CPB Discussion Paper

No 139

The Effect of Education on Smoking Behaviour

New evidence from smoking durations of a sample of twins

Pierre Koning, Dinand Webbink and Nicholas G. Martin

CPB Netherlands Bureau for Economic Policy Analysis
Van Stolkweg 14
P.O. Box 80510
2508 GM The Hague, the Netherlands

Telephone +31 70 338 33 80
Telefax +31 70 338 33 50
Internet www.cpb.nl

ISBN 978-90-5833-435-0

Abstract in English

This paper analyses the causal effect of education on starting and quitting smoking, using longitudinal data of Australian twins. The endogeneity of education, censoring of smoking durations and the timing of starting smoking versus the timing of completion of education are taken into account by using the flexible Mixed Proportional Hazard (MPH) specification. Unobserved effects in the specification are assumed to be twin specific and possibly correlated with completed education years. In addition, we use various unique control indicators reflecting the discounting behaviour of individuals that may affect both the smoking decision and the number of education years. In contrast to previous studies in our model specification, differences in the number of education years cannot explain differences in smoking behaviour at young ages. We find one additional year of education to reduce the duration of smoking with 9 months, but no significant effect of education on starting smoking. The effect of education on quitting smoking largely confines to male twins. This suggests that education policies that succeed in raising the level of education may improve public health through an increase of smoking cessation, but are not effective in preventing smoking at young ages.

Key words: Smoking, duration models, education.

JEL code: C41, I21.

Abstract in Dutch

In dit onderzoek staat de vraag centraal in hoeverre de opleiding de beslissing om met roken te starten of stoppen bepaalt. Omdat veel eigenschappen van mensen van invloed zijn op zowel de beslissing om (verder) te leren als op die om te gaan roken, maken we daarbij gebruik van gegevens van (Australische) tweelingen. Met deze gegevens kunnen we rekening houden met gemeenschappelijke omgevings- en genetische kenmerken van de ondervraagde tweelingen. Bovendien bevatten de gegevens informatie over de wijze waarop ondervraagden in de praktijk hun beslissingen nemen – bijvoorbeeld intuïtief of zonder de gevolgen ervan op lange termijn te overzien. Onze analyse leert dat starten met roken niet of nauwelijks afhangt van het niveau van opleiding. Deze beslissing wordt doorgaans genomen op jonge leeftijd, als verschillen in onderwijsjaren nog klein zijn. Onderwijs leidt wel tot het eerder stoppen met roken. Een jaar onderwijs vermindert de tijd dat men rookt met negen maanden. Dit effect is alleen gevonden bij mannen.

Steekwoorden: Roken, duurmodellen, scholing.

Contents

Summary

This paper analyses the causal effect of education on starting and quitting smoking, using longitudinal data of Australian twins. For this purpose, we estimate hazard rate models for smoking and non-smoking durations, measured from the age of 12. Our contribution to the literature is essentially twofold. First, to our knowledge the current analysis is the first to use a Mixed Proportional Hazard (MPH) specification to assess the impact of education on starting and quitting smoking, taking explicit account of the apparent correlation between unobserved (twin) effects to obtain consistent estimates. In our analysis, the twin aspect of our data is used to control for unobserved heterogeneity, reflecting unobserved genetic and family or household determinants. As we have two spells for each twin pair, this also offers the advantage that the identification of the individual heterogeneity distribution is stronger than in the case of univariate spells. Our analysis also includes age and duration effects, as well as various unique indicators reflecting the discounting behaviour of individuals.

The second contribution of this paper is with respect to the specification of education years. The conventional way of modelling this variable in the literature is to include the number of completed education years as a time invariant variable in the hazard specification. In our data, however, the vast majority of smoking durations start before schooling has been completed. As a result, taking the number of education years that is ultimately completed yields inaccurate estimates that are biased by individual ability and group behaviour ─ factors also affecting the decision to start smoking at young ages. We argue that this results in overestimation of the effect of education on the hazard of starting smoking. We therefore include the number of education years as a time variant variable that increases with age, up to the level of education years that is ultimately completed. Consequently, differences in education years are only small at young ages and cannot explain differences in smoking behaviour at young ages.

Our main finding is that a higher educational attainment increases the probability of smoking cessation. One additional year of education reduces the duration of smoking with 9 months. The effect of education on quitting smoking seems largely confined to male twins – for females the impact is only small and insignificant. In contrast to previous studies, we find no effect of education on the decision to start smoking. This difference in findings can be explained by the way we model the education variable. Previous studies used years of education as a time constant variable and without controlling for (twin) fixed effects.

We conclude that education policies that succeed in raising the level of education may improve public health through an increase of smoking cessation. Raising the level of educational attainment may be not effective in preventing smoking at young ages. The decision to start

smoking is mostly taken while attending school and seems to be determined by factors which are also important for the decision to invest in human capital, such as time preferences.

1 Introduction[1]

Tobacco smoking is the leading preventable cause of death and disease in many countries. For instance, in the US smoking causes more than 440,000 deaths per year and adults who smoke cigarettes on average die 14 years earlier than non-smokers.[2] For Australia it has been estimated that 15 % of all deaths were due to tobacco smoking and many deaths occurred before the age of 65.[3] In 2004-2005 26% of Australian men and 20% of Australian women were current smokers. The highest rates of smoking for men were reported in the 18-24 years age group (34%) and for women in the 25-34 years age group (27%). Thus policies reducing the proportion of people that start smoking or decrease the duration of smoking yield large potential returns for public health. It is often argued that education may therefore be an attractive policy, as this may lead to greater awareness of health risks in later life.

