



The effects of test scores and truancy on youth unemployment and inactivity: a simultaneous equations approach

Steve Bradley¹ · Rob Crouchley¹

Received: 15 May 2018 / Accepted: 4 April 2019 / Published online: 16 April 2019
© The Author(s) 2019

Abstract

We analyse the relationships between test scores, truancy and labour market outcomes for youths. Our econometric approach enables us to disentangle the observable direct and indirect effects of truancy and test scores on the risk of unemployment or ‘Not in Education, Employment or Training’ (NEET) from their unobserved effects. Using data for England and Wales, we show that models of youth unemployment and NEET that ignore the correlation between the unobservable determinants of test scores and truancy will lead to misleading inference about the strength of their effects. Truancy has an indirect observed effect on labour market outcomes via its effect on test scores, and a weak direct effect. The unobserved component of truancy has a direct effect on labour market outcomes. Test scores have a direct effect on those outcomes, but also mitigate the detrimental effects of truancy. Our analysis raises important implications for education policy.

Keywords Test scores · Truancy and labour market outcomes · Trivariate models

JEL Classification I21 · J64

✉ Steve Bradley
s.bradley@lancaster.ac.uk

¹ Department of Economics, Lancaster University Management School, Lancaster University, Lancaster LA1 4YX, UK

1 Introduction

The youth unemployment rate has been rising since 2004, pre-dating the 2008 recession, following a fairly predictable pattern with regard to cyclical downturns (Petrongolo and van Reenen 2011).¹ Although it is difficult to pin down the causes of the rise in youth unemployment, one possible cause highlighted by Petrongolo and van Reenen (2011) is the quality of schooling. Furthermore, what is very clear from their analysis of *UK LFS* data is that the unemployment rate for 16–17 year olds was as high in 2010 as it was in the last major recession in 1980—exceeding 30%. However, official measures of youth unemployment, and teenage unemployment in particular, are likely to understate the true magnitude of joblessness for this group given the propensity of some youths to drop out of the labour market and remain economically inactive for periods of time. A better measure of the labour market fortunes of youths is therefore likely to be the proportion of the group who are ‘Not in Education, Employment or Training’ (NEET)—the unemployed and economically inactive. Since this group of young people are not engaged in skill formation of any kind, they are most likely to be ‘scarred’ by this early labour market experience.

Previous research has, in fact, shown that a poor start to a young persons career can lead to an increased probability of unemployment, as well as a negative effect on future earnings. A lack of investment in human capital, such as vocational skills, acquired through work and training, or through education, causes such detrimental effects. For instance, Arulampalam et al. (2001) have shown that earnings can be 6% lower on re-entry to a job and 14% lower after 3 years—the most damaging spell of unemployment is the first. Clark, Georgellis and Sanfey (2001) also show that past unemployment is correlated with current life satisfaction, an additional dimension to the scarring effect, although Knabe and Ratzel (2011) have recently shown that this effect operates via a fear of future unemployment (See also Bell and Blanchflower 2011, 2012). Mroz and Savage (2007) provide a counter argument suggesting that, whilst earnings growth can be retarded, young workers who experience spells of unemployment can respond by acquiring human capital which reduces the risk of future spells of unemployment.

This paper focuses on the labour market outcomes of the teenage group (16–17 year olds) and assesses the relationships between test scores and truancy behaviour, each of which are a function of school quality, family background and personal characteristics, including a pupil’s tastes for schooling, and the risk of unemployment or NEET several months after leaving school.² There has been considerable debate in the press and amongst politicians in recent years about pupils who persistently miss school due to unauthorised absence (truancy), which has,

¹ Youths are often defined as those aged 16–24 years, however, it is often the case that a distinction is made between teenagers (aged 16–19) and the rest because the former have very little work experience or skills.

² We choose September of the year of leaving compulsory schooling since by this date those who intend to go to college, enter an apprenticeship/training scheme or take up employment will have done so.

in several high profile cases, led to fines for parents.³ Pupils who miss schooling because of truancy are likely to have lower test scores than pupils who do not miss school. When truancy reaches high levels, such that attendance at school is infrequent, then truants are more like high school dropouts, a phenomenon that has received a lot of attention in the USA and Europe. The consequences for this group of truants in terms of adverse labour market outcomes are likely to be similar to those for high school dropouts. Pupils who truant less frequently, whose behaviour is amenable to teacher interventions, are likely to have different and more positive labour market outcomes, suggesting that we need to disaggregate truants by the frequency of this event. Furthermore, it is not clear whether truancy has a direct effect on the risk of unemployment or NEET. There are several possible mechanisms through which truancy could have a direct effect on the risk of unemployment or NEET. First, a higher propensity to truant could increase the probability of unemployment or NEET insofar as it could act as a negative productivity signal to employers and training providers, providing of course that employers actually receive this signal. However, there is no guarantee that pupils would provide evidence of truancy from school during the job selection process. Second, truants may mix with other young people who are unemployed or inactive, and this peer effect may increase their risk of unemployment or NEET. Third, schooling not only creates human capital in the form of cognitive skills, it also increases soft skills which employers value. Truants could miss out on the acquisition of these soft skills which increases their risk of unemployment or NEET. Fourth, pupils who truant may fail to develop social networks which are useful for finding work, and their lack of motivation at school could be signalled via teacher references to employers, thereby increasing the risk of unemployment or NEET. Finally, a contrary argument is that, pupils who have truant, could search more intensively for jobs because they have 'switched off' school and want to work. Unfortunately, our data do not permit us to investigate these different channels. Nevertheless, if truancy does lead to lower test scores, then there is a possible indirect, and positive, effect of truancy on the risk of unemployment or NEET. In contrast, there is a large literature which demonstrates a strong link between low high school test scores and a higher probability of unemployment and NEET (see Sect. 2).

We argue that because decisions regarding truancy, which could be seen as a proxy for effort at school, and performance in tests affect the subsequent transition from school, then these behavioural outcomes (decisions) are simultaneously determined. To capture this simultaneity, a three-equation model is estimated in which

³ The 1996 Education Act introduced school attendance orders, making unauthorised pupil absence (truancy) an offence. The Act states that: "If a parent on whom a school attendance order is served fails to comply with the requirements of the order, he is guilty of an offence, unless he proves that he is causing the child to receive suitable education otherwise than at school." Over time these 'rules and regulations' have been gradually tightened (see the 2006 Education Act) and enforced more frequently. For instance, the BBC reported recently on the outcome of a court case in which a school had fined a parent for taking their child out of school to go on holiday, classed as unauthorised absence. The parent had challenged the fine but lost the case (see www.bbc.co.uk/news/education-39504338). Whilst this is an unusual case it does illustrate the steps that schools are now taking to reduce truancy.

we allow for correlation in unobservables between models for truancy, test scores, both of which are ordered categorical variables, and unemployment (or NEET), which are binary variables. Therefore, unlike virtually all the previous literature, our models acknowledge the endogeneity of truancy and test score performance for early labour market behaviour; specifically, we model the endogeneity of truancy for test scores, and we model the joint endogeneity of test scores and truancy for early labour market behaviour. By building more comprehensive models, we are able to obtain more insight into the determinants of early labour market behaviour, and get closer than other researchers have to uncovering causal effects when using cross-sectional data.⁴

To estimate our model, we use pupil-level data from the Youth Cohort Studies (YCS), specifically YCS6 to YCS12, which cover the period of the early 1990s and early 2000s. To each of these datasets, we append detailed information on the characteristics of the school attended which was obtained from the School Performance Tables and Schools Census.

Closely related papers include Boardman et al (1977), which models the educational process with six simultaneous equations for student's achievement, motivation, expectations, efficacy, and perceived parents' and teachers' expectations. However, the closest paper to our own is that of Buscha and Conte (2013) who estimate a bivariate simultaneous equation model for the discrete ordered responses of student's truancy and test scores. The Buscha and Conte (2013) paper has the advantage that it allows for a correlation between the unobservables in their model and they allow for the endogeneity of the latent value of truancy on the latent value of test scores using YCS data.

Our paper differs from Buscha and Conte (2013) in a number of respects. We extend their approach by estimating a trivariate model, primarily to explore the direct effect of truancy and the direct test scores on the risk of unemployment and NEET. Also, we present an alternative identification strategy to that proposed by Buscha and Conte (2013); their model relies on local labour market conditions, that is, part-time pay, to identify the latent variable on truancy in the latent model for test scores. It is difficult to believe, however, that local labour market conditions do not have, in their own right, a direct effect on the test scores. Working part-time (e.g. in a supermarket) whilst still at school may encourage students to work harder at school so as to avoid this kind of job post-school. Our approach to identification differs to theirs insofar as we include the actual values, or direct effects, of the endogenous variables, truancy and test scores, in our sub-models for test scores and labour market outcomes (see below for a fuller discussion of our identification strategy). A further difference between the two approaches is that we investigate the direction of 'causation' between truancy and test scores on the risk of unemployment and NEET. We also estimate a range of alternative specifications of our system of equations.

⁴ It is difficult to imagine how a true 'experiment', or 'quasi-experiment' in the UK could be conducted with respect to truancy from school and test score completion, since it is the statutory responsibility of all Local Education Authorities to provide pupils with schooling up to the age of 16–17.

Finally, we have a richer set of covariates than Buscha and Conte (2013) because we map detailed school-level data on to the pupil-level YCS data.⁵

The findings from our preferred model (Heterogeneous Model 1) suggest that truancy works primarily through test scores (i.e. an indirect effect), having a ‘weak’ direct effect on labour market outcomes. However, truancy also has an unobserved effect on the risk of unemployment and the risk of NEET insofar as the correlation between latent variables for truancy and labour market outcomes are positive and statistically significant. Test scores have a direct effect on labour market outcomes, and through the estimation of ATTs, we show a good performance in high stakes tests (i.e. GCSEs) can mitigate the effect of truanting from school on labour market outcomes. In sum, truancy need not be a significant problem for young people in terms of their post-school outcomes, so long as this behaviour does not reduce test score performance. This makes sense insofar as employers observe test score performance in the job/training selection process whereas they are less likely to observe truancy behaviour. We find no evidence of ‘reverse causality’, i.e. that test scores determine truancy. We draw out the implications for policy in our conclusions.

The remainder of the paper is structured as follows. In Sect. 2 we briefly discuss the existing literature on the determinants of test scores, truancy and unemployment or NEET. This is followed by the specification of our simultaneous model—a trivariate-ordered probit model. Section 4 provides a discussion of the data that is used in our econometric analysis, and in Sect. 5 we present our results. This is followed by our conclusions.

