

Buscha, Franz; Conte, Anna

Working Paper

The impact of truancy on educational attainment: A bivariate ordered probit estimator with mixed effects

Jena Economic Research Papers, No. 2010,062

Provided in Cooperation with:

Max Planck Institute of Economics

Suggested Citation: Buscha, Franz; Conte, Anna (2010) : The impact of truancy on educational attainment: A bivariate ordered probit estimator with mixed effects, Jena Economic Research Papers, No. 2010,062, Friedrich Schiller University Jena and Max Planck Institute of Economics, Jena

This Version is available at:

<https://hdl.handle.net/10419/56812>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



JENA ECONOMIC RESEARCH PAPERS



2010 – 062

The Impact of Truancy on Educational Attainment: A Bivariate Ordered Probit Estimator with Mixed Effects

by

**Franz Buscha
Anna Conte**

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

© by the author.

**The Impact of Truancy on Educational Attainment: A Bivariate Ordered Probit
Estimator with Mixed Effects**

Franz Buscha

Centre of Employment Research
University of Westminster, London, UK

buschaf@wmin.ac.uk

Anna Conte

Strategic Interaction Group
Max-Planck-Institut für Ökonomik, Jena, Germany

aconte@econ.mpg.de

Abstract: This paper investigates the relationship between educational attainment and truancy. Using data from the Youth Cohort Study of England and Wales, we estimate the causal impact that truancy has on educational attainment at age 16. Problematic is that both truancy and attainment are measured as ordered responses requiring a bivariate ordered probit model to account for the potential endogeneity of truancy. Furthermore, we extend the ‘naïve’ bivariate ordered probit estimator to include mixed effects which allows us to estimate the distribution of the truancy effect on educational attainment. This estimator offers a more flexible parametric setting to recover the causal effect of truancy on education and results suggest that the impact of truancy on education is indeed more complex than implied by the naïve estimator.

JEL: I20; C35; C51

Keywords: educational attainment, truancy, bivariate ordered probit, mixed effects

The authors gratefully acknowledge valuable advice and comments from Paul Latreille (Swansea), seminar participants at the International Workshop in Applied Education Economics 2010 (Catanzaro) and seminar participants at the Work, Pensions and Employment Group annual conference 2010 (Bristol).

1. Introduction

Choices made in the early years of an individual's life have long-lasting effects into adulthood and one of the most important decisions undertaken by young people in the UK is at the age of 16. This is when compulsory education ends and students sit General Certificate of Secondary Education examinations (GCSE). Those who receive low educational attainment scores are more likely to experience lower wages, higher unemployment and, in general, are likely to experience a lower quality of life (Bradley and Nguyen, 2004).

One major factor in determining educational outcomes is truancy which has been identified as a strong predictor of low educational attainment (Bosworth, 1994) and of 'poor life outcomes' (Hibbert and Fogelman, 1990; Farrington, 1996). In the UK unauthorised school absence before the age of sixteen is illegal; however, official truancy statistics generally make for grim reading with the most recent truancy statistics reporting a record rate of school sessions being missed (1.03% - DCSF, 2009). Moreover, the long term trend does not look favourable with truancy rates now higher in 2009 than they were in 1997. This is even though the UK government has spent over £885 million on anti-truancy policies during the period 1997/98 to 2003/04 and set targets to cut truancy by a third during this timer period. A report by the Select Committee on Public Accounts (2006) estimated that the cost of absent pupils was £1.6 billion in missed education in the school year 2003/04. Clearly, studies which examine the determinants of truancy and the impact that truancy may have on educational outcomes can be considered important given such a social and political context.¹

¹ Whilst this paper focuses on the UK impact of school truancy, a likewise problem exists for the US. A good overview of the US literature is provided by the National Centre for School Engagement (Heilbrunn, 2007) whose report highlights the strong detrimental effect that truancy can have and how research surrounding truancy is still in its infancy.

Romer's (1993) study was one of the first to argue that truancy is a significant phenomenon and that evidence suggested a strong association between poor exam performance and high levels of truancy. Since then a number of other studies have examined the relationship between truancy and educational attainment and have come to similar conclusions; namely that truancy is a strong indicator of poor educational outcomes (Durden and Ellis, 1995; Marburger, 2001, Kirby and McElroy, 2003). However, the majority of this evidence relates truancy to educational outcomes at a university level and relatively few studies have examined the impact of truancy on educational performance at age 16 when education is still compulsory. Indeed, the compulsion by law to attend school until the age of 16 creates an interesting scenario where truancy can be seen as a rational phenomenon for youths whose discount rates are high and have no other way of escaping additional years of schooling.² Within this context, and given that age 16 signifies the start of the 'school-to-work' transition, it then becomes an important question to determine exactly to what extent truancy causally affects compulsory educational attainment.

For the UK, Bosworth (1994) explores the determinants and effects of truancy for pupils in their final year of compulsory schooling. Using data from the Youth Cohort Study (YCS 3) his findings suggest that boys are slightly more likely to truant than girls and that individuals from single parent families are more likely to truant. Individuals who reported high levels of truancy are less likely to obtain good educational scores at age 16. However, even individuals who report fairly low levels of truancy significantly tend to obtain lower grades; suggesting that it is not so much the hours of truancy that impacts, rather the signal that truanting represents. Payne (2001), also using data from the Youth Cohort Studies (YCS 9 and 10), finds that pupils who stayed-on and went into post-compulsory education (further education, ages 16-18) are much more likely

² See Harmon, *et al.* 2003 for a discussion of optimal schooling choices

to drop-out before they reach the end (at age 18) if they truanted during their final year of compulsory schooling (age 16).

However, because truancy is an individual choice, it is likely to be motivated by a series of unobserved characteristics and one can therefore consider truancy to be an endogenous variable in any educational attainment regression. In particular, it is easy to argue that truancy is driven by unobserved factors such as motivation and satisfaction which clearly also impact on educational outcomes and choices. Unfortunately, the above cited UK studies do little to address this identification problem and their conclusions can therefore only be seen as suggestive rather than causal.

Furthermore, one could argue that education determines truancy: those not doing well at school become disillusioned and want to leave the compulsory education system as soon as possible because they believe they can gain a higher pay-off in the labour market by working early. However, because they are bound by law to remain in school until the age of 16 they can only accomplish this by truanting. Seen in such a light the causal nature of truancy on education is reversed and any causal impact of truancy on education is nil.

Therefore, in order to estimate the causal impact that truancy has on education we must take the potential endogeneity of truancy into account.³ However, many of the aforementioned studies do little to control for the possible endogeneity of truancy in their attainment regressions. Only in recent years have studies by Stanca (2006), Chen and Lin (2008) and Dobkin *et al.* (2009) attempted to use more innovative methods (such as random design, panel data and regression discontinuity) to recover a *causal* effect on education by controlling for unobservable characteristics which could influence both truancy and attainment. However, none of these

³ Be this omitted variable bias (ability) or the possible simultaneous nature of truancy and education (education causing truancy).

studies examine the impact of truancy on attainment before individuals are eligible to leave school and in this paper we argue that it is important to identify the causal nature of truancy on education before self-selection into further and higher education takes place.

Moreover, in this paper we go even further because it is our belief that, even after having dealt with truancy as an endogenous variable, the relationship linking truancy behaviour to educational attainment is more complex than can be approximated by conditional mean estimators. In other words, we think that even in a *ceteris paribus* condition there might still be substantial individual differences in the intensity of truancy effect on education. For this reason we estimate the causal impact of truancy on the educational outcomes of 16 year olds in the UK by using a modified bivariate ordered probit model which allows us to estimate the distribution of the effect of truancy on attainment. By appropriately modelling the parameter of interest according to a suitable distribution, we aim to take the heterogeneity of the truancy effect into account in any estimates of educational attainment.

Results support our intuition and suggest that the impact of truancy on educational attainment is more complex than assumed by conditional mean estimators and that the effect of truancy on educational attainment is more heterogeneous for boys than for girls. The next section describes the data used in this study. Section 3 outlines our methodology and derives the bivariate ordered probit estimator with mixed effects. Section 4 presents the results whilst section 5 concludes.

2. Data

One of the main data sources for the UK education literature which examines the schooling experiences of 16 to 18 year olds (and contains records on truancy and educational attainment) is the Youth Cohort Study of England and Wales (YCS). The YCS is specifically designed to

monitor the behaviour and decisions made by a representative sample of the UK school population as they transit from compulsory education (age 16) to further education and higher education, or to the labour market. The YCS data is a longitudinal dataset designed to follow individuals over 3 years (3 consecutive sweeps starting at age 16) and now has 13 cohorts with over 35 sweeps. In this study we make use of only the first sweep as truancy information is not available in later interviews. Moreover, we make use of the restricted version of the dataset which enables us to map local economic conditions to individuals (we have this information for YCS 10 which corresponds to student eligible to leave school in 1999). Although the YCS data uses a multi-stage stratified random sampling procedure, differences in selection and response rates may still be an issue. We therefore make use of the included sample weights which correct for differential selection probabilities, correct the ethnicity boost, and take into account non-response bias.

However, the YCS records truancy and educational attainment as ordered variables. Moreover, the measuring of ordered educational outcomes is relatively common in the UK as the UK does not use grade point averages like the United States, but instead relies on categorical targets – for example in 2008 the Department for Children Schools and Families (DCSF) set a national target of 60% of pupils achieving 5 or more A*-C grades at GCSE. This is reflected in the data sources whereby information for YCS 10 returns the following descriptive statistics:

Table 1: Truancy and Educational Attainment in the Youth Cohort Study, Sweep 1.

<i>Truancy</i>	<i>Frequency</i>	<i>Percent</i>	<i>Weighted Percent</i>	<i>Highest Educational Attainment at 16</i>	<i>Frequency</i>	<i>Percent</i>	<i>Weighted Percent</i>
Never	8,230	68.46%	65.04%	None	518	4.21%	4.18%
For the odd day or lesson	2,881	23.96%	25.08%	1-4 GCSE D-G grade	193	1.57%	3.38%
Particular days or lessons	578	4.81%	6.16%	5+ GCSE D-G grade	1,023	8.32%	18.16%
For several days at a time	177	1.47%	1.95%	1-4 GCSE A*-C grade	2,672	21.73%	25.50%
For weeks at a time	156	1.30%	1.78%	5+ GCSE A*-C grade	7,889	64.16%	48.78%
Total	12,022	100.00%	100.00%	Total	12,295	100.00%	100.00%

Source: YCS 10 (surveys those eligible to leave school in 1998-1999), weighted

Given that the data can be interpreted as a representative sample of the UK school population (in their final year of compulsory schooling) suggests that the majority of youths did not truant. However, a substantial portion (25%) engaged in minor forms of truancy whilst a small proportion (4%) engaged in more serious forms of truancy. Approximately half of all pupils achieved the highest grade category of 5 or more A*-C grade GCSE's, whilst a further 25% achieved the second highest category of 1 to 4 A*-C grades. A small proportion (6%) achieved none or poor grades.