Many studies find better educated individuals indeed to have a better health and a lower risk of mortality (Cutler & Lleras-Muney 2006). However, it is not clear whether this strong association reflects a causal effect of education. The main challenges in estimating causal effects of education on smoking behaviour in particular concern the endogeneity of education, the censoring of smoking durations and the issue that most starting decisions of smoking occur before schooling has been completed. The endogeneity of schooling is addressed in several studies by using an instrumental variable approach.[4] For instance, Sander (1995) studies the effect of education on the decision to quit smoking with parental schooling as an instrument for schooling. He finds schooling to have a substantial positive effect on quitting smoking.[5] Two recent studies exploit variation in educational attainment induced by the Vietnam draft avoidance behaviour that increased college attendance in the US (Walque, de 2007, Grimard & Parent 2007). Both find that education decreases the probability of ever having smoked substantially, but the evidence on quitting smoking is mixed.

Obviously, the major disadvantage of the instrumental variable studies on smoking is that they do not take account of the longitudinal character of smoking decisions. To our knowledge, only Douglas & Hariharan (1994) and Douglas (1998) address this issue explicitly by estimating

[1] We would like to thank Bas Straathof, Aico van Vuuren and seminar (series) participants at the CPB and the Institute for Research on Poverty (IRP) at the University of Wisconsin-Madison for their very useful comments to this paper.

[2] http://www.cdc.gov/tobacco/data_statistics/fact_sheets/health_effects/tobacco_related_mortality.htm.

[3] http://www.abs.gov.au/ausstats/abs@.nsf/mf/4831.0.55.001.

[4] Various recent studies that focus on health outcomes other than smoking also use an instrumental variable approach (Currie & Moretti 2003, Lleras-Muney 2005, Oreopoulos 2006, Kenkel et al. 2006, Lindeboom et al. 2007, Mazumder 2007, Albouy & Lequien 2008). Typically institutional differences in education systems or educational reforms are used as instruments for education. Most studies find that more schooling leads to better health.

[5] Kenkel et al. (2006) however question the validity of parents schooling as instruments.

duration models for smoking.[6] Using US data from the National Health Interview Survey (NHIS), Douglas & Hariharan (1994) find the hazard of starting smoking to decrease with about 10% for each additional year of schooling. Douglas (1998) obtains similar results for the starting decision with more recent waves of the NHIS. Still, there are two major concerns on the consistency of these effects. First, unobserved heterogeneity of hazard rates is ignored, leading to specification errors in the (genuine) duration dependency and endogeneity biases in the education effect. Second, completed education is used as a time constant explanatory variable, whereas smoking decisions will be relevant during the schooling period as well. Generally, one may expect the inclusion of completed education years to overestimate the effect of education on starting smoking.

This paper is the first that simultaneously takes into account the endogeneity of education, censoring of smoking durations and the timing of starting smoking versus the timing of completion of education. We estimate hazard rate models for smoking and non-smoking durations using longitudinal data of Australian twins. Our analysis of the effect of education on starting and quitting smoking takes explicit account of the apparent correlation between unobserved effects. To obtain consistent estimates we use a Mixed Proportional Hazard (MPH) specification (Abbring & Van den Berg 2003; Van den Berg 2001). In our analysis the twin aspect of our data is used to control for unobserved heterogeneity, reflecting unobserved genetic and family determinants (see e.g. Hougaard et al. 1992). Moreover, as we have two spells for each twin pair, this also offers the potential advantage that the estimation of the individual heterogeneity distribution is more efficient than in the case of univariate spells (Van den Berg 2001; Honoré 1993). In our analysis we also include age and duration effects and various unique indicators reflecting the discounting behaviour of individuals. These variables may affect both the smoking decision and the number of education years (see Fersterer & Winter-Ebmer 2003 and Khwaja et al. 2007).

Next to the use of twin-specific effects, a second major contribution of this paper is particularly relevant to the starting decision on smoking. The conventional way of modelling the education variable in the literature is to include the number of completed education years as a time invariant variable. However, the majority of smoking durations starts before schooling has been completed. As a result, taking the number of education years that is ultimately completed yields inaccurate estimates that are biased by individual ability and group behaviour factors that also affect the decision to start smoking at young ages. We argue that this results in overestimation of the effect of education on the hazard of starting smoking. Instead we include the number of

[6] Duration models of smoking have also been used in studies focusing on the effects of tobacco prices and tobacco regulation on the starting and quitting decision of smoking (Tauras & Chaloupka 1999, Forster & Jones 2001, Decicca et al. 2007, Malhotra & Boudarbat 2008, Kidd & Hopkins 2004).

education years as a time variant variable that increases with age, up to the level of education years that is ultimately completed.

Our main finding is that a higher educational attainment increases the probability of smoking cessation. One additional year of education reduces the duration of smoking with 9 months. This finding is robust with respect to various specification assumptions. The effect of education on quitting smoking seems largely confined to male twins – for females the impact is only small and insignificant. In contrast to previous studies, we find no effect of education on the decision to start smoking. This difference in findings can be explained by the fact that we model the education variable as a time varying variable.

The remainder of this paper is organized as follows. The next section describes the data that are used. Section 3 explains the empirical strategy that is followed, and estimation results and robustness checks are shown in Sections 4 and 5, respectively. Section 6 concludes.

2 Data description

The Canberra sample

In this study we use data from a cohort of twins of the Australian Twin Register which is called the older cohort (or the 'Canberra sample'). The data were collected in two mail surveys, in 1980-1982 and 1988-1989. The sample consists of all 5,967 twin pairs aged over 18 years enrolled in the Australian National Health and Medical Research Council Twin Registry at the time of the first survey. In the first survey 3,808 complete pairs have participated, and in the follow-up survey 2,934 twin pairs have responded (Miller et al. 1995). The surveys gathered information on the respondent's family background (parents, siblings, marital status, and children), socioeconomic status (education, employment status and income), health behaviour (body size, smoking and drinking habits), personality, and feelings and attitudes. Zygosity was determined by a combination of diagnostic questions plus blood grouping and genotyping.

