

Non-pecuniary returns to higher education: The effect on smoking in the UK

Alfonso Miranda*, Massimiliano Bratti^{†‡}

Abstract. In this paper we investigate whether higher education (HE) produces non-pecuniary returns via the reduction in the consumption of health-damaging substances. In particular, the paper focuses on studying the smoking behaviour of British individuals. We use data for current smokers from the 1970 British Cohort Study (BCS70) and estimate endogenous switching count models for cigarette consumption. Results show that HE is endogenous with respect to smoking. Moreover, HE is found to have a significant negative effect on smoking even when potential unobserved heterogeneity is properly controlled for. Once endogeneity is controlled for, HE is found to have a higher effect than in models where HE is treated as exogenous. We argue that the lower effect of HE found in models with exogenous education is caused by the omission of relevant unobserved factors that affect both HE and smoking such as an individual's unobservable "health stock".

JEL classification: C35, I12, I21.

Keywords: endogenous switching, count data, higher education, smoking, UK.

1 Introduction

Higher education (HE, hereafter) has been found to generate significant wage returns. In the case of the UK, for instance, Blundell et al. (2000) found

*School of Economic And Management Studies, University of Keele, Staffordshire ST5 5BG, UK. E-mail: A.Miranda@econ.keele.ac.uk

[†]Department of Economics, Business and Statistics, University of Milan, Via Conservatorio 7, I-20122, Milan, Italy. E-mail: Massimiliano.Bratti@unimi.it

[‡]We acknowledge the original data depositor, the UK Data Archive (UKDA), which made available data from the 1970 British Cohort Study to us, and the original data collectors, the Department of Child Health at Bristol University and the Centre for Longitudinal Studies at the Institute of Education (University of London). Neither the original data collectors nor the data depositor bear any responsibility for the analyses or interpretations presented here. We are also grateful to Gauthier Lanot for useful comments. The usual disclaimers apply.

that an undergraduate degree has, on average, ‘raw’ wage returns of around 21% for men and 39% for women in the National Child Development Study (NCDS). Similarly, Bratti et al. (2005) reported undergraduate degree returns of 14% for men and 18% for women from the 1970 British Cohort Study (BCS70).

Besides wage returns, there is a general agreement in the literature that education has important non-pecuniary returns (see, for instance, the survey in Wolfe and Haveman 2003). In the field of health, for example, education is thought to help individuals to avoid health damaging behaviour and to increase their life expectancy (Grossman 1975, Thaler and Sherfrin 1981). Several articles (see the review in section 2) have found important positive causal effects of education on health using instrumental variables (IV) techniques, a few of them, however, have investigated the pathways through which education produces health returns. One of such pathways is the effect of education on health-related behaviour, such as cigarette or alcohol consumption. Therefore, in the current paper, we focus our attention on a specific channel through which HE may affect health: the intensity of dairy cigarette consumption. We mainly focus on the UK, for which the empirical evidence on causal non-pecuniary returns to HE is much more scarce than for the US.

There are at least two reasons motivating our study. First, there is the fact that the UK has undergone a gradual transition towards a quasi-market in the HE sector. As a consequence, university student tuition fees are being progressively increased to properly reflect the large private returns that graduates from HE obtain in terms of both wages and a higher probability of employment. From this perspective, individuals who pay for HE studies are naturally interested in learning whether their investment has unaccounted non-pecuniary returns such as general improvements in health. Second, the existence of important health returns to HE implies that an increased access to HE may help to improve the health condition of the average UK citizen and decrease expenditures in public health. Hence, the government may be

interested in internalising such benefits when deciding the amount of public funds to be allocated to support HE activities.

Evaluating non-pecuniary returns to HE is substantially complicated by the fact that HE is potentially endogenous with respect to smoking behaviour (Becker and Murphy 1988, Farrell and Fuchs 1982). Such endogeneity arises whenever unobserved individual characteristics affecting the likelihood to get a HE degree are correlated with unobserved heterogeneity determining smoking decisions. The rate of time preference is an example of such unobserved characteristics that is often emphasised in the literature. On the one hand, it is argued that individuals who heavily discount the future are more likely to engage in health damaging activities such as smoking. On the other hand, impatient individuals heavily discount future incomes and are therefore less likely to invest in human capital (see, for instance, Fersterer and Winter-Ebmer 2003). Hence, if education is taken as exogenous in a smoking equation estimates may be subject to a substantial bias.

Previous work in the smoking literature typically studies the determinants of smoking status (i.e., activity participation) rather than the determinants of intensity of substance use.¹ For these reasons, binary health responses are analysed and the endogeneity of education — usually assumed to be a continuous variable — is controlled for by instrumental variable techniques (see De Walque 2004, Sander 1995a, Nayga 1999, among others).

In contrast to previous literature, the present paper aims to model the intensity of substance use of current smokers. In particular, the variable of interest is the number of cigarettes daily smoked. We condition, as a consequence, on smoking participation. Given the non negative and discrete

¹Another strand of the literature on smoking is concerned with tobacco expenditure. These studies have the disadvantage of modelling a variable that fluctuates due to price changes and that, as a consequence, introduces uninformative noise to the two relevant dimensions of analysis: (a) participation status, and (b) intensity of substance use. The expenditure approach has also the disadvantage of artificially imposing a continuous variable framework to a problem that is necessarily generated by discrete data generating mechanisms. Double-Hurdle models are commonly used in this literature (for further reference see, among others, Cragg 1971, Yen and Jensen 1996, Blaylock and Blisard 1993, Jones 1989, Labeaga 1999).

nature of the data, the usual binary-variable techniques are no longer suitable. We use instead count data techniques.

The potential endogeneity of the HE dummy is explicitly addressed by using an endogenous switching framework. Among other advantages, the endogenous dummy variable strategy allows the researcher to properly control for the presence of unobserved heterogeneity in both the count outcome and the switching equation. Besides, unlike popular IV procedures, the endogenous switching model directly delivers correct standard errors. This is major advantage since other studies fail to obtain statistically significant effects of education when using a two-stage IV approach, probably due to weak instruments (see for instance Arendt 2005, on the effect of education on health).

Our study benefits from a unique longitudinal data set, the 1970 British Cohort Study (BCS70, hereafter), that follows the same individuals at ages 0, 5, 10, 16, 26 and 30. These data are rich in background information on the sampled individuals and offer a number of potential instruments for the endogenous HE dummy.

The structure of the paper is as follows. Section 2 mainly outlines the reasons for which a positive correlation between HE and smoking might emerge, some of which reflect causal effects, and reports a brief survey of the previous literature. Section 3 briefly describes our estimation sample. Section 4 introduces the endogenous switching count data model that we use to estimate the causal effect of HE on dairy cigarette consumption. Section 5 reports the main results of our empirical analysis and the last section concludes.

2 Education and smoking: findings from the previous literature

The literature has emphasised several possible sources of correlation between education and health, some of which reflect causal explanations:

1. the *third variable* argument. As emphasised by Fuchs (1982), education

is likely to be endogenous in a smoking equation and affected by the same unobserved factors influencing smoking behaviour. Some examples of characteristics that are not observed in the data sets commonly used to estimate smoking equations are an individual's time discount rate and cognitive ability. Individuals with higher discount rates are likely to invest less in education and consume more cigarettes while abler individuals are more likely to achieve more education and use more effectively information on the harmful effects of smoking. In this case the correlation between education and smoking will stem from a third variable and will not reflect a causal relationship. Farrell and Fuchs (1982) find a negative effect of future education on current smoking for individuals aged 17 and interpret this correlation as consistent with the third variable hypothesis. Indeed, according to the authors, as individuals who acquired more education were less likely to smoke also before completing the additional years of schooling the effect of the latter cannot be causal and should reflect some form of individual unobserved heterogeneity. Sander (1998) uses future education as a proxy for an individual's discount rate and reaches the same conclusion. However, we will see later when illustrating the "economic returns to health" argument, that this interpretation is not necessarily correct;

2. the *productive efficiency* argument. This was originally introduced by Grossman (1972). The underlying idea is that education directly affects the health production function and more educated individuals are able to produce a higher stock of health even when using the same amount of inputs as the less educated ones. Empirical investigation of this hypothesis is usually made by estimating a health production function, in which the stock of health capital is measured by self-rated health, including in the right-hand-side (RHS, hereafter) education and the other inputs into the health production function and testing for the significance of the coefficient on education. Since in the current paper we analyse the decision to smoke, i.e. an input into the health pro-

duction function, we do not develop further on this argument and the interested reader can refer to Grossman (2005) for a very recent survey on the topic;

