Working Paper Series
CSER WP No. 0004

Dealing with Selection Bias in Educational Transition Models: The Bivariate Probit Selection Model

Anders Holm and Mads Meier Jæger

Centre for Strategic Educational Research
DPU, Aarhus University

Dealing with Selection Bias in Educational Transition Models: The Bivariate Probit Selection Model

Anders Holm* and Mads Meier Jæger**
This version: June, 2010

Running head: Dealing with Selection Bias in Educational Transition Models

Abstract:
This paper proposes the bivariate probit selection model (BPSM) as an alternative to the traditional Mare model for analyzing educational transitions. The BPSM accounts for selection on unobserved variables by allowing for unobserved variables which affect the probability of making educational transitions to be correlated across transitions. The BPSM is easy to estimate with standard software. We use simulated and real data to illustrate how the BPSM improves on the traditional Mare model in terms of correcting for selection bias and providing credible estimates of the effect of family background on educational success. We conclude that models which account for selection on unobserved variables and high-quality data are both required in order to estimate credible educational transition models.

* Centre for Strategic Educational Research, Danish School of Education, University of Education, Denmark, email: kbk@dpu.dk.
** Centre for Strategic Educational Research, Danish School of Education, University of Education: aholm@dpu.dk.

Total words: 8118 words (including all text, notes, and references) 4 tables 1 appendix. Keywords: Mare model; unobserved heterogeneity; educational transitions; sequential logit model; methodology Keywords: Mare model; unobserved heterogeneity; educational transitions; sequential logit model; methodology

1. Introduction

Robert Mare's (1979, 1980, 1981) model of educational transitions represents one of the major methodological contributions to the literature on family background and educational success. Instead of years of completed schooling and a linear regression model, Mare suggested to treat educational attainment as a sequence of discrete transitions from lower to higher educational levels and to use a sequential logit model. The principal advantages of Mare's educational transition model are that, first, the model is invariant to changes over time in the overall distribution of education, second, the model conforms better to the way most sociologists think about educational attainment (as a sequence of transitions) and, third, it allows researchers to model the effect of family background variables on the probability of making successive educational transitions.

The Mare model is, and for long time has been, highly influential in applied research (e.g., Garnier & Raffalovich 1984; Cobalti 1990; Heath & Clifford 1990; Shavit & Blossfeld 1993; Hansen 1997; Shavit & Westerbeek 1998; Vaid 2004; articles in *Research in Social Stratification and Mobility* vol. 28, issue 1, 2010). One of the consistent findings from applied research using the Mare model is that the effect of family background variables tends to decrease or "wane" across educational transitions. Substantive theories such as the theories of Maximally Maintained Inequality (Raftery & Hout 1993) and Effectively Maintained Inequality (Lucas 2001) have been proposed to explain this "waning coefficients" phenomenon.

However, in two influential papers Cameron and Heckman (1998, 2001) argue that the waning coefficients in the Mare model may be artifacts of, first, an arbitrary choice of functional form in the logit model and, second, selection on unobserved variables. Selection on unobserved variables means that the group of individuals "at risk" of making educational transitions becomes increasingly selective at higher transitions due to characteristics that are not observed in the data. These unobserved characteristics might relate to academic ability, motivation, or cultural resources.

For example, compared to the group that completes elementary school and faces the decision about whether or not to enter high school (the first transition), the group that eventually completes high school and now has to decide about whether or not to enter university (the second transition) has not only better academic ability but is probably also more academically motivated. It is unrealistic to assume that we are able to observe (and control for) all the relevant characteristics that make the group that faces the second transition different from the group that faces the first transition and, consequently, the problem of selection on unobserved variables is endemic in all analyses using the Mare model. It can be shown that selection on unobserved variables leads to bias in the estimates of the effect of explanatory variables on the probability of making the second and higher educational transitions. This type of bias may be the cause of the waning coefficient phenomenon reported in previous research. Mare himself (1979, 1980, 1981, 1993) noted that selection on unobserved variables might lead to bias in the Mare model.

In addition to selection on unobserved variables, the Mare model may also yield biased results because the error variances in the model's equations for each educational transition are different. Differences in error variances across model equations are potentially important in the Mare model because the coefficients which express the effects of explanatory variables in each model equation are scaled relative to the error variance in that equation (i.e., the regression coefficients themselves are not known but the regression coefficients divided by the error variance in the equation are known). Consequently, differences in the effects of explanatory variables across transitions (for example, waning coefficients) might be driven by differences in error variances rather than reflecting real differences. This type of bias is called a scaling effect.

In this paper we propose a simple approach to dealing with selection on unobserved variables in the Mare model.[1] Our approach is based on a bivariate probit selection model (BPSM) and it allows for unobserved variables which affect the probability of making lower educational transition to be correlated with unobserved variables which affect the probability of making higher educational transitions. Compared to the other approaches presented in this special issue, our approach is technically relatively simple, it is easy to understand, and it may be readily implemented using existing software (such as Stata, SAS, and R). The potential drawbacks of our approach are that, in order to be identified, the BPSM requires, first, parametric assumptions and, second, instrumental variables to provide exogenous variation in the probability of making each educational transition. Our approach is applicable is the "classic" Mare setup, i.e., in the case in which the analyst is interested in modeling a sequence of binary educational transitions. We illustrate bias in the Mare model and the applicability of our approach using simulated data and empirical data from the British National Child Development Study (NCDS). To simplify the presentation, we develop a BPSM model with only two educational transitions. However, our approach easily generalizes to a BPSM model with many transitions.

Our paper has three key messages. First, there is no technical "magic bullet" for dealing with selection on unobserved variables in educational transition models. The BPSM represents a useful tool for addressing selection bias, but the analyst's ability to address selection on unobserved variables is ultimately limited by the quality of the data. Our empirical example using

---

[1] Several studies deal with other aspects of the Mare model such as improvements in identification from repeated measurements of family background variables (Lucas 2001) and a more parsimonious formulation of the Mare model along the lines of Anderson's (1984) Stereotype Ordered regression model (Hauser & Andrew 2006). With the exception of Breen and Jonsson (2000) we are not familiar with any study in sociology that addresses selection on unobserved variables in educational transition models. In the economics of education literature selection bias in educational transition models is more frequently dealt with (e.g., Chevalier & Lanot 2002; Lauer 2003; Arends-Kuenning & Duryea 2006; Colding 2006).