For our analysis we have selected a sample of 5,378 individuals from complete twin pairs below the age of 60 for which we observe smoking behaviour and educational attainment. Table 2.1 shows the sample means and proportions for relevant background characteristics and outcome variables for this sample. The main independent variable here is educational attainment. In both surveys this variable was measured using a seven point scale: less than 7 years schooling; 8-10 years schooling; 11-12 years schooling; apprenticeship, diploma, certificate; technical or teachers' college; university, first degree; university, postgraduate degree. These categories have been recorded as 5, 9, 11.5, 13, 15 and 17 years of education, respectively (Miller et al. 1995). Other covariates for our analysis include mother's and father's education, age, birth weight and personality traits. We include birth weight to control for differences within pairs of identical twins, as recent research has shown that this variable is an important predictor of later outcomes in life (Black et al. 2007).[7]

[7] Birth weight has been measured in the first survey (1980) and information about the age of respondents has been derived from the Twin Registry.

Table 2.1 Summary statistics of covariates selected twins sample (N=5,378)

	Mean	Standard deviation
Individual characteristics		
Gender (male = 1)	0.34	0.47
Identical twin	0.49	0.50
Age (in 1980)	31.8	10.9
Birth weight (in grams)	2,503	577
Education years (in 1988)	11.8	2.5
9 years	0.27	0.44
11.5 years	0.38	0.49
13 years	0.13	0.33
15-17 years	0.08	0.26
Education years of father	9.9	3.0
Education years of mother	9.5	2.4
Smoking at time of interview ($R = 3$)	0.22	0.42
Has smoked ($R = 2$)	0.21	0.40
Never smoked ($R = 1$)	0.57	0.41

The Canberra sample includes about 13 questions on personality traits that are informative on the time preferences of respondents (see the appendix to this paper). It could be argued that both investments in education and decisions on smoking behaviour are determined by similar general measures of time preference (Khwaja et al. 2007). Respondents with high discounting rates are likely to quit schooling early, whereas they may be less inclined to stop smoking. We therefore include the following four (retained factor) indicators in our analysis, which are represented by the factors “taking decisions quickly”, “making decisions on instinct”, “having debts and no savings” and “running out of money”. The derivation of these four indicators is presented in the appendix to this paper.

The last rows in Table 2.1 show that 22 % of our sample reports being a smoker at the time of the interview and 21 % reports to have smoked. A comparison with available population statistics indicates that the proportion of smoking individuals in our sample is somewhat lower than in the population. In particular, Hill et al. (1991) report that 30 % of men and 27 % of women were smoking in 1989. In addition, the distribution of self-reported education for the total sample of 1989 respondents has been contrasted with census data from the Australian Bureau of Statistics for a sample of men and women with a comparable age range (Baker et al. 1996). This comparison showed a slight upward bias in educational attainment in the sample of 1989 respondents, especially for men. The lower smoking prevalence in our sample might therefore be attributed to this upward bias in educational attainment and age restrictions used for the estimation sample (below the age of 60).

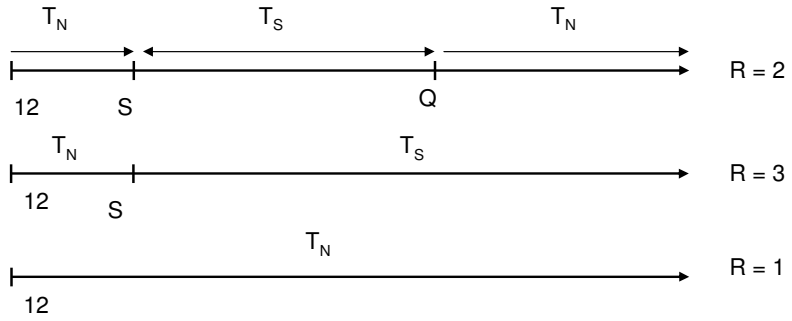
Smoking durations

Key to our analysis is the measurement of smoking behaviour. For this purpose, we use the following items:

- Smoking during lifetime: respondent has never smoked, is an ex-smoker or currently a smoker. We denote this variable by *R*, representing respondent type 1, 2, or 3, respectively. The fractions of these groups are equal to 57%, 21% and 22%, respectively (see also Table 2.1).
- Age of starting smoking (for $R = 2, 3$).
- Age of quitting smoking (for $R = 2$).
- Number of years that the respondent has smoked (for $R = 2, 3$).

With these four items, smoking durations can be derived either from the starting and quitting dates, or from the reported number of smoking years that have passed (i.e. the fourth item). In our analysis, we use the first option, allowing us to determine non-smoking durations as (intervening) spells as well.[8] Figure 2.1 shows that this results in three possible combinations of successive smoking and non smoking durations that start from the age of 12.[9] We denote these by T_s and T_n, respectively. When constructing the duration data, our key assumption is that respondents smoke or have smoked only one (major) period in their life. Thus, time intervals where respondents have stopped smoking only temporarily are not measured. We return to this issue when discussing the estimation results in Section 4.

Figure 2.1 **Combinations of smoking and non smoking durations as a function of age, censored and uncensored**



[8] We have used the third item (the reported number of smoking years) to test for the sensitivity of our estimation results with respect to measurement errors – see also footnote 10.

[9] In the data, the age of 12 is the minimum age at which smoking durations start, which is the same as in Douglas (1998).

Table 2.2 Smoking and non-smoking durations in selected sample (standard deviations in brackets)

	Smoking durations		Non-smoking durations	
	Complete	Censored	Complete	Censored
Number of observations	1,217	1,105	2,246	3,056
Duration (years)	13.4	21.1	5.6	29.1
	(9.7)	(9.5)	(3.6)	(11.1)
Age at start	17.5	17.4	12.0	12.0
	(3.5)	(3.8)	(.)	(.)
Age at end	30.9	38.4	17.6	41.1
	(10.3)	(9.9)	(3.6)	(11.1)
Self reported smoking durations	12.8	18.7		
(1,195 and 1.076 observations)	(9.5)	(9.6)		

Table 2.2 presents the sample statistics of the smoking and non-smoking durations, and figures 2.2 and 2.3 depict the observed hazard rates of starting and quitting smoking as a function of the elapsed durations. The hazard rates of starting and quitting smoking are derived from the full sample and a sub-sample of 2,322 observations, respectively. Figure 2.2 shows that smoking durations mostly start at younger ages, between the age of 12 and 22, which is consistent with other studies (Malhotra & Boudarbat 2008, Kidd & Hopkins 2007). The average starting age is 18 years. It should further be noted that the value averages of self reported smoking durations are very similar to those that are obtained from the responded beginning and starting dates. This consistency check suggests that measurement errors are not an important concern.