3. Methodology: A Bivariate Ordered Probit model with Mixed Effects

Generally, when a suspected endogenous explanatory variable is encountered in the applied micro-economics literature, instrumental variable methods are often applied to estimate a causal and consistent effect.⁴ As such, instrumental variable estimation has long been a mainstay of the econometric literature and is arguably one of the most commonly used empirical methodologies. However, when both the dependent variable and the suspected endogenous variable take the form of categorical data, standard IV techniques (such as two-stage least squares) often break down and more complicated analytical techniques are required.⁵

When both the dependent variable and the endogenous variable take a binary form, a bivariate probit model can be used (Greene, 2008: 827).⁶ When both the dependent variable and

⁴ We will not discuss the LATE implications of the IV estimator here.

⁵ For a more nuanced argument see Angrist's (2001) discussion of limited dependent variable models with dummy endogenous regressors.

⁶ These can be extended to the multivariate case when more than one binary endogenous regressor is on the 'right-hand-side' of the equation. See Cappellari and Jenkins (2003, 2006). Implemented as a Stata routine `mvprobit` by Cappellari and Jenkins (2003).

the endogenous variable take the form of ordered categorical data, then a bivariate ordered probit model can be applied (Greene and Hensher, 2009: 223).⁷ Finally, when the dependent variable is ordered with more than two choices and the endogenous variable is binary, then a semi-ordered bivariate probit model is needed to correctly estimate the system (Greene and Hensher, 2009: 225).⁸ In recent years a variety of papers have been written which make use of bivariate ordered probit and semi-ordered probit estimators and a convenient overview of this literature is provided by Greene and Hensher (2009: 226) who highlight approximately 25 different papers making use of this methodology in a variety of circumstances from 1991 to 2007.

Our variables of interest, as shown in Table 1, suggest that a bivariate ‘ordered-ordered’ methodology would be an appropriate solution. However, we modify the existing bivariate ordered probit estimator (hereafter called the naïve bivariate ordered probit estimator) to include a mixed effect. This methodology, emerging in the behavioural economics literature,⁹ aims at estimating the distribution over the population of the relevant parameters of the model under investigation instead of just reporting a point estimate of these parameters. In other words, we assume that, *ceteris paribus*, the effect of the endogenous variable on the dependent variable can differ individual by individual and that such differences are captured by an appropriate choice of the distribution function for this effect. The reason for such a hypothesis is that truancy and education are complex phenomenon that go beyond observable explanatory factors. Our approach aims to capture nuances in the truancy effect on educational attainment; nuances

⁷ Implemented as a Stata routine bioprobit by Sajaia (2008)

⁸ It should be noted that the semi-ordered bivariate probit estimator is a special case of the bivariate ordered probit estimator and does not require special modifications to the likelihood function.

⁹ See, for example, Botti *et al* (2008) and Conte *et al* (2009).

which cannot be captured by simply framing a particular individual in a specific cohort of people identified by a certain combination of observables.¹⁰

Assume that the two latent variables attitude to truancy and educational outcome, respectively T_i^* and E_i^* , are determined by the following system of equations:

$$\begin{cases} T_i^* = \mathbf{x}'_{Ti}\beta_T + \varepsilon_{Ti} \\ E_i^* = \mathbf{x}'_{Ei}\beta_E + \gamma_i T_i^* + \varepsilon_{Ei} \end{cases} \quad \text{with } \gamma_i \sim N(\mu_\gamma, \sigma_\gamma^2) \text{ and } \begin{pmatrix} \varepsilon_{Ti} \\ \varepsilon_{Ei} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right). \quad (1)$$

Here, \mathbf{x}_{Ti} and \mathbf{x}_{Ei} are vectors of observables, β_T and β_E are a vector of parameters, γ_i represents the effect that T_i^* has on E_i^* for individual i , and ε_{Ti} and ε_{Ei} are two error terms, assumed to be jointly normal with correlation coefficient ρ and uncorrelated with everything else in the model; in particular, $E(\mathbf{x}_{Ti}\varepsilon_{Ti}) = 0$ and $E(\mathbf{x}_{Ei}\varepsilon_{Ei}) = 0$.¹¹

The basic idea underlying our model is that, when trying to explain people's choices, at stake are both observed factors, represented by \mathbf{x}_{Ti} and \mathbf{x}_{Ei} (for example demographic variables) and unobserved factors (for example motivation) embedded in γ_i and in the joint distribution of ε_{Ti} and ε_{Ei} . In this model, we assume that there is heterogeneity between individuals in the way the latent variable T_i^* influences E_i^* . To capture such an effect we adopt a continuous mixture

¹⁰ The mixed effect approach has further advantages; even if a large vector of explanatory variables were available, a 'kitchen sink' approach can sometimes cause more harm than good in terms of collinear variables or introducing further endogenous regressors (which in case of simultaneous regressors can bias all other parameter estimates). Furthermore, whilst one can use interactions terms within the regression specification to model more complex truancy effects, the interpretation of these is notoriously difficult in a non-linear setting.

¹¹ The derivation of the reduced form of the system in eq. (1) follows in the Appendix A.

approach by estimating the distribution of this effect over the population. In other words, we assume that each individual draws their own γ_i from a normal distribution with mean μ_γ and standard deviation σ_γ , and what we do here is to estimate the parameters of the underlying distribution of γ_i .

However, we do not observe the realisation of the two latent variables T_i^* and E_i^* . What we observe, instead, are the two categorical variables T_i and E_i . The unobserved latent variables T_i^* and E_i^* are related to their respective observed outcomes, according to the classification in Table 1, by the following observational rules:

$$T_i = \begin{cases} 1 = \text{Never} & \text{if } T_i^* \leq j_{11} \\ 2 = \text{Odd day} & \text{if } j_{11} < T_i^* \leq j_{12} \\ 3 = \text{Particular day} & \text{if } j_{12} < T_i^* \leq j_{13} \\ 4 = \text{several days} & \text{if } j_{13} < T_i^* \leq j_{14} \\ 5 = \text{Weeks at a time} & \text{if } j_{14} \leq T_i^* \end{cases} \quad E_i = \begin{cases} 1 = \text{None} & \text{if } E_i^* \leq j_{21} \\ 2 = 1 - 4 \text{ GCSE } D - G & \text{if } j_{21} < E_i^* \leq j_{22} \\ 3 = 5 + \text{GCSE } D - G & \text{if } j_{22} < E_i^* \leq j_{23}, \\ 4 = 1 - 4 \text{ GCSE } A^* - C & \text{if } j_{23} < E_i^* \leq j_{24} \\ 5 = 5 + \text{GCSE } A^* - C & \text{if } j_{24} \leq E_i^* \end{cases} \quad (2)$$

where the $j_{..}$ are cut-points to be estimated along with the other parameters of the model.¹²

The probability of observing $T_i = l$ and $E_i = m$ for individual i is:

$$\begin{aligned} Pr(T_i = l, E_i = m) &= \Phi(j_{1l} - x'_{Ti}\beta_T, (j_{2m} - \gamma_i x'_{Ti}\beta_T - x'_{Ei}\beta_E)\lambda_i, \tilde{\rho}_i) \\ &- \Phi(j_{1l-1} - x'_{Ti}\beta_T, (j_{2m} - \gamma_i x'_{Ti}\beta_T - x'_{Ei}\beta_E)\lambda_i, \tilde{\rho}_i) \\ &- \Phi(j_{1l} - x'_{Ti}\beta_T, (j_{2m-1} - \gamma_i x'_{Ti}\beta_T - x'_{Ei}\beta_E)\lambda_i, \tilde{\rho}_i) \\ &+ \Phi(j_{1l-1} - x'_{Ti}\beta_T, (j_{2m-1} - \gamma_i x'_{Ti}\beta_T - x'_{Ei}\beta_E)\lambda_i, \tilde{\rho}_i) \end{aligned} \quad (3)$$

¹² The cut-points meet the following conditions: $j_{10} < \dots < j_{1l} < \dots < j_{1L}$, with $j_{10} = -\infty$ and $j_{1L} = \infty$; $j_{20} < \dots < j_{2m} < \dots < j_{2M}$, with $j_{20} = -\infty$ and $j_{2M} = \infty$.

where $\Phi(\dots)$ is the bivariate standard normal cumulative distribution function and λ_i and $\tilde{\rho}_i$ are respectively defined as follows: $\lambda_i = 1/\sqrt{\gamma_i^2 + 2\gamma_i\rho + 1}$ and $\tilde{\rho}_i = \lambda_i(\gamma_i + \rho)$.

The log-likelihood contribution of individual i is:

$$L_i = \int_{-\infty}^{\infty} \prod_{l=1}^L \prod_{m=1}^M Pr(T_i = l, E_i = m)^{I(T_i=l, E_i=m)} f(\gamma_i; \mu_\gamma, \sigma_\gamma^2) d\gamma_i, \quad (4)$$

where $f(\gamma_i; \mu_\gamma, \sigma_\gamma^2)$ is the normal density function for the random variable γ_i , and $I(T_i = l, E_i = m)$ is an indicator function that equals one when $T_i = l$ and $E_i = m$.

The sample log-likelihood function $\ln L = \sum_{i=1}^N \ln L_i$ is maximised using 20-point Gauss-Hermite quadrature. The program is written in STATA version 11.0, and is available from the authors upon request. To examine the small sample properties of our estimator we implement Monte Carlo simulations (Appendix B). We find that our estimator performs well even in small samples (200 observations).

The assumption that $\gamma_i \sim N(\mu_\gamma, \sigma_\gamma)$ is made *a priori*, as we assume that the impact that truancy has on an individual's attainment may be positively or negatively distributed around an unknown mean. Moreover, there is little reason to assume that this distribution is asymmetric around the mean. Finally we assume that there is 'clumping' near the mean and few individuals experience extreme positive or negative effects. The standard normal distribution suits these assumptions well. If the true effect of truancy is similar for every individual i , then σ_γ will be estimated as 0 and the estimator collapses into a naïve bivariate ordered probit estimator. The distributional assumption of γ_i must be made on a case by case basis as other relationships may require different distributions. For example, one could imagine that the impact health-checkups

on length of survival will always have positive effects which may be distributed with an exponential decay and thereby warrant a power function.¹³

In addition to using a vector of standard exogenous controls such as socio-economic background, school type, ethnicity, parental education, gender, household composition, housing and disability status¹⁴ (Bosworth, 1994; Bradley and Taylor, 2004), we have to impose an exclusion restriction on the regressors in \mathbf{x}_E to identify the recursive system of equations. In other words, it is required that at least one of the explanatory variables in the truancy equation do not enter the education equation, i.e. \mathbf{x}_E is a subset of \mathbf{x}_T .¹⁵ With this in mind, we have to look for at least one variable that in principle contributes to explain T_i^* but that can be reasonably excluded by the group of variables that are meant to explain educational attainments, E_i^* .