2 A review of the literature

There is a large literature which investigates the determinants of the school-to-work transition, including the risk of unemployment and NEET (see Bradley and Nguyen 2004 for a review). Many of these papers estimate reduced-form, single-equation models, where the role of test scores features prominently as a determinant of a successful school-to-work transition (Lynch 1987; Andrews and Bradley 1997; Crawford et al. 2011; Duckworth and Schoon 2012). Coles et al. (2010) argue that the main determinants of NEET occur pre-school leaving and refer to different forms of ‘educational disaffection and educational disadvantage’. Ermisch and Janatti (2012) go further and argue that low test scores is a key mechanism that perpetuates disadvantage across the generations.

It is also worth noting that Coles et al (2010) see a direct correlation between educational disaffection and the probability of NEET. This is important in our context because educational disaffection refers to involuntary exclusion from school as well as what they refer to as ‘self-exclusion’—truanting from school. Duckworth and Schoon (2012) also find this effect.

⁵ Dustmann et al. (1998) also estimate a system of equations, including test scores, however, their focus is upon the effect of part time work on this and the school-to-work transition.

School effects on the school-to-work transition have also been identified (Micklewright 1989; Rice 1987, 1999; Dolton et al 1999), as well as gender and ethnic differences are also evident. Non-white girls are more likely to stay on beyond compulsory school leaving age to avoid unemployment (Leslie and Drinkwater 1999), an effect which is largest for Indian and Chinese pupils than for Black Caribbeans (Bradley and Taylor 2002).

Not surprisingly, the probability of staying on at school, and hence avoiding unemployment or NEET, is higher for young people from a professional family background, and much lower if their father is a manual worker (Rice 1987, 1999; Crawford et al. 2011). Young people from single parent families and those with unemployed heads of household also tend to leave school early, partly because of financial constraints on the household, and enter NEET (Coles et al 2010). Duckworth and Schoon (2012) show that for cohorts of school leavers in the 1970s having parents with low education and living in social housing increased the likelihood of NEET. However, this finding disappears for cohorts of school leavers in the 1990s, suggesting some degree of educational mobility in more recent years.

In terms of the determinants of test scores and truancy, many studies have shown that a similar set of variables influence these outcomes. Family background is of prime importance as a determinant of test scores (Hanushek 1986, 1992). Dustmann et al. (1998) distinguish between financial and time resources allocated to the child. Financial resources enable parents to choose better schools for their child and provide a more suitable environment for studying, whereas time resources are related to the help given in explaining homework, for instance. These effects are often proxied by a wide range of parental and household variables, which also affect truancy behaviour. There are clear differences in the effect of parental occupation on test scores and truancy (Feinstein and Symons 1999; Bosworth 1994; Ermisch and Francesconi 2001; Fuchs and Woessmann 2004). Pupils with parents in professional occupations, for instance, have higher test scores and a lower probability of truanting, whereas pupils whose parents are in manual occupations are significantly more likely to be absent from school. Experience of life in a single parent family reduces test scores and increases the probability of truanting (Bosworth 1994; Ermisch and Francesconi 2001; Robertson and Symons 1996). The structure and state of the local labour market also play a part in determining test scores and truancy. For instance, McIntosh (2001) investigates the effect of labour market conditions on transitions into training and finds only a small effect, whereas expected returns to continued schooling and prior academic attainment are more important determinants.

Steele et al. (2007) investigate the effect of a school's resources on pupil test scores, and find that expenditure per pupil and the pupil-teacher ratio, which captures average class size, effect test score performance in mathematics and science. More generally, Gibbons and McNally (2013) review the evidence on the causal relationship between school resources, including class size, and test scores.

In sum, there is a considerable literature on the school-to-work transition and on the determinants of test scores, though there is less analysis of truancy behaviour and post-school outcomes. Few papers have analysed the determinants of the risk of NEET. Much of the existing literature finds that a similar set of covariates 'determine' the school-to-work transition and schooling outcomes which makes the

identification of a system of equations more challenging. However, the more recent, and smaller, literature has sought to advance the literature by estimating systems of equations, and it is in this context that the current paper should be seen.

3 The data and institutional background

3.1 Institutional background

During the period of our study, young people could leave compulsory schooling at the age of 16 and could then continue into further education, enter the labour market for work or a government sponsored training programme, or become unemployed or economically inactive. Selection for entry into jobs with training, for instance apprenticeships, or government sponsored training programmes with good prospects started before young people left school and start dates typically occurred up to September of any particular year.⁶ Similarly, young people who wished to continue onto further education also gained admission prior to leaving school and also started around September. The remaining young people who entered the labour market had to compete for available jobs, usually with relatively poor career prospects, or enter government sponsored training programmes where career development was equally uncertain. Consequently, for these young people entry to spells of unemployment were common and, in the absence, of unemployment benefits for 16–18 year olds, some would simply give up searching and become economically inactive. The inactive would often re-enter the labour market and re-commence their search for work or a ‘good’ Youth Training scheme. It is therefore important to consider this group in our analysis, as well as those registered as unemployed, who together comprise the NEET group.

3.2 The data

The data used in the following analysis have been obtained from several sources. First, pupil-level data are extracted from the Youth Cohort Study (YCS) for England and Wales, which refers to Cohorts 6–12, covering the time period 1989–90 to 2000–01. The YCS is a nationally representative, longitudinal, sample survey, typically with three waves, and respondents complete a questionnaire at each wave and the information covers the age range 15/16–18/19.⁷ The YCS contains detailed

⁶ The introduction of government sponsored training programmes for unemployed young people began in the late 1970s, and during the early 1980s included the Youth Training Scheme, which was reformed in the mid-1990s into simply Youth Training. Modern Apprenticeships were gradually rolled out over the 1990s and into the 2000s. Bell and Jones (2002) provide a full description of the history of youth training programmes in the UK.

⁷ Since YCS 6, the YCS has been a single-stage 10% random sample, conducted between 2 and 4 sweeps in the versions we use. Response rates to Sweep 1, which we use in this analysis, are between 65 and 77%. Attrition is an issue in YCS, especially between Sweeps 1 and 2 when young people have left school. Since we use Sweep 1, the effect of attrition bias should be minimised in our analysis.

Table 1 The relationship between pupil test scores and truancy

Truancy	Test scores						
	None	1–4 D-G	5+ D-G	1–4 A*–C	5–9 A*–C	10+ A*–C	All
<i>Females</i>							
Never	0.012	0.014	0.069	0.222	0.489	0.194	22,874
Odd days	0.022	0.022	0.106	0.333	0.418	0.099	11,143
Particular days	0.058	0.062	0.171	0.405	0.271	0.032	2516
Several days	0.166	0.093	0.176	0.355	0.194	0.016	808
Weeks at a time	0.376	0.118	0.160	0.244	0.097	0.005	595
Total	1028	874	3411	10,255	16,732	5636	37,936
<i>Males</i>							
Never	0.018	0.019	0.113	0.246	0.476	0.129	20,327
Odd days	0.036	0.040	0.167	0.328	0.372	0.056	8371
Particular days	0.084	0.080	0.236	0.362	0.215	0.024	1711
Several days	0.221	0.130	0.213	0.264	0.158	0.014	493
Weeks at a time	0.453	0.130	0.160	0.187	0.066	0.004	470
Total	1122	984	4284	8584	13,261	3137	31,372

information on the young person's family background, personal characteristics as well as their propensity to truant, their test scores in GCSE subjects and their destination post-school, that is, whether they are employed, unemployed, in training or further education, or whether they are economically inactive. The latter is a heterogeneous group including those young people who are caring for family members, for instance. We regard the NEET group as a joint category for the economically inactive and unemployed young people.

Second, we map additional school information not present in the YCS to each pupil, which is obtained from the School Performance Tables and the School Census, both of which were obtained from the Department for Education and Skills (DfES). The School Performance Tables contain information about the type of school, the number of pupils and the gender composition, whereas the Schools Census provides additional information on the proportion of qualified teachers, support staff hours and the proportion of pupils on free school meals. From this data, we are able to construct measures of school background and quality, as well as the pupil's peer group.

The dataset also contains information on truancy at school, test score and labour market outcomes for nearly 70,000 young people, which is a major strength of these data when compared to other survey-based datasets.

Test scores are recorded for all of the GCSE subjects that a young person studies, not all of which are eventually examined, and graded from 'non-exam/fail' to 'A*'. We combine the grade and number of GCSE subjects studied to form an ordinal scale of test scores, and our classification system has the advantage that it covers the full range of the ability distribution, including the category '5 or more GCSE grades A* to C'. At the pupil level, this is a very important threshold because performance

Table 2 Truancy behaviour and labour market outcomes

Truancy	Females			Males		
	Risk of			Risk of		
	Unemployment	NEET	<i>n</i>	Unemployment	NEET	<i>n</i>
Never	0.014	0.035	22,874	0.022	0.043	20,327
Odd days	0.028	0.062	11,143	0.044	0.070	8371
Particular days	0.072	0.126	2516	0.093	0.137	1711
Several days	0.140	0.209	808	0.136	0.176	493
Weeks at a time	0.180	0.277	595	0.255	0.328	470
Total	0.028	2147	37,936	1161	1942	31,372

The proportions of unemployed and NEET pupils are computed relative to the non-unemployed categories and the non-NEET categories

Table 3 Test scores and labour market outcomes

Test scores	Females			Males		
	Risk of			Risk of		
	Unemployment	NEET	<i>n</i>	Unemployment	NEET	<i>n</i>
None	0.211	0.330	1028	0.210	0.260	1122
1–4 D-G	0.136	0.193	874	0.149	0.189	984
5+ D-G	0.066	0.113	3411	0.071	0.103	4284
1–4 A*–C	0.034	0.068	10,255	0.039	0.068	8584
5–9 A*–C	0.008	0.029	16,732	0.009	0.029	13,261
10+ A*–C	0.001	0.014	5636	0.004	0.018	3137
Total	0.028	0.057	37,936	0.037	0.062	31,372

at this level, in addition to successful study at A Level, permits entry to University, whereas at a school level the higher the proportion achieving in this category, the more ‘successful’ the school is deemed to be. The propensity of a pupil to truant is measured on an ordinal scale ranging from ‘never truant’ to ‘truants for weeks at a time’. Table 1 shows the relationship between the frequency of truancy and test scores for males and females separately.

There is an almost monotonic increase in the level of test score performance as the frequency of truancy decreases and there appears to be a significant break in this relationship between ‘Particular days’ and ‘Several days’. In general, Table 1 does suggest a very clear negative relationship between the frequency of truancy and test scores—higher truancy is associated with lower test scores. It should be noted, however, that the number of observations in some of the cells in Table 1 is relatively small—for males, see category ‘Never’ truant and test score ‘None’, and for females truanting ‘Weeks at a time’ and test scores ‘10+ A*–C’. We address this issue in Sect. 4.