As the variation in starting age is limited, the separate (non-parametric) identification of duration and age-effects is more cumbersome for non-smoking durations than for the smoking durations. Figure 2.2 also suggests that observed hazard rates are strongly driven by selection effects, i.e. almost all those respondents that were likely to start smoking anyhow have started doing this by the age of 22. Most respondents are interviewed at older ages than at the start of smoking durations, so it seems that underreporting at younger ages is not very important here.

When considering the pattern of quitting hazards in Figure 2.3, the picture is mixed. During the first 15 years of smoking, the likelihood of quitting gradually increases, whereas there is a gradual decrease in the years thereafter. This finding is similar to e.g. Kidd & Hopkins (2003) and Douglas (1998). Essentially, this hump shaped pattern may result from three sources: habit formation, selection effects and age effects. When modelling the quitting hazard, we therefore allow for all these effects in the MPH specification.

Figure 2.2 **Observed (non-parametric) hazard rate of starting smoking.**

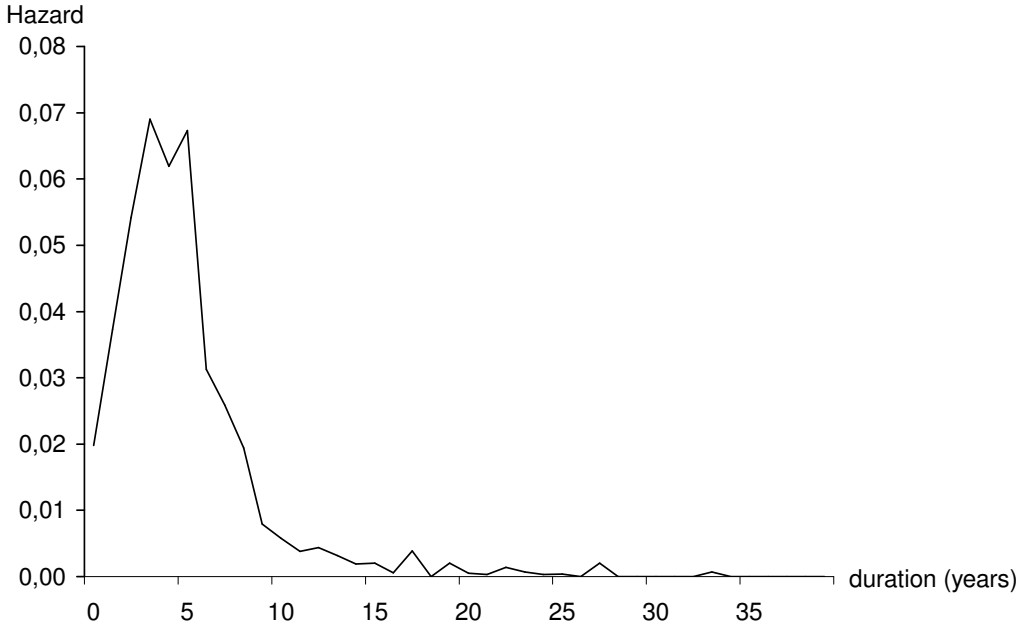
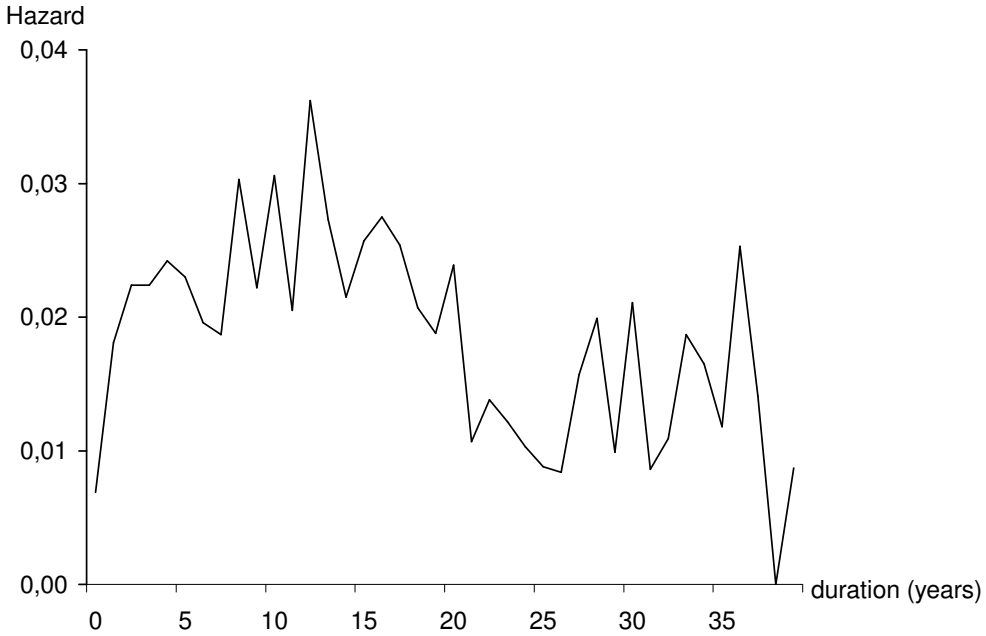


Figure 2.3 **Observed (non-parametric) hazard rates of quitting smoking**



3 Empirical strategy

The MPH model

Our research strategy takes advantage of the longitudinal information in our data. That is, we use hazard rate or — stated differently — duration models to examine the impact of education on smoking and non-smoking spells. Within the context of our analysis, the hazard rate is defined as the rate at which the event of starting or quitting smoking takes place over a short period of time $[T,T+dt]$, given that this event has not occurred so far, up to time T.