NCDS data and exploiting different amounts of information in this data illustrate this point. Second, selection on unobserved variables always leads to biased estimates of the effect of explanatory variables when analysts use a selective sample. For example, an analyst may be interested in higher education only (a later educational transition) and disregard earlier transitions. Our simulation results illustrate that not explicitly modeling selection processes in previous transitions leads to biased results at higher educational transitions. Third, in empirical applications it is typically not possible to separate bias from selection on unobserved variables from bias from scaling effects (Mare 2006). In our theoretical analysis we do, however, show that the likely direction of either type of bias can be assessed under a set of reasonable assumptions.

2. Selection in Educational Transition Models

2.1 The Baseline Model

Mare's (1979, 1980, 1981) educational transition model consists of a sequence of binary logit models in which the dependent variables are dummy variables for making the *j*'th educational transition conditional on previously having made the *j*-1'th transition. Our model consists of only two transitions and, instead of the logit specification used by Mare, we use the probit specification.

Define two latent stochastic variables y_1^* and y_2^* which capture the propensity to make the first and second transition in an educational system. As described previously, the first transition represents the transition from elementary school to high school (or, equivalent, to upper secondary education such as *A* levels in the UK), and the second transition represents the transition from high school to higher education (for example, college, university, or university-college education). These types of transitions exist in most Anglo-American and Western European educational systems. We assume that in order for individuals to make the second transition, they must first successfully make the first transition. We do not observe the latent variables y_1^* and y_2^* (the propensities to make the two transitions) but instead two binary variables *y* indicating if individuals actually make each of the two transitions. These binary variables are defined as $y_j = 1$ if $y_j^* > 0$ and 0 otherwise, *j = 1,2*, with *j* indexing transitions.

The likelihood that an individual makes each educational transition depends on a set of observed explanatory variables (for example, parents' education and income, academic ability, and motivation) and some unobserved variables. The sequence of educational transitions can be represented by the following system of linear regression equations

$$\begin{aligned} y_1^* &= \beta_1 x_1 + e_1 \\ y_2^* &= \beta_2 x_2 + e_2 \end{aligned} \quad (1),$$

where x_j, *j = 1,2* represents observed variables for each transition and e_j, *j = 1,2* represent error terms that capture the effect of unobserved variables. Because the observed transitions are binary variables, we cannot identify the variance of the error terms *e* and, as is always the case in the probit model, they are both normalized to 1. In order to obtain the bivariate probit selection model (BPSM), we assume that the error terms follow a bivariate normal distribution

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \sim N(\mathbf{0}, \mathbf{\Sigma}),$$

where

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

In the BPSM model the parameter ρ captures the correlation between the unobserved variables in each educational transition. In other words, the parameter ρ is a summary measure of the importance of selection on unobserved variables. Our reason for choosing the probit specification over the logit specification is that the probit specification allows us to estimate ρ and thus to take into account that unobserved variables that affect the propensity to make the first transition are likely to be correlated with the unobserved variables that affect the propensity to make the second transition. There is no bivariate logistic distribution and, consequently, it is not possible to estimate ρ with the logit specification used in the traditional Mare model.

2.2 Bias from Selection on Unobserved Variables

The fundamental problem in analyzing educational transitions is that the probability of making the second transition depends on whether or not individuals have previously made the first transition; i.e., the group at risk at the second transition always represents a selected sample. In this section we explain how sample selection might lead to bias in educational transition models. First, using simulated data we give an intuitive explanation of how selection leads to bias is the estimated effects of explanatory variables on the probability of making the second transition. Second, we provide a formal statistical account of selection bias.

We use simulated data to show how selection effects operate and illustrate their consequences. The advantage of simulated data is that we control the data generating process and know all the true relationships in the data. The selection problem arises because the group of students eligible for making the second transition differs from the group that is eligible for making the first transition both in terms of observed and unobserved characteristics. The Mare model does not take selection on unobserved variables into account and, consequently, it yields biased estimates of the effect of observed explanatory variables on the probability of making the second (and higher) transition(s). But how strong is this bias? We estimate two types of models on the simulated data: a traditional Mare model but, for comparability, using the probit rather than the logit specification (we label this model the Mare probit model) and our BPSM.

We generate the simulated data on the basis of the following model with two educational transitions

$$\begin{aligned} y_1^* &= x_1 + x_2 + v_1 + \varepsilon_1; y_1 = 1 \text{ iff } y_1^* > 0 \\ y_2^* &= x_1 + x_3 + v_2 + \varepsilon_2; y_2 = 1 \text{ iff } y_1^*, y_2^* > 0 \end{aligned} \qquad (2)$$

In this model $y_1 = 1$ represents making the first transition (for example, completing high school) and $y_2 = 1$ represents making the second transition (for example, completing college), and $x_1 - x_3$ represent explanatory variables. Furthermore, the terms v_1 and v_2 summarize unobserved variables which are correlated across transitions and ε_1 and ε_2 summarize the random error variance. We generate the data such that $x_1, x_2 \overset{\text{iid}}{\sim} N(0,1)$, $x_3 = x_2 + \upsilon; \upsilon \sim N(0,1)$, $\varepsilon_1, \varepsilon_2 \overset{\text{iid}}{\sim} N(0,1)$ and $\nu_1, \nu_2 \sim N(\mathbf{0}, \mathbf{\Sigma})$, $\mathbf{0} = (0,0)'$ and $\mathbf{\Sigma} = \begin{pmatrix} 1 & \tau \\ \tau & 1 \end{pmatrix}$. Note that τ represents the correlation between ν_1 and ν_2, that is the correlation between the unobserved variables, $\tau = cor(\nu_1, \nu_2)$, while $\rho = corr(e_1, e_2) = corr(v_1 + \varepsilon_1, v_2 + \varepsilon_2) = \frac{\tau}{2}$ in Equation (1) represents the *total* correlation between both the unobserved variables and the random error variance. The difference between Equation (1) and Equation (2) is that in Equation (1) we do not distinguish between the effect of unobserved variables systematic that are correlated across transitions (v) and unobserved variables that are not correlated across transitions (ε). In Equation (2) x_1 is an observed family background variable of substantive interest, for example father's education. The variables x_2 and x_3 are observed explanatory variables that vary over transitions and which act as instrumental variables which help to identify the BPSM (for example, x_2 and x_3 may be grades or test scores from exams taken before the first and second transition). As we discuss in more detail below, these instrumental variables are required in order for the BPSM not to be identified solely from the parametric assumption of bivariate normality.