Table 4 The relationship between truancy and test scores by labour market status

Truancy	Test scores						All
	None	1–4 D-G	5+ D-G	1–4 A*–C	5–9 A*–C	10+ A*–C	
<i>Panel A: unemployed female</i>							
Never	0.073	0.085	0.241	0.384	0.198	0.018	328
Odd days	0.123	0.107	0.240	0.385	0.142	0.003	317
Particular days	0.204	0.155	0.254	0.326	0.061	0.000	181
Several days	0.434	0.159	0.142	0.230	0.035	0.000	113
Weeks at a time	0.636	0.103	0.065	0.168	0.028	0.000	107
Total	0.207	0.114	0.214	0.336	0.122	0.007	1046
<i>Panel B: NEET female</i>							
Never	0.064	0.051	0.161	0.331	0.328	0.064	807
Odd days	0.103	0.073	0.206	0.352	0.237	0.029	688
Particular days	0.164	0.123	0.217	0.371	0.116	0.009	318
Several days	0.373	0.142	0.166	0.225	0.083	0.012	169
Weeks at a time	0.612	0.091	0.091	0.170	0.030	0.006	165
Total	0.158	0.079	0.179	0.323	0.225	0.036	2147
<i>Panel C: unemployed males</i>							
Never	0.098	0.082	0.293	0.347	0.160	0.020	450
Odd days	0.143	0.134	0.282	0.337	0.099	0.006	365
Particular days	0.226	0.189	0.277	0.270	0.038	0.000	159
Several days	0.358	0.224	0.164	0.164	0.090	0.000	67
Weeks at a time	0.667	0.133	0.125	0.050	0.025	0.000	120
Total	0.203	0.127	0.263	0.292	0.106	0.010	1161
<i>Panel D: NEET males</i>							
Never	0.067	0.060	0.232	0.319	0.271	0.050	878
Odd days	0.105	0.104	0.248	0.338	0.190	0.015	589
Particular days	0.201	0.162	0.244	0.295	0.094	0.004	234
Several days	0.310	0.195	0.184	0.230	0.081	0.000	87
Weeks at a time	0.630	0.110	0.130	0.091	0.033	0.007	154
Total	0.150	0.096	0.228	0.300	0.198	0.028	1942

The risk of unemployment or NEET doubles as the level of truancy increases (see Table 2). For instance, with respect to risk of unemployment, compare truancy category ‘Particular days’, where the proportion unemployed is 7.2% for females, whereas the category ‘Several days’ is 14%. Similar effects are found for males. However, for both males and females, at higher levels of truancy the rate of increase in the risk of unemployment and NEET slows down. The risk of unemployment and NEET does differ between males and females and is almost always greater for males. For instance, for those pupils who truant for weeks at a time, the risk of unemployment is 8 percentage points higher for males, whereas for NEET it is 5 percentage points higher.

Table 3 shows a clear negative relationship between the level of test scores and the risk of unemployment and NEET. In fact, these risks fall close to zero for the very highly qualified simply because they have more options after leaving school, such as college or employment. This is not the case for the unqualified where the risk of unemployment or NEET after leaving school is between 20–33%.

In Table 4, we show the combined effect of truancy and test scores on labour market status, separately for females (Panels A and B) and males (Panels C and D). Panels A and C report the risk of unemployment, where the risks are calculated row-wise implying a direct relationship between truancy and test scores. Panels B and D show the risk of NEET. The pattern of risks in Table 4 now differs to those reported in Tables 2 and 3. For instance, the risk of unemployment for the unqualified (Test scores = ‘None’) who have never truanted is 7.3% for females and 9.8% for males, as compared with 1.4% and 2.2% in Table 2. In contrast, unqualified females who truant for weeks at a time have a risk of unemployment of 63.6% and a risk of NEET of 61.2%, which are far higher than those reported in Tables 2 and 3 for the two-way cross-tabulations. The corresponding figures for males are slightly higher—66.7% and 63%. As we move up the test score distribution, there is wider variation in terms of the risk of unemployment and NEET than is implied by estimates in Table 3, simply because of the additional effect of truancy behaviour on those risks. For instance, Table 3 shows that the average risk of unemployment for young people with 5–9 GCSE grades A*–C is around 0.9%, and 2.9% for the risk of NEET. However, Table 4 shows that for females the risk of unemployment with 5–9 GCSE grades A*–C ranges from 2.8% to 19.8%, depending on whether the young person had never truanted versus those who truanted for weeks at a time. In terms of the risk of NEET, the corresponding figures are 3% to 32.8%. Similar findings are observed for males. These findings suggest that the relationship between truancy, test scores and labour market status are complex, and there are differences in the effects of truancy and test scores on the risk of unemployment and the risk of NEET.

“Appendix A”, Table 9, contains the sample proportions for the explanatory variables used in the statistical models. The control variables used in our analysis refer to personal, family, school and cohort effects, reflecting the education production process. The analysis of the data has been carried out for males and females separately (see below for a justification).

4 Statistical methodology

4.1 The relationship between test scores, truancy and unemployment or NEET

Our literature review shows that very few studies have examined the effects of test scores and truancy on the risk of youth unemployment or NEET in a simultaneous equations framework. In this section, we discuss the possible relationships between these three variables.

Two effects of truancy on test scores can be identified. There is a *direct* effect, whereby repeated absence from school leads to the acquisition of less knowledge, culminating in lower test scores. Since we observe in the data the incidence and

duration of truancy, we can measure this effect on test scores. However, it is likely that truancy also reflects a latent, *unobservable*, negative attitude to schooling, such as a dislike of studying and of school discipline or school ethos. Moreover, whilst it is highly likely that truancy will reduce test scores, its effect on the risk of unemployment or NEET is ambiguous (see the Introduction). Truancy could be treated as a negative signal of productivity, hence increasing the risk of unemployment and inactivity, or truants dislike of school could reflect a strong desire to work or train (see Mroz and Savage 2007), leading to increased search effort and hence a lower of risk of unemployment or NEET. Nevertheless, truancy could still affect the risk of unemployment and NEET *indirectly* via its effect on test scores. As the literature review shows, the effect of lower test scores on the risk of unemployment is well documented, but less so with respect to the risk of NEET. Our modelling strategy attempts to identify these direct, indirect and unobserved effects on the risk of unemployment and NEET.

The general specification of the model with latent variables $(Y_{ti}^*, Y_{ei}^*, Y_{ni}^*)$ is as follows:

$$Y_{ti}^* = \beta'_{1t}(\mathbf{Family}_i, \mathbf{Personal}_i, \mathbf{School}_i, \mathbf{Place}_i) + \epsilon_{ti} = \eta_{ti} + \epsilon_{ti} \tag{1}$$

$$Y_{ei}^* = \beta'_{1e}(\mathbf{Family}_i, \mathbf{Personal}_i, \mathbf{School}_i, \mathbf{Place}_i) + \sum_{k=2}^{k=3} \gamma_{ek} Y_{ti}^k + \epsilon_{ei} = \eta_{ei} + \sum_{k=2}^{k=3} \gamma_{ek} Y_{ti}^k + \epsilon_{ei}, \tag{2}$$

$$\begin{aligned} Y_{ni}^* &= \beta'_{1n}(\mathbf{Family}_i, \mathbf{Personal}_i, \mathbf{School}_i, \mathbf{Place}_i) + \sum_{k=2}^{k=3} \gamma_{nk} Y_{ti}^k + \sum_{l=2}^{l=4} \theta_{nl} Y_{ei}^l + \epsilon_{ni} \\ &= \eta_{ni} + \sum_{k=2}^{k=3} \gamma_{nk} Y_{ti}^k + \sum_{l=2}^{l=4} \theta_{nl} Y_{ei}^l + \epsilon_{ni} \end{aligned} \tag{3}$$

where Y_{ti}^* refers to a pupils latent propensity to truant, Y_{ei}^* is the latent test score performance of the young person, and Y_{ni}^* represents the latent probability of becoming unemployed or entering NEET on completion of schooling. The $(\epsilon_{ti}, \epsilon_{ei}, \epsilon_{ni})$ are from a trivariate standard normal distribution with correlation matrix Σ_{tel} , implying that the observed responses (Y_{ti}, Y_{ei}, Y_{ni}) are from a trivariate-ordered probit model. We have used η_{zi} to represent the linear predictors, ($z = t, e, n$) of the exogenous covariates $(\mathbf{Family}_i, \mathbf{Personal}_i, \mathbf{School}_i, \mathbf{Place}_i)$. The exogenous covariates in the model represent a standard set of variables included in many specifications of the so-called education production function (see “Appendix A” for variables and descriptive statistics). The linear predictors do not contain constants as these are not identified.

4.2 The trivariate model

The YCS data contain 5 levels of school truancy (Y_t) at age 16 for student i at school, as follows:

Response Y_{ti}	Description
1	Never truant
2	Odd days
3	Particular days
4	Several days
5	Weeks at a time

Truancy from school is self-reported. The ordered nature of the observed Y_{ti} suggests we treat it as an ordered response with 5 categories. In addition to modelling truancy as a response, we also treat truancy as an endogenous variable in the linear predictors for the models for test scores at age 16 (Y_{ei}) and subsequent unemployment and NEET (Y_{ni}) models. To simplify the joint estimation of the endogenous truancy effects and the correlations in the random effects, we use a reduced number of endogenous dummy variables in the linear predictors for Y_{ei} and Y_{ni} . Specifically, we combine categories 2 and 3 and also categories 4 and 5, with Y_{ti}^1 taken as the reference category. Given the relatively low number of observations in some categories (see Sect. 3), this should ensure that we obtain more precise estimates. Note, we do not collapse the categories for the response variable because these are representations of our underlying latent variable—the propensity to truant.

In the UK, a pupil's performance at school is typically measured by the level of attainment in public examinations. In this paper, test scores refer to the number of, and grade in, the General Certificate of Secondary Education (GCSE), obtained from the YCS, which is classified into one of six levels l of educational attainment at age 16 (Y_e), these are as follows⁸:

Response Y_{ei}	Description
1	no GCSEs
2	1–4 D-G GCSEs
3	5+ D-G GCSEs
4	1–4 A*–C GCSEs
5	5–9 A*–C GCSEs
6	10+ A*–C GCSEs

The nature of Y_{ei} suggests that we treat it as an ordered response with six categories. Again, to simplify the joint estimation of the endogenous test score effects and correlation in the random effects of the other responses, and to ensure that we obtain more precise estimates, we also use a reduced number of dummy variables for GCSE scores in the linear predictors for Y_{ni} . Specifically, for Y_{ei} we combine categories 1 and 2 together, which becomes Y_{ei}^1 , and categories 5 and 6 together, where

⁸ The category 'no GCSEs' requires explanation. This category refers to those pupils who did not take any examinations, or they did so but achieved fail grades in all subjects.

Y_{ei}^1 is treated as the reference category. As for truancy, we do not collapse the categories of the response variable.

