be a linear function function of x_{ij} and an alternative-specific random error term ξ_{ia}:

$$y^*_{ia} = \sum_{j=1}^{J} b_{aj} x_{ij} + \xi_{ia}, \text{ where } \qquad sd(\xi_{ia}) = \sigma_a \tag{1}$$

σ_a is the standard deviations of the alternative-specific residuals (and therefore captures the variance of each alternative not explained by the observed variables). b_{aj} is the effect of x_{ij} on the latent propensity. In (1) each individual has an unobserved propensity to choose the *a*th alternative, but we only observe which of the A alternatives the individual *actually* chooses. To identify the model, we need to assume that

$$y_{ia} = a \;\; if \;\; y^*_{ia} > y^*_{ia'} \qquad for\, all\, a \neq a'. \tag{2}$$

In other words, we assume that the individual chooses the alternative for which he or she has the largest propensity. Moreover, we assume that the random error term, ξ_{ia}, is uncorrelated across alternatives and that it follows a standard type-I extreme value distribution. The assumption of uncorrelated error terms is also known as the assumption of Independence from Irrelevant Alternatives (IIA). IIA implies that if we remove one alternative individuals who would have chosen this alternative are randomly distributed among the remaining alternatives (McFadden, 1974; Wooldridge, 2002:501-2).[3] Given the assumptions on the error terms, the probability of choosing *a* can be written as

[3] To fix ideas, imagine an educational system with three choice options in the transition from primary to secondary education: exit, vocational track, and academic track. If IIA holds, then the consequence of closing

$$\Pr(y_i = a \mid x_{ij}) = \frac{\exp(\sum_{j=1}^{J} \frac{b_{aj}}{\sigma_a} x_{ij})}{\sum_{a=1}^{A} \exp(\sum_{j=1}^{J} \frac{b_{aj}}{\sigma_a} x_{ij})}\ , \tag{3}$$

where $\sum_{a=1}^{A} \Pr(y_i = a) = 1$. We normalize the model such that $\frac{b_{1j}}{\sigma_1} = 0$, which is equivalent to stating that alternative $a = 1$ is a reference category that defines the contrast alternative against which the other alternatives are defined. To simplify notation we also let $\beta_{aj} = \frac{b_{aj}}{\sigma_a}$, where β_{aj} is the well-known logit coefficient (i.e., the log odds-ratio).[4] We may thus rewrite equation (3) such that

$$\Pr(y_i = a \mid x_{ij}) = \frac{\exp(\sum_{j=1}^{J} \beta_{aj} x_{ij})}{1 + \sum_{a=2}^{A} \exp(\sum_{j=1}^{J} \beta_{aj} x_{ij})} \quad \text{for } a > 1. \tag{4}$$

Taking the log-odds of the probability in (4) returns the familiar multinomial logit model:

$$logit[\Pr(y_i = a \mid x_{ij})] = \sum_{j=1}^{J} \beta_{aj} x_{ij} \quad \text{for } a > 1. \tag{5}$$

The multinomial transition model is an extension of the model in (4) and (5).

---

down the academic track (i.e., removing that choice alternative) is that individuals that would have chosen the academic track (i.e., individuals with a high propensity for doing so) would be distributed randomly across the two remaining tracks. This assumption is not realistic in this example because we would expect the affected individuals to have a higher propensity to enroll in the vocational track than to exit the educational system. The multinomial transition model with unobserved heterogeneity I present below relaxes the IIA assumption.

[4] Notice that these logit coefficients are identified relative to the scale of the residual variances of each alternative. If these variances vary across alternatives (i.e., $\sigma_a \neq \sigma_{a'}$), we cannot know whether differences in regression coefficients across alternatives are due to differences in residual variance or in the underlying regression coefficients (from the model in (1)) (cf. Allison, 1999). I return to this issue in the results section.

2.2 The multinomial transition model with unobserved heterogeneity

In this subsection I extend the model in (4) and (5) to include two or more transitions and to accommodate unobserved heterogeneity. Let y^*_{iak} be a continuous latent propensity of individual *i* associated with choice of the *a*th educational alternative at the *k*th transition, where $k = 1,\ldots,K$. We define x_{ij} as before. I now decompose the error term similar to the one in (1) into a systematic and random component: $\xi_{iak} = u_{akw} + \varepsilon_{iak}$. u_{akw} is drawn from a discrete distribution with *W* latent classes, where $w = 1,...,W$, and where π_w is the share in class *w* and where $\sum_{w=1}^{W} \pi_w = 1$. These latent classes can be thought of as groups of individuals that have similar unobserved characteristics which lead them to make similar educational choices.

Following the specification of the multinomial logit model defined in (1)-(4), I write the conditional multinomial probability of the *a*'th choice on the *k*'transition as:

$$\Pr(y_{ik} = a \mid x_{ij}, \nu_{akw}) = \frac{\exp\left(\sum_{j=1}^{J} \beta_{akj} x_{ij} + \nu_{akw}\right)}{1 + \sum_{a=2}^{A} \exp(\sum_{j=1}^{J} \beta_{akj} x_{ij} + \nu_{akw})},$$

where β_{akj} is the logit coefficient of x_{ij} for alternative *a* at transition *k*, and ν_{akw} captures the effect of the unobserved variable for the *w*'th latent class (for alternative *a*, transition *k*).[5] In the analysis I model two transitions (i.e., $K = 2$), and I therefore define the joint probability of making two consecutive transitions as

$$\Pr(y_{i1} = a \mid x_{ij}, \nu_{a1w}) \times \Pr(y_{i2} = a' \mid x_{ij}, \nu_{a'2w}).$$

Finally, I write the multivariate probability unconditional on unobserved variables (i.e., they are averaged or integrated out), $\Pr(y_i = a, a')$, as a finite mixture model:

[5] Note that, because I use a latent variable formulation, it holds that: $\nu_{akw} = \frac{u_{akw}}{sd(\varepsilon_{iak})}$. In other words, the effect of the unobserved variable is only identified up to scale.