In each simulation of the model in Equation (2) we use a sample of 1000 observations and 200 replications. We use different values for τ to emulate different levels of selection on unobserved variables. For each simulation and fixed value of τ, we estimate the Mare probit model and the BPSM for the probability of making the second transition, $y_2 = 1$. The parameter of interest is the coefficient on the observed explanatory variable x_1. We set up the simulations so that the true value of this coefficient is 1. We then calculate the absolute bias (compared to the true coefficient of 1) of the estimates of the effect of x_1 for different values of τ in the Mare probit model and the BPSM.[2] Figure 1 shows the magnitude of the absolute bias in both models and for values of τ ranging from 0 and to 0.95.

-- FIGURE 1 ABOUT HERE --

[2] We need to rescale the estimates of the effect of x_1 when calculating the magnitude of the bias. The rescaling of the effect of x_1 in the Mare probit model and the BPSM is given by $\frac{\hat{b}}{\sqrt{(1-\rho^2)}}$, where $\hat{b}$ is the estimated coefficient on x_1. The rescaling is necessary because the true model has error variance $\mathrm{var}(\nu_1 + \varepsilon_1) = \mathrm{var}(\nu_2 + \varepsilon_2) = 1 + 1 = 2$ but the models which use the selective sample that has made the first educational transition has error variance $1 - \rho^2$.

Figure 1 shows that, first, in the Mare probit model bias in the coefficient of x_1 increases more or less monotonically with the correlation between the unobserved variables in each transition (e_1 and e_2) and, second, in the BPSM the estimated coefficient of x_1 is largely unaffected by the size of the correlation between the unobserved variables. These simulations illustrate that, although the substantive magnitude of the parameter bias does not relate to any real-world application (other simulations would have led to other estimates of this bias), the Mare probit model yields increasingly biased estimates of the effect of observed explanatory variables with increasing levels of selection on unobserved variables.

-- TABLE 1 ABOUT HERE --

Our main result from the simulations is that the Mare probit model is highly susceptible to selection on unobserved variables. We further illustrate this point in Table 1 in which, again based on our simulated data, we report probit coefficients for x_1 from three model specifications of $y_2 = 1$. In the simulations the correlation between the unobserved variables (v_1 and v_2) is set to 0.99 (τ), which means that the correlation between all error terms (v_1 and ε_1 and v_2 and ε_2 u_2 - ρ) is 0.5. The first of the three specifications is a Mare probit model (which disregards selection on unobserved variables entirely). The true coefficient on x_1 is 1, but the estimate from the Mare probit model is 0.840 and downwardly biased. The second specification is a Mare probit model which includes v_2; i.e., a Mare model which controls for unobserved variables in the second transition. The reason why we can control for the unobserved variables in the Mare probit model is that we use simulated data in which we know all true relations in the data (this would not be the case with real data). The Mare probit model which includes v_2 estimates the effect of x_1 at 1.019 and thus very close to the true value of 1. Finally, the third specification is the BPSM which estimates the effect of x_1 very accurately

(1.023). The main conclusion from this simulation example is that the traditional Mare model yields biased results of explanatory variables at the second (and higher) transition(s) because it fails to control for unobserved variables which are correlated across transitions and which lead to an increasingly selected sample. Consequently, analysts who analyze selective samples, for example students in secondary or higher education, and who do not explicitly model the selectivity of these samples arising from earlier educational transition, end up with biased estimates of explanatory variables. This is an important point and entails that analysts should model the entire educational career even if they are only interested in a later educational transition.

2.3 Bias from Scaling Effects

In addition to selection on unobserved variables, we might also experience bias in the estimated effects of explanatory variables at higher educational transitions because the error variance in the model equation for the first educational transition (which includes the whole sample) is different from the error variance in the model equation for the second transition (which includes a selected sample). It turns out that the variance of the error term in the selected sample is smaller than the variance in the whole sample (intuitively, this happens because students become more similar on unobserved as well as on observed characteristics at higher educational transitions, cf. Table 2). Unlike the linear regression model, in binary probability models such as the Mare model the variance of the error term is not identified and must be normalized (in the probit model the error variance is normalized to 1 and in the logit model the error variance is normalized to $\pi^2/3$). The reason why we do not know the error variance is that the dependent variable is binary and we assume (rather than know) a functional form for the underlying probability distribution. Furthermore, in binary probability models the actual regression coefficients associated with explanatory variables are not identified. Rather, the regression coefficients divided by the error variance in the probability (probit/logit) model are identified (intuitively, the regression coefficients must be expressed relative to some "unit" or scale for the dependent variable). However, since the error variance in the selected sample is lower than the error variance in the whole sample the regression coefficients in the second transition are upwardly biased when we analyze the selected sample because the denominator (in the case of the probit) is smaller than 1 (assuming that the variance in the selected sample is 0.8 and the true regression coefficient is 1, it is easy to illustrate the upward bias from scaling. For the first transition we get: $\tilde{\beta} = 1/1 = 1$ and for the second transition we get: $\tilde{\beta} = 1/0.8 = 1.25$). We describe bias from scaling formally in Appendix 1.

2.4 Selection Effects and Scaling Effects in a Formal Model

The previous sections have illustrated the intuition behind selection and scaling effects in the Mare model. We have also illustrated how the BPSM deals with the problem of selection on unobserved variables. In this section we recast the BPSM model in a formal context which addresses both selection and scaling effects. We also illustrate how selection and scaling effects most likely bias estimates of effects of explanatory variables.

The problems of selection on unobserved variables and scaling effects exists at the second (and higher) educational transition(s). We decompose the probability of making the second educational transition into three components: (1) the true effects of the observed explanatory variables (the *x*'s), (2) the selection effect, and (3) the scaling effect. Our decomposition is based on an approximation developed by Nicoletti and Peracchi (2001) which has been shown to work well for ρ correlation coefficients up to 0.8 (see Appendix 2 for a derivation of the approximation). The approximation is a convenient way of representing *attenuation bias* (the combined effect of selection and scaling bias) and has the following form

$$P\left(Y_2 = 1 \mid Y_1 = 1\right) \approx \Phi\left(\frac{\beta_2 x_2 + \rho\lambda\left(\beta_1 x_1\right)}{\sqrt{1-\rho^2\left\{\beta_1 x_1 \lambda\left(\beta_1 x_1\right) + \lambda\left(\beta_1 x_1\right)^2\right\}}}\right). \text{ (3)}$$