In this analysis, we will use two levels of response for post 16 labour market outcomes (Y_n) at age 16 for individual i , as follows:

Response Y_{ni}	Description
1	Education, employment or training
2	Unemployed or NEET

The nature of Y_{ni} means we can treat it as an ordered response with just 2 categories (binary). Clearly, we do not treat labour market outcomes as an endogenous variable in the models for truancy and educational attainment.

There are various joint models that can be used for trivariate-ordered responses, the most widely used assumes that observed responses (Y_{ii}, Y_{ei}, Y_{ni}) are obtained from underlying normally distributed variables ($Y_{ii}^*, Y_{ei}^*, Y_{ni}^*$). The continuous latent variables, e.g. Y_{ii}^* , are observed in one of the (in this case $K = 5$) categories through a censoring mechanism, that is:

$$\begin{aligned}
 Y_{ii} = 1 & \quad \text{if} \quad c_{t0} < Y_{ii}^* \leq c_{t1} \\
 Y_{ii} = 2 & \quad \text{if} \quad c_{t1} < Y_{ii}^* \leq c_{t2} \\
 Y_{ii} = 3 & \quad \text{if} \quad c_{t2} < Y_{ii}^* \leq c_{t3} \\
 Y_{ii} = 4 & \quad \text{if} \quad c_{t3} < Y_{ii}^* \leq c_{t4} \\
 Y_{ii} = 5 & \quad \text{if} \quad c_{t4} < Y_{ii}^* \leq c_{t5}
 \end{aligned}$$

where the c_{ik} , $k = 1, \dots, 5$ are finite cut points or thresholds of the latent variable Y_{ii}^* , with $c_{t0} = -\infty$, and $c_{t5} = \infty$. In this paper, we assume that the cut points (c_{ik}, c_{ei}, c_{nm}) do not vary across individuals (i). Ordered responses based on latent variables can be given a utility maximisation interpretation, see Bhat and Pulugurta (1998).

As suggested earlier, there are likely to be unobserved effects that determine truancy, test scores and the transition from school, such as attitudes to school discipline and ethos, summarised as ‘tastes for schooling’ and motivation. These unobserved effects may bias the estimates of the variables of interest—truancy and test scores. Therefore, to disentangle the observable direct and indirect effects from the unobservable effect requires the simultaneous estimation of Eqs. 1–3 where test scores and truancy are treated as endogenous variables in our models of the risk of unemployment or NEET. This is shown in Eq. 4.

The probabilities of the observed responses for truancy, Y_{ii} , test scores, Y_{ei} and labour market outcomes, Y_{ni} , are given by a triple integral which does not have a closed form, so for example if $Y_{ii} = 2$, $Y_{ei} = 3$, $Y_{ni} = 2$ then this individual’s contribution to the likelihood is given by:

$$L_i = \Pr [Y_{ii} = 2, Y_{ei} = 3, Y_{ni} = 2] = \int_{c_{11}-\eta_{i1}}^{c_{12}-\eta_{i1}} \int_{c_{e2}-\eta_{ei}-\gamma_{e2}}^{c_{e3}-\eta_{ei}-\gamma_{e2}} \int_{c_{n1}-\eta_{ni}-\gamma_{n2}-\theta_{n3}}^{c_{n2}-\eta_{ni}-\gamma_{n2}-\theta_{n3}} \phi(\epsilon_t, \epsilon_e, \epsilon_n; \Sigma_{ten}) d\epsilon_n d\epsilon_e d\epsilon_t.$$

where $\phi(\epsilon_t, \epsilon_e, \epsilon_n; \Sigma_{ten})$ is a trivariate standard normal density function with the 3×3 correlation matrix Σ_{ten} . Note that γ_{ek} and γ_{nk} refer to the direct effects of truancy on test scores and the direct effect of truancy of labour market outcomes, respectively. θ_{ne} refers to the direct effect of test scores on labour market outcomes. The log likelihood for all individuals is then

$$\log L = \sum_{i=1} \log L_i \tag{4}$$

The log likelihood is maximised to provide the parameter estimates using CMP in Stata 14 (Roodman 2011). To evaluate the three-dimensional cumulative normal distributions, CMP uses simulated likelihood methods, specifically the Mata function, that is, the Geweke–Hajivassiliou–Keane algorithm (*ghk2()*) with Haldane sequences (Geweke 1989, Hajivassiliou and McFadden 1998). For instance, for Heterogeneous Model 1 for the NEET and unemployment outcomes, this involves 390 draws per observation for females, and 355 draws per observation for males. We let CMP decide on the number of draws to be used. As a check on the adequacy of the simulated likelihood approach, we also evaluated the log-likelihoods at the solutions using the NAG Fortran Library.⁹

4.3 Alternative specifications

Equations 1–3 are estimated several times, each with a slightly different specification, and we refer to these as models. In the Homogenous Model, we actually estimate Eqs. 1–3 separately which is consistent with much of the existing literature. We then estimate several different versions of Eq. 4. In model 1, we allow for correlation between the random effects of each sub-model and include all direct effects of endogenous variables. In model 2, we drop the direct effect of truancy on youth unemployment (and NEET), which means that the impact of truancy behaviour at school on labour market outcomes is picked up via its effect on test scores (the indirect effect) and through the unobserved effects. Model 3 drops the unobserved effect in the truancy equation. By dropping the direct effect of truancy on unemployment (and NEET) and the correlation between Eqs. 1 and 3, we can determine whether test scores play a more important role than truancy. Model 4 takes a different approach. In this model, we re-introduce the direct effect of truancy in Eq. 3 and the correlation between the unobservables in Eqs. 1 and 3; however, we explore the

⁹ See The Numerical Algorithms Group (NAG), Oxford, United Kingdom www.nag.com. Specifically, we used routine G01HBF, which calls D01AJF as we have a three dimensional integral. D01AJF adaptively uses Gauss 10 point and Kronrod 21 point rules for the outer integral, the two inner integrals are evaluated by routine G01HAF, a method described by Divgi (1979). We found that the total log likelihoods, for males and females, obtained by CMP and NAG differed only in the first decimal place.

possibility of reverse causation between test scores and truancy. In this case, Y_e is inserted in Eq. 1 and Y_t is dropped from Eq. 2.

Reverse causality could arise if pupils who systematically fail at school eventually reduce effort and start to truant. This is plausible given the ‘teaching to test’ that has arisen since the introduction of competition between schools and School Performance Tables following the 1988 Education Reform Act. However, we argue that this reverse causality should be less of an issue in our data for two reasons. First, our measure of test score is a summative statement of performance measured primarily at the end of compulsory schooling at age 16 when pupils sit for their GCSE examinations, whereas our measure of truancy refers to behaviour between the ages of 14 and 16. Second, it is more likely that poor performance in coursework could increase the incidence of truancy because this does contribute to the final GCSE grades. But, performance in tests in GCSE subjects is still weighted heavily and this implies that truancy behaviour will therefore affect overall performance in GCSE examinations at age 16. Nevertheless, we do investigate the issue of the direction of causation between Y_{ti} and Y_{ei} in our modelling.

As we show below, our preferred model is Heterogenous Model 1.

4.4 Identification

Identifiability in structural equation models with discrete outcomes has been widely discussed. Wilde (2000) shows that the existence of one varying exogenous regressor in each equation is sufficient to avoid small variation identification problems in multiple equation probit models with endogenous dummy regressors. Thus, in our models we include the actual measures of truancy, Y_{ti}^k , albeit in collapsed form for reasons explained above, for our model of educational attainment Y_{ei}^* and early labour market behaviour Y_{ni}^* , and similarly with the effect of dummy endogenous variables for the actual test scores or GCSE levels, Y_{ei}^l , in our models for early labour market behaviour Y_{ni}^* .¹⁰ Wilde also notes that his result applies to other distributions besides the normal (multivariate probit), but these models have some additional constants in them. Our approach implies that we rule out reverse causality between Y_{ti}^k and Y_{ei}^l from the outset; however, there are two reasons why this is not the case and hence why it is important to explore the issue of reverse causality in these data. First, we can regard Y_{ti}^k and Y_{ei}^l as joint, or mutually dependent, measures of educational experience and outcomes. The GCSE outcome is the final result of studying between the ages of 14–16, in the same way that truancy reflects behaviour which translates into an ‘educational experience’ between the ages of 14–16. Ideally, we would like to allow for reciprocal dependence, rather than recursive dependence, as in our models. However, reciprocal multivariate probit models with dummy endogenous variables cannot be identified whatever

¹⁰ This contrasts with the approach taken by Buscha and Conte (2013) who include a latent variable for truancy in their test score model, but crucially in our view, do not include the actual values of truancy outcomes. There is some ambiguity in the literature, e.g. Roodman (2011), *Stata Journal*, 11, number 2, pp 159–206, notes ‘Wilde (2000) shows that a general (recursive) multi-equation probit model is identified as long as each equation contains one varying predetermined variable. Despite the theoretical results, identification might still be more robust if exclusion restrictions were imposed—that is, if the classical order condition, though theoretically unnecessary, were still met’.

instruments one might use (Heckman 1978). As a compromise, we estimate alternative models in which either Y_{it}^k or Y_{ei}^l are treated as predetermined. However, our prior is that Y_{ei}^l occurs at the end of the education production process, whereas Y_{it}^k is an intermediate part of that process. Second, although we do not base our identification strategy on this, there are measures of school-level and neighbourhood-level truancy in our model for Y_{it}^k which are excluded from our model for Y_{ei}^l .

In sum, our approach to identification is analogous to that of Elbers and Ridder (1982) and Heckman and Singer (1984) for heterogeneity in duration data. Other papers in the education economics literature that adopt the same identification strategies as ours are Evans and Schwab (1995) and Neal (1997), whereas Goldman et al (2001) is an example from the statistics literature.