3. the *allocative efficiency* argument. Education alters the input mix in the health production function. As Rosenzweig and Schultz (1983) put forward, this argument in its strongest form maintains that education will have no impact on health unless it changes inputs in the health production function, and the coefficient on education in this function would be zero if all inputs were included.² According to this argument the main mechanism through which education affects the input mix is by increasing health related knowledge, e.g. on the harmful effects of smoking, or the speed of adoption of health enhancing inputs. However, the evidence supporting this causal link is mixed. Farrell and Fuchs (1982), for instance, found that the the smoking-education relationship became stronger when the information on the harmful effects of smoking became widespread. Similarly, De Walque (2004) found that as information about the health hazards of smoking diffused, smoking declined earlier and most impressively among college graduates. By contrast, Kenkel (1991) observes that the coefficient on education in a smoking equation remains highly significant and is only marginally reduced in size when health knowledge is expressly controlled for. Since education is treated as exogenous the author is not able to conclude whether the effect of education is causal or not. A negligible role of health knowledge is also found by Nerín et al. (2004) who study the smoking habits of a sample of first-year medical students and observe an increase of smoking prevalence when they were in their third year. As the authors state: “An improvement in the awareness of smoking as a risk factor was observed for medical students, but no change in

²Rosenzweig and Schultz (1983) argue: “It not clear...how education can actually alter marginal products of inputs...unless inputs are omitted from [the production function]. That is, it is doubtful that schooling can affect the production of...[health] without it being associated with some alteration in an input” (p.19).

attitude was observed.” (p. 341). However, in our sample almost all individuals were informed about the damaging effects of smoking already at age 10, and therefore later education and particularly HE is unlikely to have had a substantial impact on improving health knowledge related to smoking;³

4. the *economic returns to health* argument. According to this argument rational utility-maximising individuals take into account both the current and the future health consequences of their current smoking decisions. Adult workers with more education, and therefore higher wages, have higher economic returns to good health (lower morbidity and higher longevity) and are more likely to refrain from current smoking. Similarly, young individuals who plan to invest more in education, and will have higher future wages, will also have higher future costs of current cigarette consumption, due to future loss of work days or a shorter life expectancy, and might smoke less even before completing education. This can provide an alternative explanation for the evidence in Farrell and Fuchs (1982) that highly educated individuals smoke less even before completing education, which the authors interpret as supporting the third variable hypothesis;
5. other causal explanations. There might be other channels through which education has a causal effect on smoking behaviour. Examples include the positive effect of education on self-control⁴(see Ross and Mirowsky 1999), and the effect of education on a person’s time discount rate (Becker and Mulligan 1997).

It should also be noted that in the economic literature some theoretical models explain smoking behaviour by emphasising the “addictive” charac-

³In particular, 77.29% of our sample declared that they were sure that smoking can damage health, 7.31% were not sure while 0.99% declared that they were sure that smoking does not damage health (the remaining individuals did not answer).

⁴Although there might be a problem of reverse causality since self-control is likely to affect educational achievement as shown by Feinstein (2000).

ter of cigarette smoking, the so-called “rational addiction” models. Becker and Murphy (1988) model, for instance, assumes that past consumption of cigarettes affects current consumption through a process of learning by doing. Non-economic explanations also stress the addictive power of nicotine (“irrational addiction”). In all such models individuals who smoked in the past are also more likely to smoke in the future. Empirically this is equivalent to including past cigarette consumption among the regressors in a current smoking equation and testing for state dependence. However, we have decided not to emphasise here the addictive power of cigarettes since past cigarette consumption is likely to be endogenous with respect to both current consumption and an individual’s education and we do not have good instruments to address its endogeneity.⁵ We nevertheless recognise that adolescent smoking is probably one of the most powerful predictors of adult smoking habits by including among the regressors in the smoking equation early family socio-economic characteristics and peer group variables that are likely to predict adolescent smoking initiation.

The most recent economic literature has attempted to tackle the issue of the endogeneity of education with smoking behaviour or health status by using instrumental variables techniques and has mostly used US data. In these papers education is usually measured by years of completed education. Sander (1995a;b) uses father’s schooling, mother’s schooling, rural residence at age 16 and number of siblings as instruments for schooling. He generally finds a negative causal effect of education on smoking participation and a positive effect on the likelihood of quitting smoking, so his results are not compatible with the third variable hypothesis. Other papers use as instruments ancestry, average real per capita income and average real per capita

⁵Clark and Etilé (2002) using several waves of the British Household Panel Survey (BHPS), i.e. panel data, instrument changes in past cigarette consumption with lags of past levels and changes of cigarette consumption. In our case it is not possible to do the same since we only have information on cigarette consumption when individuals were 26 years old (from the 26-year follow up of the BCS70). Moreover, unlike Clark and Etilé (2002), who estimate a model of rational addiction, in our case the main objective is to address the endogeneity of higher education.

expenditures on education in the state in which an individual resided when he/she was a child (Berger and Leigh 1989), IQ and Knowledge or Work test scores (Berger and Leigh 1989), compulsory education laws (Adams 2002, Lleras-Muney 2005, Arendt 2005), local unemployment rates during a person's teenage years (Arkes 2003), the risk of draft induction during the Vietnam war (De Walque 2004), inter-state variation in education testing and compulsory schooling policies (Kenkel et al. 2004). The common findings of these articles is that education has sizeable causal effects on health and health related behaviour.

As to the UK, a number of researchers have already investigated the effect of education on smoking, although most of them have focused on smoking participation or quitting rather than on smoking intensity and have treated education as exogenous. Chandola and Bartley (2004) used data from the British Household Panel Survey (BHPS) for the period 1991-2000 to estimate multilevel logistic models and did not find any effect of education on the probability to quit smoking. Balia and Jones (2005) used data from the British Health and Lifestyle Survey (HALS) for the period 1984-85 and estimated a multivariate probit model where one of the dependent variables is the decision of being a non-smoker. The authors found that individuals with less than O-level education were more likely to smoke than those with O-level education, while did not find any additional effect of HE and A-level with respect to O-level. Contoyannis and Jones (2004) used the HALS data set for the period 1984-1991 and estimated a multivariate probit model where being a non-smoker is one of the outcome variables. The authors found that education has a monotonic decreasing effect on the probability of being a non-smoker. Forster and Jones (2003) used HALS data for the 1984 to estimate a Probit model of the decision to start smoking and find that individuals with HE are less likely to start smoking than those with O-level education and that the effect is stronger for females. All the articles reviewed up to now have treated education as exogenous. Therefore, it is not possible to establish whether in the case significant effects of education

were found they only reflect spurious correlations. The only article treating education as exogenous of which we are aware is Chandola et al. (2006). The authors used National Child Development Data (NCDS) which collects data on the 1958 British Cohort, to estimate a structural model of the effect of education on health. They distinguish between ‘direct’ and ‘indirect’ causal effects of education where the latter are those produced by third variables that affect health and that are in turn affected by education. Education was simultaneously modelled with health status and health-related behaviour, among other things. The authors found a strong effect of education on the probability of being a non-smoker, which represents one of the main indirect causal effects of education on health. With respect to Chandola et al. (2006) we consider here a different British cohort, born in 1970, and focus on the effect of HE on smoking intensity rather than on smoking status. Moreover, we adopt a different econometric methodology to account for endogeneity of HE. Therefore, we will be able to add some new empirical evidence to complement early findings from Chandola et al. (2006), which showed the existence of important causal effects of education on smoking in the UK.

Besides education, the literature has emphasised other important correlates or determinants of individual smoking decisions. Powell et al. (2005) found strong peer effects in youth smoking decisions. As the authors state: “moving a high-school student from a school where no children smoke to a school where one quarter of the youths smoke is found to increase the probability that the youth smokes by about 14.5 percentage points.” (p. 950). Moreover, the authors found no evidence of endogeneity of peer group’s smoking with an individual’s smoking behaviour, no effect of parents’ education and a significant effect of communication between parents and children.