Equation (3) shows the approximation of the probability of making the second educational transition conditional on having made the first transition. The true effects of the observed explanatory variables on the probability of making the second transition are represented by the term $\beta_2 x_2$. However, because of selection and scaling we do not estimate these true effects but instead biased effects. The selection term is $\rho\lambda(\beta_1 x_1)$ and arises from selection in the first transition. The scaling term is $\sqrt{1-\rho^2\left\{\beta_1 x_1 \lambda\left(\beta_1 x_1\right) + \lambda\left(\beta_1 x_1\right)^2\right\}}$ and captures the difference in the standard deviation (the square root of the variance) in the selected sample compared to in the whole sample. In empirical applications we estimate a combination of the true and the attenuation effect. That is, we estimate $P(Y_2 = 1 \mid Y_1 = 1) = \Phi\left(\alpha x_2\right)$ where

$$\alpha x_2 \approx \frac{\beta_2 x_2 + \rho\lambda\left(\beta_1 x_1\right)}{\sqrt{1-\rho^2\left\{\beta_1 x_1 \lambda\left(\beta_1 x_1\right) + \lambda\left(\beta_1 x_1\right)^2\right\}}}. \text{ (4)}$$

Since in most cases we do not have any information about the actual magnitude of the selection and the scaling effect, we cannot determine the severity of the parameter bias in the Mare model. This is the fundamental problem. However, by applying results from statistical theory and plausible assumptions about the relationships in the model, we may learn more about the likely direction of the selection and the scaling bias.

We begin by assessing selection bias. Following Wooldridge (2002), we note that the differential coefficient of the selection term wrt. x_1 is

$$\partial\lambda(\beta_1 x_1) \big/ \partial x_1 = -\beta_1 \lambda(\beta_1 x_1)\left(\beta_1 x_1 + \lambda(\beta_1 x_1)\right). \quad (5)$$

The differential coefficient measures the rate of change in the selection term when x_1 changes, and it provides a convenient tool for evaluating what happens when we manipulate parameters in the model. However, in order to make inferences on the direction of the selection bias, we need to apply assumptions about the relationships in the data. We therefore invoke the four plausible assumptions

1. $\beta_1, \beta_2 > 0$. Assume that x_1 and x_2 are family background variables which both have a positive effect on the probability of making both transitions (for example, father's education and mother's education).
2. The explanatory variables x_1 and x_2 are positively correlated. This would make sense if they both measure some aspect of family background such as parents' education.
3. The variable x_1 takes on a positive value larger than 0 (for example, father's education). This assumption is included for expository reasons and may be relaxed.
4. The correlation between the unobserved variables in the two transitions is positive, i.e., $\rho > 0$. This assumption makes intuitive sense and implies that if a person has a high (low) value in distribution of unobserved effect in the first transition he or she is also likely to have a high (low) value in the distribution of unobserved effects in the second transition.

Assumption 1 and 3 ensure that $\partial\lambda(\beta_1 x_1)\big/\partial x_1 < 0$, cf. Equation (5), as $\lambda(s) > 0; \forall s$. In other words, the selection effect tends to be small when x_1 is large and vice versa. Consequently, there is an inverse relationship between the selection term and x_1. When also invoking assumptions 2 and 4, we find that $\operatorname{cov}\left(\beta_2 x_2, \rho\lambda(\beta_1 x_1)\right) = \underset{>0}{\beta_2\rho}\underbrace{\operatorname{cov}\left(x_2, \lambda(\beta_1 x_1)\right)}_{<0} < 0$ because x_2 and the selection term are negatively correlated. In other words, the selection effect drives down the estimate of β_2 and, consequently, if our assumptions are valid selection on unobserved variables most likely leads to a downward bias in the estimated effect of explanatory variables at the second transition. This result might help to explain the waning coefficients reported in empirical research using the Mare model.

With regard to the scaling effect we note that $Var\left(Y_2^* \mid Y_1 = 0\right) = 1 - \rho^2\left\{\beta_1 x_1 \lambda(\beta_1 x_1) + \lambda(\beta_1 x_1)^2\right\}$ (cf. Appendix 1). In other words, the variance in the selected sample at the second transition is always smaller than the variance in the whole sample at the first transition. Consequently, the scaling effect, the denominator in Equation (3), tends to inflate the estimate of the combined attenuation effect α compared to the true effect β_2.

In summary, two interrelated processes lead to attenuation bias in the estimated effects of explanatory variables at the second transition: *selection effects* which typically lead to downward bias and *scaling effects* which lead to upward bias. Furthermore, we also need to apply parametric assumptions on the model governing selection and true effects to distinguish between true and attenuation effects. The BPSM addresses selection effects by allowing for the unobserved effects to be correlated across transitions. The BPSM does not, however, address scaling effects.

In practice, it is very difficult to deal effectively with selection effects. Equation (3) shows that the selection effect enters the model as a nonlinear function of the explanatory variables in the first transition. If these explanatory variables are the same as those also included in the second transition (i.e. $x_1 = x_2$, this is typically the case when analysts use time-invariant family background variables), the only difference between the selection effect $\rho\lambda(\beta_1 x_1)$ in Equation (3) and the true effects of the explanatory variables in the second transition, $\beta_2 x_2$ is the nonlinear specification of $\lambda(.)$. In other words, the only reason why we can separate true effect from selection effect is the functional form assumption imposed on the selection function $\lambda(.)$. This situation is unsatisfactory because, first, we deal with selection bias by imposing functional form assumptions that are completely exogenous to the data we analyze and, second, in practice there is often so little variation in $\lambda(.)$ compared to the variation in x_2 that the BPSM is unidentified.

One way of improving identification of the BPSM is to include extra explanatory variables which appear in the models for each educational transition. These variables act as "instrumental variables" which induce exogenous variation in the probability of making each transition. The instrumental variables should affect the probability of making a specific educational transition but should not have any direct effect on other educational transitions. By including these variables one insures that the selection effect varies independently of the true effect and that the BPSM is not exclusively identified from functional form assumptions. In practice, it is often difficult to find credible instrumental variables. This challenge adds further complexity to the problem of addressing selection on unobserved variables in educational transition models. In our empirical example, we include transition-specific instrumental variables which we argue affect only the first and second educational transitions students make in the British educational system.

3. Empirical Example

The first part of the paper has shown how selection on unobserved variables and scaling effects may lead to bias in the Mare model. Furthermore, we have introduced the BPSM as a simple, alternative approach which addresses selection on unobserved variables.