4.5 Measuring the average treatment effect on the treated

It is not straightforward to estimate marginal effects holding everything else constant for the endogenous covariates in our modules, because of the correlation between responses. To aid the interpretation of the endogenous effects, we compute the average treatment effect on the treated (ATT). In our case, the levels of test scores (Y_{it}^l) and truancy (Y_{it}^k) are different treatment effects for unemployment and NEET (Y_{ni}). To obtain the treatment effects for the *expanded levels of truancy and test scores*, we use the joint model for the various observable treatments (Y_{it}, Y_{ei}), and the unobservable counterfactual treatments for the same unemployment response. This is given by setting the parameters for the endogenous effects (γ_{nk}, θ_{nl}) to zero. For our example, with $Y_{it} = 2, Y_{ei} = 3, Y_{ni} = 2$ we have:

$$\begin{aligned} & \Pr [Y_{it} = 2, Y_{ei} = 3, Y_{ni} = 2 | \gamma_{n2} = 0, \theta_{n3} = 0] \\ &= \int_{c_{11}-\eta_t}^{c_{12}-\eta_t} \int_{c_{e2}-\eta_e-\gamma_{e2}}^{c_{e3}-\eta_e-\gamma_{e2}} \int_{c_{n1}-\eta_n}^{c_{n2}-\eta_n} \phi(\epsilon_t, \epsilon_e, \epsilon_n; \Sigma_{ten}) d\epsilon_n d\epsilon_e d\epsilon_t. \end{aligned}$$

The joint probability of the (Y_{it}, Y_{ei}) treatment is:

$$\Pr [Y_{it} = 2, Y_{ei} = 3] = \int_{c_{11}-\eta_t}^{c_{12}-\eta_t} \int_{c_{e2}-\eta_e-\gamma_{e2}}^{c_{e3}-\eta_e-\gamma_{e2}} \phi(\epsilon_t, \epsilon_e; \Sigma_{te}) d\epsilon_e d\epsilon_t.$$

where Σ_{te} is the 2×2 correlation matrix for (ϵ_t, ϵ_e) . The treatment effect on the treated, i.e. when $Y_{it} = 2$, and $Y_{ei} = 3$, for individual i is

$$TT_{23i} = \frac{\Pr [Y_{it} = 2, Y_{ei} = 3, Y_{ni} = 2] - \Pr [Y_{it} = 2, Y_{ei} = 3, Y_{ni} = 2 | \gamma_{n2} = 0, \theta_{n3} = 0]}{\Pr [Y_{it} = 2, Y_{ei} = 3]}$$

This estimate of the treatment effect varies by young person, i , because the exogenous covariates vary with i . The sample average of the treatment effects (e.g. when $Y_{it} = 2$, and $Y_{ei} = 3$) gives the average treatment effect for unemployment or NEET

(Y_n) on the treated (in this example ATT_{23}). The reference groups for the endogenous dummy variables are Y_{it}^1, Y_{ei}^1 and Y_{ei}^2 , so $ATT_{11} = ATT_{12} = 0$. The reference categories act as a control group.

5 Econometric results

We begin our discussion of the results by first discussing our preferred models and the reasons for this selection. This will allow us to focus our discussion of the results on particular models, whilst also comparing the findings from our preferred models with other models. Table 10 in “Appendix B” compares the various models and shows the results of a number of likelihood ratio tests. Based on the likelihood ratio tests, it is clear that Heterogeneous Model 1 substantially outperforms the Homogenous Model for males and females, and for the unemployment and NEET outcomes. This is a significant finding in the sense that much of the existing literature has ignored the simultaneous nature of the relationship between truancy, test scores and early labour market outcomes. However, Table 10 also compares each of the other heterogenous models with Heterogeneous Model 1, and the model which comes closest is Heterogeneous Model 2. In this model, we drop the direct effect of truancy on unemployment and NEET; however, the likelihood ratio test shows that Model 2 is still rejected against Model 1, albeit marginally for the male unemployment model and the female NEET models. Models 3 and 4 are easily rejected when compared with Heterogeneous Model 1, and so we focus our discussion of the results on this model drawing a particular comparison with Model 2, where appropriate.

Table 11 in “Appendix B” tests for differences between the male and female models. It is clear from the likelihood ratio tests that there are statistical differences between the male and female models in all cases, but the combined Heterogenous Model 1 still outperforms the combined Homogenous Model. Together with the well-established behavioural differences between male and female students vis-a-vis test score performance (see Andrews et al. 2004), it is clearly important to also perform this analysis separately for males and females.

5.1 The determinants of test scores

The main focus of this paper is on the effects of test scores and truancy on the probability of unemployment or NEET several months after leaving school. However, given that we estimate a system of equations it is important to briefly assess the sub-model for test scores.

Table 12, “Appendix B”, reports the effects of truancy on test scores for males/females and by NEET/unemployment outcomes. This Table shows that the endogenous truancy indicators on test scores change from negative to positive when we allow for a correlation in the unobservables of truancy and test scores, that is, in models 1, 2 and 3. This feature of these models may seem counter-intuitive; however, the correlation between the unobservable components of the latent variables

Table 5 Unemployment Ordered Probit Model Results

Var.	Homogenous model			Heterogeneous Model 1			Heterogeneous Model 2			Heterogeneous Model 3			Heterogeneous Model 4		
	Est.	s.e.	p value	Est.	s.e.	p value	Est.	s.e.	p value	Est.	s.e.	p value	Est.	s.e.	p value
<i>Panel A: unemployment females</i>															
Yt23	0.241	0.033	0.000	-0.028	0.080	0.727							-0.046	0.082	0.574
Yt45	0.641	0.053	0.000	0.123	0.155	0.424							0.080	0.157	0.610
Ye3	-0.420	0.050	0.000	-0.396	0.061	0.000	-0.444	0.058	0.000	-0.663	0.049	0.000	-0.446	0.063	0.000
Ye4	-0.702	0.045	0.000	-0.657	0.084	0.000	-0.737	0.077	0.000	-1.137	0.053	0.000	-0.687	0.085	0.000
Ye56	-1.268	0.052	0.000	-1.178	0.137	0.000	-1.300	0.126	0.000	-2.047	0.083	0.000	-1.219	0.138	0.000
Rho_te				-0.512	0.017		-0.513	0.017		-0.514	0.017		-0.262	0.014	
Rho_in				0.194	0.054		0.202	0.020		0.000	-		0.202	0.052	
Rho_en				-0.099	0.048		-0.071	0.041		0.207	0.024		-0.071	0.045	
LogL	-3902.2			-87,555.6			-87,559.0			-87,607.5			-87,699.5		
n	37,936			37,936			37,936			37,936			37,936		
<i>Panel B: unemployment males</i>															
Yt23	0.221	0.032	0.000	-0.101	0.083	0.221							-0.132	0.085	0.119
Yt45	0.597	0.056	0.000	-0.043	0.166	0.795							-0.091	0.167	0.584
Ye3	-0.414	0.046	0.000	-0.420	0.060	0.000	-0.461	0.057	0.000	-0.674	0.046	0.000	-0.485	0.063	0.000
Ye4	-0.669	0.044	0.000	-0.680	0.088	0.000	-0.747	0.080	0.000	-1.126	0.055	0.000	-0.725	0.088	0.000
Ye56	-1.221	0.051	0.000	-1.223	0.140	0.000	-1.326	0.130	0.000	-2.000	0.085	0.000	-1.282	0.141	0.000
Rho_te				-0.465	0.022		-0.466	0.021		-0.470	0.021		-0.253	0.019	
Rho_in				0.224	0.056		0.178	0.021		0.000	-		0.239	0.055	
Rho_en				-0.071	0.048		-0.028	0.043		0.223	0.027		-0.041	0.046	
LogL	-4157.5			-73,030.7			-73,033.3			-73,070.0			-73,115.7		
n	31,372			31,372			31,372			31,372			31,372		

is large and negative, and for Heterogenous Model 1 the correlation ranges from -0.464 to -0.512 for the male/female—unemployment/NEET models (see Table 5). This implies that the overall relationship, combining the direct and correlated effects between truancy and test scores will still be negative in our preferred model. Stated differently, the positive direct effects of the endogenous truancy indicators are not large enough to dominate the large negative correlation in the unobservables for our preferred model, and indeed for all heterogenous models.

It is worth noting that in model 4, we include our test score variable to investigate whether test score performance affects truancy, i.e. causality runs in the opposite direction. In general, our estimates (not reported) are statistically significant and suggest that pupils who (ultimately) achieve higher test scores are less likely to truant. This implies that causation could run from test scores to truancy, however, as suggested in Table 10, “Appendix B”, this model does not fit the data as well as Heterogenous Model 1, hence we reject model 4.

5.2 The direct effects of test scores and truancy on the risk of unemployment and NEET

Tables 5 and 6 show the estimated effects of truancy and test scores on the probability of a young person becoming unemployed (Table 5) or entering the NEET category (Table 6) several months after leaving school. The estimated effects for the homogenous model are fairly standard findings in the cross-sectional literature and so are a good place to start the discussion of our findings.

In the Homogenous Model, higher levels of truancy increase the probability of unemployment and NEET for both males and females. Conversely, the higher the pupils test scores the lower the likelihood of unemployment and NEET, which is consistent with the view that these pupils have more choices after leaving compulsory schooling insofar as they can continue their education, enter a training programme or get a job. These effects can be regarded as direct effects of truancy and test scores on labour market outcomes. However, for the Homogenous Model, truancy also reduces test scores (not shown), and so there is also an additional indirect effect of truancy on the probability of a young people becoming unemployed or NEET. The total effect of truancy on labour market outcomes would therefore be underestimated by simply looking at its direct effect.

Of course these homogenous models do not take account of the effect of unobservables, which imply correlations in the latent variables of truancy, GCSE, and unemployment and NEET. Tables 5 and 6 also report the estimated direct effects of truancy and test scores on youth unemployment and NEET for the various heterogeneous models. Recall, our preferred model is Heterogeneous Model 1. For this model, the direct effect of test scores remains negative and statistically significant, but for females in particular, the estimates are smaller than the equivalent estimates from the homogenous model. Allowing for the correlations in the unobservables of models 1–4, (Σ_{ten}) , may therefore produce more precise estimates of the direct effect of test scores on the risk of entering unemployment or NEET on leaving school. The

Table 6 NEET-ordered probit model results

Var.	Homogenous			Heterogeneous Model 1			Heterogeneous Model 2			Heterogeneous Model 3			Heterogeneous Model 4		
	Est.	s.e.	p value	Est.	s.e.	p value	Est.	s.e.	p value	Est.	s.e.	p value	Est.	s.e.	p value
<i>Panel A: NEET females</i>															
Yt23	0.248	0.024	0.000	-0.025	0.064	0.692							-0.040	0.066	0.545
Yt45	0.639	0.045	0.000	0.135	0.127	0.287							0.079	0.129	0.540
Ye3	-0.467	0.044	0.000	-0.418	0.049	0.000	-0.461	0.047	0.000	-0.665	0.043	0.000	-0.467	0.051	0.000
Ye4	-0.717	0.039	0.000	-0.624	0.061	0.000	-0.692	0.057	0.000	-1.054	0.045	0.000	-0.653	0.061	0.000
Ye56	-1.089	0.041	0.000	-0.914	0.093	0.000	-1.013	0.087	0.000	-1.679	0.065	0.000	-0.953	0.094	0.000
Rho_te				-0.511	0.017		-0.512	0.017		-0.517	0.017		-0.261	0.014	
Rho_in				0.198	0.042		0.207	0.015		0.000	-		0.206	0.042	
Rho_en				-0.127	0.032		-0.107	0.027		0.141	0.019		-0.102	0.030	
LogL	-7236.0			-90,884.2			-90,889.8			-90,988.2			-91,027.1		
n	37,936			37,936			37,936			37,936			37,936		
<i>Panel B: NEET males</i>															
Yt23	0.179	0.025	0.000	-0.201	0.069	0.004							-0.235	0.071	0.001
Yt45	0.579	0.051	0.000	-0.177	0.142	0.213							-0.230	0.142	0.106
Ye3	-0.371	0.042	0.000	-0.382	0.050	0.000	-0.427	0.048	0.000	-0.598	0.042	0.000	-0.453	0.052	0.000
Ye4	-0.570	0.039	0.000	-0.590	0.066	0.000	-0.659	0.062	0.000	-0.955	0.049	0.000	-0.635	0.067	0.000
Ye56	-0.923	0.041	0.000	-0.940	0.101	0.000	-1.043	0.096	0.000	-1.562	0.071	0.000	-0.995	0.102	0.000
Rho_te				-0.464	0.022		-0.468	0.022		-0.475	0.021		-0.251	0.019	
Rho_in				0.261	0.046		0.158	0.016		0.000	-		0.280	0.046	
Rho_en				-0.079	0.035		-0.019	0.031		0.175	0.022		-0.049	0.033	
LogL	-6552.8			-75,419.0			-75,427.6			-75,475.1			-75,503.4		
n	31,372			31,372			31,372			31,372			31,372		

estimated effects for test scores in Heterogenous Model 1 do also differ in the unemployment and NEET models.