Bantle and Haisken-DeNew (2002) emphasise the intergenerational links in smoking behaviour. There are several reasons why there might be intergenerational transmission in smoking behaviour. As explained by von Laffert (1998), parents act as role models for children and children may infer from parents’ smoking that they can derive utility from cigarettes (“utility

conjecture”) or that the benefits from smoking exceed its harmful effects (“illusion of high utility”). However, parents might also attempt to impose smoking bans on their children and if the parents themselves are smokers they might not be credible. The authors included several potential determinants of both parents’ and children’s smoking behaviours such as parents’ highest education, household size, income and income per capita and did not find a significant effect of these factors over and above the positive effect of parents’ smoking on children’s smoking. Similar results are found by Emery et al. (2001) who observed an insignificant effect of parental education and incomes on smoking behaviour of individuals aged 14 or more, while finding a strong positive effect of exposure to smoking within the family. Blow et al. (2005) investigated the effect of parental income on children’s smoking behaviour using BHPS data and found no effect of parental income when including a dummy for the presence of an adult smoker in the family and a significant positive effect on children’s smoking only of mothers with no formal education. Their conclusion is that “a large part of the relationship between children’s smoking behaviour and parental socio-economic status is transmitted via the smoking status of the adults” (p. 10).⁶

3 Data

In this paper we use data drawn from the BCS70. The BCS70 began in 1970 when data were collected on the births and families of 17,198 babies born in England, Wales, Scotland and Northern Ireland from the 5th to the 11th of April.⁷ There are currently five complete follow-up surveys available: 5-year, 10-year, 16-year, 26-year and 30-year. As to the variables included in our empirical analysis, data on smoking habits and the highest educational qualification were collected in the 30-year follow-up survey while all the other

⁶An insignificant effect of father’s education on a child’s smoking is also found in Farrell and Fuchs (1982).

⁷Subjects from Northern Ireland were included in the birth survey, but have been excluded from all subsequent sweeps.

contextual variables were provided in the 10-year follow-up survey. We use, therefore, a sample of individuals who matched between the two waves. The BCS70 was affected by same panel attrition. Although nothing ensures that panel attrition is at random with respect to the variables representing the main focus of our paper, i.e. education and smoking, in previous work (see, e.g., Bratti et al. 2005) we have noted that individuals in the matched samples generally have very similar characteristics to those in the initial waves and that therefore attrition across waves should not affect the individual characteristics which we are interested in.⁸

In the age 30 follow-up we have information on smoking behaviour. In particular as far smoking is concerned the relevant questions are:

would you say that: (a) you've never smoked cigarettes; (b) you used to smoke cigarettes but not at all now; (c) you now smoke cigarettes occasionally but not every day; (d) you smoke cigarettes every day?

and

how many cigarettes a day do you usually smoke?

The distribution of valid answers to the first question is shown in Table 1. About 29% of individuals in our sample smoke cigarettes every day, while 7.7% smoke occasionally. Every-day smokers are asked the number of cigarettes smoked a day (i.e. the second question). In the present paper we consider all current smokers and set the number of cigarettes smoked a day at zero for occasional smokers.

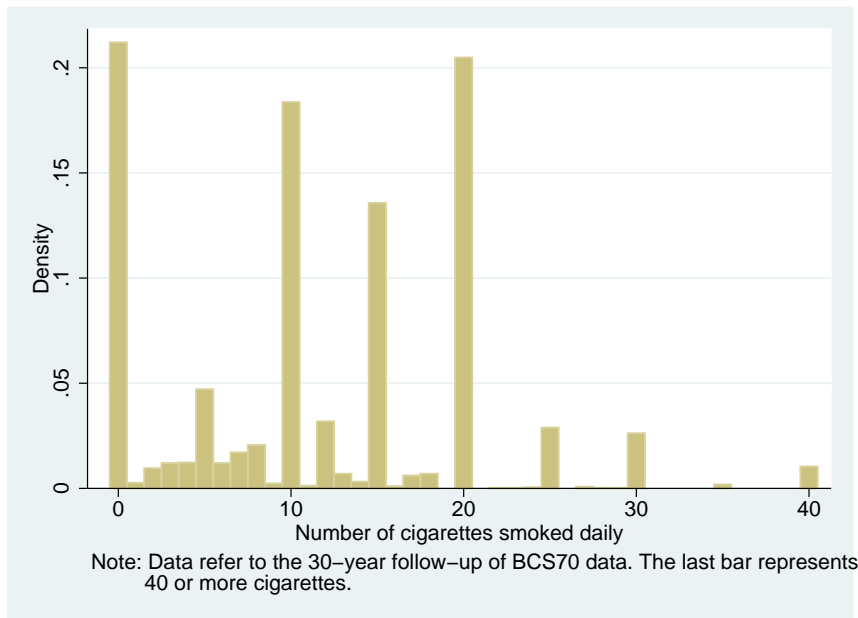
Self-reported data on dairy cigarette consumption may contain errors. In particular, when looking at the distribution of the number of cigarettes

⁸The Office for National Statistics (1999) writes “Analysis of differential response comparing achieved samples and target samples for any follow-up, using data gathered during the birth and earlier follow-ups, show that the achieved sample are broadly representative of the target sample. However, as in other surveys, some groups (eg those from minority ethnic, low social class, and atypical family backgrounds) are under-represented” (p. 11).

Table 1. Smoking habits in the 30-year follow-up of BCS70

Smoking habits	Frequency	Percent	Cumulate
never smoked cigarettes	4,937	44.06	44.1
used to smoke but don't at all now	2,125	18.97	63
smoke cigarettes occasionally	863	7.7	70.7
smoke cigarettes every day	3,279	29.27	100
Total	11,204	100	

Note. The distribution refers to valid answers only.

Figure 1. Number of cigarettes smoked daily (current smokers)

smoked daily in our sample we observe that fractions and multiples of a packet are more likely to occur. This is not a distinctive feature of our sample, since this phenomenon is often found in studies of cigarette consumption (see, for instance Clark and Etilé 2002, who use BHPS data).

Table 2 reports the average number of cigarettes smoked a day by level of education. Male smokers with HE smoke about four less cigarettes a day than those with less than HE (-41%).⁹ The difference in the number of cigarettes

⁹HE refers to individuals with a degree level qualification or more. In particular, HE includes individuals with levels of qualification 4 or 5 in the standard UK classification

Table 2. Number of cigarettes daily smoked by education (current smokers)

Education at age 30	Males			Females		
	Mean	Std. Dev.	Freq.	Mean	Std. Dev.	Freq.
lower than HE	13.45	9.45	1529	11.37	7.94	1350
HE	9.53	8.89	461	7.68	7.80	408
Total	12.54	9.47	1990	10.51	8.06	1758

smoked a day between female smokers with HE and those with less than HE is similar and amounts to almost four cigarettes (-48%).

Table 3. Highest educational qualification of the lower than HE group

Composition	Males			Females		
	Freq.	Percent	Cum.	Freq.	Percent	Cum.
No formal qualification	350	23.01	23.01	336	24.96	24.96
Level 1: GSCE D-G, CSE 2-5,..	180	11.83	34.85	167	12.41	37.37
Level 2: O-level equiv.	667	43.85	78.7	639	47.47	84.84
Level 3: A-level equiv.	324	21.3	100	204	15.16	100
Total	1,521	100		1,346	100	

In this paper we will compare dairy cigarette consumption of individuals with HE with those with lower educational equalifications. Hence, the “treatment” is having HE and the control group is represented by individuals with lower levels of education. The composition of the control group (lower than HE) by highest educational qualification achieved is reported in Table 3. Table 3 clearly shows that the control group is mostly composed of individuals who did not go on in post-compulsory schooling (75% for males and 80% for females). Hence, the effect of HE that we will estimate can be roughly interpreted as the effect with respect to individuals with no more than compulsory education.

of educational qualifications (see, for instance, Bynner et al. 2002, p. 45).