Generally, the determinants of teenage truancy have not received much academic attention in the UK although some previous literature by Bosworth (1994), Dustman *et al.* (1997) and Burgess *et al.* (2002) does exist. From these a variety of factors have been found to influence the truancy decision and a convenient (non-academic) overview of many of such factors is provided by the Illinois State Board of Education 'Truants' Alternative & Optional Education Program:

¹³ For comparative purposes we also estimate a specification with a log-normally distributed effect:

$$\gamma_i \sim \ln N(\mu_\gamma, \sigma_\gamma)$$

¹⁴ An table with descriptive statistics for these controls is provided in Appendix B

¹⁵ For example, suppose that $\mathbf{x}_{Ti} = \mathbf{x}_{Ei}$; then the system in (1) becomes:

$$\begin{cases} T_i^* = \mathbf{x}'_{Ti}\beta_T + \varepsilon_{Ti} \\ E_i^* = \mathbf{x}'_{Ti}(\beta_E + \gamma_i\beta_T) + \gamma_i\varepsilon_{Ti} + \varepsilon_{Ei} \end{cases}$$

Consequently, only β_T is identified, and β_E , γ_i and ρ cannot be estimated consistently.

Table 2: Possible causes of Teenage Truancy

Child	Family
1. Poor self-concept; low self-esteem	1. Parents lack appreciation for value of education
2. Low academics; particularly behind in reading and math	2. Parents are high-school dropouts
3. Socially isolated; poor inter-personal skills	3. Financial difficulties (lack of adequate nutrition, transportation; inability to pay fees)
4. Feeling of not belonging at school, of being different	4. Ineffective parenting; lack of parental control and supervision
5. Feeling of lack of control over life	5. Familial instability (one-parent household; poor parent-child relationship)
6. Little or no extra-curricular involvement	6. Inappropriate role models
7. Mental and/or emotional instability; childhood depression	7. Child abuse and/or neglect
8. Unidentified learning disabilities	8. Substance abuse; alcoholic parent
9. Vision and/or auditory problems that have gone undiagnosed	9. Parental convenience (older sibling kept home to babysit)
10. Language barriers	
11. Poor health	
12. Negative peer relationships; older, non-school oriented friends	
13. Substance abuse	
14. Fear of school, teachers, and/or administrators	
15. Experienced recent traumatic event (divorce, death of a loved one)	
16. Fear of physical protection going to or at school	
17. Teenage pregnancy and/or parenting	
School	Community
1. Uninteresting and irrelevant curriculum	1. Lack of support for school
2. Improper class placement (above or below abilities)	2. Lack of, or unresponsive community service agencies
3. Failure to identify and provide services for problem students	3. Community upheaval and social change
4. Poor pupil-teacher relationships	4. Loss of neighbourhood schools and sense of 'ownership' of schools
5. Insufficient counselling and guidance staff	5. Negative peer influences
6. High student/teacher ratio	6. High incidence of substance abuse, criminal activity
7. Low teacher expectations	7. Gang activity
8. Lack of parent - school communication and involvement	8. Inadequate provision for transportation
9. Too weak, or too rigid administration of policies.	

Source: Illinois State Board of Education's Truants' Alternative and Optional Education Program

Table 2 highlights the conceptual problem we are faced with; the overall majority of factors that influence truancy also influence the education decision. Virtually all child, family and school specific factors are related not only to truancy but also to education making them invalid choices for a possible instrument. We therefore decide to go to a community/regional level and make use of the fact that we have local education authority (LEA)¹⁶ codes available in our data (restricted version).

¹⁶ The term LEA has been defunct since the Education and Inspections Act of 2006 and been replaced by Local Authority (LA). However, the term is still used informally and we use the term in that capacity.

In particular we argue that peer group effects are a likely candidate for an instrument.¹⁷ Specifically, the UK government has since 1994 published National Pupil Absence tables every year which measure the level of authorised and unauthorised pupil absence in each LEA. We argue that these *local* truancy rates are likely to be related to *individual* truancy through peer group effects. Pupils who live in high truancy areas are more exposed to truancy behaviour in their neighbourhood and are thus more likely to truant themselves.

Important is that we argue that the peer group effect we are using occurs at a relatively aggregated regional level and not at a school level – where peer groups effects have strong influences on individual educational outcomes. Moreover, Angrist and Pischke (2008: 146) show that using average school level information (with the individual included in that average) cannot possibly be informative about causal effects for an individual within that school. Excluding the individual from the average is better but still unsatisfactory as the presence of group-level random shocks induces spurious peer group effects. Mankiw (1993) refers to this as “correlated” effects where individuals and peer groups have similar unobserved characteristics and any peer group effect identified could therefore be spurious.

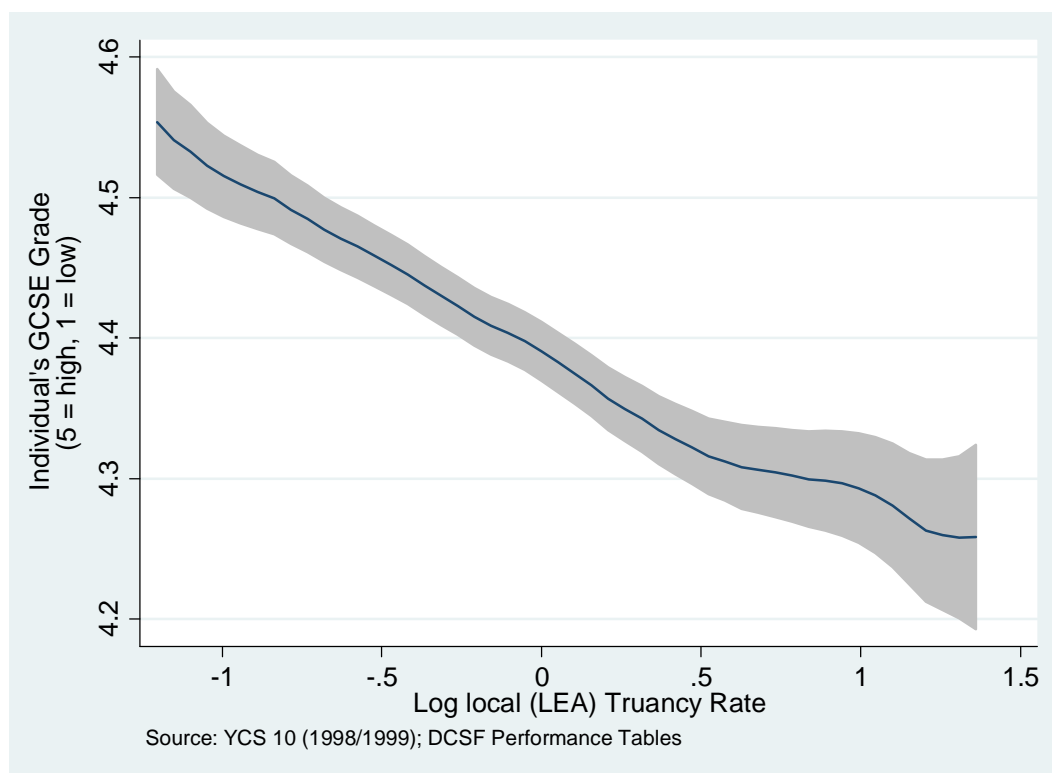
¹⁷ Earlier attempts to find a suitable instrument included local part-time pay rates which have been shown to influence truancy (Dustmann, *et al.*, 1997). However, this proved to be a *weak* instrument giving unreliable results. Various regional crime rates from the Home Office were also investigated but were found to be non-significantly correlated with individual truancy. Regional unemployment indicators showed good results but were also strongly associated with educational performance.

We argue that local truancy rates, as measured at an LEA level, are unlikely to suffer from such problems because the regional aggregation of this variable is too large.¹⁸ At such an aggregated level individuals cannot influence the average local truancy rate and there are unlikely to be spurious peer group effects due to group-level unobservables. Moreover, we cannot think of a rational theoretical transmission mechanism which would link local truancy rates to individual educational performance. For example, local unemployment rates at an LEA level have been shown to influence educational decisions (Andrews and Bradley, 1997; Rice 1999). However, these effects are transmitted to education through the cost-benefit analysis of additional years of schooling (high unemployment leads to a low probability of obtaining a job reducing the opportunity cost of an additional year of schooling). No such transmission exists for local truancy rates as local truancy rates should not affect the probability of post-school economic outcomes for individuals.

Unfortunately, an initial examination of the relationship between local truancy rates and individual educational performance suggests a strong negative relationship (Figure 1) – contrary to our argued position that local truancy rates should not affect individual educational performance.

¹⁸ LEA populations are generally measured in 100,000s and ignoring the City of London and the Isles of Scilly (which were removed from our data-set) population estimates for 1999 suggest an average of 19,872 youths aged 15 to 19 per LEA. Values range from 2,700 (Rutland) to 80,300 (Kent). *Source: Nomis*

Figure 1: Non-parametric correlation between individual educational performance and local truancy rates