In this second part of the paper we provide an empirical illustration of how the BPSM may be used to address selection on unobserved variables. We analyze data from the National Child Development Study (NCDS). This data set is well suited for our analysis because, first, the British educational system at the time when the NCDS respondents completed their education (the 1960s and 1970s) had a "ladder" structure similar to that implied in the Mare model and, second, the NCDS includes variables which may be used as credible instrumental variables which help to identify the BPSM. Our example is loosely built around trying to distinguish between a "waning coefficients" hypothesis and a "constant inequality" hypotheses. The former hypothesis states that the effect of family background decreases across educational transitions and the latter hypothesis states that the effect is constant across transitions. Distinguishing between these two different hypotheses is important for theoretical and substantive reasons and has been a recurring theme in the literature using the Mare model (e.g., Raftery & Hout 1993; Shavit & Blossfeld 1993; Lucas 2001).

-- TABLE 2 ABOUT HERE --

3.1 Data

We analyze data from the National Child Development Study (NCDS). The NCDS is an ongoing longitudinal study of all children (approximately 17500) born during the first week of March 1958 in the United Kingdom (UK; see Plewis et al. 2004 for more information on the NCDS). The NCDS respondents have been followed since birth and surveys have been carried out in 1965, 1969, 1974, 1981, 1991, 1999/2000, 2004, and 2008-2009. We use a sample from the NCDS, and our sample with non-missing information on all variables is 3955. Table 2 shows descriptive statistics for all variables used in the analysis for the whole sample and for the selective sample which makes the first transition.

3.2 Variables

3.2.1 Dependent Variables: Educational Transitions

We construct two dummy variables to indicate whether respondents have completed each of two educational transitions: (1) A (advanced) level examinations and (2) higher education. In the UK students finish elementary school at around age 16. After finishing elementary school, students may choose to pursue A level examinations which, similarly with high school in the United States and the *Gymnasium* in Germany, provide access to higher education. Upon successful completion of their A levels, students may choose to enroll in higher education, for example university or university-college. In our NCDS sample 38 percent of the respondents complete A levels and, of those who complete A levels, more than 80 percent complete some type of higher education. These frequencies suggest that in the UK the first transition into A levels is highly selective while the second transition into higher education is not very selective.

3.2.2 Explanatory Variables

We include three family background variables which have often been used in previous research: father’s education (measured by years of completed schooling), mother’s education (years of completed schooling), and a dummy variable indicating whether respondents grew up in a single-parent household.

In addition to these family background variables, we also include variables measuring respondents' sex, cognitive ability, and exam performance. Our measure of cognitive ability is the first, standardized principal component extracted from a Principal Component Analysis of respondents' scores on seven different math, reading, and general ability tests taken from age 7 to age 16. This variable (which accounts for 68.7 percent of the total variance in the seven ability tests) is corrected for random measurement error and is a good proxy for respondents' "true" cognitive ability. We also include two measures of exam performance. The first measure is the respondent's performance on the General Certificate of Education (GCE) exams taken around age 16 at the end of elementary school. The NCDS includes equivalent scales of 21 O-level/CSE exams (with the codes: 1 = O-level, grade A or B; 2 = O-level, grade C and CSE grade 1; 3 = O-level, grade D or E and CSE grade 2 or 3; 4 = CSE grade 4 or 5; 5 = other result; 6 = no entry). Similarly with Breen and Yaish (2006), we invert these codes (meaning that higher values signify a better grade) and summarize respondents' total score across the 21 exams. The second measure of exam performance is the respondent's achievement at the A level examination taken at around age 18. The NCDS includes a variable which measures A level grades in the form of a 15-point scale formed by summing the three best A-level grades.

-- TABLE 3 ABOUT HERE --

3.3 Results from the Empirical Example

Table 3 shows results from the empirical example using NCD data. We estimate two different types of models: the Mare probit model and the bivariate probit selection model (BPSM). We used Stata's *probit* command to estimate the Mare probit models and the *heckprob* command to estimate the

BPSM.[3] To compare results using different information in the data we estimate three versions of each model: (1) a baseline model with the family background variables and gender; (2) a model which adds cognitive ability (which is invariant over the two educational transitions); and finally (3) a model which adds the two instrumental variables GCE performance (for the first transition) and A level performance (for the second transition).

Table 3 shows probit regression coefficients and average partial effects (APE) for each model specification. The APE expresses the population-averaged change in the probability of making an education transition resulting from a unit change in the explanatory variable of interest while holding all other explanatory variables constant, and it is easier to interpret than probit (or logit) coefficients. Table 3 also shows the estimated correlations between the unobserved variables in the BPSM, ρ.

From the baseline Mare probit model we find that parents' education has highly significant and positive effects on the probability of completing A levels (the first transition). Each additional year of parents' education increases the probability of completing A levels by 6-7 percent. Being raised in a broken family decreases the likelihood of completing A level by 16.5 percent. We also observe that the effects of parents' education and family type are much lower at the second transition into higher education than in the first transition. The APEs for parents' education in the second transition are less than half of those in the first transition, and the dummy for broken family is no longer significant. This result is consistent with the waning coefficients hypothesis. In the baseline BPSM we first observe that the estimated correlation between the unobserved variables ρ is negative and insignificant. Since ρ is the only additional (but

[3] The BPSM may also be estimated in SAS using PROC QLIM and in R using the sampleSelection package. Our NCDS sample data and Stata code for running the models presented in this section can be downloaded from the Research *in Social Stratification and Mobility* web site.

insignificant) parameter in the BPSM compared to the Mare probit model, we expect only trivial differences between the two models (the log-likelihoods for the two models are also almost identical). The main difference between the two model specifications is that parents' education is only marginally significant in the second transition in the BPSM. However, the negative estimate of ρ is highly counterintuitive and may signal a poorly identified model, especially because the baseline BPSM does not include any instrumental variables.

In the second set of models we add cognitive ability. In both the Mare probit model and the BPSM we find that, in addition to cognitive ability having a highly significant and positive effect on the probability of completing both educational transitions, the effects of parents' education and (in the first transition) the dummy for being raised in a broken family decrease dramatically. This result is not surprising since parents' education is known to affect cognitive ability indirectly (e.g., Jackson et al. 2007). More importantly, we also find that the change in the effects of parents' education across transitions are now less pronounced compared to the previous set of models in which cognitive ability was not included. This result tends to give more support to the constant inequality rather than the waning coefficient hypothesis. Including cognitive ability, arguably an important predictor of educational transitions which is often unobserved, improves our BPSM model. The estimated ρ is 0.271 but still far from being significant. Consequently, although we would expect improved identification of the unobserved variables in the model in which we are better able to control for important determinants of educational success such as cognitive ability, the BPSM is still unable to detect selection effects.