$$
\begin{aligned}
&\Pr(y_{ik} = a, a' \mid x_{ij}) = \\
&\sum_{w=1}^{W} \Pr(y_{i1} = a \mid x_{ij}, \nu_{a1w}) \times \Pr(y_{i2} = a' \mid x_{ij}, \nu_{a'2w})^{I} \pi_w = \\
&\sum_{w=1}^{W} \frac{\exp(\sum_{j=1}^{J} \beta_{a1j} x_{ij} + \nu_{a1w})}{1 + \sum_{a=2}^{A} \exp(\sum_{j=1}^{J} \beta_{a1j} x_{ij} + \nu_{a1w})} \times \left(\frac{\exp(\sum_{j=1}^{J} \beta_{a'2j} x_{ij} + \nu_{a'2w})}{1 + \sum_{a'=2}^{A'} \exp(\sum_{j=1}^{J} \beta_{a'2j} x_{ij} + \nu_{a'2w})} \right)^{I} \pi_w \qquad (6)
\end{aligned}
$$

The unconditional joint probability in (6) is given by a weighted (over the *W* latent classes) product of the conditional probability of completing choice *a* at transition *k* (cf. Wedel & DeSarbo, 1995, 2002; McLachlan & Peel, 2000:145f). *I* is an indicator taking on the value 1 for those who survive to face the second transition (k = 2), and taking on the value 0 for those who do not survive. In other words, the MTMU in (6) jointly models two transitions and accommodates for unobserved heterogeneity. If the model does not correct for unobserved heterogeneity, then the initial sorting of individuals on observed and unobserved characteristics results in bias of the estimates at later transitions (Cameron & Heckman, 1998; Holm & Jæger, this issue).[6] Moreover, correcting for unobserved heterogeneity in the multinomial case also relaxes the assumption of IIA (cf. Henscher & Greene, 2003).[7] I nonparametrically identify the parameters in model (6) that I estimate in my analysis by including alternative-specific instrumental variables at the first transition (see the data description). This identification strategy provides me with the necessary exclusion restrictions for identifying the unobserved variables (without relying on arbitrary parametric

[6] In the duration model literature, this dynamic selection problem is known as *frailty*. It refers to the identification problem that in duration models changes in survival probabilities can be a mixture of unobserved population heterogeneity and state dependence (cf. Vaupel & Yashin, 1985; Trussel & Richards, 1985:245; Yamaguchi, 1987:78; Lancaster, 1990:64). In my example, the problem can also be conceived of as a sample selection problem, because only select individuals experience later transitions (see Heckman, 1979; Berk, 1983; Winship & Mare, 1992).

[7] Notice that the model in (6) also corrects for rescaling bias induced by not including ν_{akw} (Cameron & Heckman, 1998:282; Nicoletti & Rondinelli, 2006). This problem is related to the fact that logit coefficients are identified up to scale and thus depends on the included variables in the model (Amemiya, 1975; Winship & Mare, 1984; Yatchew & Griliches, 1985).

assumptions).[8] I estimate the model with *gllamm* for Stata (Rabe-Hesketh *et al.*, 2004). A worked example with sample data and dofile is available from the author.

2.3 Intuition behind the MTMU

Before I proceed to the data description, I give an account of the intuition behind the model, in particular the role played by the unobserved variable. Imagine two transition points, each with three choice alternatives: primary to secondary education (exit, vocational track, academic track) and secondary to tertiary education (exit, short-cycle track, university track). Assume that only students who complete the academic track at the first transition are allowed to make the second transition (i.e., they "survive" the first transition). Similar to Cameron and Heckman (1998:296), I assume that the population of students can be divided into two mutually exclusive types (or classes). The first type is characterized by low educational aspirations, while the other type is characterized by high educational aspirations. The researcher does not observe whether an individual belongs to one type or the other, i.e., the aspiration variable is unobserved. In addition, imagine that the researcher observes the social class membership of the student (low/high), i.e., the social class variable is observed.

We expect that, compared to students with low aspirations, students with high aspirations are more likely to complete the academic track in the first transition and, if they "survive," also more likely to complete the university track in the second transition. We also expect that, compared to lower class students, higher class students are more likely to complete these tracks. From these expectations and assumptions, it follows that students who survive to face the choices of the second transition have higher aspirations and tend to come more from the high social class than those who do not survive. This selection or sorting

[8] Another identification strategy that establishes the necessary exclusion restrictions is the inclusion of time-varying covariates (see Holm & Jæger and Lucas, this issue). In the multinomial case, alternative specific variation is also a necessary condition for identification.

mechanism induces a *negative* correlation between aspirations and social class in the sample of those who survive (Cameron & Heckman, 1998:276; cf. Mare, 1980:298f, 1981:82). Because omitted variables (aspirations) which are correlated with both the observed variables (social class) and the outcome (completing the university track in the second transition) give rise to bias in the estimates of the observed variables, the selection mechanism obscures the estimates of the influence of social class on the second transition (cf. Heckman 1979). Cameron & Heckman (1998) refer to this kind of selection bias as dynamic selection bias.

The consequences of dynamic selection bias on the estimates at later transitions may be severe (depending on the magnitude of the induced correlation between the observed and unobserved variables and on the magnitude of the effect of the unobserved variable on the outcome). However, in the multinomial case matters are even more complicated. We identify a standard multinomial logit model as in (5) through the assumption of IIA. If this assumption does not hold, we expect bias to arise in the estimates of the multinomial logit model. A similar logic holds for a multinomial transition model. We would expect that, compared to students with low aspirations, students with high aspirations are more likely to complete the academic track in the first transition and university track in the second transition. A consequence of this expectation is that students would not distribute themselves randomly across the remaining alternatives if one of the choice alternatives was removed. The IIA assumption is thus violated: If the academic track at the first transition was removed then we would expect those with high aspirations to opt for the vocational track more so than those with low aspirations. We would expect the same with respect to social class, thereby inducing a correlation between the unobserved (aspirations) and observed (social class) variables that may result in biased estimates of the observed variables. Thus, unobserved heterogeneity may, if not corrected for, bias estimates both through dynamic selection and through violations of the IIA assumption. The model in (6) corrects for both sources of bias.

3. Data and variables

I analyze data from the Danish Longitudinal Survey of Youth (DLSY) (Hansen, 1995). I refer to Jæger and Holm (2007) and Jæger (2007) for a detailed data description. The DLSY follows the life course of 3,151 children born in or around 1954 who were all attending the 7th grade of comprehensive school when they were first interviewed in 1968. The DLSY is based on cluster sampling and respondents were sampled from 151 complete school classes. The survey contains information on family background and ability, and the longitudinal data structure enables me to reconstruct the educational careers of the individuals. My final sample consists of 2,199 individuals, i.e., 30 percent of the original sample is set to missing. This non-response is a consequence of drop-out of the survey and of total non-response on both dependent and explanatory variables. Because of the low non-response rate, I take the sample to be representative of the 1954 birth cohort.