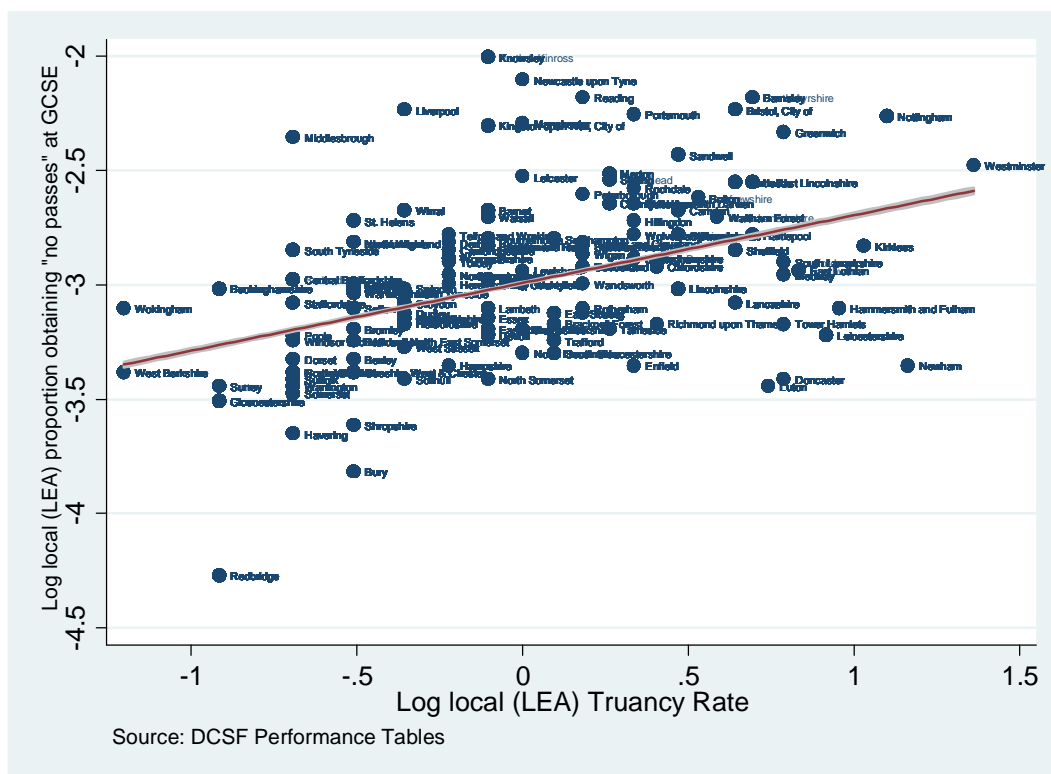
However, further investigation reveals that such initial descriptives can be misleading as there is a strong relationship between the *quality* of an LEA and local truancy rates. It must be remembered that in the UK LEA's have substantial power to allocate funding to schools, hire teachers and in general determine educative policies. LEA's also have the power to set truancy strategies as these are implemented at a local level in conjunction with local police forces.¹⁹ We

¹⁹ LEA's are responsible for implementing truancy policy and since 2002 the UK government has prescribed two 'truancy sweeps' per year. These are three week events where local Education Welfare Services, in conjunction with local police forces, conduct truancy sweeps to try and 'catch' truanting pupils on the streets. However, the first truancy sweep occurred in 2000 although in reality local anti-truancy measures did not begin in earnest until after 2002. Either way, given that we are examining data from 1998/99 the implication is that our observations are unlikely to be affected by differential LEA truancy strategies. This also implies that we cannot use local policy interventions as instruments.

must therefore be sure that the above described relationship is not simply a reflection of the underlying relationship between *local educational quality* and individual exam performance. Moreover, if this is the case, does any relationship between local truancy rates and individual truancy remain or would the correlation between local and individual truancy disappear once local educational factors are controlled for?

To test for this we make use of data from the School Performance Tables (1999) and impute the LEA averages of 5+A*-C passes, 5+A*-G passes and no passes (measures which proxy the educational quality of an LEA). Figure 2 clearly shows that there is a positive relationship between high LEA truancy rates and poor LEA educational performances, suggesting that the correlation between individual exam performance and local truancy rates could simply be a reflection of individual exam performance and LEA quality.

Figure 2: Relationship between LEA educational performance and LEA truancy rates



Examining this in a multivariate context highlights our point; by introducing local educational averages as proxies for LEA performance into both the truancy and educational regressions we find that the impact that local truancy rates have on the individual propensity to truant remains positive and statistically significant. This suggests that there is a *local* truancy peer group effect, above and beyond any influence that LEA quality may have. Moreover, local educational indicators do not influence individual truanting probabilities. This is very important as our argument rests on the case that random variations in local truancy causally influence individual truancy.

Furthermore, by introducing LEA performance indicators into the educational attainment regression we find that the estimated impact of local truancy rates on individual educational performance decreases. This implies that a large part of the negative relationship between local truancy and individual educational performance as depicted in Figure 1 was actually due to LEA performance. Depending on which indicator we use, this decrease changes the coefficient of local truancy from statistically significant to statistically insignificant. The best results are obtained using the local average 'no pass' variable and to some extent this stands to reason; the 'no pass' rate is likely the best indicator of the local *educational quality* provided to the 'lesser able' who in turn are most likely to be susceptible to truancy behaviour.

Table 3: Independent Ordered Probits of the determinants of Truancy and Educational Attainment

Variables		I			II			III			IV		
Dependent	Independent	Coef	S.E	P-val	Coef	S.E	P-val	Coef	S.E	P-val	Coef	S.E	P-val
Truancy													
	Vector of X's	included			included			included			included		
	Log local ed. rate - 5+ A to C	not included			0.059	0.090	0.514	not included			not included		
	Log local ed. rate - 5+ A to G	not included			not included			0.248	0.409	0.544	not included		
	Log local ed. rate - No pass	not included			not included			not included			-0.009	0.046	0.839
	Log local truancy rate	0.093	0.026	0.000	0.102	0.029	0.000	0.101	0.029	0.000	0.096	0.029	0.000
Education													
	Vector of X's	included			included			included			included		
	Log local ed. rate - 5+ A to C	not included			0.270	0.080	0.001	not included			not included		
	Log local ed. rate - 5+ A to G	not included			not included			1.259	0.364	0.001	not included		
	Log local ed. rate - No pass	not included			not included			not included			-0.216	0.042	0.000
	Log local truancy rate	-0.103	0.025	0.000	-0.062	0.028	0.026	-0.061	0.028	0.028	-0.038	0.028	0.174

Source: YCS 10 (surveys those eligible to leave school in 1998-1999), weighted, vector of X's are the same as for the main models.

We are aware that the results in Table 2 are conditional on observable characteristic and we must ask to what extent local truancy rates are random and whether other factors influence local truancy rates, which in turn also affect individual educational performance. Clearly the most important factor would be LEA quality, which we control for (although the nature of ‘quality’ is certainly debatable). In this context we cannot think of other factors that might drive the local truancy rate and also determine individual educational performance. We therefore argue that, conditional on LEA performance indicators, the conditions for an appropriate instrument are fulfilled.

4. Results

For comparison purposes we present results from a naïve independent ordered probit model (IOP), the naïve bivariate ordered probit (BIOP) and bivariate ordered probit with mixed effects

(BIOPmixed). In addition to estimating a mixed effect with a Gaussian distribution we also estimate a mixed effect with a log-normal distribution. Results are presented in Table 4.

Before examining the coefficients and implications of our findings, a quick comparison of the naïve bivariate ordered probit model and the ‘mixed’ bivariate ordered probit models suggests that the mixed effects model significantly increases the estimated log-likelihood, thereby resulting in model with better explanatory power. The log pseudo-likelihood of the mixed (Gaussian) model is -20034.573 whilst the log pseudo-likelihood of the naïve model is -20216.206. Unfortunately, the bivariate models cannot be run in a ‘intercept only’ mode and we are thus unable to gauge the impact of the mixed effect on Pseudo R^2 indicators. However, a increase in the log likelihood function of nearly 200 points seems substantial given that we are not adding extra variables into the modelling structure. This suggests that that the mixed model substantially improves the explanatory power of estimation at little cost (other than computational).

Examining the coefficients on the determinants of truancy reveals that girls are more likely to be truant, non-whites are less likely to be truant, pupils from selective/independent schools are less likely to be truant, those living in rented housing are more likely to be truant, disabled pupils are more likely to be truant, those from lower socio-economic backgrounds are more likely to truant and those who have both parents in the household are less likely to commit truancy. The log local truancy rate has a positive impact on the probability to truant. Examining the determinants of educational attainment we find that; girls, non-whites, those with educated parents, those from better schools, those in owned houses, individuals from high socio-economic backgrounds and with both parents do better in school.

Examining the impact of truancy on educational attainment reveals that independent ordered probit estimation estimates a negative effect of -0.325^{20} which is statistically significant. The naïve bivariate ordered probit estimator suggests a larger effect of truancy on educational attainment of -0.768 . Although this estimate is associated with a higher standard error it remains statistically significant at a 0.1% level.²¹ The estimated value of ρ is 0.459 and statistically significant at a 5% level suggesting that the use of a bivariate model is justified and that using an independent ordered probit model could lead to biased results.

²⁰ Where the IOP method uses y_T as a continuous regressor.

²¹ The large standard errors in the bivariate models are likely due to the fact that our instrument is based on LEA regions of which we have 132. Whilst we argue this is enough to identify the model it only offers limited variation over the 11,942 individuals in the sample. A better instrument (perhaps based on individual characteristics) with more data points is likely to offer a much better estimated coefficient. However, we have earlier highlighted the strong problems associated with finding a relevant instrument for truancy.

Table 4: The Impact of Truancy on Educational Attainment at Age 16 (GCSEs) – Bivariate Ordered Probit with Mixed Effects

Variables		IOP			BIOP			BIOPmixed (Gaussian)			BIOPmixed (Log-normal)		
Dependent	Independent	Coef	S.E	P-val	Coef	S.E	P-val	Coef	S.E	P-val	Coef	S.E	P-val
Truancy	Female	0.064	0.026	0.015	0.087	0.023	0.000	0.062	0.023	0.007	0.093	0.023	0.000
	Non-white	-0.096	0.043	0.024	-0.091	0.037	0.014	-0.115	0.036	0.001	-0.082	0.037	0.026
	Neither parents employed	-0.115	0.045	0.011	-0.097	0.038	0.011	-0.113	0.037	0.002	-0.102	0.037	0.006
	At least one parent with A-level (ref: degree level)	0.045	0.042	0.291	0.033	0.038	0.381	-0.025	0.038	0.516	0.036	0.039	0.352
	Neither parent with A-level (ref: degree level)	-0.025	0.036	0.492	-0.026	0.031	0.413	-0.090	0.031	0.004	-0.024	0.032	0.444
	Selective school	-0.306	0.058	0.000	-0.270	0.054	0.000	-0.305	0.055	0.000	-0.260	0.056	0.000
	Independent school	-0.413	0.049	0.000	-0.348	0.046	0.000	-0.427	0.047	0.000	-0.339	0.049	0.000
	House: rented (ref: owned)	0.300	0.038	0.000	0.316	0.032	0.000	0.300	0.032	0.000	0.317	0.032	0.000
	House: other (ref: owned)	0.139	0.092	0.128	0.113	0.082	0.170	0.105	0.083	0.204	0.115	0.083	0.168
	Disabled	0.102	0.063	0.107	0.167	0.054	0.002	0.118	0.052	0.023	0.150	0.052	0.004
	SEC: lower professional/higher technical	0.032	0.039	0.412	0.012	0.035	0.735	-0.073	0.034	0.032	0.007	0.036	0.845
	SEC: intermediate occupations	0.106	0.042	0.013	0.110	0.038	0.004	0.035	0.038	0.367	0.103	0.039	0.009
	SEC: lower supervisory occupations	0.200	0.053	0.000	0.186	0.047	0.000	0.115	0.047	0.013	0.183	0.047	0.000
	SEC: semi-routine routine/routine	0.170	0.052	0.001	0.173	0.046	0.000	0.102	0.045	0.023	0.179	0.046	0.000
	SEC: other	0.233	0.064	0.000	0.202	0.055	0.000	0.118	0.053	0.027	0.193	0.054	0.000
	Both parents in household	-0.315	0.033	0.000	-0.315	0.029	0.000	-0.394	0.027	0.000	-0.314	0.029	0.000
	Log local truancy rate (unauthorised absence)	0.093	0.026	0.000	0.091	0.023	0.000	0.093	0.023	0.000	0.086	0.023	0.000
		j11	0.271	0.048	0.000	0.346	0.044	0.000	0.141	0.036	0.000	0.354	0.045
	j12	1.212	0.049	0.000	1.343	0.045	0.000	1.146	0.036	0.000	1.364	0.047	0.000
	j13	1.737	0.053	0.000	1.869	0.049	0.000	1.692	0.042	0.000	1.923	0.052	0.000
	j14	2.077	0.058	0.000	2.213	0.053	0.000	2.098	0.051	0.000	2.335	0.061	0.000
Education	Female	0.334	0.025	0.000	0.332	0.031	0.000	0.341	0.050	0.000	0.332	0.033	0.000
	Non-white	0.095	0.044	0.030	0.052	0.048	0.275	0.056	0.057	0.327	0.054	0.046	0.244
	Neither parents employed	-0.064	0.039	0.097	-0.107	0.042	0.011	-0.131	0.045	0.003	-0.114	0.042	0.007