$$\theta = \Pr(T < t < T + dt \mid t \geq T) \tag{3.1}$$

In the (non-)smoking model, the time interval dt is normalized to one year. For both the starting (S) and the quitting (Q) decision d, we specify the hazards as a mixed proportional hazard (MPH) rate model (see e.g. Van den Berg 2001):

$$\theta^{d}_{ijt,\tau} = \lambda_{0d}(t)\ \exp\{\ \alpha\ educ_{ij\tau} + \boldsymbol{X}_{ij\tau}\boldsymbol{\beta}^{d}\ \}\ \psi^{d}(\tau)\ \upsilon_{j}^{d} \tag{3.2}$$

where i indicates the individual ($i = 1..I$), j indicates the twin pair ($j = 1..$ I/2), t is the elapsed duration and τ indicates calendar time. Equation (2) shows that the MPH specification consists of four parts, representing the genuine (or state) duration dependence λ_0, variation in hazards due to observed individual and twin specific characteristics $\boldsymbol{X}$, education years ($educ$), calendar time effects ψ and unobserved twin pair specific characteristics υ, respectively.

Duration dependence in the (non-)smoking decision is specified by the baseline hazard, $\lambda_{0d}(t)$. A sufficiently flexible baseline specification is needed to take account of habit formation in the quitting hazard. We model genuine duration dependence in the quitting hazard as a (semi-parametric) polynomial function of the elapsed duration. With only one polynomial, this specification is equivalent to the familiar Weibull model for duration dependence. We perform Likelihood ratio tests on additional polynomials. For the starting hazard rate, we abstract from duration effects, as habit formation is less relevant here and most smoking durations start in only a relatively short time span, causing problems with respect to the separate identification of duration and age effects.

Obviously, the parameter of interest of our model is the number of education years ($educ$) for individual i per twin pair j, measured at (calendar) time τ. Variation in observed values of education years thus essentially comes from three sources: variation in completed education years between twin pairs, variation in completed education years within twin pairs, and variation per individual in the number of education years over time. The third source of variation results from the fact that durations are measured from 12 years of age, when schooling

has not been completed yet. We further include various other time variant and invariant independent variables in our model, both for the starting and quitting decision. Variables that do not vary over time are cohort dummies indicating the period the respondent has been born (before 1945; between 1945 and 1955; and after), gender, birth weight and the four proxies for the discounting behaviour of the respondents. The age of respondents varies with calendar time τ. Finally, calendar time effects itself are modelled as dummies affecting all respondents equally at the same time intervals. We distinguish between three periods: until 1970, 1970 to 1980, and the years after.

Identification of the effect of education

For both the starting and quitting hazards unobserved twin effects are taken into account by the time-invariant random effect υ^d. In order to allow for correlation between this effect and education per twin pair, we use the modified random effects (RE) framework proposed by Mundlak (1978) and Chamberlain (1982). Within the context of the current analysis, the intuition behind this approach is that the smoking and education decisions for both individuals of the same twin pair are driven by similar time invariant unobserved factors. Including the average completed education per twin pair in the regression would then control for potential endogeneity biases that are due to these unobserved twin effects. This approach requires the strict exogeneity assumption to hold for education with respect to smoking – that is, the decision of starting or quitting smoking (itself) cannot affect the (future) number of education years. Thus we specify the twin specific effects in the following auxiliary regression:

$$\ln \upsilon^d{}_j = (\gamma^d / 2) \sum_{i=1,2} \max_\tau (educ_{ij,\tau}) + \ln \xi^d{}_j \qquad (3.3)$$

with $d = Q, S$. In equation (3), the maximum value of education years per individual i equals the number of completed education years. So our key assumption is that the average value of this variable per twin controls for any correlation between twin fixed effects in the (completed) education variable and the smoking hazard rates, while the residual term ξ^d is assumed to be uncorrelated with education years.

Due to the multiplicative MPH structure, the average value of completed education years per twin pair can simply be added to the other control variables in our model. By adding average values of education years as a controls for unobserved heterogeneity, we disentangle the well known "within" from the "between" estimators of both coefficients. Thus the coefficient estimate of education years, α, is identified from variation "within" twin pairs – both in completed education years and variation in education years per individual over time.

Maximum likelihood estimation

In order to estimate the model in equations (2) and (3), we need to make closing assumptions on the distribution of the twin random effects $\boldsymbol{\xi}^d$ for both hazard rates. We do so by modelling the distribution of $\boldsymbol{\xi}$ in a non-parametrical fashion, assuming K mass points for $\boldsymbol{\xi}^d$, with probability weights $P_1, P_2, \ldots, 1 - P_1 - \ldots - P_{K-1}$, respectively (Heckman & Singer 1984). Thus, the unknown distribution of $\boldsymbol{\xi}^d$ is represented by a distribution with a finite number of points of support, where the first point of support is normalized to $\{0, 0, 0, 0\}$.

The parameters of interest in our model include the polynomials for duration effects, the vector value of $\boldsymbol{\beta}^d$, the calendar time dummies and the points of support and the respective weights of $\boldsymbol{\xi}^d$. All these parameters are estimated with Maximum Likelihood. Conditional upon the points of support $\boldsymbol{\xi}^d$ ($d = S,Q$) and for respondent type R, there are three possible outcomes for the individual log likelihood contribution Λ :

$$\Lambda_{ij}(T_{N1}, T_s, T_{N2}, R \mid \boldsymbol{\xi}^S, \boldsymbol{\xi}^Q) = L_{ij}(T_{N1} \mid \boldsymbol{\xi}^S) \times L_{ij}(T_S \mid \boldsymbol{\xi}^Q)^{\mathrm{I}(R \neq 3)} \times L_{ij}(T_{N2} \mid \boldsymbol{\xi}^S)^{\mathrm{I}(R = 1)} \qquad (3.4)$$

where T_{N1}, T_{N2} and T_s indicate the (two) non-smoking and smoking durations and I is an dummy indicator representing whether the respondent has smoked ($R = 1$), is currently smoking ($R = 2$) or has never smoked ($R = 3$). Note that two non-smoking durations are observed only for R=2. L indicates the likelihood of the observed durations (in parentheses) and equals the product of the survival probability of the duration and the hazard rate (if no censoring applies). The joint likelihood $\boldsymbol{\Lambda}$ is defined as the product of all likelihood contributions per twin pair, integrated over the (non-parametric) mass point distribution of unobserved effects:

$$\boldsymbol{\Lambda} = \prod_{\mathrm{j}} \sum_{\mathrm{k}} P_k \left\{ \Lambda_{1j}(\,.\mid \xi_k^S, \xi_k^Q) + \Lambda_{2j}(\,.\mid \xi_k^S, \xi_k^Q) \right\}. \qquad (3.5)$$