3.1 Dependent variables

Figure 1 shows the institutional structure of the Danish educational system and the flow of students born in 1954 as they progress through the educational system (see Table 1 for the marginal distributions).[9] Students first complete comprehensive school after 7-10 years of

[9] The presentation in Figure 1 is simplified. According to the Danish Education Act of 1958 (“Skoleloven 1958”) those students who did not leave comprehensive school after 7 years of schooling were divided into two tracks (of two to three years of length): a theoretically oriented track (“Realafdelingen”) or a practically oriented track. Completion of either track gave the opportunity to choose the academic track in secondary education. To keep my transition model as simple as possible and to keep it comparable to the one in Breen and Jonsson (2000), I do not include this early tracking. Moreover, at that time students who completed the theoretically oriented track in comprehensive school were allowed to enroll in short-cycle tertiary education programs. Thus, a fraction of the birth cohort did enroll in tertiary education without completing the academic track in secondary education,

schooling. After completing comprehensive school, at ages 14-17, the individual can choose between three alternatives in secondary education: Leave school, enroll in a vocational track (apprenticeship based education, typically three-four years), or enroll in an academic track (*Gymnasium*, a three year program). Of the total sample, around one fifth leaves the educational system after ending primary education, around half completes the vocational track, and around one third completes the academic track. Those who complete the academic track face the tertiary education decision, around ages 19-20: Leave school, enroll in a short-cycle track (typically aiming at the professions such as teacher or nurse, two-four year programs), or enroll in a university track (five year programs). Of the 718 individuals completing the academic track, around one fourth leaves the educational system with the degree, around half completes a short cycle education, and around one third completes a university education.

-- FIGURE 1 HERE –

-- TABLE 1 HERE --

3.2 Explanatory variables

I include both family background characteristics and individual characteristics as explanatory variables in my analysis. *Parental highest social class* is measured with the EGP scheme divided into five classes (EGP-5) (Halpin, 1999; Jæger, 2007; cf. Erikson & Goldthorpe, 1992): I/II (professional and managerial employees and self-employed with 10 or more employees), III (routine non-manual professionals), IV (self-employed and small employers (1-9 employees), V/VI (skilled workers), and VII (unskilled and semi-skilled workers). To

thereby compromising the logic of the academic track in secondary education being an absorbing state. However, they account for a minor fraction of the total sample. I exclude these individuals in my analysis.

avoid too large non-response, I include a category indicating missing information in the EGP-5 variable (12 percent of the total sample). *Parental highest education* is the number of years of completed schooling for the parent with the highest level of education. Because 22 percent of the parents in total sample have not reported their educational attainment, I replace these missing values with the average number of years in the total sample and I include a dummy variable indicating whether the parents are missing or not. *Non-intact family* is a dummy variable indicating whether the child did not live with both biological parents at age 14. *Boy* indicates the gender of the child. *Ability* is a measure of the academic skills of the student at age 14 and is constructed as the principal component from a principal component analysis on three test scores in a verbal test, spatial test, and inductive test (each measured by the number of correct answers on the test). The principal component accounts for 70.5 percent of the total variation in the three items. Ability is transformed into a scale from 0 to 100 in the total sample, where higher scores indicate higher ability.

Table 2 shows descriptive statistics for the explanatory variables for the total sample and the sample that survives to face the tertiary education decision. We see that the sample becomes more selective as respondents progress through the educational system. For example, 11.5 percent originates in social classes I and II in the total sample, while 21.7 percent does so in the selected sample. We also see that average number of years of parental highest education changes from around 10 years to around 11 years. These changes show that students from socioeconomically well-off families have a higher propensity to complete the academic track in secondary education. Moreover, the average ability score is 13 points larger for the selected sample than for the total sample (from 53 to 66 points), and the standard deviation of ability decreases 4 points (from 18 to 14 points). This selection pattern suggests that students with higher ability have a higher propensity to complete the academic track, and that these students are more homogeneous than the total population in terms of academic ability.

-- TABLE 2 HERE --

3.3 Instrumental variables

I also include two instrumental variables for the first transition in the MTMU, one for each choice alternative (vocational and academic). These variables ensure nonparametric identification of the MTMU model, and their distributions are described in the bottom columns of Table 2. The instrumental variable for the vocational track is the share of the respondent's school class in comprehensive school that chooses the vocational track. Similarly, for the academic track I use the share of the respondent's school class that chooses the academic track. In the construction of both variables 151 school classes are used, and the respondent is omitted in the calculation of the class mean (thereby avoiding tautological inferences). I thus exploit the cluster design of the DLSY in which respondents are nested in school classes. I interpret my two instruments as indicators of the *influence of peers* on the educational decision.[10] This influence operates through the revealed preferences for secondary education choices of the school class peers in comprehensive school (i.e., in primary education). Given the sociological evidence of the influence of peers on educational attainment (e.g., Sewell *et al.*, 1969), I find this assumption plausible. Moreover, because I control for family background and ability, the peer influence operates net of these potentially

[10] Because these two variables work as instruments in my model, I assume that they (A) directly affect their respective choice alternatives on the first transition, but do not, directly or indirectly, (B) either affect the choices made in the second transition or the other choice alternative in the first transition. Other identification strategies would have been possible to pursue. For example, time-varying covariates combined with covariates varying across choice alternatives (and not individuals) would provide nonparametric identification.

confounding characteristics. Thus, I net out the potential sorting into school classes on parental characteristics and child characteristics.

I only include instrumental variables at the first transition in order to establish the necessary exclusion restrictions. Because school classes in comprehensive school dissolve after the completion of comprehensive school, and the influence of those peers therefore markedly decreases later in the educational career, this model strategy appears credible. Moreover, at later points in the educational career we expect other peer groups to have formed, and we expect these peer groups, rather than the old ones, to influence the later educational decisions. In addition to this, I find it credible that each instrument only affects the chosen alternative on the first transition (i.e., that the instruments are alternative-specific). Thus, the instrument for the vocational track only affects the respondent's propensity to choose the vocational track, not the academic track, and vice versa for the academic track. This assumption may be violated if school class spillover effects exist, but given the control for family background and ability, I find this assumption credible.