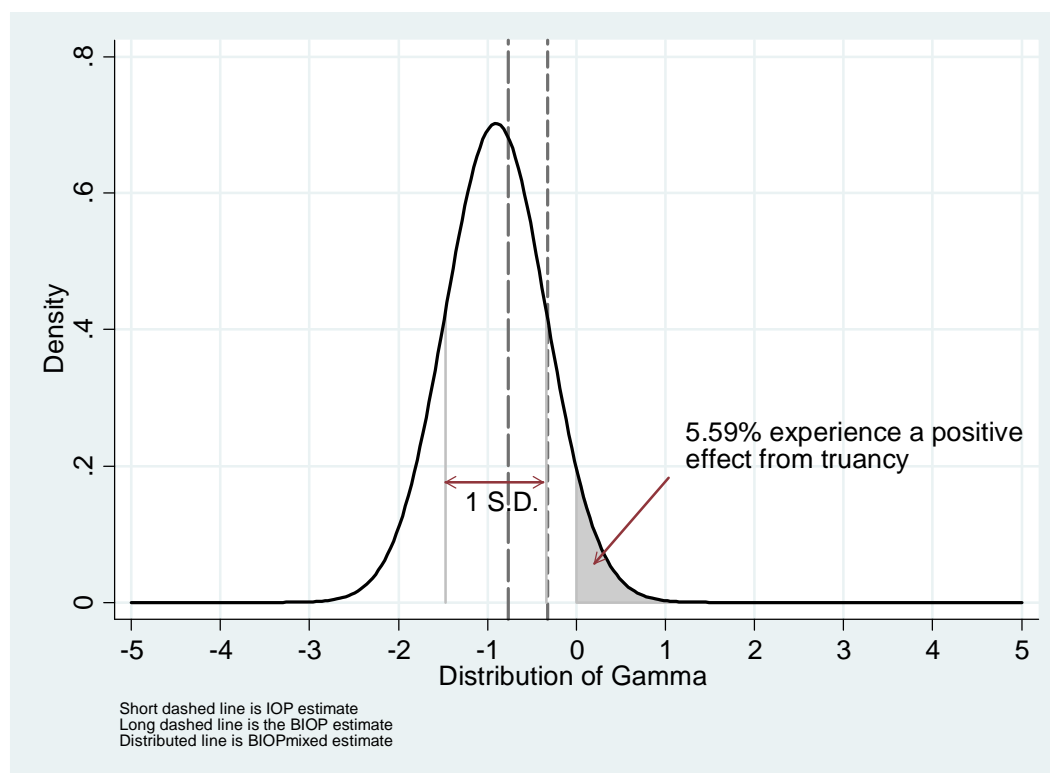
Variables		IOP			BIOP			BIOPmixed (Gaussian)			BIOPmixed (Log-normal)		
Dependent	Independent	Coef	S.E	P-val	Coef	S.E	P-val	Coef	S.E	P-val	Coef	S.E	P-val
	At least one parent with A-level (ref: degree level)	-0.237	0.048	0.000	-0.191	0.056	0.001	-0.242	0.060	0.000	-0.180	0.059	0.002
	Neither parent with A-level (ref: degree level)	-0.370	0.039	0.000	-0.346	0.053	0.000	-0.406	0.061	0.000	-0.335	0.058	0.000
	Selective school	1.198	0.126	0.000	1.019	0.238	0.000	1.150	0.326	0.000	0.986	0.251	0.000
	Independent school	0.621	0.094	0.000	0.598	0.194	0.002	0.712	0.272	0.009	0.577	0.202	0.004
	House: rented (ref: owned)	-0.396	0.032	0.000	-0.219	0.126	0.083	-0.236	0.140	0.091	-0.192	0.131	0.144
	House: other (ref: owned)	-0.211	0.080	0.009	-0.164	0.102	0.107	-0.167	0.118	0.157	-0.152	0.111	0.170
	Disabled	-0.420	0.058	0.000	-0.275	0.102	0.007	-0.326	0.112	0.004	-0.269	0.103	0.009
	SEC: lower professional/higher technical	-0.223	0.047	0.000	-0.170	0.049	0.000	-0.240	0.049	0.000	-0.170	0.049	0.001
	SEC: intermediate occupations	-0.396	0.049	0.000	-0.281	0.082	0.001	-0.356	0.088	0.000	-0.272	0.084	0.001
	SEC: lower supervisory occupations	-0.614	0.053	0.000	-0.420	0.122	0.001	-0.495	0.136	0.000	-0.399	0.128	0.002
	SEC: semi-routine routine/routine	-0.749	0.051	0.000	-0.595	0.140	0.000	-0.679	0.163	0.000	-0.561	0.149	0.000
	SEC: other	-0.701	0.060	0.000	-0.538	0.145	0.000	-0.628	0.163	0.000	-0.517	0.151	0.001
	Both parents in household	0.027	0.032	0.394	-0.143	0.079	0.071	-0.224	0.087	0.010	-0.157	0.079	0.047
	Log local education rate - No pass	-0.103	0.025	0.000	-0.162	0.048	0.001	-0.175	0.057	0.002	-0.157	0.049	0.001
	j21	-3.129	0.071	0.000	-1.807	0.233	0.000	-1.950	0.339	0.000	-1.754	0.261	0.000
	j22	-2.775	0.067	0.000	-1.633	0.212	0.000	-1.744	0.304	0.000	-1.586	0.238	0.000
	j23	-1.824	0.063	0.000	-1.066	0.153	0.000	-1.104	0.202	0.000	-1.034	0.166	0.000
	j24	-0.975	0.059	0.000	-0.246	0.111	0.026	-0.225	0.121	0.062	-0.238	0.110	0.030
	μ_γ	-0.325	0.014	0.000	-0.768	0.190	0.000	-0.905	0.145	0.000	-0.213	0.225	0.345
	σ_γ	n/a	n/a	n/a	n/a	n/a	n/a	0.568	0.110	0.000	0.035	0.045	0.432
	ρ	n/a	n/a	n/a	0.459	0.231	0.047	0.588	0.201	0.003	0.503	0.229	0.028
	Log pseudolikelihood	-10032.449/-10160.923			-20216.206			-20034.573			-20045.156		
	N	11942			11942			11942			11942		

Source: YCS 10 (surveys those eligible to leave school in 1998-1999), weighted, reference category for SEC is "higher professional occupations"

However, results from the bivariate ordered probit model with mixed effects suggest that the impact of truancy on educational attainment is more complex than assumed by the naïve bivariate ordered probit model – which only estimates a common mean effect, γ , for every individual i in the sample. The mixed model estimates an additional parameter; σ_γ , which represents the standard deviation of a normal/log-normal distribution of the effect of truancy on educational attainment. A statistically significant estimate of 0.568 for σ_γ indicates that there is considerably heterogeneity in the impact of truancy on educational outcomes across the population which a naïve estimator cannot capture. In addition, the estimated value of ρ increases somewhat to 0.568, although the standard error suggests no significant difference compared the naïve estimate of ρ .²² Figure 3 depicts our estimates in graphical form which further highlights the varied effect that truancy can have on educational attainment.

²² We would like to stress the difference between ρ and $\gamma_i - \rho$ measures the correlation between the error terms in the two regressions. The error terms are comprised of any unobservable characteristics which we do not observe and any statistically significant correlation between the two error terms is indicative that that the exogeneity condition of T_i in regression E_i cannot be accepted. This is because error term ε_{T_i} now enters the second equation and if it is correlated with error term ε_{E_i} , the expected value of both error terms will not be zero, $E(\varepsilon_{T_i}, \varepsilon_{E_i}) \neq (0,0)$. This violates the zero-conditional mean assumption of the error term. Gamma (γ) estimates the effect that T_i has on E_i and should not be mistaken for endogeneity. Gamma could influence the estimated value of rho by correctly ascribing a more nuanced impact of T_i on E_i , and hence reducing the amount of ‘unobservables’ in the error term. However, this is generally a very unlikely scenario.