Turning to the direct effect of truancy, we observe that for males in both the unemployment and NEET models the effects are negative and generally insignificant (see Heterogeneous Model 1). For females at least one of the estimated effects of truancy, Y_{145} , is positive but is statistically insignificant. Interestingly, for the lower levels of truancy, Y_{123} , the effects are negative and statistically significant for females in the case of unemployment, which is also the case for males with respect to the risk of NEET. Thus, the effects for truancy are more mixed when compared with the test score results. A more parsimonious model may therefore be one that drops the direct effects of truancy (see Heterogeneous Model 2), which implies that the effect of truancy only works through test scores (an indirect effect) and through the correlations (the unobservables). There is little evidence for this indirect effect in Heterogenous Model 2 for females, where the correlations for Rho_tn are almost identical when compared to those for Heterogeneous Model 1, whereas in the case of males the size of the correlations actually fall. Similarly, the size of the estimates for truancy in the test score models (see Table 12) are almost identical in Heterogenous Models 1 and 2, suggesting that the indirect effects of truancy on labour market outcomes are unchanged. In view of this, and given our likelihood ratio tests where we reject Model 2, we argue that the direct truancy effects are at least jointly significant and should be retained.

Tables 5 and 6 also report the correlations between the unobservables in the various branches of the model, which pick up the effect of unobservable differences between students, e.g. differences in motivation. We compare the correlations for the unemployment outcome for females and males, respectively, in Table 5. Table 6 shows the equivalent results for NEET. What is clear for Heterogenous Model 1 is that, when one compares the estimates of the correlations for the unemployment (Table 5) and the NEET (Table 6) outcomes for each gender, there are only small differences.

In both Tables 5 and 6 for Model 1, there is a negative and statistically significant correlation between the unobserved effects on truancy and test score sub-models (see Rho_{te}). Pupils who are unobservably more likely to truant, perhaps because they are demotivated by school, are also unobservably less able and so their test scores are lower. (This effect is almost identical in terms of magnitude for models 2 and 3.) With regard to the correlations between unobservables for the truancy and unemployment and NEET models (see Rho_m), the estimates for Heterogeneous Model 1 are positive and statistically significant, suggesting that students who are unobservably more likely to truant are more likely to become unemployed or NEET. A lack of motivation at school translates into poor entry into the job market, possibly because of poor motivation to find a job or training place, or because employers are able to screen out such youngsters during the selection process. (Again this result is consistent—see models 2 and 4.) Finally, we consider the correlations between the unobservables for examination scores and the probability of unemployment or NEET (see Rho_{en}). There is some variation in the estimated correlations; however, the correlations between the unobserved effects are negative and statistically significant.

Table 7 Estimated ATTS for Model 1, unemployment

Truancy	Test scores						Total
	None	1–4 D-G	5+ D-G	1–4 A*–C	5–9 A*–C	10+ A*–C	
<i>Panel A: females</i>							
Never	0	0	–0.054	–0.069	–0.071	–0.059	–0.065
	279	325	1567	5088	11,176	4439	22,874
Odd days	0	0	–0.069	–0.092	–0.107	–0.100	–0.093
	244	247	1176	3716	4659	1101	11,143
Particular days	0	0	–0.082	–0.115	–0.137	–0.124	–0.102
	147	157	431	1019	682	80	2516
Several days	0	0	–0.101	–0.153	–0.193	–0.180	–0.112
	134	75	142	287	157	13	808
Weeks at a time	0	0	–0.113	–0.166	–0.224	–0.236	–0.082
	224	70	95	145	58	3	595
Total	0	0	–0.066	–0.085	–0.085	–0.068	–0.077
	1028	874	3411	10,255	16,732	5636	37,936
<i>Panel B: males</i>							
Never	0	0	–0.065	–0.086	–0.098	–0.089	–0.086
	357	383	2306	4998	9666	2617	20,327
Odd days	0	0	–0.081	–0.111	–0.138	–0.143	–0.109
	300	339	1395	2749	3118	470	8371
Particular days	0	0	–0.099	–0.138	–0.190	–0.193	–0.119
	143	137	403	619	368	41	1711
Several days	0	0	–0.119	–0.174	–0.238	–0.252	–0.112
	109	64	105	130	78	7	493
Weeks at a time	0	0	–0.130	–0.192	–0.271	–0.278	–0.076
	213	61	75	88	31	2	470
Total	0	0	–0.076	–0.100	–0.111	–0.099	–0.095
	1122	984	4284	8584	13,261	3137	31,372

Unobservably more able students are less likely to become unemployed or economically inactive.

In this analysis, unobservable effects do matter, and in terms of the magnitudes of the effects that we estimate, it is Rho_m that has the largest effect. Disentangling these effects allows us to tell a richer story of the relationships between truancy behaviour and test score performance at school on the one hand, and their impact on labour market outcomes on the other.

5.3 Calculating the magnitude of the main effects of truancy and test scores

We turn now to the estimation of the ATTS for our preferred model—Heterogeneous Model 1—in order to gauge how important, from a quantitative perspective, the effect of truancy and test scores are for the risk of unemployment and NEET. Our

Table 8 Estimated ATTS for Model 1, NEET

Truancy	Test scores						Total
	None	1-4 D-G	5+ D-G	1-4 A*-C	5-9 A*-C	10+ A*-C	
<i>Panel A: Females</i>							
Never	0	0	-0.086	-0.105	-0.111	-0.092	-0.101
	279	325	1567	5088	11,176	4439	22,874
Odd days	0	0	-0.103	-0.131	-0.150	-0.140	-0.131
	244	247	1176	3716	4659	1101	11,143
Particular days	0	0	-0.118	-0.155	-0.185	-0.175	-0.139
	147	157	431	1019	682	80	2516
Several days	0	0	-0.141	-0.186	-0.234	-0.216	-0.140
	134	75	142	287	157	13	808
Weeks at a time	0	0	-0.148	-0.203	-0.259	-0.249	-0.100
	224	70	95	145	58	3	595
Total	0	0	-0.100	-0.123	-0.127	-0.103	-0.113
	1028	874	3411	10,255	16,732	5636	37,936
<i>Panel B: Males</i>							
Never	0	0	-0.078	-0.105	-0.126	-0.114	-0.109
	357	383	2306	4998	9666	2617	20,327
Odd days	0	0	-0.089	-0.124	-0.162	-0.167	-0.125
	300	339	1395	2749	3118	470	8371
Particular days	0	0	-0.110	-0.156	-0.217	-0.216	-0.134
	143	137	403	619	368	41	1711
Several days	0	0	-0.124	-0.181	-0.252	-0.278	-0.118
	109	64	105	130	78	7	493
Weeks at a time	0	0	-0.136	-0.198	-0.278	-0.308	-0.078
	213	61	75	88	31	2	470
Total	0	0	-0.087	-0.117	-0.138	-0.124	-0.114
	1122	984	4284	8584	13,261	3137	31,372

approach allows us to calculate the ATTs for the observed levels of truancy (5 levels) and test scores (6 levels). For example, the ATT for the unemployed is the probability of a flow into unemployment and is obtained by estimating a model where the test score variable is initially nonzero. We then estimate a model where the test score variable is set to zero, and the ATT is the conditional difference between the two. This is repeated for the NEET category. The ATTs are calculated for the categories of the response variables, they reflect the direct, indirect and unobservable effects of the endogenous effects on these categories.

Tables 7 and 8 report the estimated ATTs for the unemployment and NEET models, for males and females separately. There are differences in the ATT effects between these groups, however, in all cases the effects of test scores and truancy are negative. The effects can be interpreted as follows: for a particular level of test score and truancy, and when compared with the control group, the negative effect is to

reduce the risk of unemployment when compared to the base category. This implies that the effect of test score dominates the effect of truancy, that is, we can think of test score performance as compensating for poor attendance. This is best seen by looking at low levels of truancy (e.g. 'Odd days'). For females, Table 7, Panel A reports the ATTs for unemployment, and the compensating effect of test scores is modest, as expected, because truancy at this level is not so much of a problem (e.g. the ATT for test scores = '5+D-G' is -0.069 versus test scores = '10+ A*-C' is -0.100), whereas for truancy = 'Weeks at a time' and test scores = '5+ D-G' the ATT is -0.113 versus an ATT of -0.236 where test scores = '10+ A*-C'.

The ATT effects for unemployed males (Table 7, Panel B) are higher than those for females, whereas for NEET models (Table 8) the ATTs are similar in magnitude for males and females. Note, however, that truancy also impacts indirectly via its effect on test scores and through the correlation in the unobservables. Nevertheless, it is still the case that the test score effect dominates the truancy effect in terms of labour market outcomes. Of course, this is not to deny that reducing truancy is important; reducing truancy is likely to improve test scores which in turn improves labour market prospects.

6 Conclusion

In this paper, we investigate the effects of test scores and truancy behaviour on the labour market outcomes of teenagers in England and Wales. We also investigate the interdependencies, and the direction of causation, between truancy behaviour and test score performance. This is because it may be that truancy has a direct effect on the risk of unemployment or NEET amongst young people as well as an indirect effect via the effect of truancy on test scores. Truancy could be regarded as a proxy for effort at school, and performance in tests, affects the subsequent transition from school, then all three behavioural outcomes (decisions) are jointly determined. Consequently, to capture this joint nature, a three-equation model was estimated in which we allow for correlation in the unobservables between the latent variable models for truancy, test scores, for the ordered categorical variables, and similarly for unemployment (or NEET), which is a binary variable. Several models are estimated which allow for different specifications of the relationships between the three outcome measures. To estimate our models, we use detailed pupil and school-level data from the Youth Cohort Studies (YCS), specifically YCS6 to YCS12, as well as school performance and school census data, which cover the period of the late 1990s and early 2000s.