4. Results

In this section I present the results from my MTMU model with two latent classes and compare the results with a standard MTM (i.e., a model without unobserved heterogeneity). I first report the results in logit coefficients (i.e., log odds-ratios), and thereafter report average partial effects as defined by Wooldridge (2002:22-24). I pay particular attention to the influence of parental social class, parental education, and the student's academic ability on the choices alternatives at each of the two branching points (see Figure 1 to recall the institutional structure of the educational system).

4.1 The MTMU

Table 3 shows the logit coefficients from a MTM and a MTMU. Although the logit coefficients from the two models cannot be directly compared (because they are measured on different scales), in the final column of Table 3 I also report whether the MTM coefficients differ substantially from the MTMU coefficients. In general the MTM underestimates the effects compared to the MTMU, although exceptions exist.[11] For example, the effect of social classes I/II (relative to class VII) on the university track in tertiary education (Panel D in Table 3) is about 75 percent larger for the MTMU (2.255) than for the MTM (1.295). Such difference reflects considerable underestimation of the MTM estimates. Moreover, in some cases the significance of estimates changes, thereby returning qualitatively different conclusions. For example, the gender effect on the short-cycle track (Panel C in Table 3) is statistically significant in the MTM (-0.477), but insignificant in the MTMU (-0.339), and the effect is around 30 percent smaller in numerical terms. Compare this change to the change in the gender effect on the university track (Panel D in Table 3), where the effect almost doubles from the MTM (0.916) to the MTMU (1.868). Thus, had I not controlled for unobserved heterogeneity, I would have drawn somewhat erroneous conclusions with respect to the effects of many of the included variables in my model. I return to this issue below, when I report average partial effects.

For now, however, I report the results from the MTMU (and not the MTM), because I consider this model to be my preferred model (i.e., the model on which I will base my inferences). For the vocational track at the first transition (Panel A in Table 3), all social

[11] Given the nature of rescaling bias in logit models, we would expect—all other things being equal—the coefficient to increase between MTM and MTMU, because we divide the logit coefficients with a smaller number in the MTMU than in the MTM (because the MTMU explains “more variation” in the outcome and thus reduces the underlying residual standard deviation). Thus, whether the percent change from MTM to MTMU reflects a change in the underlying “causal” effects, or simply is a consequence of rescaling, is not possible to confirm here (for a thorough discussion of this identification problem, see Karlson, Holm, & Breen 2010).

classes are more likely than class VII to complete the track (a joint Wald-test confirms that the joint social class effect is statistically significant). Parental education is, on the other hand, insignificant, indicating that the socioeconomic family influence on this decision runs through social class. The effect of ability is highly significant and positive. Thus, the vocational track appears to be both socially and academically selective.

For the academic track at the first transition (Panel B in Table 3), the effects of parental social class, parental education, and student ability are all positive and significant. We see that the higher the social class, the higher the likelihood of completing the academic track, indicating a linear trend. It would now be informative to investigate at which track (vocational or academic) the effects are largest. However, because track-specific logit coefficients are identified up to different scales (cf. footnote 4), we cannot compare the coefficients from the two tracks. Thus, although the social selectivity (i.e., the family background effects) and the influence of ability appear larger for the academic track than for the vocational track, we cannot ascertain such conclusion. To overcome this issue, I later report average partial effects, which are less sensitive to scale identification.

-- TABLE 3 HERE --

For the short-cycle track in tertiary education (Panel C in Table 3), the effect of each social class is insignificant and their joint contribution is also insignificant (confirmed by a joint Wald-test). In addition, the effect of parental education is negative, although insignificant. This pattern suggests that the social selectivity in completing the short-cycle track is negligible,[12] if not even reversed in such a way that students of well-educated parents are less

[12] Note that the main reason for the effects being insignificant is the low number of individuals surviving to face the tertiary education decision (N = 718). More data would provide me with more efficient estimates.

likely to complete the track (relative to leaving school) than students of less-educated parents. Moreover, the effect of ability on the short-cycle track is insignificant, indicating that ability does not matter for completing the track. One explanation of these negligible effects is that completing the short-cycle track in tertiary education is just as difficult (if not less than) as completing the academic track in secondary education for the cohort under study. Thus, the selective nature of the academic track renders the influence of family background and ability less important for completing the short-cycle track. This finding supports the conclusions drawn for the somewhat more selective university track to which I now turn.

For the university track in tertiary education (Panel D in Table 3), the effects of social classes I/II and IV are positive and statistically significant, while parental education is insignificant. Consequently, the socioeconomic family influence on the completion of this track appears to run through social class. Contrary to what might be expected, the effect of ability is not statistically significant, once again indicating the selective nature of the academic track in secondary education. Thus, while we cannot trace any social or academic selectivity in the short-cycle track, we do trace some social selectivity in the university track. The relative sizes of the effects across tracks are, as already mentioned, difficult to evaluate, because of the scale identification of logit coefficients. Before I therefore return to reporting the results in average partial effects, however, I briefly dwell on the unobserved variable in the MTMU.

-- TABLE 4 HERE --

Table 4 describes the distribution and effects of the binary unobserved variable in the MTMU. Each category in this variable can be equated with the unobserved types described in a previous section (e.g., unobserved aspirations). The first type or class comprises around 44 percent of the population, while the second type comprises the remaining 56 percent. Members in the first type have a persistently lower likelihood of completing the four educational tracks in the model (two on each transition) relative to the baseline group defined by the intercepts in the MTMU. By way of contrast, members in the second type have a higher likelihood of completing the tracks. Thus, the interpretation of the binary unobserved variable as capturing educational aspirations might be appropriate. The population consists of low-aspiring individuals (type 1) and high-aspiring individuals (type 2), who differ in success rates at the different tracks.[13] Omitting the variable capturing these types may have consequences for the estimates of a MTM and should therefore be included as in the MTMU.