In the final set of models which include transition-specific instrumental variables: GCE exam performance in the first transition and A level performance in the second transition, we get more reliable results. The idea behind our instruments is that GCE and A level performance, net of observed cognitive ability and family background characteristics, capture respondents' contemporaneous investments (time and effort) in making a particular educational transition. The GCE and A level exams are high-stake exams. Performing well at the GCE examinations improves the likelihood of making it into A levels but, conditional on having made it into A levels and observed A level performance, GCE performance arguably does not have any direct effect on the likelihood of making it into higher education. The validity of our exam performance instruments is further reinforced by the fact that in the NCDS we are also able to control reasonably well for respondents' cognitive ability. This situation reduces the risk that we conflate the effect of cognitive ability (which is correlated with GCE and A level performance) with that of transition-specific investments (which is the exogenous and transition-specific variation we want to isolate).[4]

Table 3 shows that both in the Mare probit model and the BPSM the exam performance measures have highly significant and positive effects on the probability of making the first and second educational transitions net of cognitive ability and family background. Consequently, our transition-specific instrumental variables appear to be working well. From the Mare probit model we find that the effect of parents' education is declining across transitions and being raised in a broken family has a negative effect on the probability of making the first transition but no effect on the probability of making the second transition. Results are somewhat different for the second transition in the BPSM. Here, we find that the substantive effect of father and mother's education do not differ much across transitions (measured by APEs) and significance levels indicate much stronger effects of parents' education on the probability of making the second transition in the BPSM compared to the Mare Probit Model. These results suggest that, as might be expected because only 38 percent of NCDS respondents in our sample make the first transition, the Mare

---

[4] It would of course have been ideal to observe respondents' cognitive ability both at age 16 (just prior to making the first transition) and at age 18 (just prior to making the second transition). If this were the case we could clearly separate the effect of cognitive ability from that of transition-specific investments. Unfortunately, there is no information on cognitive ability around age 18 in the NCDS.

probit model underestimates the true effect of family background on the probability of making the second transition due to the selective nature of the sample in the second transition. (See also Table 2 which shows that respondents who make the first transition and who are eligible for the second transition come from much more privileged family backgrounds compared to the total sample) In other words, consistent with the constant inequality hypothesis there are strong effects of family background on the second educational transition as well as on the first transition. In the BPSM we find that the estimate of ρ is 0.550 and now statistically significant, thereby indicating that selection on unobserved variables is identified and present in the model.

4. Conclusion

The Mare model is a major contribution to the applied researcher's toolkit when analyzing determinants of educational success. However, despite its many advantages the Mare model is susceptible to bias from selection on unobserved variables and from scaling effects. These sources of bias may be the reason why applied research using the Mare model has often reported that the effect of family background appears to "wane" across educational transitions.

In this paper we propose that a bivariate probit selection model (BPSM) may be preferable to the traditional Mare model. The BPSM deals with selection on unobserved variables by allowing for the unobserved variables which affect the probability of making lower educational transitions to be correlated with the unobserved variables which affect the probability of making higher educational transitions. Furthermore, the BPSM is conceptually easy to understand and it can be estimated using standard software. The potential drawbacks of the BPSM are that is requires parametric assumptions and instrumental variables to be properly identified. Also, the BPSM does not remedy scaling effects.

We use simulated data and data from the National Child Development Study (NCDS) to illustrate how selection on unobserved variables leads to bias in the traditional Mare model and how our BPSM copes with selection. Our simulations show that the Mare model yields increasingly biased results as the strength of selection on unobserved variables increases. The BPSM yields consistent results even in the presence of selection on unobserved variables. Our results using NCDS data show that identification of the BPSM and, consequently, our ability to effectively address the problem of selection on unobserved variables is highly contingent upon data quality and the amount of information that is available in the data. “Naïve” Mare models which do not address selection on unobserved variables and which do not control for cognitive ability (or other important determinants of educational success which are often unobserved) suggest that the effect of family background decreases across the two educational transitions we study. This finding is consistent with the waning coefficients hypothesis. However, our most sophisticated BPSM which includes cognitive ability, transition-specific instrumental variables, and which provides credible estimates of the correlation between the unobserved variables, suggests that the effect of family background is largely constant across educational transitions. This result is consistent with the constant inequality hypothesis.

Our analysis has two implications for future research. First, we provide strong evidence that the traditional Mare model is highly susceptible to bias from selection on unobserved variables. Consequently, analysts who wish to study educational transitions should use modeling approaches that deal with the inherent selectivity issues in these types of analyses. This result also applies to analysts who study later educational transitions but ignore earlier transitions. Our findings suggest that these types of analyses also yield biased results unless the selectivity of earlier transitions is taken into account. Second, we argue that data quality and the extent to which the data is informative about the processes that explain selection is crucial for dealing effectively with selection on unobserved variables. There is no methodological “magic bullet” which fixes bias from selection on unobserved variables in educational transition models, and empirical results ultimately depend on data quality. The different approaches and results presented in this special issue clearly demonstrate this point. Consequently, improving data quality is an important task for future research analyzing educational transitions.

Appendix 1.

The Effect of Scaling on the Estimated Parameters in the Second Educational Transition

Define the Inverse Mills' Ratio as

$$\lambda(\beta_1 x_1) = \varphi(\beta_1 x_1) \big/ \Phi(\beta_1 x_1),$$

where $\varphi(.)$ and $\Phi(.)$ are the standard normal density and distribution functions (Heckman 1979).

Note that $Var\left(Y_2^*\right) = \sigma^2 \neq Var\left(Y_2^* \mid Y_1^* > 0\right)$ because

$$Var\left(Y_2^* \mid Y_1^* > 0\right) = Var\left(Y_2^* \mid Y_1 = 0\right) = \sigma^2 \left(1 - \rho^2 \left\{\beta_1 x_1 \lambda(\beta_1 x_1) + \lambda(\beta_1 x_1)^2\right\}\right)$$

(Heckman 1979). Also note that $\sigma^2 \geq Var\left(Y_2^* \mid Y_1 = 0\right)$ since the term

$0 \leq \beta_1 x_1 \lambda(\beta_1 x_1) + \lambda(\beta_1 x_1)^2 \leq 1$. Finally, note that this relationship only exists because of the assumption of joint normality.