The findings from our preferred model (Heterogeneous Model 1) suggest that truancy works through test scores (i.e. an indirect effect) with only a weak direct effect on labour market outcomes. However, truancy also has an unobserved effect on the risk of unemployment and the risk of NEET insofar as the correlation between latent variables for truancy and labour market outcomes are positive and statistically significant. Test scores have a direct effect on labour market outcomes, and through the estimation of ATTs, we show a good performance in high stakes tests (i.e. GCSEs) can mitigate the effect of truanting from school on labour market outcomes. Truancy

need not be a significant problem for young people in terms of their post-school outcomes so long as this behaviour does not reduce test score performance. This makes sense insofar as employers observe test score performance in the job/training selection process whereas they are less likely to observe truancy behaviour.

The popular view that truancy is universally bad for young people is open to question according to our findings. The story is more complex and it is important to simultaneously track academic performance (an ‘intervening’ variable) rather than focus in on truancy *per se*. This is not to say that the government, schools, and parents should ignore truancy behaviour; it matters where test score performance will be adversely affected because this will lead to poor labour market outcomes. We also expect that the determinants of truancy behaviour and its effect on academic performance, and hence test scores, goes back further into the educational careers of young people than we are able to control for. Nevertheless, our analysis of the latter part of the educational process between ages 14–16 has helped to shed some light on the complex interaction between truancy behaviour, test score performance and early labour market outcomes.

Acknowledgements We would like to thank Dave Stott for research assistance, we also acknowledge financial support from Lancaster University Management School. We are also grateful to the ESRC and DfE for providing the data. We wish to acknowledge the valuable comments from two anonymous reviewers and the Associate Editor. All remaining errors remain our responsibility.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Data and computer code availability This study uses pupil-level data from the Youth Cohort Studies (YCS) which is available from the UK Data Service, together with data on the characteristics of the school attended from the School Performance Tables and Schools Census which is available from Department for Education (gov.uk). All the information needed to proceed from the raw data to the results of the paper (including code) are available from the authors upon request.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix A

See Table 9.

Table 9 Summary statistics for explanatory variables, male and female

Variable	Males				Females			
	Mean	SD	Min	Max	Mean	SD	Min	Max
<i>Cohort</i>								
7	0.144	0.352	0	1	0.143	0.350	0	1
8	0.127	0.333	0	1	0.124	0.329	0	1
9	0.138	0.345	0	1	0.134	0.340	0	1
10	0.131	0.338	0	1	0.130	0.336	0	1
11	0.152	0.359	0	1	0.159	0.365	0	1
12	0.130	0.336	0	1	0.141	0.348	0	1
<i>Ethnic background</i>								
Afro-Caribbean	0.015	0.123	0	1	0.021	0.145	0	1
Indian	0.009	0.093	0	1	0.008	0.087	0	1
Bangladeshi/Pakistani	0.021	0.143	0	1	0.020	0.139	0	1
Other race	0.031	0.174	0	1	0.031	0.174	0	1
Unknown	0.025	0.156	0	1	0.026	0.159	0	1
<i>Fathers occupation</i>								
Professional/Managerial	0.167	0.373	0	1	0.166	0.372	0	1
Skilled non-manual	0.095	0.293	0	1	0.099	0.299	0	1
Skilled manual	0.131	0.337	0	1	0.123	0.328	0	1
Unskilled non-manual	0.182	0.385	0	1	0.188	0.390	0	1
Unknown	0.304	0.460	0	1	0.314	0.464	0	1
<i>Mothers occupation</i>								
Professional/Managerial	0.106	0.308	0	1	0.114	0.317	0	1
Skilled non-manual	0.109	0.312	0	1	0.111	0.314	0	1
Skilled manual	0.296	0.456	0	1	0.312	0.463	0	1
Unskilled non-manual	0.083	0.276	0	1	0.090	0.286	0	1
Unknown	0.333	0.471	0	1	0.305	0.460	0	1
<i>Household status</i>								
Father only	0.041	0.197	0	1	0.038	0.192	0	1
Mother only	0.126	0.332	0	1	0.151	0.358	0	1
<i>School characteristics</i>								
school size (pupil nos)	0.876	0.446	0.002	2.382	0.867	0.454	0.002	2.382
Pupil-teacher ratio	16.373	1.465	8.429	27.86	16.424	1.449	9.058	27.86
Eligibility for FSM	0.151	0.123	0	0.905	0.154	0.124	0	0.902
Voluntary-aided/control	0.155	0.362	0	1	0.155	0.362	0	1
Grant maintained	0.150	0.357	0	1	0.140	0.347	0	1
Secondary modern	0.038	0.191	0	1	0.039	0.194	0	1
Selective (i.e. Grammar)	0.033	0.179	0	1	0.036	0.186	0	1
Single sex	0.089	0.285	0	1	0.121	0.327	0	1
truand	0.010	0.012	0	0.135	0.010	0.011	0	0.163
truand	0.010	0.006	0.001	0.030	0.011	0.006	0.001	0.030
<i>Region</i>								
North/North East	0.085	0.279	0	1	0.085	0.279	0	1

Table 9 (continued)

Variable	Males				Females			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Yorkshire & Humberside	0.116	0.320	0	1	0.112	0.315	0	1
North West	0.138	0.345	0	1	0.131	0.337	0	1
East Midlands	0.087	0.281	0	1	0.081	0.272	0	1
West Midlands	0.119	0.324	0	1	0.118	0.323	0	1
East Anglia/Eastern	0.088	0.283	0	1	0.090	0.286	0	1
South East (exc G. London)	0.174	0.379	0	1	0.179	0.383	0	1
South West	0.089	0.284	0	1	0.090	0.287	0	1
Sample size (<i>n</i>)	31,372				37,936			

Appendix B

See Tables 10, 11 and 12.

Table 10 Comparison of models using likelihood ratio tests

Unemployment		Log(Likelihood)	#params	Chi-square	df	<i>p</i> value
<i>Female</i>						
Homog Model	Truancy	-36,679.2	45			
	GCSE	-47,202.3	46			
	Unemp/NEET	-3902.2	37			
	Full Homog Model	-87,783.7	128			
Model 1	Full Heterog Model 1 (Vs Homog)	-87,555.6	131	456.2	3	<0.00001
Model 2	No truancy on Unemp (Vs Model 1)	-87,559.0	129	6.8	2	0.0333
Model 3	$\rho_{13} = 0$ (Vs Model 1)	-87,607.5	130	103.8	1	<0.00001
Model 4	Alt full model	-87,699.5	131			
<i>NEET</i>						
Homog Model	Truancy	-36,679.2	45			
	GCSE	-47,202.3	46			
	Unemp/NEET	-7236.0	37			
	Full Homog Model	-91,117.5	128			
Model 1	Full Heterog Model 1 (Vs Homog)	-90,884.2	131	466.5	3	<0.00001
Model 2	No truancy on Unemp (Vs Model 1)	-90,889.8	129	11.1	2	0.0039
Model 3	$\rho_{13} = 0$ (Vs Model 1)	-90,988.2	130	207.8	1	<0.00001
Model 4	Alt full model	-91,027.1	131			

Table 10 (continued)

Unemployment		Log(Likelihood)	#params	Chi-square	df	p value
<i>Male</i>						
Homog Model	Truancy	-28,121.0	45			
	GCSE	-40,887.9	46			
	Unemp/NEET	-4219.6	37			
	Full Homog Model	-73,228.5	128			
Model 1	Full Heterog Model 1 (Vs Homog)	-73,030.7	131	395.62	3	<0.00001
Model 2	No truancy on Unemp (Vs Model 1)	-73,033.3	129	5.24	2	0.073
Model 3	rho_13 = 0 (Vs Model 1)	-73,070.0	130	78.68	1	<0.00001
Model 4	Alt full model	-73,115.7	131			
<i>NEET</i>						
Homog Model	Truancy	-28,121.0	45			
	GCSE	-40,887.9	46			
	Unemp/NEET	-6552.8	37			
	Full Homog Model	-75,561.7	128			
Model 1	Full Heterog Model 1 (Vs Homog)	-75,419.0	131	285.36	3	<0.00001
Model 2	No truancy on Unemp (Vs Model 1)	-75,427.6	129	17.07	2	0.000
Model 3	rho_13 = 0 (Vs Model 1)	-75,475.1	130	112.17	1	<0.00001
Model 4	Alt full model	-75,503.4	131			

Table 11 Likelihood ratio tests of no difference between male and female models

Unemployment	Models	Log(likelihood)	Chi-square	df	p value
Homog Model	Truancy	-64,921.2			
	GCSE	-88,719.8			
	Unemp	-8095.8			
	Pooled Homog Model	-161,736.8	1449.2	128	<0.00001
Model 1	Pooled Heterog Model 1	-161,383.1	1593.7	131	<0.00001
Model 2	Model with no truancy on Unemp	-161,389.4	1594.3	129	<0.00001
Model 3	Pooled model with rho_13 = 0	-161,396.3	1437.6	130	<0.00001
Model 4	Pooled Alt full model	-161,605.3	1580.3	131	<0.00001
<i>NEET</i>					
Homog Model	Truancy	-64,921.2			
	GCSE	-88,719.8			
	NEET	-13,821.6			
	Pooled Homog Model	-167,462.5	1566.6	128	<0.00001
Model 1	Pooled Heterog Model 1	-167,097.9	1589.3	131	<0.00001
Model 2	Model with no truancy on NEET	-167,111.5	1588.3	129	<0.00001
Model 3	Pooled model with rho_13 = 0	-167,122.7	1318.8	130	<0.00001
Model 4	Pooled Alt full model	-167,318.1	1575.1	131	<0.00001

Table 12 Ordered probit models of the truancy on test scores, homogenous and heterogeneous models