Figure 3: The Impact of Truancy on Educational Attainment at Age 16



Results from the mixed estimator suggest that the average effect of truancy on educational attainment is negative (-0.905) and statistically significant and that there exists a significant amount of heterogeneity around this average. For example, 68% of individuals experience a negative effect of truancy in the range of -1.473 and -0.337. Alternatively, 5.59% of individuals experience a positive (greater than 0) effect of truancy on educational attainment. This result is likely an artefact of the distributional assumption we make as by definition a small but finite percent of individuals will experience a positive effect if we impose an unrestricted normal distribution.²³ Indeed to test the distributional assumption we make by using a normal distribution we also impose log-normal distributed mixed effect which has the advantage of

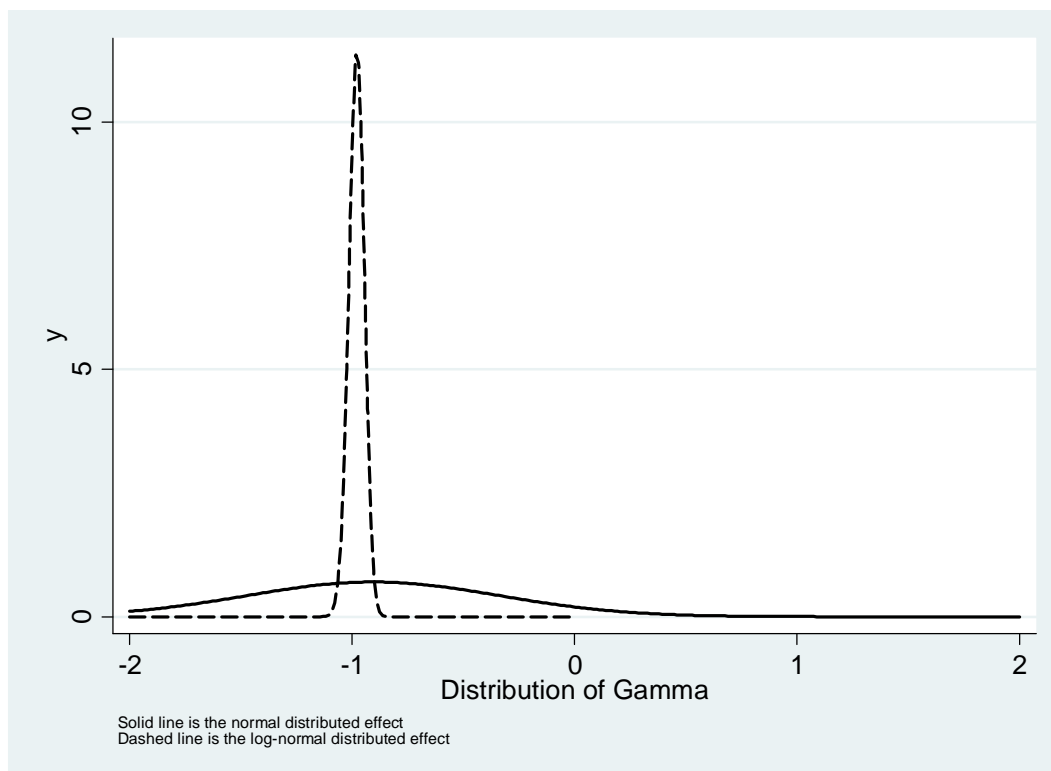
²³ An alternative hypothesis for the positive impact of truancy on educational attainment is that we are unable to examine 'why' students truant. One could imagine that youths who truant in the 'last hour of school' would not have received much added educational value anyway. Moreover, if they truant because they participate in extra-curricular activities or hold part-time jobs then the marginal benefit of an hour outside school to educational skills may be higher than inside school.

being asymmetrical (perhaps a bulk of individuals experience a small effect and a few individuals experience very large effects) and bounded between $-\infty$ and 0.²⁴

Regression results indicate that the log-normal distribution does not improve the pseudo-likelihood and that the moments of the underlying normal distribution are statistically insignificant from zero ($\mu_\gamma = -0.213$ with a p-value of 0.345 and $\sigma_\gamma = 0.035$ with a p-value of 0.432). However, it is prudent to plot the log-normal distribution anyway as the nature of the log-normal distribution can involve the estimation of rather small effects (as these are later exponentiated) which may not be easily seen. We plot both the normal and log-normal distribution in Figure 4 for comparative purposes and find that both estimators estimate approximately the same mean effect, μ_γ . However, it appears that the log-normal estimator attempted to impose a symmetrically distributed effect around this mean – which in the case of log-normal distribution can only be done with a very ‘tight’ standard deviation. Such results therefore confirm our initial hypothesis that a symmetric distribution is the right choice and we accept the normally distributed mixed effect.

²⁴ The alternative bounding is between 0 and ∞ which is unlikely to be correct in the case of truancy on attainment.

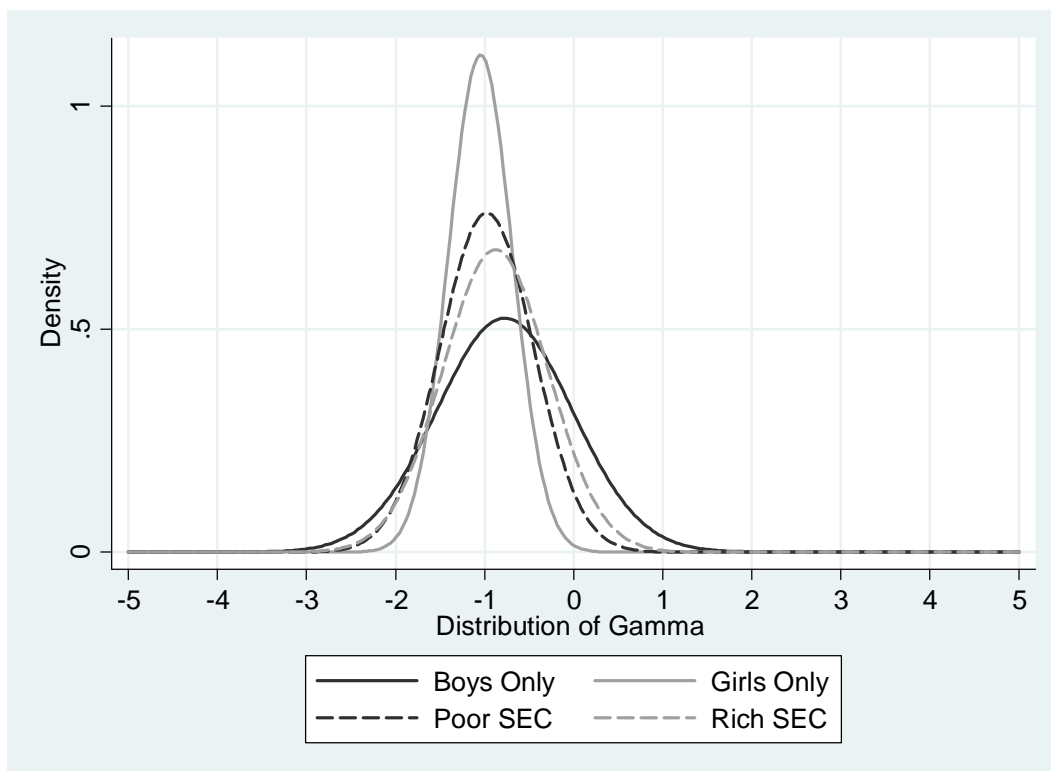
Figure 4: Normal vs. log-normal distributed effect



Our estimator has shown that the effect of truancy on educational outcomes at 16 is more nuanced than originally supposed. We have estimated a distributional effect of truancy over the schooling population (aged 16) and found that substantial heterogeneity exists. However, our mixed effect estimator is best used when multiple distributional effects of sub-groups are compared with each other in order to examine how the effect of truancy on attainment differs by different groups of individuals. We therefore also estimate the effect of truancy on educational attainment by gender and socio-economic class. For convenience we have suppressed the full regression output for these results and present our results in a graphical format in Figure 5.²⁵

²⁵ Regression results are available upon request.

Figure 5: The Impact of Truancy on Educational Attainment at Age 16 by Gender and Socio-Economic Class



Results from Figure 5 suggest that there is little difference in the impact of truancy on educational outcomes at 16 by socio-economic class. Both individuals from poor and rich socio-economic background experience the same detrimental effect of truancy on education; in terms of the estimated average effect and how the effect is distributed around this average. However, results for gender indicates that the estimated distribution for boys is significantly different compared to the estimated distribution for girls. In particular, girls experience a higher negative mean effect (-1.046 for girls compared to -0.776 for boys) and the distribution for girls is much ‘tighter’ than that for boys (0.358 for girls compared to 0.761 for boys). The impact of truancy on educational attainment for boys is therefore more heterogeneous than for girls.

5. Conclusion

In this paper we examined the effect that truancy has on the educational outcomes of 16 year old pupils in the UK. Using a modified bivariate ordered probit estimator with mixed effects we found that the causal effect of truancy on GCSE outcomes is a) more detrimental to educational attainment than would be estimated by a naïve estimator which does not take the endogeneity of truancy into account; b) truancy is a distributional effect as opposed to a common mean effect; c) this distributional effect is likely to be a normally distributed effect and d) the distributional effect appears to differ by gender.²⁶

That truancy is found to be an endogenous variable and that its effect on educational attainment is highly negative is perhaps not surprising – after all, one would hardly expect a positive effect. Nonetheless, no study has so far attempted a causal estimation of the effect of truancy on educational attainment at age 16 in the UK and our result further strengthens the urgent need for policy measures which prove to be *effective* at dealing with the truancy problem. What such policy measures could entail, why recent policy measures have been deemed to fail and what the determinants of truancy might be are important questions for further research which we strongly encourage. Given the importance of education and the gravity of the decision that 16 year old pupils face in the UK it is surprising that so little research has been carried out in examining one of the most important determinants of educational attainment.

In addition to contributing to the education literature this paper also offers an important empirical contribution by outlining a bivariate ordered probit estimator with mixed effects. Such an estimator can have many different types of applications and is valid whenever both dependent

²⁶ These findings mirror research done by Arulampalam *et al.* (2008) who also use IV techniques in conjunction with quantile regression to argue that a standard conditional mean estimation is an oversimplification and that the impact of truancy (at university) on attainment is more complex than originally assumed.

variable and a suspected endogenous variable take an ordered form (including binary). In particular our estimator allows one to recover a sense of the heterogeneity surrounding the point estimates. Such estimation can be particularly useful when the number of explanatory variables is low or there is a danger of over-complicating the regression specification with too many, possibly endogenous, variables being included as additional regressors. Removing some and adding a mixed effect could be a viable option in such a case.²⁷ Moreover, the notoriously difficult interpretation of non-linear models (in terms of interaction terms and marginal effects) can make the mixed effects estimator an attractive proposition by visualising differing heterogeneities among sub-groups.

²⁷ Although this comes at the cost of additional assumptions about the distributional properties of γ

Bibliography

Angrist, J.D. (2001) "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice", *Journal of Business and Economic Statistics*, Vol. 19, No. 1, pp. 2-28

Angrist, J.D. and Pischke, J.S. (2008) *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, Princeton, New Jersey

Arulampalam, W., Naylor, R. and Smith, J. (2008) "Am I Missing Something? The Effects of Absence from Class on Student Performance", IZA Discussion Paper No. 3749

Bosworth, D.L. (1994) "Truancy and Pupil Performance" *Education Economics*, Vol. 2, No. 3, pp. 243-265

Botti, F., Conte, A., Di Cagno, D.T. and C. D'Ippoliti (2008) "Risk attitude in real decision problems", *The B.E. Journal of Economic Analysis & Policy – Advances*, Vol. 8: Iss. 1, Article 6.