In order to determine the number of mass points for both models, we start by estimating the model without any unobserved twin effects ($K = 1$). Subsequently, we increase the number of points of support K iteratively, so as to improve the fit of the model. We perform a Likelihood Ratio test to determine the optimal K, that is, the number of points of support where the inclusion of an additional point of support, together with an additional weight, improves the likelihood significantly.

4 Main estimation results

Table 4.1 shows the Maximum Likelihood estimation results of equations (2) and (3) with two mass points for the twin unobserved effects in both the quitting and starting hazard. When specifying the model with two mass points, we started with imposing the restriction that there is no correlation between the two hazard rates. It turns out that both MPH models with $K = 3$ (three mass points) or without restrictions on the correlation between the unobserved effects do not improve the goodness of fit substantially. We therefore restrict the attention to the model outcomes with two uncorrelated mass points for both the starting and quitting hazard.

Table 4.1 Estimation results MPH model (non-)smoking durations (standard errors in parentheses; *, ** and * denote significance at the level of 10%, 5% and 1%).**

	Starting hazard		Quitting hazard	
Baseline Hazard				
Constant	– 6.188	(0.354)***	– 10.341	(1.014)***
ln(duration)			– 1.110	(0.322)***
Idem, squared			0.001	(0.092)
Individual and twin characteristics				
Education years	– 0.009	(0.018)	0.100	(0.022)***
Completed education, average per twin pair	– 0.068	(0.013)***	0.032	(0.021)*
ln (age-11)	7.611	(0.281)***	2.365	(0.926)***
Idem, squared	– 2.147	(0.074)**	0.035	(0.192)
Education years father	0.006	(0.013)	–0.027	(0.020)*
Idem, missing dummy	0.358	(0.224)*	0.436	(0.325)*
Education years mother	– 0.003	(0.016)	0.007	(0.027)
Idem, missing dummy	– 0.172	(0.237)	0.262	(0.323)
Born 1945-1955	– 0.762	(0.108)***	– 0.173	(0.145)
Born after 1955	– 1.013	(0.126)***	– 0.904	(0.219)***
Female	– 0.409	(0.065)***	0.036	(0.101)
Birth weight (kg)	0.206	(0.052)***	0.086	(0.077)
Idem, missing dummy	0.211	(0.075)***	0.482	(0.108)***
Discounting variables				
Decide quickly	0.186	(0.044)***	– 0.027	(0.075)
Decide instinctively	0.183	(0.059)***	– 0.193	(0.096)**
Debts, no savings	0.198	(0.099)**	– 0.132	(0.175)
Out of money	0.191	(0.036)***	– 0.099	(0.060)**
Calendar time effects				
1970-1980	– 0.406	(0.093)***	0.232	(0.128)**
> 1980	– 1.566	(0.184)***	1.103	(0.169)***
Mass point distribution parameters				
P	0.714	(0.051)***	0.361	(0.179)**
ln(v)	– 2.218	(0.059)***	2.063	(0.145)***
N = 5,378				
Log likelihood	– 8,677.6		– 3,728.5	

The estimates in table 4.1 show that the starting decision of smoking is unaffected by the number of education years. This contrasts to Douglas (1998), who finds the impact on starting to be negative, significant and equal to 14%. It is likely that our result can be explained by our estimation method, that exploits the "within-twin" variance, rather than cross sectional variation in completed education years only. For instance, if the culture of starting smoking among students is more prevalent in schools that prepare for low educated jobs, cross sectional estimation is likely to lead to overestimation of the education effect. Using our method, however, these effects cannot be picked up by our education variable, as in the relevant time period education differences are low. Especially at young ages the influence of peers on smoking – where we control for – is substantial (Harris & Lopez-Valcarel 2004). We also re-estimated the model for sub samples of twin couples with two brothers or sisters only. The impact of education on starting is somewhat higher for males – with a coefficient value of 0.045 (0.041) of male twins compared to 0.006 (0.027) for female twins – but insignificant in both cases.

Our model results suggest that various covariates have an effect on the decision to start smoking. Smoking durations are more likely to start at young ages (with a peak at 18 years of age), younger cohorts, women and individuals with high birth weight. Moreover, all four indicators for the time preferences have the expected sign and are significant. We also find the decision to start smoking to have become less likely as from 1970. Unobserved twin heterogeneity is captured by a mass point for twins with a relatively high starting hazard (with a probability weight of 71%) and those with a starting hazard that is close to zero (with a probability weight of 29%).

As to the quitting decision, we find a significant effect of education on quitting smoking. For each additional year of schooling, the quitting hazard increases with about 10%. This effect implies a reduction in the expected smoking duration with about 9 months, with an average smoking duration of 21 years in our sample. The coefficient estimate of education years on the quitting is somewhat smaller than that of Douglas (1998), who finds a coefficient value equal to 12% with US data. In contrast to the starting decision, most quitting decisions are made when education is complete. Thus it seems that education explains (future) smoking decisions, rather than the decision of starting smoking. When estimating the model for sub samples of male and female twin couples, we find this effect to be confined to males only – with coefficient estimates of 0.131 (0.039)*** and 0.024 (0.035) for male and female twins, respectively. This finding is in line with previous studies on gender differences in smoking. For instance, Bauer et al. (2007) find a strong effect of education on smoking for males and no effect for women.[10]

[10] They also report that 86% of the gender difference in the number of cigarettes smoked per day is due to differences in the estimated coefficients and only 14% due to different characteristics.