4 Econometric methodology

Endogenous switching models for count variables are used for our empirical analysis (see, for instance, Terza 1998, Kenkel and Terza 2001, Miranda and Rabe-Hesketh 2005). Denote by y_i the number of cigarettes diary consumed by the i – *th* individual. By assumption y_i takes on non-negative integer values and is a function of a potentially endogenous dummy, E_i , and a $K \times 1$ vector of individual characteristics \mathbf{x}_i — including the constant term. In the present context E_i represents an indicator for successful completion of a HE degree. We suppose that y is distributed as a Poisson variate,

$$Pr(y_i; \mathbf{x}_i, E_i, u_i) = \frac{\exp[\mathbf{x}_i' \boldsymbol{\beta} + \theta E_i + u_i]^{y_i} \exp\{-[\mathbf{x}_i' \boldsymbol{\beta} + \theta E_i + u_i]\}}{y_i!}, \quad (1)$$

where u_i is an unobserved individual heterogeneity term. As usual, the endogenous dummy is modelled as a latent response,

$$\begin{aligned} E_i^* &= \mathbf{z}_i' \boldsymbol{\gamma} + v_i \\ E_i &= \begin{cases} 1 & \text{if } S_i^* > 0 \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \quad (2)$$

with \mathbf{z}_i representing a $L \times 1$ vector of explanatory variables (including the constant), $\boldsymbol{\gamma}$ the vector of associated coefficients, and v_i a random variate. The model is identified by functional form (Kenkel and Terza 2001). As a consequence, vectors \mathbf{x}_i and \mathbf{z}_i may contain some or all common elements.

The main objective of the researcher is to obtain consistent estimators of the $K \times 1$ vector of parameters $\boldsymbol{\beta}$ and the coefficient on E_i , $\theta \in \mathbb{R}$. Matters are essentially complicated by the fact that u_i and v_i may be correlated, leading to a classic problem of endogeneity of E_i in equation (1). Clearly, in such a case estimation of equation (1) under the assumption of exogenous education will deliver biased and inconsistent estimators. To avoid this problem the relationship between u_i and v_i must be explicitly modelled and equations (1) and (2) should be estimated as a system.

The standard approach to allow correlation between u_i and v_i is to assume that the two random terms are jointly normally distributed (see Terza 1998).

Here we follow the same tradition. However, for simplicity, direct dependence between u_i and v_i is induced. Namely, we suppose that

$$v_i = \lambda u_i + \zeta_i, \quad (3)$$

with u_i and ζ_i distributed as independent normal variates with 0 mean. The $\lambda \in \mathbb{R}$ term in equation (3) represents a ‘factor loading’ that is estimated along the other parameters of the model.

Given that the variance in a Probit is only identified up to a constant, we set $\text{Var}(\zeta_i)$ to 1 without loss of generality. In the case of u_i there is no need to impose any restriction on $\sigma^2 = \text{Var}(u_i)$ because the Poisson distribution in this model can accommodate overdispersion (see below). Under the structure imposed by equation (3) the covariance matrix of u_i and v_i is given by:

$$\text{Cov}[(u_i, v_i)'] \equiv \Sigma = \begin{bmatrix} \sigma^2 & \lambda\sigma^2 \\ \lambda\sigma^2 & \lambda^2\sigma^2 + 1 \end{bmatrix},$$

and the correlation is

$$\rho = \frac{\lambda\sigma^2}{\sqrt{\sigma^2(\lambda^2\sigma^2 + 1)}}.$$

Equations (1) through (3) define a reparametrisation of the model introduced by Terza (1998). In Terza’s model, however, the variance of v_i is set to 1 while the total variance of v_i is $(\lambda^2\sigma^2 + 1)$ here. Hence, it is possible to recover Terza’s parameterisation by dividing our estimates for the switching equation by $\sqrt{\lambda^2\sigma^2 + 1}$.

The model is estimated by maximum likelihood and at convergence the negative of the inverse of the Hessian matrix can be used as an estimator for the covariance matrix. To evaluate the likelihood function, the random term u_i must be integrated out. Gauss-Hermite quadrature may be used to perform this task. Here, however, we use adaptive quadrature (Rabe-Hesketh and Skrondal 2002). Adaptive quadrature is a numerical integration technique that, at each iteration, updates the location and weights of the Gauss-Hermite quadrature points using the posterior distribution of u_i . After update, locations are centred around the posterior mean and equally

spread out according to the posterior standard deviation. Adaptive quadrature has proven to achieve higher accuracy with fewer integration points than the ordinary Gauss-Hermite quadrature.

One should be aware that a model with exogenous switching (EXS) is nested within the endogenous switching (ES) framework. This is so because if $\rho = 0$ the random terms u_i and v_i are independent and the likelihoods for the count and the dummy variables are separable — which is what is actually meant by exogenous switching in the econometrics literature (see Winkelmann 1998). This implies then that a test for the endogeneity of E_i in equation (1) can be performed on the basis of a simple likelihood ratio test for $\rho = 0$.

A final remark about the ES Poisson model is necessary. Namely, that despite a Poisson distribution imposes the equality of mean and variance (known as the equidispersion property of the Poisson distribution) the model defined in equation (1) does not necessarily implies that, given the covariates, the conditional mean of y should be equal to its conditional variance. In fact, it is easy to show that

$$\text{Var} [y_i | \mathbf{x}_i, E_i] = \mu_i + k\mu_i^2$$

where $\mu_i = E[y_i | \mathbf{x}_i, E_i]$ and $k = (\exp(\sigma^2) - 1)$. Variable y will therefore exhibit overdispersion as long as $\sigma^2 \neq 0$. Notice, furthermore, that σ^2 may be statistically different from zero even in the case that a $\rho = 0$ is found. Hence, unobserved heterogeneity may be present — and controlled for — even if the higher education dummy, E_i , turns out to be exogenous in the smoking equation.

We will use the evidence from the previous literature to propose some candidate instrumental variables to identify the effect of higher education on smoking. In the light of the findings reported in section 2 one can reasonably assume that the main influence of a parent's education on his/her children smoking amounts to the transmission of health knowledge, differences in parenting styles, and the role model transmitted through his/her own smoking habits. As a consequence, a parent's education should affect smoking be-

haviour only through its effect on children's education once health awareness and parental smoking status are controlled for. Following this argument we exclude parental education from the smoking equation. Clearly, imposing this restriction delivers an additional instrument for identifying the effect of HE. The effect of parental education on a child's education is already well established in the literature (see for instance Micklewright 1989, Ermisch and Francesconi 2001, Chevalier and Lanot 2002, for some UK related evidence). The validity of these instruments was in any case tested by performing formal Wald tests for the exclusion of parental education variables from the smoking and higher education equations in the ES poisson model including parents' education in both equations.

On the grounds of the literature review in section 2, our econometric model will enable us to distinguish among some alternative hypotheses on the effect of HE on smoking. In particular, we will be able to distinguish among these four different situations:

1. The correlation coefficient ρ is not statistically different from zero and the coefficient on education in the smoking equation is statistically significant. In this case education is exogenous with respect to smoking behaviour and its effect is causal;
2. The correlation coefficient ρ is statistically significant while the coefficient on education in the smoking equation is not. In this case education is endogenous and the correlation between education and smoking behaviour is driven by unobserved heterogeneity: it is the so-called "third variable hypothesis" emphasised by Farrell and Fuchs (1982);
3. Both the correlation coefficient ρ and the coefficient on education in the smoking equation are significant. In this case although education is endogenous with smoking, it also has a causal impact on smoking behaviour. The estimates of ρ and of the causal effect of education will also give an idea of the relative importance of the two alternative explanations, i.e. "third variable hypothesis" vs. causal effects;

4. The correlation coefficient and the coefficient on education in the smoking equation are both insignificant. In this case our analysis will not support any of the hypotheses outlined in the literature review.

Like most of the previous literature based on IV, our analysis will only shed light on the presence of causal effects of education on smoking behaviour while we will not be able to distinguish among the different pathways through which this influence takes place. As we have said, different channels could be involved, such as those based on health-related information, endogenous time preference or the economic returns to health. It should also be noted that by excluding from the RHS of the smoking equation current or past job-related variables — such as qualification, sector or income that are likely to be affected by education — our study is necessarily constrained to estimate the “total effect” of education on smoking.¹⁰

5 Empirical Results

We start our empirical analysis by estimating count models treating the higher education dummy as exogenous. Since these models represent a benchmark for our later analysis we include all the variables that we use in the ES model and that are selected using Wald tests.¹¹ Given the model specification in section 4 an incidence rate (IR) is defined as:

$$IR = \exp(\mathbf{x}_i' \boldsymbol{\beta} + \theta E_i). \quad (4)$$

The IR is the rate at which an event occurs, and in our case can be roughly interpreted as the “speed” at which smoking takes place in the unit of time, i.e. a day in our case. IRs multiplied by the time of exposure (e.g., number

¹⁰However, Grossman and Kaestner (1997) and Grossman (2000) conclude from their surveys that a remarkable portion of the gross schooling effect is not accounted for by income or occupation, which are the other two main components of socio-economic status.