Appendix 2. The Approximation of the Conditional Probability of Making the Second Educational Transition

In this appendix we derive the expression in Equation (1) which we use to show analytically the effect of selection on unobserved variables and scaling on the conditional probability of making the second transition. The joint probability of making both the first and the second transition is

$$\begin{aligned} P\left(Y_2=1 \mid Y_1=1\right) &= \int_0^\infty\int_0^\infty \varphi\left(y_1^*-\beta_1x_1, y_2^*-\beta_2x_2\right)\partial y_2^*\partial y_1^* \\ &= \int_{-\beta_1 x}^{\infty}\int_{-\beta_2 x_2}^{\infty} \varphi\left(e_1\right)\varphi\left(\frac{e_2 \mid e_1}{\sqrt{1-\rho^2}}\right)\partial e_2\partial e_1, \end{aligned}$$

where $e_2 \mid e_1 = e_2 + \rho\left(e_1\right)$. The mean and variance of $e_2 \mid e_1$ depends on the value of the outer integrand and has no closed form solution. However, noting that $E\left(Y_2^* \mid Y_1^* > 0\right) = \beta_2 x_2 + \rho\lambda\left(\beta_1 x_1\right)$ and $\operatorname{var}\left(Y_2^* > 0 \mid Y_1^* > 0\right) = 1-\rho^2\left\{\beta_1 x_1 \lambda\left(\beta_1 x_1\right) + \lambda\left(\beta_1 x_1\right)^2\right\}$ suggests that we may approximate the conditional distribution of $e_2 \mid e_1$ with

$$\Phi\left(\frac{\beta_2 x_2 + \rho\lambda\left(\beta_1 x_1\right)}{\sqrt{1-\rho^2\left\{\beta_1 x_1 \lambda\left(\beta_1 x_1\right) + \lambda\left(\beta_1 x_1\right)^2\right\}}}\right).$$

References

Anderson, J. A. (1984). Regression and ordered categorical variables (with discussion). *Journal of the Royal Statistical Society, Series B, 46*: 1-30.

Arends-Kuenning, M & Duryea, S. (2006). The Effect of Parental Presence, Parents' Education, and Household Headship on Adolescents' Schooling and Work in Latin America. *Journal of Family and Economic Issues, 27*, 263-286.

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

Breen, R. & Yaish, M. (2006). Testing the Breen-Goldthorpe Model of Educational Decision Making. In S. L. Morgan, D. B. Grusky, & G. S. Fields. (Eds.), *Frontiers in Social and Economic Mobility* (pp. 232-258). Stanford: Stanford University Press.

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.

Cameron, S. V. & Heckman, J. J. (2001). The Dynamics of Education Attainment for Black, Hispanic, and White Males. *Journal of Political Economy, 109*, 455-499.

Chevalier, A. & Lanot, G. (2002). The Relative Effect of Family Characteristics and Financial Situation on Educational Achievement. *Education Economics, 10*, 165-81.

Cobalti, A. (1990). Schooling inequalities in Italy: Trends over time. *European Sociological Review, 6*, 199-214.

Colding, B. (2006). A dynamic analysis of educational progression of children of immigrants. *Labour Economics, 13*: 479-492.

Garnier, M. A. & Raffalovich, L. E. (1984). The Evolution of Educational Opportunities in France. *Sociology of Education, 57*: 1-10.

Hansen, M. N. (1997). Social and Economic Inequality in Educational Career: Do the Effects of Social Background Characteristics Decline? *European Sociological Review, 13*: 305-320.

Hauser, R. M. & Andrew, M. (2006). Another Look at the Stratification of Educational Transitions: The Logistic Response Model with Partial Proportionality Constraints. *Sociological Methodology, 36*: 1-26.

Heath, A. F. & Clifford, P. (1990). Class inequalities in education in the twentieth century. *Journal of the Royal Statistical Society A, 153*: 1-16.

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

Jackson, M., Erikson, R. E., Goldthorpe, J. H. & Yaish, M. (2007). Primary and Secondary Effects in Class Differentials in Educational Attainment. *Acta Sociologica, 50*: 211-229.

Lauer, C. (2003). Family background, cohort and education: A French-German comparison based on a multivariate ordered probit model of educational attainment. *Labour Economics, 10*: 231-251.

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

Mare, R. D. (1979). Social Background Composition and Educational Growth. *Demography, 16*: 55-71.

Mare, R. D. (1980). Social 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.

Mare, R. D. (1993). Educational Stratification on Observed and Unobserved Components of Family Background. In Y. Shavit & H.-P. Blossfeld (Eds.), *Persistent Inequality. Changing Educational Attainment in Thirteen Countries* (pp. 351-376). Boulder, CO: Westview Press.

Mare, R. D. (2006). Statistical Models of Educational Stratification: Hauser and Andrew's Models of School Transitions. *Sociological Methodology, 36*: 27-37.

Nicoletti, C. & Perracchi, F. (2001). Two-step estimation of binary response models with sample selection. Working paper, Faculty of Economics, Tor Vergata University, Rome.

Plewis, I., Caldergood, L., Hawkes, D. & Nathan, G. (2004). *National Child Development Study and 1970 British Cohort Study Technical Report.* London: Centre for Longitudinal Studies, Institute of Education, University of London.

Raftery, A. E. & Hout, M. (1993). Maximally Maintained Inequality: Expansion, Reform, and Opportunity in Irish Education, 1921-75. *Sociology of Education*, *66*: 41-62.

Research in Social Stratification and Mobility (2010), vol. 28(1).

Shavit, Y. & Blossfeld, H-P. (Eds.) (1993). *Persistent Inequality. Changing Educational Attainment in Thirteen Countries.* Boulder, CO.: Westview Press.

Shavit, Y. & Westerbeek, K. (1998). Educational Stratification in Italy. Reforms, Expansions, and Equality of Opportunity. *European Sociological Review, 14*: 33-47.

Vaid, D. (2004). Gendered Inequality in Educational Transitions. *Economic and Political Weekly*: 3927-3938.

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

Table 1

Bias in Different Model Specifications Using Simulated Data.

	Model	Rescaled Coefficient	Standard Deviation
1	Mare probit model	0.840	0.150
2	Mare probit model with v_2	1.019	0.120
3	BPSM	1.023	0.168

The reported coefficient is the rescaled coefficient of x_1 in the model for $y_2 = 1$ in Equation (3).