4.2 Average partial effects

Table 5 presents average partial effects of the MTM and MTMU (Wooldridge 2002).[14] This effect measure has two notable properties. First, it states the effects on the probability scale (from zero to one), i.e., how a one unit change in x changes the probability of the outcome, Pr(y=1). Second, it is less sensitive to the scale identification than logit coefficients.[15] Since Table 5 contains as many estimates as Table 3, I pay special attention to the average partial effects of social class (indicating the social selectivity of the tracks). Using estimates from the

[13] Notice that the unobserved variable involves counterfactual statements. For example, a type-1 student that in fact chose vocational track would have performed poorly on the academic track, had he chosen the academic track. Similar counterfactuals can be constructed. However, the general conclusion to be drawn here is that no matter which track factually completed, the type-1 (type-2) students would have had lower (higher) completion rates on the other tracks, had they pursued them, than the baseline group defined by the intercepts in the MTMU.
[14] The expected probability used in the calculation average partial effects for the MTMU is given with respect to the prior distribution of the unobserved variable. The *gllamm* post estimation command, *gllapred*, predicts these probabilities (Rabe-Hesketh *et al*., 2004). The authors of the command recommend using these predicted probabilities (rather than those given with respect to the posterior distribution of the unobserved variable).
[15] Average partial effects are not fully insensitive to changes in the scale parameter (for a formal discussion, see Karlson, Holm, & Breen, 2010). Moreover, through the expected probability of the model it depends on the marginal distribution of the observed binary outcome variable. Thus, average partial effects are not foolproof effect measures for comparing the effects across tracks and transitions.

MTMU, I first report differences across tracks both within and across transitions. Thereafter I emphasize how the MTM considerably underestimates the effects for the university track given by the MTMU.

Looking at the first transition, we see that the social class effects are substantially larger for the academic track than for the vocational counterpart (compare Panels A and B in Table 5). For example, social classes I/II have about 7 percent larger probability of completing the vocational track than social class VII (the reference). For completing the academic track, social classes I/II have about 32 percent larger probability than social class VII, reflecting a large difference in the social selectivity of the two tracks. We reach a similar conclusion for the second transition. Here the social class effects are much more pronounced for the university track than for the short-cycle track (compare Panels C and D in Table 5). For example, students originating in social class IV are about 10 percent more likely to complete the short-cycle track than students originating in social class VII, while this difference is about 32 percent for the university track. Thus, the overall conclusion to be drawn from the MTMU estimates stated as average partial effects is that the academic track at the first transition and the university track are more socially selective than the other tracks at the respective transitions. Among all tracks, the university track appears to be the most socially selective.

-- TABLE 5 HERE --

The average partial effects reported in Table 5 clearly show that controlling for unobserved heterogeneity increases the influence of family background at later transitions (see the final column which contains the difference between the MTM and MTMU average partial effects).

The MTM tends to underestimate the “true” parameters (given by the MTMU). For example, the social class effects for the university track (Panel D in Table 5) are underestimated with between eight and 15 percentage points. Such differences are considerable. Thus, had we not controlled for unobserved heterogeneity, we would have reported that social class V/II are about six percent more likely to complete the university track than social class VII. However, correcting for unobserved heterogeneity reveals a 21 percent difference, indicating much more pronounced social selectivity. Such a difference in effects between the MTM and the MTMU must be considered to “make a difference” in terms of the social selectivity of the university track.

5. Discussion

The multinomial transition model with unobserved heterogeneity is a flexible tool for modeling the influence of family background characteristics and individual characteristics on complex educational pathways in diversified educational systems. Compared with a standard SLM, the model allows for branching points with more than two choice alternatives (e.g., exit or continue), it controls for the selection bias at later transitions induced by the selective nature of educational systems, and it relaxes the often unrealistic assumption of IIA on which the standard multinomial logit model is based. Thus, the MTMU provides researchers with insight into the social and individual heterogeneity in educational decision making in diversified systems and with better (i.e., unbiased *ceteris paribus*) estimates.

In the paper I estimate the MTMU on longitudinal survey data from a cohort born in 1954 in Denmark. I find marked social selectivity for the academic track in secondary education and for the university track in tertiary education, while the social selectivity is less pronounced for the vocational track in secondary education and more or less non-existent for the short-cycle track in tertiary education. Moreover, academic skills appear to matter for completion of the secondary education tracks, in particular for the academic track, but not for the completing the tracks in tertiary education. Thus, compared to the standard SLM, multinomial transition models have the potential of revealing important heterogeneity in the social and academic selectivity across tracks and transitions in diversified educational systems. Moreover, using a MTMU compared to a MTM provides estimates that are controlled for the potential bias caused by selection bias and violations of the IIA assumption. The MTM generally underestimates the true estimates (provided by the MTMU), in particular the social class estimates for the university track in tertiary education. In terms of average partial effects, these social class effects are underestimated with between eight and 15 percentage points. If stratification researcher are to inform policy-makers such considerable differences may “make a difference” in terms of policy interventions to be constructed and implemented. Thus, researchers may have good reasons for adopting the MTMU rather than the conventional MTM.

Despite the apparent advantages of the MTMU and the fact that Breen and Jonsson (2000) presented the model more than a decade ago, the model has not diffused into mainstream stratification research. In this paper I have tried to address this problem by applying a MTMU on the educational careers of a Danish cohort born 1954. However, although the Danish educational system at that time had a specific institutional structure (cf. Figure 1), the MTMU can be accommodated to almost any diversified educational system. Future research on educational transitions should therefore exploit the opportunities and flexibility of the model to study the selectivity of educational decisions in diversified educational systems.

References

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

Amemiya, T. (1975). Qualitative Response Models. *Annals of Economic and Social Measurement*, 4, 363-388.

Ayalon, H. & Shavit, Y. (2004). Educational Reforms and Inequalities in Israel: The MMI Hypothesis Revisited. *Sociology of Education*, 77, 103-120.

Ballarino, G. & Schadee, H. (2010). Allocation and distribution: A discussion of the educational transition model, with reference to the Italian case. *Research in Stratification and Social Mobility*, 28, 45-58.

Becker, R. (2003). Educational Expansion and Persistent Inequalities of Education: Utilizing Subjective Expected Utility Theory to Explain Increasing Participation Rates in Upper Secondary School in the Federal Republic of Germany. *European Sociological Review,* 19, 1-24.

Becker, R. & Hecken, A.E. (2008). Why are Working-class Children Diverted from Universities?—An Empirical Assessment of the Diversion Thesis. *European Sociological Review*, 25, 233-250.

Berk, R.A. (1983). An Introduction to Sample Selection Bias in Sociological Data. *American Sociological Review*, 48, 386-398.