¹¹The Wald tests are available upon request from the authors. We selected regressors by estimating the smoking and HE equations separately and keeping only statistically significant regressors at 10% level or more.

of days) gives the number of occurrences of smoking in the period of interest, in our case the number of cigarettes smoked. Differences in rates of smoking between two individuals i and j with different characteristics can be evaluated using Incidence Rate Ratios (IRR) that are defined as:

$$IRR = \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta} + \theta E_i)}{\exp(\mathbf{x}'_j \boldsymbol{\beta} + \theta E_j)}. \quad (5)$$

If two individuals have the same characteristics and, therefore, smoke at the same speed the IRR equals one. Hence, a way to evaluate the effect of individual characteristics on smoking is to compute their impact on IRRs. In the case of a dichotomic characteristic, such as the dummy for higher education E_i , the effect on smoking for a generic individual i can be evaluated by comparing the change in the IRR produced by E_i :

$$\Delta IRR(E_i) = \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta} + \theta)}{\exp(\mathbf{x}'_i \boldsymbol{\beta})} - \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta})}{\exp(\mathbf{x}'_i \boldsymbol{\beta})} = \exp(\theta) - 1. \quad (6)$$

In the case of continuous variables we computed the effect determined by a one-unit increase in the value of the variables.

We estimated two models, the first one is a simple poisson model and the second one a poisson model with unobserved heterogeneity (see section 4, it is the smoking equation in the ES model with $\rho = 0$). Since, as in the case of the wage returns to education, also non-pecuniary returns are likely to differ by gender, we estimated separate gender-specific models. The estimates from the simple poisson models reported in the first column of Tables 4 and 5 show that HE is significantly associated with a reduction in smoking rates both for males and females. The estimated effects are very similar for both genders. The change in the IRR is -0.246 for females and -0.247 for males, i.e. females with HE smoke 24.6% less cigarettes a day than those with lower education, while the equivalent effect is 24.7% for males. However, the simple poisson model is based on the restrictive assumption that the mean of the distribution is equal to its variance and cannot account for overdispersion. Therefore, a first generalisation of the poisson model is to introduce unobserved heterogeneity. As we already said in section 4 unobserved het-

erogeneity can be present and controlled for even in the case we consider a poisson model with exogenous HE. The second column of Tables 4 and 5 reports changes in IRRs. When controlling for unobserved heterogeneity, the effect of HE on IRRs remains highly statistically significant. However, the effect of HE increases for females and decreases for females compared to the simple poisson model. Indeed, women with HE smoke 44.6% less than those with lower education while men with HE smoke 17.6% less than those with lower educational qualifications.

Although considering only current smokers reduces the degree of unobserved heterogeneity among individuals, which is more likely to be severe when pooling in the same analysis both smokers and non-smokers, a correlation between the unobservable variables driving HE decisions and smoking habits cannot be excluded on theoretical grounds. In section 2 we have seen some possible sources for this correlation, some examples are differences in individual discount rates or in innate ability. In general any unobserved variable affecting both investments in education and smoking decisions is likely to make it the HE endogenous with respect to smoking behaviour. For this reason, estimating a model where HE is treated as an endogenous variable represents a major improvement with respect to the poisson regression models that we already estimated. In particular, by estimating our model we will be able to assess whether the effect of HE on smoking can be interpreted as a “causal effect” or simply reflects a spurious correlation induced by omitted variables.

Although our model is formally identified through functional form, in order to have also an “economic identification” at least one variable affecting HE must be excluded from the smoking equation. The covariates to be included in the HE and in the smoking equations were selected estimating the two equations separately, performing Wald tests and excluding all variables not significant at the 10% level. Therefore, in general the model is identified by more than one exclusion restriction. However, as we already said

Table 4 Effect of explanatory variables on the IRR of smoking - Females

Variable	poisson			poisson with unobserved heterogeneity		
	ΔIRR		s.e.	ΔIRR		s.e.
<i>Education at age 30</i>						
HE	-0.246	***	0.015	-0.446	***	0.041
<i>Parents</i>						
no mother	0.368	***	0.118	1.584	*	0.950
no father	0.055		0.046	0.145		0.138
<i>Friends smoke</i>						
some of them	-0.026		0.052	-0.071		0.285
none of them	-0.190	***	0.040	-0.264		0.189
missing info.	-0.183	***	0.050	-0.450	**	0.189
<i>Mother's interest in child's education</i>						
little interest	0.120	***	0.026	0.319		0.276
cannot say	0.138	***	0.025	0.467	***	0.097
missing info.	0.050		0.031	0.401		0.354
<i>Father's smoking habits</i>						
non-smoker	-0.159	***	0.014	-0.143		0.208
missing info.	-0.167	***	0.047	-0.288	*	0.167
<i>Mother's smoking habits</i>						
non-smoker	-0.073	***	0.017	-0.128	*	0.068
missing info.	-0.155	***	0.038	-0.182		0.117
<i>Other smokers in the family</i>						
no	-0.108	***	0.021	-0.198	***	0.076
missing info.	0.088	**	0.041	0.064		0.333
<i>Siblings</i>						
no. elder siblings	0.026	***	0.008	-0.005		0.025
no. younger sibilings	0.020	**	0.010	0.063		0.075
<i>Region of residence</i>						
North East	0.009		0.078	0.121		0.362
North West	0.032		0.074	-0.116		0.169
Yorkshire and Humberside	-0.167	***	0.062	-0.345	**	0.150
East Midlands	-0.168	***	0.065	-0.363	*	0.217
West Midlands	0.082		0.079	-0.063		0.258
East	-0.101		0.066	-0.365		0.325
London	-0.086		0.067	-0.376	***	0.099
South East	-0.050		0.068	-0.158		0.219
South West	-0.050		0.072	-0.192		0.124
Wales	-0.045		0.073	-0.163		0.228
Scotland	0.075		0.078	0.021		0.177
<i>School intake</i>						
% father's from social class I	-0.005	***	0.001	-0.007		0.006
missing info.	0.022		0.037	0.158		0.147
σ^2	-		-	1.216	***	0.093
No. obs.	1754			1754		
Wald test ^(a)	1195.74(0.00)			722.22(0.00)		
Log-likelihood	-91142			-61470		

Note. Standard errors are computed using the Delta method. All the explanatory variable, except HE, refer to age 10. The reference individual did not specify the residence in the 10-year follow-up survey, has less than HE at age 30, most friends smoking at age 10, a mother very intested in her education, a father and a mother who smoked and other smokers in her family at age 10. ^(a) Overall significance test for the exclusion of all regressors but the constant from the smoking equation (p-value in brackets). ***significant at 1%; **significant at 5%; *significant at 10%.

Table 5. Effect of explanatory variables on the IRR of smoking - Males

Variable	poisson		poisson with unobserved heterogeneity			
	ΔIRR		s.e.	ΔIRR	s.e.	
<i>Education at age 30</i>						
he	-0.247	***	0.013	-0.176	***	0.066
<i>Parents</i>						
no mother	0.286	**	0.118	1.185	***	0.372
no father	0.051		0.046	0.009		0.103
<i>Friends smoke</i>						
some of them	0.016		0.038	0.071		0.116
none of them	-0.014		0.035	-0.017		0.074
missing info.	-0.009		0.048	-0.056		0.095
<i>Mother's interest in child's education</i>						
little interest	0.075	***	0.021	0.026		0.047
cannot say	0.048	**	0.020	0.093		0.064
missing info.	0.117	***	0.029	0.194		0.125
<i>Father's smoking habits</i>						
non-smoker	-0.098	***	0.013	-0.062	*	0.032
missing info.	-0.240	***	0.040	-0.250	***	0.054
<i>Mother's smoking habits</i>						
non-smoker	-0.104	***	0.013	0.004		0.036
missing info.	-0.146	***	0.038	-0.084		0.101
<i>Other smokers in the family</i>						
no	0.065	***	0.024	0.052		0.045
missing info.	0.251	***	0.047	0.573	***	0.142
<i>Siblings</i>						
no. elder siblings	-0.002		0.007	0.002		0.024
no. younger sibilings	0.022	**	0.009	0.002		0.038
<i>Region of residence</i>						
North East	-0.072		0.057	0.055		0.156
North West	-0.254	***	0.044	0.034		0.149
Yorkshire and Humberside	-0.145	***	0.051	0.141		0.189
East Midlands	-0.131	**	0.054	0.429		0.337
West Midlands	-0.140	***	0.051	0.014		0.167
East	-0.153	***	0.052	0.086		0.198
London	-0.137	***	0.052	-0.003		0.159
South East	-0.156	***	0.049	0.090		0.175
South West	-0.222	***	0.049	-0.023		0.203
Wales	-0.158	***	0.052	0.052		0.163
Scotland	-0.194	***	0.048	0.012		0.183
<i>School intake</i>						
% father's from social class I	-0.002	***	0.000	-0.002		0.001
missing info.	-0.044		0.031	-0.049		0.075
σ^2	-		-	0.976	***	0.062
No. obs.		1980		1980		
Wald test ^(a)		846.14(0.00)		170.41(0.00)		
Log-likelihood		-11694		-74070		