Table 2
Descriptive Statistics for NCDS Sample. Means and Standard Deviations (SD)

Educational Attainment:	Mean	SD
Proportion of sample that completes A levels	0.384	-
Proportion of sample that completed A levels which also completed higher education	0.825	-
Explanatory Variables:		
Whole Sample:		
Father's education	10.094	2.089
Mother's education	10.024	1.603
Broken family	0.094	-
Gender (= female)	0.555	-
Cognitive ability	0.000	1.000
GCE performance	14.958	12.657
Sample Which Completes A levels:		
Father's education	10.996	2.788
Mother's education	10.698	2.134
Broken family	0.054	-
Gender (= female)	0.525	-
Cognitive ability	0.837	0.689
A level performance	2.727	4.155

Table 3
Results from Mare Probit Models and Bivariate Probit Selection Models (BPSM). Probit Regression Coefficients, Standard Errors in Parenthesis, and Average Partial Effects in Brackets.

	Mare Probit Model (1)	BPSM (1)	Mare Probit Model (2)	BPSM (2)	Mare Probit Model (3)	BPSM (3)
A LEVELS						
Father's education	0.173 (0.014)*** [0.057]	0.173 (0.014)*** [0.057]	0.104 (0.016)*** [0.025]	0.105 (0.016)*** [0.025]	0.096 (0.016)*** [0.022]	0.097 (0.016)*** [0.022]
Mother's education	0.221 (0.019)*** [0.072]	0.221 (0.019)*** [0.072]	0.144 (0.021)*** [0.034]	0.145 (0.021)*** [0.034]	0.137 (0.021)*** [0.032]	0.137 (0.021)*** [0.032]
Broken family	-0.550 (0.080)*** [-0.165]	-0.550 (0.080)*** [-0.165]	-0.312 (0.093)*** [-0.073]	-0.313 (0.093)*** [-0.073]	-0.286 (0.094)*** [-0.065]	-0.290 (0.093)*** [-0.065]
Gender (= female)	-0.134 (0.043) [0.044]	-0.134 (0.043)** [0.044]	-0.135 (0.050)** [0.032]	0.133 (0.050)* [0.032]	-0.165 (0.050)*** [-0.039]	-0.164 (0.050)*** [-0.039]
Cognitive Ability			1.016 (0.034)*** [0.243]	1.015 (0.034)*** [0.243]	0.868 (0.038)*** [0.202]	0.861 (0.038)*** [0.195]
GCE performance					0.020 (0.002)*** [0.005]	0.020 (0.002)*** [0.005]
HIGHER EDUCATION						
Father's education	0.089 (0.020)*** [0.025]	0.062 (0.033)* [0.014]	0.063 (0.022)*** [0.018]	0.073 (0.041)* [0.020]	0.047 (0.023)** [0.014]	0.074 (0.023)*** [0.017]
Mother's education	0.069 (0.026)*** [0.019]	0.036 (0.043) [0.008]	0.056 (0.027)** [0.016]	0.073 (0.062) [0.018]	0.056 (0.029)* [0.017]	0.098 (0.032)*** [0.023]
Broken family	-0.107 (0.165) [-0.031]	0.002 (0.200) [0.000]	-0.044 (0.174) [-0.012]	-0.095 (0.225) [-0.044]	-0.030 (0.183) [-0.009]	-0.141 (0.171) [-0.033]
Gender (= female)	-0.126 (0.078)* [-0.035]	-0.095 (0.080) [-0.021]	-0.075 (0.084) [-0.021]	-0.092 (0.095) [0.001]	-0.064 (0.089) [-0.019]	-0.110 (0.083) [-0.026]
Cognitive ability			0.815 (0.062)*** [0.230]	0.934 (0.388)** [0.250]	0.621 (0.066)*** [0.184]	0.916 (0.108)*** [0.217]
A level performance					0.271 (0.041)*** [0.080]	0.257 (0.041)*** [0.061]

ρ (and *p*-value for test of $\rho = 0$)		-0.278 (0.487)		0.271 (0.639)		0.550 (0.051)
Log-likelihood	-2961	-2961	-2551	-2251	-2171	-2169

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The number of observations in all models is 3955. The estimates of the BPSM for the second transition into higher education are adjusted by $\sqrt{1-\hat{\rho}^2}$ in order to be comparable with the corresponding estimates for the Mare probit model. The reported log-likelihood values for the Mare probit models summarize the log-likelihoods for the first and second transition.

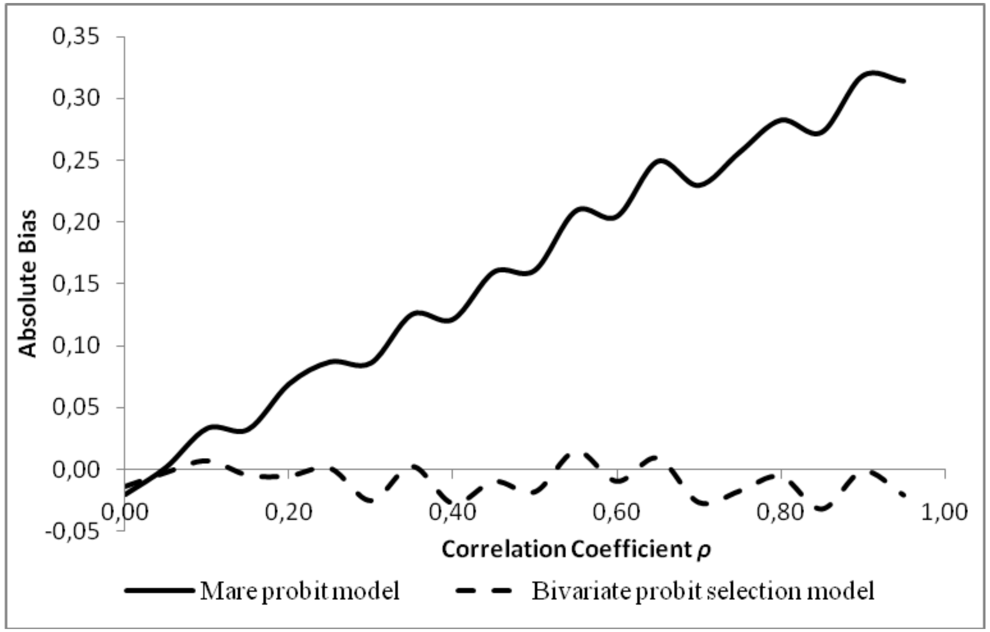

Fig 1. Bias in the Mare Probit and Bivariate Probit Selection Model

In order to obtain ρ higher than 0.5 we weigh down u_1, u_2 and weigh up e_1, e_2 in the simulations using weights that ensure constant error variance in all simulations.

CSER

CENTRE FOR STRATEGIC
RESEARCH IN EDUCATION