Boudon, R. (1974). *Education, Opportunity and Social Inequality*. New York: Wiley.

Breen, R. & Jonsson, J.O. (2000). Analyzing Educational Careers: A Multinomial Transition Model. *American Sociological Review*, 65, 754-772.

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

Erikson, R. & Goldthorpe, J.H. (1992). *The Constant Flux*. Oxford: Oxford University Press.

Gambetta, D. (1987). *Where they pushed or did they Jump? Individual decision mechanisms in education*. Cambridge: Cambridge University Press.

Halpin, B. (1999). Is Class Changing? AWork-life History Perspective on the Salariat. *Sociological Research Online*, 4.

Hansen, E. J. (1995). *En generation blev voksen*. [A Generation Grew Up]. Copenhagen: SFI – Danish Institute of Social Research.

Hansen, M.N. (2007). Rational Action Theory and Educational Attainment: Changes in the Impact of Economic Resources. *European Sociological Review*, 24, 1-17.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47, 153-161.

Hensher, D.A. & Greene, W.H. (2003). The Mixed Logit model: The state of practice. *Transportation*, 30, 133-176.

Jao, J.C. & McKeever, M. (2006). Ethnic Inequalities and Educational Attainment in Taiwan. *Sociology of Education*, 79, 131-152.

Jæger, M.M. (2007). Educational Mobility Across Three Generations: The Changing Impact of Parental Social Class, Economic, Cultural and Social Capital. *European Societies*, 9, 527-550.

Jæger, M.M. (2009). Equal Access but Unequal Outcomes: Cultural Capital and Educational Choice in a Meritocratic Society. *Social Forces*, 87, 1943-1971.

Jæger, M.M. & Holm, A. (2007). Does parent's economic, cultural, and social capital explain the social class effect on educational attainment in the Scandinavian mobility regime? *Social Science Research*, 36, 719-744.

Karlson, K. B., Holm, A., & Breen, R. (2010). Comparing Regression Coefficients Between Models using Logit and Probit: A New Method. *Unpublished paper*. [Draft available from authors].

Kreidl, M. (2004). Politics and Secondary School Tracking in Socialist Czechoslovakia, 1948-1989. *European Sociological Review*, 20, 123-139.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data*. Cambridge: Cambridge University Press.

Lucas, S.R. (2001). Effectively Maintained Inequality: Education Transitions, Track Mobility, and Family background Effects. *The American Journal of Sociology*, 106, 1642-1690.

Maastekaasa, A. (2006). Educational Transitions at Graduate Level: Social Origins and Enrolment in PhD Programmes in Norway. *Acta Sociologica*, 49, 437-453.

Mare, R. D. (1980). Family background and School Continuation Decisions. *Journal of the American Statistical Association*, 75, 295-305.

Mare, R.D. (1981). Change and Stability in Educational Stratification. *American Sociological Review*, 46, 72-87.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105-142). New York: Academic Press.

McLachlan, G. & Peel, D. (2000). *Finite Mixture Models*. New York: Wiley.

Need, A. & de Jong, U. (2001). Educational Differentials in the Netherlands: Testing Rational Action Theory. *Rationality and Society*, 13, 71-98.

Nicoletti, C. & Rondinelli, C. (2006). The (mis)specification of discrete time duration models with unobserved heterogeneity: a Monte Carlo Study. *ISER Working Paper* (2006-53). Colchester: Institute for Social & Economic Research, University of Essex.

Powers, D. & Xie, Y. (2000). *Statistical Methods for Categorical Data Analysis*. Boston: Academic Press.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2004). GLLAMM Manual. *U.C. Berkeley Division of Biostatistics Working Papers Series, WP 160*. California: University of California, Berkeley.

Sewell, W.H., Haller, A.O., & Portes, A. (1969). The Educational and Early Occupational Attainment Process. *American Sociological Review*, 34, 82-92.

Tieben, N. & Wolbers, M.H.J. (2010). Transitions to post-secondary and tertiary education in the Netherlands: a trend analysis of unconditional and conditional socio-economic background effects. *Higher Education*, 60, 85-100.

Tolsma, J., Need, A., & de Jong, U. (2010). Explaining Participation Differentials in Dutch Higher Education: The Impact of Subjective Success Probabilities on Level Choice and Field Choice. *European Sociological Review*, 26, 235-252.

Trussel, J. & Richards, T. (1985). Correcting for Unmeasured Heterogeneity in Hazard Models Using the Heckman – Singer Procedure. *Sociological Methodology,* 15, 242-276.

Vaupel, J. W. & Yashin, A.I. (1985). Heterogeneity's Ruses: Some Surprising Effects of Selection on Population Dynamics. *American Statistician*, 39, 176-185.

Wedel, M. & DeSarbo, W.S. (1995). A Mixture Likelihood Approach for Generalized Linear Models. *Journal of Classification*, 12, 21-55.

Wedel, M. & DeSarbo, W.S. (2002). Mixture Regression Models. In J.A. Haganaars, & A.L. McCutcheon (Eds.), *Applied Latent Class Analysis* (pp. 366-382). Cambridge: Cambridge University Press.

Winship, C. & Mare R.D. (1984). Regression Models for Ordinal Variables. *American Sociological Review*, 49, 512-525.

Winship, C. & Mare R.D. (1992). Models for Sample Selection Bias. *Annual Review of Sociology,* 18, 327-350.

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Yamaguchi, K. (1987). Event-History Analysis: Its Contributions to Modeling and Causal Inference. *Sociological Theory and Methods*, 2, 61-82.

Yatchew, A. & Griliches, Z. (1985). Specification Error in Probit Models. *The Review of Economics and Statistics,* 67, 134-139.