Note. Standard errors are computed using the Delta method. All the explanatory variable, except HE, refer to age 10. The reference individual did not specify the residence in the 10-year follow-up survey, has less than HE at age 30, most friends smoking at age 10, a mother very intested in her education, a father and a mother who smoked and other smokers in his family at age 10. ^(a) Overall significance test for the exclusion of all regressors but the constant from the smoking equation (p-value in brackets). ***significant at 1%; **significant at 5%; *significant at 10%.

in section 4, we consider as our main identifying variables parents' education, mainly for two reasons. Firstly, the literature on education has shown strong intergenerational links between parents' and children's education and therefore parents' education is likely to be a good predictor of children's education. Secondly, we have seen from the literature review in section 2 that parents' education is generally not a significant predictor of children's smoking behaviour once parents' smoking behaviour is controlled for. Since we will include in the smoking equation some covariates for parents' smoking habits, our exclusion restrictions are likely to be valid.

Table 6 reports for females the results of the ES model using parents' education as exclusion restrictions. The coefficient on parents' educational qualifications show that they are strong predictors of their children's levels of education (see Table 8 in Appendix). Women whose mothers have a university degree are 38.52 percentage points more likely to get a HE with respect to those whose mothers have less than O-level education, while women whose fathers have HE are 17.1 percentage points more likely to obtain HE. There are other factors that are significantly correlated with HE such as ethnicity, mother's interest in a child's education (assessed at age 10) and parents' social class. Our model shows a significant positive correlation between the unobservables affecting HE and those affecting smoking. Since it is hard to say what is generating this correlation, we can only provide here some speculative answers. The economic literature have emphasised two main reasons for finding a negative, rather than a positive, correlation between the unobservables affecting education and smoking. A first example is unobserved ability. Ability affects a child's likelihood to go on in HE and for several reasons can affect smoking, e.g. by increasing the amount of medical information to which an individual has access or an individual's labour incomes and consequently the potential loss due to the diseases caused by smoking. A second example is an individual's discount rate. Individuals with high discount rates are likely to invest less in education and engage in health-damaging behaviour, such as smoking. However, it must be noted that a

Table 6 Effect of explanatory variables on the IRR of smoking, ES model
- Females

Variable	ES poisson model		
	ΔIRR		s.e.
<i>Education at age 30</i>			
HE	-0.561	***	0.084
<i>Parents</i>			
no mother	0.502		0.428
no father	0.071		0.161
<i>Friends smoke</i>			
some of them	-0.089		0.153
none of them	-0.281	**	0.111
missing info.	-0.288	*	0.153
<i>Mother's interest in child's education</i>			
little interest	0.103		0.106
cannot say	0.127		0.101
missing info.	0.087		0.121
<i>Father's smoking habits</i>			
non-smoker	-0.217	***	0.051
missing info.	-0.199		0.167
<i>Mother's smoking habits</i>			
non-smoker	-0.147	**	0.060
missing info.	-0.253	**	0.123
<i>Other smokers in the family</i>			
no	-0.182	***	0.070
missing info.	0.109		0.152
<i>Siblings</i>			
no. elder siblings	0.036		0.028
no. younger sibilings	0.040		0.036
<i>Region of residence</i>			
North East	0.083		0.320
North West	0.187		0.328
Yorkshire and Humberside	-0.198		0.236
East Midlands	-0.158		0.250
West Midlands	0.129		0.322
East	-0.155		0.256
London	-0.053		0.265
South East	-0.062		0.261
South West	-0.065		0.279
Wales	-0.070		0.284
Scotland	0.241		0.348
<i>School intake</i>			
% fathers from social class I	-0.006	**	0.002
missing info.	-0.051		0.146
ρ	0.192	**	0.086
σ^2	1.096	***	0.043
No. obs.	1754		
Wald test ^(a)	467.89(0.00)		
Wald test parents' education ^(b)	9.66(0.47)		
Log-likelihood	-6903		

Note. Standard errors are computed using the Delta method. All explanatory variables, except HE, refer to age 10. The reference individual did not specify the residence in the 10-year follow-up survey, has less than HE at age 30, most friends smoking at age 10, a mother very intested in her education, a father and a mother who smoked and other smokers in her family at age 10. ^(a) Overall significance test for the exclusion of all regressors but the constant from both equations of the ES model (p-value in brackets). ^(b) Test for the exclusion of parents' education from the smoking equation, performed in the ES poisson model including parents' education in both equations ($\chi^2(10)$, p-value in brackets). ***significant at 1%; **significant at 5%; *significant at 10%.

Table 7 Effect of explanatory variables on the IRR of smoking, ES model
- Males

Variable	ES poisson model		
	ΔIRR		s.e.
<i>Education at age 30</i>			
he	-0.621	***	0.112
<i>Parents</i>			
no mother	0.454		0.674
no father	0.020		0.171
<i>Friends smoke</i>			
some of them	0.008		0.143
none of them	-0.095		0.120
missing info.	-0.095		0.166
<i>Mother's interest in child's education</i>			
little interest	0.035		0.105
cannot say	0.064		0.099
missing info.	0.213		0.144
<i>Father's smoking habits</i>			
non-smoker	-0.158	***	0.053
missing info.	-0.418	***	0.147
<i>Mother's smoking habits</i>			
non-smoker	-0.123	**	0.058
missing info.	-0.141		0.159
<i>Other smokers in the family</i>			
no	0.045		0.098
missing info.	0.452	**	0.210
<i>Siblings</i>			
no. elder siblings	-0.006		0.032
no. younger sibilings	0.021		0.037
<i>Region of residence</i>			
North East	-0.033		0.220
North West	-0.322	**	0.152
Yorkshire and Humberside	-0.149		0.191
East Midlands	-0.026		0.223
West Midlands	-0.193		0.188
East	-0.114		0.203
London	-0.123		0.205
South East	-0.146		0.185
South West	-0.261		0.181
Wales	-0.175		0.198
Scotland	-0.200		0.180
<i>School intake</i>			
% father's from social class I	-0.004		0.002
missing info.	-0.080		0.124
ρ	0.281	**	0.127
σ^2	1.149	***	0.047
No. obs.	1980		
Wald test ^(a)	354.66(0.00)		
Log-likelihood	-8245		

Note. Standard errors are computed using the Delta method. All explanatory variables, except HE, refer to age 10. The reference individual did not specify the residence in the 10-year follow-up survey, has less than HE at age 30, most friends smoking at age 10, a mother very intested in her education, a father and a mother who smoked and other smokers in his family at age 10. ^(a) Overall significance test for the exclusion of all regressors but the constant from both equations of the ES model (p-value in brackets). ***significant at 1%; **significant at 5%; *significant at 10%.

positive correlation between ability and smoking cannot be excluded on theoretical grounds, as greater earnings ability may increase the demand for cigarette consumption through an income effect or increase access to higher quality medical services in case of bad health.¹²In our HE equation we control for a proxy of an individual's ability (the British Ability Scales score)³, while some included variables may be proxies of an individual's discount rate, such as parents' social class. Then, the effect of these unobservable variables may be partly captured by the variables included in our econometric specification. More importantly, there is a variable omitted from the two equations that could generate a positive correlation between their error terms. This variable is an individual's "health stock", another crucial component of her amount of human capital. Healthy individuals are likely to invest more in education and to achieve better educational results. At the same time, those individuals who are in better health can smoke more being at risk of compromising their health as compared to individuals who are in bad health. Evidence consistent with this interpretation is provided, for instance, in Clark and Etilé (2002) who found that past improvements in health while smoking are positively correlated with current cigarette consumption.¹⁴Table 6 shows the effect on the IRR of HE. Women with HE consume 56.1% less cigarettes than those with lower education. The effect turns out to be higher than that estimated in the models treating HE as exogenous and this depends on the already observed positive correlation between the unobservable variables affecting HE and those affecting smoking, respectively.