Bradley, S. and Nguyen, A.N. (2004) "The School-to-work transition" in Johnes, G. and Johnes, J. *International Handbook on the Economics of Education* (pp. 484-521), Edward Elgar, Cheltenham

Bradley, S. and Taylor, J. (2004) "Ethnicity, Educational Attainment and the Transition from School" *The Manchester School*, Vol. 74, No. 3, pp. 317-346

Burgess, S., Gardiner, K., Propper, C. (2002) "The Economic Determinants of Truancy", Centre for Analysis of Social Exclusion, CASE paper 61

Cappellari, L. and Jenkins, S.P. (2003) "Multivariate probit regression using simulated maximum likelihood", *The Stata Journal*, No. 3, pp. 278-294

Cappellari, L. and Jenkins, S.P. (2006) "Calculation of Multivariate Normal Probabilities by Simulation, with Applications to Maximum Simulated Likelihood Estimation", *The Stata Journal*, No. 6, pp. 156-189

Chen, J. and Lin, T. (2008) “Class Attendance and Exam Performance: A Randomized Experiment”, *Journal of Economic Education*, Vol. 39, No. 3, pp. 213-227

Conte, A., Hey, J.D. and Moffatt, P.G. (2009) “Mixture models of choice under risk”, *Journal of Econometrics*, doi:10.1016/j.jeconom.2009.10.011 (in press).

DCSF (2009) “Pupil Absence in Spring 2009 Term”, Statistical Release, 26th August 2009, London

Dobkin, C., Gil, R. and Marion, J. (2009) “Skipping Class in College and Exam Performance: Evidence from a Regression Discontinuity Classroom Experiment” *Economics of Education Review*, In Press, Corrected Proof, Available online 23rd September 2009

Durden, G. and Ellis, L. (1995) “The Effects of Attendance on Student Learning in Principles of Economics”, *American Economic Review*, Vol. 85, No. 2, pp. 343–46.

Dustman, C., Rajah, N. and Smith, S. (1997). “Teenage Truancy, Part-time Working and Wages”, *Journal of Population Economics*, Vol. 10, No. 4, pp. 425-442

Farrington, D. (1996). “Later Life Outcomes of Truants in the Cambridge Study”, in *Unwillingly to School 4th edition*, (eds: Ian Berg & Jean Nursten). Gaskell Press, pp. 96-118.

Greene, W.H. (2008) *Econometric Analysis 6th Edition*, Prentice Hall, New Jersey, USA.

Greene, W.H. and Hensher, D.A. (2009) “Modeling Ordered Choices” Mimeo, Stern School of Business, New York (January 2009)

Halsey, K., Johnson, A., Kinder, K. and Fletcher-Morgan, C. (2003) “Evaluation of truancy sweep follow-ups: a Summary of Approaches Used by Seven LEAs and Factors to Consider When Running a Truancy Sweep” National Foundation for Educational Research report, commissioned by Department for Education and Skills

Harmon, C., Oosterbeek, H. and Walker, I. (2003) "The Returns to Education: Microeconomics", *Journal of Economic Surveys*, Vol. 17, No. 2, pp. 115-155

Heilbrunn, J.Z. (2007) "Piece of the 'Truancy Jigsaw: A Literature Review", an initiative by the Colorado Foundation for Families and Children, available at www.schoolengagement.org

Hibbert, A. and Fogelman, K. (1990), "Future Lives of Truants: Family Formation and Health Related Behaviour" *British Journal of Educational Psychology*, Vol. 60, No. 2, pp. 171-179.

Marburger, D.R. (2001) "Absenteeism and Undergraduate Exam Performance", *Journal of Economic Education*, Vol. 32, No. 2, pp. 99-110

Kirby, A. and McElroy, B. (2003). "The effect of attendance on grade for first year Economics students in University College Cork". *The Economic and Social Review*, Vol. 34, No. 3, pp. 311-326.

Manski, C.F. (1993) "Identification of Endogenous Social Effects: The Reflection Problem", *Review of Economic Studies*, Vol. 60, pp. 531-542

Payne, J. (2001) "Patterns of Participation in Full-Time Education after 16: An Analysis of the England and Wales Youth Cohort Study", Department for Education and Skills Research Report, RR307

Romer, D. (1993) "Do Students Go To Class? Should They?" *Journal of Economic Perspectives*, Vol. 7, pp. 167-174

Sajaia, Z. (2008) "Maximum Likelihood Estimation of a Bivariate Ordered Probit Model: Implementation and Monte Carlo Simulations", World Bank Applied Economic Research Papers

Stanca, L. (2006) "The Effects of Attendance on Academic Performance: Panel Data Evidence for Introductory Microeconomics", *Journal of Economic Education*, Vol. 37, No. 3, pp. 251-266

Select Committee on Public Accounts (2006) *Progress in improving school attendance* House of Commons Public Accounts – Eighteenth Report

Appendix A – Derivation of the reduced form

Let us consider the following transformation of the recursive system in eq. (1), where $y_{1i}^* = T_i^*$

and $y_{2i}^* = E_i^*$:

$$\begin{cases} y_{1i}^* = x_{1i}\beta_1 + \varepsilon_{1i} \\ y_{2i}^* - \gamma_i y_{1i}^* = x_{2i}\beta_2 + \varepsilon_{2i} \end{cases} \quad (\text{A1})$$

that in matrix form is:

$$\Gamma Y^* = B X + E, \text{ with } E \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma\right) \text{ and } \Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

(A2)

where:

$$Y^* = \begin{bmatrix} y_{1i}^* \\ y_{2i}^* \end{bmatrix}; \Gamma = \begin{bmatrix} 1 & 0 \\ -\gamma_i & 1 \end{bmatrix}; X = \begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix}; B = \begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{bmatrix}; E = \begin{bmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{bmatrix}. \quad (\text{A3})$$

By pre-multiplying both sides by Γ^{-1} , we get:

$$Y^* = \Lambda X + \Upsilon,$$

(A4)

where $\Lambda = \Gamma^{-1}B$, $\Upsilon = \Gamma^{-1}E$, and $\Upsilon \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Omega\right)$, with

$$\Omega = (\Gamma^{-1})\Sigma(\Gamma^{-1})' = \begin{bmatrix} 1 & \rho + \gamma_i \\ \rho + \gamma_i & 1 + 2\gamma_i\rho + \gamma_i^2 \end{bmatrix}.$$

Let us define a matrix Θ having on the principal diagonal the inverse squared root of the terms in the principal diagonal of Ω and zero somewhere else. By pre-multiplying eq. (A4) for Θ , we get:

$$\Pi Y^* = \Pi \Lambda X + \Pi \Upsilon. \quad (\text{A5})$$

It is worth noting that the error term of the system transformed as such is now:

$$\Pi \Upsilon \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Pi \Omega \Pi\right), \text{ where } \Pi \Omega \Pi = \begin{bmatrix} 1 & \tilde{\rho}_i \\ \tilde{\rho}_i & 1 \end{bmatrix} \text{ with } \tilde{\rho}_i = \frac{\rho + \gamma_i}{\sqrt{1 + 2\gamma_i\rho + \gamma_i^2}}. \quad (\text{A6})$$

Basically, we have transformed the system of recursive equations in eq. (1) in the following way in order to fill all its terms in a bivariate standard normal probability distribution function that is available in all the principal statistical packages. The system of equations after our transformation appears then:

$$\begin{cases} y_{1i}^* = x_{1i}\beta_1 + \varepsilon_{1i} \\ \frac{y_{2i}^*}{\sqrt{1+2\gamma_i\rho+\gamma_i^2}} = \frac{\gamma_i y_{1i}^* + x_{2i}\beta_2}{\sqrt{1+2\gamma_i\rho+\gamma_i^2}} + \frac{\gamma_i\varepsilon_{1i} + \varepsilon_{2i}}{\sqrt{1+2\gamma_i\rho+\gamma_i^2}} . \end{cases} \quad (A7)$$

Appendix B – Descriptive statistics

Variable	Type		Obs.	Freq	Wght. Freq
Gender	Binary	0 = boys	5,413	45.33	50.49
		1 = girls	6,529	54.67	49.51
Ethnicity	Binary	0 = white	10,560	88.43	87.59
		1 = non-white	1,382	11.57	12.41
Parental Employment	Binary	0 = otherwise	10,143	84.94	82.38
		1 = neither parent empl.	1,799	15.06	17.62
Parent Education	Categorical	0 = 1 parent has degree level	3,071	25.72	22.04
		1 = 1 parent has a-level	1,877	15.72	15.05
		2 = both parents have no a-level	6,994	58.57	62.90
School Type	Categorical	0 = comprehensive school	10,167	85.14	89.16
		1 = selective/grammar school	710	5.95	4.17
		2 = independent school	1,065	8.92	6.66
Housing	Categorical	0 = owned by parents	9,709	81.30	77.25
		1 = rented	2,008	16.81	20.77
		2 = other	225	1.88	1.98
Disability	Binary	0 = non-disabled	11,427	95.69	94.84
		1 = disabled	515	4.31	5.16
Socio-Economic Class	Categorical	0 = higher professional	2,449	20.51	17.43
		1 = lower professional	3,421	28.65	26.43
		2 = intermediate	2,489	20.84	20.90
		3 = lower supervisory	1,155	9.67	10.88
		4 = semi-routine and routine	1,408	11.79	13.98
		5 = other	1,020	8.54	10.39
Household Composition	Binary	0 = otherwise	2,442	20.45	22.04
		1 = both parents in household	9,500	79.55	77.96
Local Truancy Rate	Continuous	132 LEA local truancy averages	<i>Mean</i> 1.074	<i>Std.</i> 0.599	<i>Min / Max</i> 0.30 / 3.90
Local Education Rate (5ABCs)	Continuous	132 LEA local education averages	0.471	0.074	0.23 / 0.61
Local Education Rate (5A-to-Gs)	Continuous	132 LEA local education averages	0.898	0.033	0.77 / 0.95
Local Education Rate (No Passes)	Continuous	132 LEA local education averages	0.052	0.019	0.01 / 0.14

354 observations were removed from the full sample due to missing information in either the truancy or education variable. The above descriptives refer to the remaining sample of 11,942 students

Appendix C – Monte Carlo Simulations

To examine the small sample properties of our estimator we implement Monte Carlo simulations.

The latent variables y_{1i}^* and y_{2i}^* for the bivariate ordered probit with mixed effects model are generated by the following data generating process:

$$\begin{cases} y_{1i}^* = 1x_{1i} + 1z_{1i} + \varepsilon_{1i} \\ y_{2i}^* = \gamma_i y_{1i}^* - 2.5x_{1i} + \varepsilon_{2i} \end{cases} \quad (B1)$$

where $\gamma_i \sim N(\mu_\gamma = 0.5, \sigma_\gamma = 0.5)$. We generate x_{1i} and z_{1i} as independent standard normal random variables and ε_{1i} and ε_{2i} as bivariate standard normal random variables with correlation ρ . The observational rule is:

$$y_{1i} = \begin{cases} 1 & \text{if } y_{1i}^* \leq -2 \\ 2 & \text{if } -2 < y_{1i}^* \leq 0.5 \\ 3 & \text{if } 0.5 < y_{1i}^* \end{cases} \quad y_{2i} = \begin{cases} 1 & \text{if } y_{2i}^* \leq -3 \\ 2 & \text{if } -3 < y_{2i}^* \leq -1 \\ 3 & \text{if } -1 < y_{2i}^* \leq 1 \\ 4 & \text{if } 1 < y_{2i}^* \end{cases} \quad (B2)$$

We run 1000 replications for values of $\rho = \{-0.9, -0.5, 0, 0.5, 0.9\}$ for observations, $Obs = \{200, 500, 1000, 5000\}$. We report the sample mean estimates of μ_γ , σ_γ and ρ in addition to reporting the root mean square error (RMSE). Our simulation results indicate that our mixed estimator performs well in recovering the true parameters, even in small samples. The authors are happy to share and explain the associated Stata .do files in order to replicate the simulation results and/or to modify our estimator for personal use.