TABLE 1. Marginal distribution of dependent variables

	Frequency	Percent
Secondary education		
Leave school	406	18.46
Vocational	1,075	48.89
Academic (Gymnasium)	718	32.65
Total	2,199	100.00
Tertiary education		
Leave school	170	23.68
Short cycle	336	46.80
University	212	29.53
Total	718	100.00

TABLE 2. Descriptive statistics: Means and standard deviations

	Total sample		Sample completing academic secondary education	
	Mean	SD	Mean	SD
Explanatory variables				
Parental highest social class				
I/II	0.115	-	0.217	-
III	0.089	-	0.141	-
IV	0.269	-	0.266	-
V/VI	0.120	-	0.093	-
VII (reference)	0.286	-	0.142	-
Missing	0.121	-	0.141	-
Parental highest education (years)	10.083	2.531	11.287	2.876
Parental highest education, missing dummy (reference: not missing)	0.216	-	0.117	-
Non-intact family (reference: intact family)	0.128	-	0.102	-
Boy (reference: girl)	0.509	-	0.500	-
Ability (0-100)	53.498	18.124	66.163	14.154
Instrumental variables				
Share in school class in comprehensive school completing vocational secondary education (0-1)	0.488	0.161	-	-
Share in school class in comprehensive school completing academic secondary education (0-1)	0.327	0.211	-	-
N	2,199		718	

TABLE 3. Multinomial transition model without and with unobserved heterogeneity captured by two latent classes: Logit coefficients with t-statistics in parentheses.

	Standard MTM	**MTMU**	**Coefficient change[a]**
FIRST TRANSITION: Primary to secondary (reference: leave school)			
PANEL A: Vocational track			
Parental highest social class			
I/II	0.431 (1.50)	0.521 (1.67)	
III	0.572 (1.93)	0.609* (1.97)	x I
IV	0.447* (2.88)	0.491* (2.83)	
V/VI	0.337 (1.66)	0.336 (1.60)	
VII (reference)	-	-	-
Missing	0.007 (0.03)	0.054 (0.25)	-
Parental highest education (years)	0.036 (1.04)	0.048 (1.20)	
Parental highest education, missing dummy (reference: not missing)	-0.215 (-1.50)	-0.250 (-1.57)	-
Non-intact family (reference: intact family)	-0.280 (-1.63)	-0.299 (-1.66)	
Boy (reference: girl)	0.565* (4.64)	0.572* (4.50)	
Ability (0-100)	0.025* (6.45)	0.029* (4.22)	
Share in school class in comprehensive school completing vocational secondary education (0-1)	-0.892* (-2.56)	-0.925* (-2.43)	
PANEL B: Academic track			
Parental highest social class			
I/II	1.305* (4.13)	1.674* (3.97)	I
III	0.979* (2.92)	1.200* (2.89)	I
IV	1.064* (5.24)	1.333* (4.66)	I
V/VI	0.229 (0.86)	0.222 (0.67)	
VII (reference)	-	-	-
Missing	0.814* (3.22)	1.083* (3.10)	-
Parental highest education (years)	0.251* (6.52)	0.343* (5.30)	I
Parental highest education, missing	-0.915	-1.253	-

dummy (reference: not missing)	(-4.70)	(-4.19)	
Non-intact family (reference: intact family)	-0.297 (-1.30)	-0.315 (-1.08)	
Boy (reference: girl)	0.162 (1.06)	0.106 (0.52)	D
Ability (0-100)	0.089* (16.96)	0.115* (8.16)	I
Share in school class in comprehensive school completing academic secondary education (0-1)	1.762* (5.18)	2.286* (4.46)	I
SECOND TRANSITION: Secondary to tertiary (reference: leave school)			
PANEL C: Short-cycle track			
Parental highest social class			
I/II	0.676 (1.81)	0.847 (1.42)	I
III	0.296 (0.78)	0.356 (0.80)	I
IV	0.441 (1.35)	0.542 (1.19)	I
V/VI	-0.329 (-0.84)	-0.233 (-0.51)	D
VII (reference)			
Missing	-1.447* (-4.10)	-1.508* (-2.49)	-
Parental highest education (years)	-0.088* (-2.21)	-0.085 (-1.94)	x
Parental highest education, missing dummy (reference: not missing)	-0.673* (-2.02)	-0.582 (-1.39)	x
Non-intact family (reference: intact family)	0.064 (0.19)	-0.053 (-0.13)	*Changes direction*
Boy (reference: girl)	-0.477* (-2.34)	-0.339 (-0.67)	x D
Ability (0-100)	-0.007 (-1.01)	-0.006 (-0.81)	
PANEL D: University track			
Parental highest social class			
I/II	1.295* (3.02)	2.255* (2.52)	I
III	0.682 (1.53)	1.035 (1.40)	I
IV	0.944* (2.43)	1.534* (2.15)	I
V/VI	0.266 (0.58)	0.956 (0.96)	I
VII (reference)			
Missing	-0.918* (-2.14)	-1.098 (-1.02)	-

Parental highest education (years)	-0.009 (-0.19)	0.022 (0.30)	*Changes direction*
Parental highest education, missing dummy (reference: not missing)	-0.005 (-0.02)	0.868 (0.94)	-
Non-intact family (reference: intact family)	-0.488 (-1.22)	-0.972 (-1.43)	D
Boy (reference: girl)	0.916* (4.01)	1.868* (2.65)	I
Ability (0-100)	0.009 (1.09)	0.009 (0.65)	
MODEL INFORMATION			
Number of observations at first transition	2,199	2,199	
Number of observations at second transition	718	718	
-2LogL	4,881.88	4,875.40	
Pseudo-R^2	19.29 %	19.40 %	

Note: * Statistically significant on a 5 percent level. [a] x indicates that the coefficient from MTMU is statistically significant, while the counterpart from MTM is not. I indicates numerically increasing coefficient, while D indicates numerically decreasing coefficient. The criterion of a large increase or decrease is when the MTMU coefficient is either more or less than one fifth of the MTMU.