We have very similar results for males. Like for women, we estimate a statistically significant positive correlation between the error terms of the smoking and the HE equations and the effect of HE estimated in the ES

¹²In the smoking equation we do not include an individual's income since it is likely to be endogenous with respect to smoking and education.

¹³See Elliot and Pearson (1979).

¹⁴Although the BCS70 gathers information on an individual's current subjective health status, we do not include it in the smoking equation since it is likely to be endogenous with respect to an individual's education, which may increase an individual's "health stock".

model is higher than that obtained in the poisson models treating HE as an exogenous variable. For men HE reduces dairy cigarette consumption by 62.1% (see Table 7). The estimates of the HE equation are reported in Table 9

6 Concluding remarks

In this paper we have used an endogenous switching count model to estimate the effect of higher education (HE) on cigarette consumption. Unlike most of the previous literature which investigated the effect of education on smoking in the UK and treated education as an exogenous variable, we focused on HE and accounted for its potential endogeneity with smoking in order to recover ‘causal’ effects. To the best of our knowledge, there is only one other paper which has estimated ‘causal’ effects of education on smoking for the UK, that is Chandola et al. (2006) who found a strong effect of education on the probability of being a non-smoker using data from the 1958 British cohort. With respect to that paper, our analysis differs in several respects: we focus on current smokers and estimate the effect of HE on dairy cigarette consumption, use a different econometric strategy and data from a more recent British cohort born in 1970.

Our main findings are as follows:

- HE is endogenous with smoking, i.e. omitted variables are likely to induce correlation between the error terms of the smoking and higher education equations. Therefore, when assessing the effect of education on smoking an ES framework must be preferred to simple count models where educational attainment is exogenous;
- unlike what is emphasised by the previous literature, we found a statistically significant positive correlation between the error terms of the smoking and higher education equations. Indeed, the previous literature has mostly emphasised unobservable variables that are likely to

induce a negative correlation such as an individual's discount rate or innate ability. We argue that in our analysis these effects are likely to be partly captured by observed and included variables, such as an individual's social class or the British Ability Scales score. The explanation that we put forward for the observed positive correlation is that an individual's unobserved 'health stock' is likely to positively affect his/her educational attainment and his/her cigarette consumption, since healthier individuals can smoke at higher rates being less at risk of compromising their health status. Evidence consistent with this interpretation is found by Clark and Etilé (2002).

- we estimate sizeable causal effects of HE on cigarette consumption. HE reduces dairy cigarette consumption by 56.1% for women and 62.1% for men.

We think that our findings on the existence of important 'health returns' to HE may be important to inform the current debate on the financing of HE in the UK. Indeed, although the recent increase in student fees and in the incidence of private funding of the HE system have been motivated by the high private economic returns to HE, our analysis shows important non-pecuniary returns to HE. If these 'health returns' are not considered by individuals who plan to invest in HE, the amount of education chosen could be lower than the optimal level. Moreover, these 'health returns' are likely to produce a positive externality for the collectivity as a healthier population also means important savings in public health expenditures. Therefore, the optimal amount of public funding to HE should be determined also considering these important non-pecuniary returns.

Appendix

Table 8. Probability of having HE - Females

Variables	m.e.		s.e.
<i>Parents</i>			
nomother	-0.074		0.090
nofather	-0.087		0.063
<i>Time spent talking to parents</i>			
not very often	0.132		0.126
quite a lot	0.178	*	0.104
missing info.	0.116		0.147
<i>Mother's interest in child's education</i>			
little interest	-0.114	***	0.028
cannot say	-0.056	*	0.029
missing info.	-0.081	**	0.040
<i>Mother's smoking habits</i>			
non-smoker	0.003		0.025
missing info.	0.221	**	0.101
<i>Other smokers in the family</i>			
no	0.111	***	0.032
missing info.	0.119		0.078
<i>Mother's education</i>			
O-level	0.033		0.033
A-level	0.256	***	0.071
Professional	0.180	***	0.067
Degree	0.385	***	0.104
missing info.	-0.042		0.047
<i>Father's education</i>			
O-level	-0.015		0.035
A-level	0.003		0.045
Professional	0.144	*	0.085
Degree	0.171	***	0.063
missing info.	-0.092	**	0.045
<i>Ethnic group</i>			
non-European	0.448	***	0.115
missing info.	-0.056		0.121
<i>British Ability Scales</i>			
score	0.006	***	0.001
missing info.	0.533	***	0.097
<i>Social class^(a)</i>			
II	-0.011		0.062
III NM	-0.107	*	0.056
III M	-0.121	**	0.059
IV	-0.154	***	0.046
V	-0.110	*	0.066
unemployed	-0.191	***	0.036
OLF	-0.107	*	0.059
missing info.	-0.035		0.112
<i>Siblings</i>			
no. elder sibilings	0.010		0.011
no. younger sibilings	-0.021		0.018
<i>Days absent from school</i>			
no. of days	0.001		0.002
missing info.	0.112	*	0.068
No. obs.			1754
Wald test parents' education ^(b)			56.57(0.00)

Note. This Table reports marginal effects (m.e.) on the probability of acquiring HE estimated using the ES poisson model. Standard errors are computed using the Delta method. All explanatory variables refer to age 10. The reference individual has a non-smoking mother with less than O-level education and very interested in her daughter's education, a father with less than O-level education, comes from the professional social class, is of European ethnicity and lives in a family where other members, different from the parents, smoke. ^(a) Social classes are: I (professional), II (intermediate), III NM (skilled non-manual), III M (skilled manual), IV (partly skilled), V (unskilled), unemployed, OLF (out of the labour force). ^(b) Test for the exclusion of parents' education from the HE equation, performed in the ES poisson model including parents' education in both equations ($\chi^2(10)$, p-value in brackets). ***significant at 1%; **significant at 5%; *significant at 10%.

Table 9. Probability of having HE - Males

Variables	m.e.	
<i>Parents</i>		
nomother	-0.06	
nofather	-0.07	
<i>Time spent talking to parents</i>		
not very often	0.23	
quite a lot	0.24	*
missing info.	0.27	
<i>Mother's interest in child's education</i>		
little interest	-0.07	***
cannot say	-0.05	*
missing info.	-0.05	
<i>Mother's smoking habits</i>		
non-smoker	0.00	
missing info.	0.16	
<i>Other smokers in the family</i>		
no	-0.01	
missing info.	0.01	
<i>Mother's education</i>		
O-level	0.03	
A-level	-0.07	
Professional	0.14	**
Degree	0.08	
missing info.	-0.03	
<i>Father's education</i>		
O-level	0.02	
A-level	0.10	**
Professional	0.06	
Degree	0.18	***
missing info.	-0.01	
<i>Ethnic group</i>		
non-European	0.17	**
missing info.	0.09	
<i>British Ability Scales</i>		
score	0.01	***
missing info.	0.64	***
<i>Social class^(a)</i>		
II	-0.11	***
III NM	-0.14	***
III M	-0.14	***
IV	-0.16	***
V	-0.16	***
unemployed	-0.20	***
OLF	-0.17	***
missing info.	-0.20	***
<i>Siblings</i>		
no. elder sibilings	-0.02	
no. younger sibilings	-0.01	
<i>Days absent from school</i>		
no. of days	0.00	*
missing info.	-0.01	
No. obs.	1980	

Note. This Table reports marginal effects (m.e.) on the probability of acquiring HE estimated using the ES poisson model. Standard errors are computed using the Delta method. All explanatory variables refer to age 10. The reference individual has a non-smoking mother with less than O-level education and very interested in her daughter's education, a father with less than O-level education, comes from the professional social class, is of European ethnicity and lives in a family where other members, different from the parents, smoke. ^(a) Social classes are: I (professional), II (intermediate), III NM (skilled non-manual), III M (skilled manual), IV (partly skilled), V (unskilled), unemployed, OLF (out of the labour force). ***significant at 1%; **significant at 5%; *significant at 10%.