Var.	Homogeneous			Heterogeneous Model 1			Heterogeneous Model 2			Heterogeneous Model 3			Heterogeneous Model 4		
	Est.	s.e.	p value	Est.	s.e.	p value	Est.	s.e.	p value	Est.	s.e.	p value	Est.	s.e.	p value
<i>Panel A: unemployment, females</i>															
Yt23	-0.396	0.012	0.000	0.405	0.030	0.000	0.406	0.030	0.000	0.405	0.030	0.000	na		
Yt45	-1.390	0.030	0.000	0.212	0.065	0.001	0.214	0.065	0.000	0.202	0.065	0.002	na		
/cut1	-3.408	0.081		-2.713	0.088	0.000	-2.712	0.088	0.000	-2.711	0.087	0.000	-3.043	0.080	0.000
/cut2	-3.063	0.080		-2.406	0.086	0.000	-2.405	0.086	0.000	-2.405	0.086	0.000	-2.730	0.079	0.000
/cut3	-2.388	0.079		-1.804	0.083	0.000	-1.803	0.083	0.000	-1.802	0.083	0.000	-2.100	0.079	0.000
/cut4	-1.369	0.079		-0.889	0.081	0.000	-0.889	0.081	0.000	-0.888	0.080	0.000	-1.127	0.078	0.000
/cut5	0.144	0.079		0.491	0.078	0.045	0.491	0.079	0.000	0.491	0.078	0.000	0.339	0.078	0.000
<i>Panel B: unemployment, males</i>															
Yt23	-0.395	0.013	0.000	0.338	0.037	0.000	0.341	0.037	0.000	0.342	0.037	0.000	na		
Yt45	-1.408	0.036	0.000	0.034	0.080	0.420	0.039	0.080	0.628	0.033	0.079	0.680	na		
/cut1	-3.168	0.088		-2.687	0.094	0.000	-2.684	0.094	0.000	-2.677	0.093	0.000	-2.904	0.087	0.000
/cut2	-2.801	0.087		-2.352	0.092	0.000	-2.350	0.092	0.000	-2.343	0.092	0.000	-2.569	0.086	0.000
/cut3	-2.014	0.086		-1.634	0.089	0.000	-1.631	0.089	0.000	-1.625	0.089	0.000	-1.829	0.086	0.000
/cut4	-1.113	0.086		-0.804	0.087	0.000	-0.802	0.087	0.000	-0.796	0.087	0.000	-0.960	0.086	0.000
/cut5	0.466	0.086		0.676	0.086	0.000	0.677	0.086	0.000	0.680	0.085	0.000	0.580	0.085	0.000
<i>Panel C: NEET, Females</i>															
Yt23	-0.396	0.012	0.000	0.403	0.030	0.000	0.405	0.030	0.000	0.407	0.030	0.000	na		
Yt45	-1.390	0.030	0.000	0.208	0.066	0.001	0.213	0.065	0.001	0.207	0.065	0.001	na		
/cut1	-3.408	0.081		-2.720	0.088	0.000	-2.717	0.088	0.000	-2.701	0.087	0.000	-3.047	0.080	0.000
/cut2	-3.063	0.080		-2.413	0.086	0.000	-2.410	0.086	0.000	-2.395	0.086	0.000	-2.734	0.079	0.000
/cut3	-2.388	0.079		-1.810	0.083	0.000	-1.808	0.083	0.000	-1.793	0.083	0.000	-2.103	0.079	0.000
/cut4	-1.369	0.079		-0.896	0.081	0.000	-0.894	0.081	0.000	-0.880	0.080	0.000	-1.131	0.078	0.000
/cut5	0.144	0.079		0.485	0.078	0.000	0.487	0.078	0.000	0.497	0.078	0.000	0.335	0.078	0.000

References

- Andrews MJ, Bradley S (1997) Modelling the school-to-work transition and the demand for training in the United Kingdom. *Economica* 64:387–413
- Andrews MJ, Bradley S, Stott D, Taylor J (2004) The evolution and determinants of the educational gender gap, mimeo. Department of Economics, Lancaster University, Lancaster
- Arulampalam W, Gregg P, Gregory M (2001) Unemployment scarring. *Econ J* 111:F577–F584
- Bell DNF, Blanchflower DG (2011) Youth unemployment and the great recession. *Oxf Rev Econ Policy* 27:243–264
- Bell R, Jones G (2002) Youth policies in the UK: a chronological map, 2nd edn. Mimeo, New York
- Bhat C, Pulugurta V (1998) A comparison of two alternative behavioral mechanisms for car ownership decisions. *Transp Res Part B* 32(1):61–75
- Boardman AE, Davis OA, Sanday PR (1977) A simultaneous equation model of the educational process. *J Publ Econ* 7:23–49
- Bosworth D (1994) Truancy and pupil performance. *Educ Econ* 2:243–264
- Bradley S, Nguyen A (2004) The school-to-work transition. Chapter 13. In: Johnes G, Johnes J (eds) *International handbook on the economics of education*. Edward Elgar Publishing, Cheltenham
- Bradley S, Taylor J (1998) The effect of school size on exam performance in secondary schools. *Oxf Bull Econ Stat* 60:291–324
- Bradley S, Taylor J (2002) Ethnicity, educational attainment and the transition from school. *Manch Sch* 72:317–346
- Buscha F, Conte A (2013) The impact of truancy on educational attainment during compulsory schooling: a bivariate ordered probit estimator with mixed effects. *Manch Sch* 82:103–127
- Card D, Krueger A (1992) Does school quality matter? Returns to education and the characteristics of public schools in the United States. *J Polit Econ* 100:1–40
- Casey B, Smith D (1995) England and Wales youth cohort study: truancy and youth transitions. Department for Education and Employment, Research Series Youth Cohort Report No. 34
- Clark A, Georgellis Y, Sanfey P (2001) Scarring: the psychological impact of past unemployment. *Economica* 68:221–241
- Coles B, Godfrey C, Keung A, Parrott S, Bradshaw J (2010) Estimating the life-time cost of NEET: 16–18 year olds not in education, employment or training. University of York, final draft report for the Audit Commission, July 21st 2010
- Crawford C, Duckworth K, Vignoles A, Whyness G (2011) Young people's education and labour market choices aged 16–17 to 18–19. Department of Education Research Brief DFE-RB182, London
- Divgi DR (1979) Calculation of univariate and bivariate normal probability functions. *Ann Stat* 7:903–910
- Duckworth K, Schoon I (2012) Beating the odds: exploring the impact of social risk on young people's school-to-work transitions during recession in the UK. *Natl Inst Econ Rev* 222:R38
- Dustmann C, Soest A (2007) Part-time work, school success and school leaving. *Empir Econ* 32:277–299
- Dustmann C, Rajah N, Soest A (1998) School quality, exam performance and career choice. ESRC Research Centre, Institute for Fiscal Studies, London
- Education Act (1996) HMSO: London
- Education Act (2006) HMSO: London
- Elbers C, Ridder G (1982) True and spurious duration dependence: the identifiability of the proportional hazard model. *Rev Econ Stud* 49:403–411
- Ermisch J, Francesconi M (2001) Family matters: the impact of family background on educational attainment. *Economica* 68(270):137–156
- Ermisch J, Janatti M (2012) Socioeconomic gradients in children's outcomes. Chapter 2. In: Ermisch J, Janatti M, Smeeding M (eds) *From parents to children: the intergenerational transmission of advantage*. Russell Sage Foundation, New York
- Evans WN, Schwab RM (1995) Finishing high school and starting college: do catholic schools make a difference. *Q J Econ* 110:941–974
- Feinstein L, Symons J (1999) Attainment in secondary schools. *Oxf Econ Pap* 51:300–321
- Fuchs T, Woessmann L (2004) What accounts for international differences in student performance? A re-examination using PISA data, CESIFO working paper no. 1235, July 2004
- Geweke J (1989) Bayesian inference in econometric models using Monte Carlo integration. *Econometrica* 57:1317–1339

- Gibbons S, McNally S (2013) The effects of resources across school phases: a summary of recent evidence. Centre for Economic Performance discussion paper CEPDP 1226, Centre for Economic Performance, LSE, London UK
- Goldman D, Bhattachaharya J, McCaffrey D, Duan N, Leibowitz A, Joyce G, Morton S (2001) Effect of insurance on mortality in a HIV-positive population in care. *J Am Stat Assoc* 96:883–894
- Hajivassiliou VA, McFadden DL (1998) The method of simulated scores for the estimation of LDV models. *Econometrica* 66:863–896
- Hanushek EA (1986) The economics of schooling: production and efficiency in public schools. *J Econ Lit* 24:1141–1177
- Hanushek EA (1992) The trade-off between child quantity and quality. *J Polit Econ* 100:84–117
- Hanushek EA, Rivkin SG, Taylor LL (1996) Aggregation and the estimated effects of school resources. *Rev Econ Stat* 78:611–627
- Heckman JJ (1978) Dummy endogenous variables in a simultaneous equation system. *Econometrica* 46:931–959
- Heckman JJ, Singer B (1984) The identifiability of the proportional hazard model. *Rev Econ Stud* 51(2):231–241
- Keane MP (1992) A note on identification in the multinomial probit model. *J Bus Econ Stat* 10:193–200
- Keane MP (1994) A computationally practical simulation estimator for panel data. *Econometrica* 62:95–116
- Knabe A, Ratzel S (2011) Scarring or scaring? The psychological impact of past unemployment and future unemployment risk. *Economica* 78:283–293
- Leslie D, Drinkwater S (1999) Staying on in full time education: reasons for higher participation rates among ethnic minority males and females. *Economica* 66:63–77
- Loeb S, Bound J (1996) The effect of measured school inputs on academic achievement: evidence from the 1920s, 1930s and 1940s birth cohorts. *Rev Econ Stat* 78:653–663
- Lynch L (1987) Individual differences in the youth labour market: a cross sectional analysis of London youths. Chapter 9. In: Junankar P (ed) *From school to unemployment?*. MacMillan, The Labour Market for Young People, London
- Marcel D, Montmarquette C, Viennot-Briot N (2001) Dropout, school performance and working while in school: an econometric model with heterogeneous groups, CIRANO working papers 2001s-63, CIRANO
- McIntosh S (2001) Measuring and assessing the impact of basic skills on labour market outcomes. *Oxf Econ Pap* 53(3):453–481
- Micklewright J (1989) Choice at sixteen. *Economica* 56:23–39
- Micklewright J, Rajah N, Smith S (1994) Labouring and learning: part-time work and full-time education. *Natl Inst Econ Rev* 2:73–85
- Mroz TA, Savage TH (2007) The long-term effects of unemployment. *J Hum Resour* XLI:259–293
- Neal DA (1997) The effects of catholic secondary schooling on educational achievement. *J Labor Econ* 15(Part 1):98–123
- Petrongolo B, van Reenen J (2011) Youth unemployment. CentrePiece, Centre for Economic Performance, Summer, London
- Rice P (1987) The demand for post-compulsory education in the UK and the effects of the educational maintenance allowance. *Economica* 54:465–475
- Rice P (1999) The impact of local labour markets on investment in further education: evidence from England and Wales youth cohort studies. *J Popul Econ* 12:287–312
- Robertson D, Symons J (1996) Do peer groups matter? Peer group versus schooling effects on academic attainment. *Stata J* 11:159–206
- Roodman D (2011) Estimating fully observed recursive mixed-process models with cmp. Working paper 168. Center for Global Development, Washington, DC
- Steele F, Vignoles A, Jenkins A (2007) The effect of school resources on pupil attainment: a multilevel simultaneous equation modelling approach. *J R Stat Soc Ser A* 170(3):801–824
- Wilde J (2000) Identification of multiple equation probit models with endogenous dummy regressors. *Econ Lett* 69:309–312