TABLE 4. Estimated unobserved component from MTMU: Two latent classes

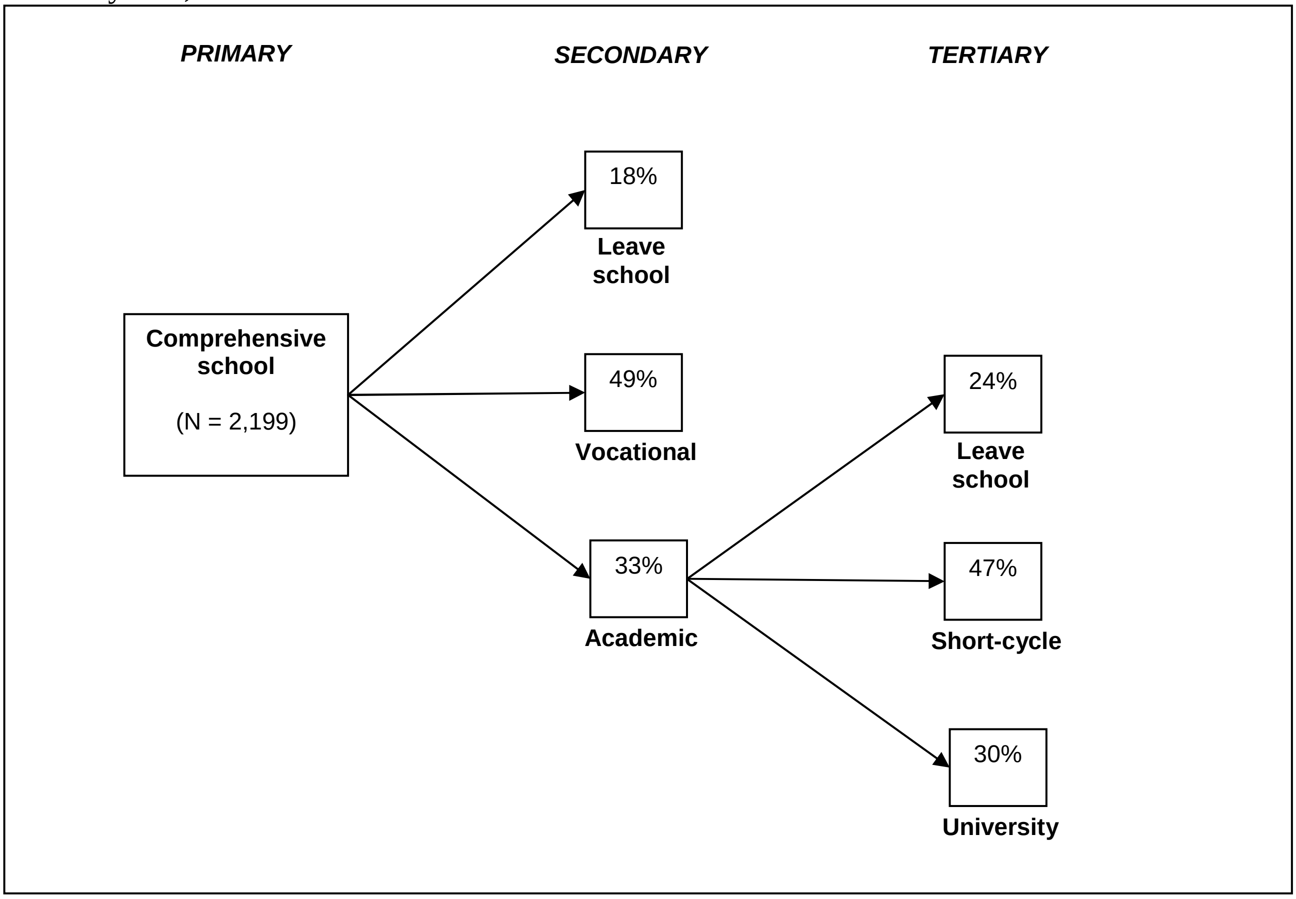
	Type 1	**Type 2**
First transition, vocational	-0.374	0.290
First transition, academic	-2.167	1.682
Second transition, short-cycle	-0.578	0.448
Second transition, university	-12.298	9.542
Weight (share in latent class)	0.437	0.563

TABLE 5. Average partial effects derived from multinomial transition model without and with unobserved heterogeneity captured by two latent classes.

	Standard MTM	MTMU	Difference (MTMU-MTM)
FIRST TRANSITION: Primary to secondary (reference: leave school)			
PANEL A: Vocational track			
Parental highest social class			
I/II	0.057	0.069	0.012
III	0.077	0.081	0.004
IV	0.059	0.065	0.005
V/VI	0.044	0.043	-0.001
VII (reference)	-	-	-
Missing	0.001	0.007	-
Parental highest education (years)	0.004	0.006	0.001
Parental highest education, missing dummy (reference: not missing)	-0.025	-0.029	-
Non-intact family (reference: intact family)	-0.033	-0.035	-0.002
Boy (reference: girl)	0.076	0.076	0.000
Ability (0-100)	0.003	0.004	0.000
Share in school class in comprehensive school completing vocational secondary education	-0.110	-0.113	-0.003
PANEL B: Academic track			
Parental highest social class			
I/II	0.255	0.316	0.061
III	0.196	0.238	0.042
IV	0.212	0.261	0.049
V/VI	0.047	0.046	-0.001
VII (reference)			
Missing	0.165	0.216	-
Parental highest education (years)	0.051	0.070	0.019
Parental highest education, missing dummy (reference: not missing)	-0.173	-0.226	-
Non-intact family (reference: intact family)	-0.060	-0.063	-0.004
Boy (reference: girl)	0.033	0.022	-0.011
Ability (0-100)	0.018	0.024	0.005
Share in school class in comprehensive school completing academic secondary education	0.359	0.468	0.109
SECOND TRANSITION: Secondary to tertiary (reference: leave school)			
PANEL C: Short-cycle track			
Parental highest social class			
I/II	0.128	0.160	0.032
III	0.053	0.064	0.011

IV	0.081	0.099	0.018
V/VI	-0.052	-0.038	0.015
VII (reference)			
Missing	-0.175	-0.179	-
Parental highest education (years)	-0.015	-0.014	0.001
Parental highest education, missing dummy (reference: not missing)	-0.099	-0.087	-
Non-intact family (reference: intact family)	0.011	-0.009	*Direction change*
Boy (reference: girl)	-0.074	-0.054	0.020
Ability (0-100)	-0.001	-0.001	0.000
PANEL D: University track			
Parental highest social class			
I/II	0.279	0.432	0.153
III	0.151	0.227	0.076
IV	0.207	0.323	0.116
V/VI	0.058	0.210	0.152
VII (reference)			
Missing	-0.177	-0.206	-
Parental highest education (years)	-0.002	0.005	Direction change
Parental highest education, missing dummy (reference: not missing)	-0.001	0.192	-
Non-intact family (reference: intact family)	-0.100	-0.186	-0.086
Boy (reference: girl)	0.201	0.378	0.177
Ability (0-100)	0.002	0.002	0.000

FIGURE 1. Flow-chart showing educational pathways (with percentages) in the Danish school system, cohort born around 1954.



NOTE: The percentages for tertiary education sum to 101 because of rounding.

CSER

CENTRE FOR STRATEGIC
RESEARCH IN EDUCATION