The psychological literature suggests that traditional sex roles can explain gender differences in smoking (Waldron 1991).

When considering the other covariates, quitting smoking is more likely among respondents that have been born after 1955. Respondents that are more prone to make decisions on their instinct show a smaller hazard of quitting smoking and for all respondents quitting has increased after 1970. Unobserved twin effects are controlled for by one mass point for twins that are unlikely to quit (with a probability of 36%) and those who are likely to do so (with 62% probability). As we have argued earlier, in our specifications we allow for genuine duration dependence in the hazard of quitting smoking only. We find such habit (or addiction) effects to be important – that is, the likelihood of quitting decreases strongly with the smoking duration. At the same time, the likelihood of ongoing smoking durations decreases as a result of ageing. This can be explained by increased health problems, making quitting smoking more likely. To increase the flexibility of the age profile and the duration effects, we also estimated specifications with third order polynomials, but this did not improve the likelihood substantially.

5 Robustness checks

To test the robustness of our findings we performed various sensitivity checks on our duration models. Table 5.1 presents the estimated effect of years of schooling on quitting and starting smoking for various specifications, with the attention predominantly being focussed on the identification assumptions we make on the twin effects.[11]

Unobserved heterogeneity

Obviously, the identification following our models relies upon the assumption that unobserved and correlated heterogeneity effects in smoking and education can be controlled for by including the average value of completed education years per twin pair. In addition to this, we assumed unobserved effects in the hazard rates only to vary among twin pairs. As a first robustness check, we changed this second assumption, modelling the unobserved heterogeneity distribution as individual effects instead of twin effects. Thus, the alternative hypothesis is that there is no correlation of individual effects within twin couples (see model variant (i)). As the table shows, this decreases the efficiency of our estimates, particularly for the quitting hazards. In particular, the number of repeated spells per stratum decreases substantially, causing the efficiency of the estimated distribution of unobserved effects to reduce.[12] At the same time, we find the new assumption to increase the fit of the smoking duration model to the data substantially, suggesting that the assumption that individual and twin effects are fully correlated is probably too strong here. It thus appears that the effects of twin pairs are less relevant for the quitting decision, which is made at higher ages. Still, for both the non-smoking and smoking durations this model variant does not change our estimated coefficients substantially.

We also tested the robustness of our results by zooming into the sub sample of identical twins in our data. The assumption that unobserved effects are equal per twin pair is probably less restrictive for this sub-sample. Again, this did not result in significant or substantial differences from the outcomes of the benchmark model.

As a third robustness check on the unobserved effects assumptions, we followed a non-parametric approach proposed by Ridder & Tunali (1999) and Griffith et al. (2007). This approach entails the use of a fixed effects specification for (MPH) duration models, without making further *a priori* distributional assumptions on unobserved effects. The key aspect of this

[11] We also tested the robustness of our model to measurement errors in reported smoking and completed education years, but the impact was negligible. For the quitting hazard we replaced the smoking durations that were inferred from the reported starting and ending dates by those directly reported by the respondents ("how many years have you smoked during your life"). We also replaced the reported education measures of twins by those that were reported by the other twin brother or sister. This also led tot virtually the same estimation results.

[12] With modelling unobserved heterogeneity as individual effects, repeated spells are observed only for the sub-sample of respondents that have quitted smoking. For this group, we observe an uncensored non-smoking duration prior to the smoking duration and a censored non-smoking spell after the smoking duration.

approach is that information on the duration data per twin pair is transformed, so as to eliminate twin fixed effects. It can be shown that the probability that the first individual for each twin pair starts (or quits) smoking first, does not depend on the twin specific effect. Instead, this probability can be explained by a Conditional Logit model where only within-twin differences in values of $\boldsymbol{X}$ are included. To use this Conditional Logit specification for estimating this probability, however, two practical conditions have to be met. First, only pairs of uncensored twin durations can be used. Second, as calendar time effects are present in our model, both durations need to start at the same date. Thus, the fixed effects approach can be used only on (1,470) uncensored non-smoking durations of twin pairs that all start at the age of 12. For this sub sample we find the estimated impact of education on the starting hazard to be insignificant and almost equal to the estimate of the benchmark model with random effects. We therefore conclude that this outcome is robust with respect to distributional assumptions on the unobserved twin effects.[13]

We also checked the sensitivity of our results by restricting the observed variation in education years to cross sectional variation – between twin pairs – in completed education years only (model variant (iv)). For the quitting model outcomes, this restriction hardly affects the coefficient estimates of education. This is not surprising, as smoking durations mostly take place when education is in fact completed. As to the starting decision, the coefficient estimate however increases substantially and becomes –0.12 (0.013). This coefficient estimate is remarkably close to that obtained by Douglas (1998), who (also) exploits cross sectional variation in education years only. We thus conclude that cross sectional (twin) variation alone – measured in completed education years – leads to inconsistent estimates of the smoking effect on starting.

As a final robustness check, we re-estimated our model without the discounting variables as controls. In particular, if discounting behaviour would be affected by education years – and this may also be the intended mechanism to affect smoking decisions – the inclusion of these variables may cause the education effect to be underestimated. The estimation results that result from such a specification however are virtually equivalent to those obtained from the benchmark model, suggesting that the discounting measures are exogenous.

[13] To estimate such a fixed effects model, the likelihood contributions of twin pairs need to be weighed with the inverse of the probability that the observation would be unobserved due to censoring (Griffith et al., 2007). In particular, we performed a two step procedure to obtain estimates. First, we used the (non-parametric) Kaplan-Meier estimates to determine the probability of censoring. Second, in the estimation of the Conditional Logit model for the sub-sample twin pairs with uncensored durations we weighed observations by the inverse of the estimated probability of censoring.