Bivariate ordered probit with mixed effects													
<i>Obs = 200</i>							<i>Obs = 500</i>						
Rho	μ_γ	RMSE	σ_γ	RMSE	ρ	RMSE	Rho	μ_γ	RMSE	σ_γ	RMSE	ρ	RMSE
-0.9	0.4971	0.0729	0.4723	0.1082	-0.8943	0.0502	-0.9	0.4997	0.0436	0.4994	0.0647	-0.9021	0.0324
-0.5	0.5092	0.1000	0.5121	0.1599	-0.5015	0.1166	-0.5	0.5043	0.0587	0.5068	0.0915	-0.5029	0.0721
0	0.5346	0.1455	0.5647	0.2184	0.0124	0.1518	0	0.5067	0.0757	0.5084	0.1091	0.0015	0.0964
0.5	0.5256	0.1489	0.5254	0.1781	0.5259	0.1387	0.5	0.5116	0.0924	0.5104	0.1066	0.5081	0.0785
0.9	0.5006	0.1504	0.4589	0.1387	0.8823	0.0591	0.9	0.4965	0.0928	0.4858	0.0845	0.9008	0.0404
<i>Obs = 1000</i>							<i>Obs = 5000</i>						
Rho	μ_γ	RMSE	σ_γ	RMSE	ρ	RMSE	Rho	μ_γ	RMSE	σ_γ	RMSE	ρ	RMSE
-0.9	0.5007	0.0324	0.5018	0.0468	-0.9033	0.0234	-0.9	0.4996	0.0140	0.4979	0.0198	-0.8997	0.0098
-0.5	0.5022	0.0434	0.4997	0.0642	-0.5027	0.0517	-0.5	0.4995	0.0186	0.4976	0.0281	-0.4996	0.0230
0	0.5043	0.0528	0.4998	0.0760	-0.0030	0.0656	0	0.4998	0.0243	0.4980	0.0341	-0.0005	0.0288
0.5	0.5047	0.0590	0.4994	0.0750	0.5001	0.0575	0.5	0.5008	0.0278	0.4989	0.0327	0.4991	0.0248
0.9	0.5066	0.0666	0.5004	0.0610	0.9029	0.0321	0.9	0.5009	0.0292	0.5000	0.0276	0.9016	0.0138

Monte Carlo Simulations: 1000 Replications

Appendix D – Supplemental Material

For completeness we attach the full results for the estimations by gender and socio-economic class.

Table C1: Bivariate Ordered Probit with Mixed Effects estimation - Supplemental results

Variables		Boys			Girls			Poor SEC			Rich SEC		
Dependent	Independent	Coef	S.E	P-val	Coef	S.E	P-val	Coef	S.E.	P-val	Coef	S.E.	P-val
Truancy	Female	n/a	n/a	n/a	n/a	n/a	n/a	0.079	0.028	0.005	0.044	0.040	0.273
	Non-white	-0.055	0.055	0.325	-0.158	0.048	0.001	-0.093	0.048	0.049	-0.154	0.056	0.006
	Neither parents employed	-0.183	0.055	0.001	-0.056	0.050	0.259	-0.133	0.053	0.013	-0.124	0.051	0.016
	At least one parent with A-level (ref: degree level)	-0.082	0.058	0.154	0.023	0.051	0.650	-0.036	0.042	0.382	0.017	0.095	0.860
	Neither parent with A-level (ref: degree level)	-0.127	0.047	0.007	-0.059	0.041	0.152	-0.057	0.033	0.090	-0.213	0.073	0.003
	Selective school	-0.226	0.084	0.007	-0.372	0.074	0.000	-0.284	0.060	0.000	-0.330	0.141	0.019
	Independent school	-0.422	0.071	0.000	-0.429	0.063	0.000	-0.391	0.049	0.000	-0.617	0.188	0.001
	House: rented (ref: owned)	0.280	0.048	0.000	0.312	0.042	0.000	0.320	0.046	0.000	0.277	0.045	0.000
	House: other (ref: owned)	0.152	0.128	0.235	0.074	0.109	0.497	0.033	0.113	0.770	0.171	0.126	0.175
	Disabled	0.135	0.076	0.075	0.105	0.072	0.142	0.157	0.068	0.020	0.077	0.082	0.346
	SEC: lower professional/higher technical	-0.092	0.050	0.069	-0.051	0.046	0.262	n/a	n/a	n/a	n/a	n/a	n/a
	SEC: intermediate occupations	0.058	0.057	0.309	0.017	0.052	0.749	n/a	n/a	n/a	n/a	n/a	n/a
	SEC: lower supervisory occupations	0.136	0.071	0.055	0.101	0.062	0.102	n/a	n/a	n/a	n/a	n/a	n/a
	SEC: semi-routine routine/routine	0.105	0.067	0.120	0.097	0.061	0.112	n/a	n/a	n/a	n/a	n/a	n/a
	SEC: other	0.224	0.079	0.005	0.033	0.073	0.650	n/a	n/a	n/a	n/a	n/a	n/a
	Both parents in household	-0.319	0.041	0.000	-0.442	0.036	0.000	-0.403	0.034	0.000	-0.325	0.045	0.000
	Log local truancy rate (unauthorised absence)	0.145	0.034	0.000	0.053	0.031	0.086	0.079	0.027	0.004	0.117	0.040	0.003
	j11	0.199	0.047	0.000	0.046	0.049	0.351	0.189	0.043	0.000	-0.107	0.071	0.128
	j12	1.159	0.048	0.000	1.088	0.050	0.000	1.247	0.044	0.000	0.812	0.070	0.000
	j13	1.720	0.058	0.000	1.626	0.058	0.000	1.821	0.052	0.000	1.336	0.077	0.000
	j14	2.085	0.073	0.000	2.064	0.071	0.000	2.279	0.070	0.000	1.711	0.085	0.000
Education	Female	n/a	n/a	n/a	n/a	n/a	n/a	0.328	0.074	0.000	0.349	0.069	0.000
	Non-white	0.135	0.073	0.063	-0.050	0.085	0.556	-0.027	0.061	0.655	0.183	0.107	0.086
	Neither parents employed	-0.167	0.077	0.031	-0.079	0.055	0.149	-0.137	0.065	0.035	-0.198	0.061	0.001

Table C1: Bivariate Ordered Probit with Mixed Effects estimation - Supplemental results

Variables		Boys			Girls			Poor SEC			Rich SEC		
Dependent	Independent	Coef	S.E	P-val	Coef	S.E	P-val	Coef	S.E.	P-val	Coef	S.E.	P-val
	At least one parent with A-level (ref: degree level)	-0.345	0.075	0.000	-0.108	0.095	0.257	-0.270	0.086	0.002	0.023	0.113	0.840
	Neither parent with A-level (ref: degree level)	-0.486	0.064	0.000	-0.277	0.125	0.026	-0.411	0.114	0.000	-0.226	0.096	0.018
	Selective school	1.311	0.290	0.000	0.770	0.690	0.264	0.961	0.448	0.032	1.318	0.484	0.006
	Independent school	0.902	0.276	0.001	0.356	0.487	0.464	0.540	0.354	0.127	1.076	0.516	0.037
	House: rented (ref: owned)	-0.308	0.140	0.029	-0.054	0.240	0.822	-0.219	0.217	0.313	-0.218	0.163	0.182
	House: other (ref: owned)	-0.112	0.164	0.494	-0.135	0.180	0.452	-0.251	0.163	0.123	-0.075	0.171	0.661
	Disabled	-0.376	0.129	0.003	-0.199	0.194	0.305	-0.316	0.184	0.087	-0.264	0.134	0.049
	SEC: lower professional/higher technical	-0.253	0.069	0.000	-0.189	0.087	0.031	n/a	n/a	n/a	n/a	n/a	n/a
	SEC: intermediate occupations	-0.358	0.095	0.000	-0.279	0.173	0.108	n/a	n/a	n/a	n/a	n/a	n/a
	SEC: lower supervisory occupations	-0.620	0.146	0.000	-0.296	0.240	0.218	n/a	n/a	n/a	n/a	n/a	n/a
	SEC: semi-routine routine/routine	-0.794	0.155	0.000	-0.446	0.316	0.158	n/a	n/a	n/a	n/a	n/a	n/a
	SEC: other	-0.733	0.191	0.000	-0.451	0.281	0.109	n/a	n/a	n/a	n/a	n/a	n/a
	Both parents in household	-0.130	0.102	0.200	-0.353	0.104	0.001	-0.226	0.117	0.054	-0.244	0.095	0.010
	Log local education rate - No pass	-0.250	0.081	0.002	-0.096	0.071	0.177	-0.144	0.074	0.051	-0.184	0.083	0.026
	j21	-2.206	0.295	0.000	-1.574	0.847	0.063	-1.765	0.529	0.001	-1.030	0.260	0.000
	j22	-1.953	0.272	0.000	-1.438	0.773	0.063	-1.582	0.474	0.001	-0.810	0.236	0.001
	j23	-1.164	0.219	0.000	-1.023	0.548	0.062	-1.032	0.311	0.001	-0.111	0.213	0.602
	j24	-0.165	0.210	0.431	-0.405	0.227	0.074	-0.260	0.144	0.071	0.854	0.311	0.006
	μ_γ	-0.776	0.224	0.001	-1.046	0.085	0.000	-0.975	0.141	0.000	-0.879	0.224	0.000
	σ_γ	0.761	0.119	0.000	0.358	0.201	0.075	0.525	0.171	0.002	0.589	0.138	0.000
	ρ	0.382	0.269	0.156	0.834	0.200	0.000	0.705	0.224	0.002	0.506	0.297	0.089
	Log pseudolikelihood	-9425.7358			-10576.755			-12163.673			-7815.2429		
	N	5413			6529			8359			3583		

Source: YCS 10 (surveys those eligible to leave school in 1998-1999), weighted, reference category for SEC is "higher professional occupations"