References

- Adams, S. (2002). Educational attainment and health: Evidence from a sample of older adults. *Education Economics*, 10:97–109.
- Arendt, J. (2005). Does education cause better health? a panel data analysis using school reforms for identification. *Economics of Education Review*, 24(2):149–160.
- Arkes, J. (2003). Does schooling improve health? RAND Health Working Papers DRU-3051.
- Balia, S. and Jones (2005). Mortality, lifestyle and socio-economic status. HEDG Working Paper 05/02.
- Bantle, C. and Haisken-DeNew, J. (2002). Smoke signals: The intergenerational transmission of smoking behavior. DIW Discussion Papers no. 277.
- Becker, G. and Mulligan, C. (1997). The endogenous determination of time preference. *Journal of Political Economy*, 112:729–758.
- Becker, G. and Murphy, K. (1988). A theory of rational addiction. *Journal of Political Economy*, 96:675–700.
- Berger, M. and Leigh, J. (1989). Schooling, self-selection, and health. *Journal of Human Resources*, 24:433–455.
- Blaylock, J. and Blisard, W. (1993). Women and the demand for alcohol: Estimating participation and consumption. *The Journal of Consumer Affairs*, 27(2):319–334.
- Blow, L., Leicester, A., and Windmeijer, F. (2005). Parental income and children’s smoking behaviour: Evidence from the British household panel survey. IFS Working Papers, W05/10.

- Blundell, R., Dearden, L., Goodman, A., and Reed, H. (2000). Policy feature. the returns to higher education in Britain: Evidence from a British cohort. *Economic Journal*, 110(461):82–100.
- Bratti, M., Naylor, R., and Smith, J. (2005). Variations in the wage returns to a first degree: Evidence from the British cohort study 1970. IZA DP No. 1631.
- Bynner, J., Butler, N., Ferri, E., Shepherd, P., and Smith, K. (2002). The design and conduct of the 1999-2000 surveys of the national child development study and the 1970 British cohort study. CLS Cohort Studies Working Paper 1. London: Centre for Longitudinal Studies.
- Chandola, T., H. J. and Bartley, M. (2004). Socio-demographic predictors of quitting smoking: How important are household factors? *Addiction*, 90:770–777.
- Chandola, T., Clarke, P., Morris, J., and Blane, D. (2006). Pathways between education and health: A causal modelling approach. *Journal of the Royal Statistical Society, Series A*, forthcoming.
- Chevalier, A. and Lanot, G. (2002). The relative effect of family characteristics and financial situation on educational achievement. *Education Economics*, 10:165–182.
- Clark, A. and Etilé, F. (2002). Do health changes affect smoking? evidence from British panel data. *Journal of Health Economics*, 21:533–562.
- Contoyannis, P. and Jones, A. (2004). Socio-economic status, health and lifestyle. *Journal of Health Economics*, 23:965–995.
- Cragg, J. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, 39(5):829–844.

- De Walque, D. (2004). Education, information, and smoking decisions: Evidence from smoking histories, 1940–2000. World Bank Policy Research Papers No. 3362.
- Elliot, CD, M. D. and Pearson, L. (1979). *British Ability Scales, Manual 4: Tables of Abilities and Norms*. Windsor: NFER.
- Emery, S., White, M., and Pierce, J. (2001). Does cigarette price influence adolescent experimentation? *Journal of Health Economics*, 20:261–270.
- Ermisch, J. and Francesconi, M. (2001). Family matters: Impacts of family background on educational attainments. *Economica*, 68(270):137–156.
- Farrell, P. and Fuchs, V. (1982). Schooling and health : The cigarette connection. *Journal of Health Economics*, 1(3):217–230.
- Feinstein, L. (2000). The relative economic importance of academic, psychological and behavioural attributes developed in childhood. Centre for Economic Performance Discussion Papers no. 0443.
- Fersterer, J. and Winter-Ebmer, R. (2003). Smoking, discount rates, and returns to education. *Economics of Education Review*, 22(6):561–566.
- Forster, M. and Jones, A. (2003). Corrigendum: The role of tobacco taxes in starting and quitting smoking: Duration analysis of british data. *Journal of the Royal Statistical Society, Series A*, 166(3):441–442.
- Fuchs, V. (1982). Time preference and health: An explanatory study. In Fuchs, V., editor, *Economic Aspects of Health*, pages 93–120. University of Chicago Press.
- Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80:223–255.
- Grossman, M. (1975). *Household Production And Consumption*, chapter On The Correlation Between Health and Schooling, pages 147–223. Columbia University Press.

- Grossman, M. (2000). The human capital model. In Culyer, A. and Newhouse, J., editors, *Handbook of Health Economics, vol. 1A*, pages 347–408. Elsevier, Amsterdam.
- Grossman, M. (2005). Education and nonmarket outcomes. NBER Working Papers No. 11582.
- Grossman, M. and Kaestner, R. (1997). Effects of education on health. In Behrman, J. and Stacey, N., editors, *The Social Benefits of Education*, pages 69–123. University of Michigan Press, Ann Arbor (Michigan).
- Jones, A. (1989). A double-hurdle model for cigarette consumption. *Journal of Applied Econometrics*, 4(1):23–39.
- Kenkel, D. (1991). Health behavior, health knowledge, and schooling. *Journal of Political Economy*, 99:287–305.
- Kenkel, D., Lillard, D., and Mathios, A. (2004). The roles of high school completion and ged receipt in smoking and obesity. mimeo.
- Kenkel, D. and Terza, J. (2001). The effect of physician advice on alcohol consumption: Count regression with an endogenous treatment effect. *Journal of Applied Econometrics*, 16(2):165–184.
- Labeaga, J. (1999). A double-hurdle rational addiction model with heterogeneity: Estimating the demand for tobacco. *Journal of Econometrics*, 93:49–72.
- Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *Review of Economic Studies*, 72:189–221.
- Micklewright, J. (1989). Choice at sixteen. *Economica*, 56:25–39.
- Miranda, A. and Rabe-Hesketh, S. (2005). Estimation of ordinal response models, accounting for sample selection bias. 11th UK Stata Users Group Meeting Proceedings.

- Nayga, R. (1999). A note on schooling and smoking: The issue revisited. *Education Economics*, 7(3):253–258.
- Nerín, I., Guillén, D., Mas, A., and Crucelaegui, A. (2004). Evaluation of the influence of medical education on the smoking attitudes of future doctors. *Arch Bronconeumol*, 40(8):341–347.
- Office for National Statistics (1999). *Tracking People: A guide to longitudinal social sources*. London: Office for Statistics.
- Powell, L., Tauras, J., and Ross, H. (2005). The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior. *Journal of Health Economics*, 24:950–968.
- Rabe-Hesketh, S. and Skrondal, A. (2002). Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal*, 2(1):1–21.
- Rosenzweig, M. and Schultz, T. (1983). Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *Journal of Political Economy*, 91(?):723–746.
- Ross, C. and Mirowsky, J. (1999). Refining the association between education and health: The effects of quantity, credential and selectivity. *Demography*, 36:445–460.
- Sander, W. (1995a). Schooling and quitting smoking. *Review of Economics and Statistics*, 92:344–348.
- Sander, W. (1995b). Schooling and smoking. *Economics of Education Review*, 14(1):23–33.
- Sander, W. (1998). The effects of schooling and cognitive ability on smoking and marijuana use by young adults. *Economics of Education Review*, 17(3):317–324.

- Terza, J. (1998). Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects. *Journal of Econometrics*, 84:129–154.
- Thaler, R. and Sherfrin, H. (1981). An economic theory of self control. *Journal of Political Economy*, 89:392–406.
- von Laffert, G. (1998). Rauchen, gesellschaft und staat-konsumanomalien, wohlfahrtseffekte und staatlicher regulierungsbedarf im zusammenhang mit dem zigarettenkonsum. Deutscher Universitaets-Verlag, Wiesbaden.
- Winkelmann, R. (1998). Count data models with selectivity. *Econometric Reviews*, 17:339–359.
- Wolfe, B. and Haveman, R. (2003). Social and nonmarket benefits from education in an advanced economy. In Kodrzycki, Y., editor, *Education in the twenty-first century: Meeting the challenges of a changing world*, number no. 47 in Conference Series, pages 97–131. Boston: Federal Reserve Bank of Boston.
- Yen, S. and Jensen, H. (1996). Determinants of household expenditures on alcohol. *Journal of Consumer Affairs*, 30(1):48–67.