Table 5.1	Estimated education effect: robustness checks (standard errors in parentheses; *, ** and *** indicate significance at 10%, 5% and 1%)	
	Starting hazard	Quitting hazard
Benchmark model: unobserved twin effects	– 0.009	0.100
	(0.018)	(0.022)***
	LL = 8,677.6	LL = 3,728.5
(i) Unobserved individual effects (i.e. no twin correlation)	0.020	0.075
	(0.019)	(0.037)**
	LL = 8,917.5	LL = 3,697.9
(ii) Sub-sample of monozygotic twins (N=2,732)	0.017	0.101
	(0.029)	(0.034)***
(iii) MPH with twin fixed effects (N=1,470)	– 0.010	
	(0.022)	
(iv) Completed education as (time invariant) variable	– 0.120	0.095
	(0.013)***	(0.021)***
(v) Discounting variables excluded as controls	– 0.012	0.102
	(0.019)	(0.022)***

Implied effect on smoking incidence

From the previous findings we may conclude that the effect of education on smoking runs through the quitting decision, rather than the initiation of smoking. We find the implied average effect of one additional year of schooling on the expected smoking duration of respondents to be equal to 8.6 months, which is a reduction of 4.1% (the average expected duration is 21 years and 3 months). As about one fifth of the interviewed twins smokes at the time of the interview, the corresponding decrease in the incidence of smoking is 0.9%-point. We compared this outcome with the effect of education on smoking incidence that follows from the estimation of a twin Fixed Effects Linear Probability Model (LPM) for smoking incidence at the time of the interview. This yielded a (significant) parameter estimate of 1.3%-point, which does not differ significantly from the effect that is inferred from the duration models.[14]

Compared to previous studies, the implied effect on smoking incidence we find seems low. Grimard & Parent (2007) find that one year of education reduces the incidence of smoking with approximately 8%-points.[15] However, this estimate should be interpreted as a local average treatment effect that applies for a group that had not started smoking upon completion of high school and who decided to attend college in order to avoid being drafted. This seems a special

[14] It should be noted here that the LPM estimates are identified from cross-sectional, within twin variation only and not exploiting the variation in education levels when smoking starts. Thus, it does not come as a surprise that the LPM estimates are higher than the incidence estimate that is inferred from the duration models.

[15] Moreover, Parent & Grimard (2007) find the (total) effect of high school completion on different measures for smoking to amount to 40 to 76%-point.

group as most individuals start smoking between the age of 12 and 18. Grimard & Parent also argue that peer effects may have been particularly strong for this group, as going to college and being with non-smokers might have been unanticipated and very different experience than not going to college. Our estimates are based on all levels of education and it seems less likely that peer effects play a role.

6 Conclusions

The main conclusion of this paper is that a higher educational attainment increases the probability of smoking cessation, rather than decreasing the probability of starting smoking. One additional year of education reduces the duration of smoking with 9 months. This finding is robust with respect to different identifying assumptions and seems largely confined to male twins. In contrast to previous studies, we find no effect of education on the decision to start smoking. This difference in findings can be explained by the modelling of the education variable, enabling us to exploit both within twin education differentials in completed education years and individual variation in education years over time. Compared to the quitting model outcomes, this additional variation over time strengthens the identification of the model considerably. Previous studies that used years of education as a time constant variable found larger effects. However, as the majority of smoking decisions is made between the age of 12 and 18, when the accumulation of human capital has not been completed yet, using education as a time constant variable is questionable.

A cautionary note on the twin aspect of the data in our study may be in order. The proportion of individuals in our sample that reported being a smoker at the time of the interview is somewhat lower than in the population and the educational attainment in our sample is slightly higher than in the population. Although various studies find samples of twins to be representative to the population at large on outcomes – such as educational attainment, IQ, psychiatric symptoms or personality (Baker et al. 1996, Calvin et al. 2009, Kendler et al. 1986, Webbink et al. 2008 – it is possible that our results might therefore not be fully transferable to the population at large.

The main findings from this paper suggest that education policies that succeed in raising the level of education may improve public health through an increase of smoking cessation. Raising the level of educational attainment may be not effective in preventing smoking at young ages. The decision to start smoking is mostly taken while attending school and seems to be determined by factors which are also important for the decision to invest in human capital, such as time preferences.

References

Abbring, J.H. and G. J. Van den Berg, 2003, The non-parametric identification of treatment effects in duration models, *Econometrica*, 71, 1491-1517.

Albouy, V. and L. Lequien, 2008, Does compulsory education lower mortality? *Journal of Health Economics*, forthcoming.

Arendt, J.N., 2005, Does education cause better health? A panel data analysis using school reforms for identification, *Economics of Education Review*, 24(2), 149-160.

Baker, L.A., S.A. Treloar, C.A. Reynolds, A.C. Heath and N.G. Martin, 1996, Genetics of educational attainment in Australian Twins: Sex differences and secular changes, *Behaviour Genetics*, 26(2), 89-102.

Bauer, T., S. Göhlman and M. Sinning, M., 2007, Gender differences in smoking behaviour, *Health Economics*, 2007, 16(9), 895-909.

Berg, van den, G. J., 2001, Duration Models: Specification, Identification, and Multiple Duration, in *Handbook of Econometrics*, Vol. 5, ed. by J. J. Heckman and E. Leamer. Amsterdam: North-Holland.

Boudarbat, B. and N. Malhotra, 2008, The Hazards of Starting the Cigarette Smoking Habit, *International Journal of Economic Perspectives*, 2(4).

Calvin, C., C. Fernandes, P. Smith, P.M. Visscher and I.J. Deary I.J., 2009, Is there still a cognitive cost of being a twin in the UK?, *Intelligence*, in press.

Chamberlain, G., 1982, Multivariate regression models for panel data, *Journal of Econometrics* 18, 5-46.

Currie, J. and E. Moretti, 2003, Mother's education and the intergenerational transmission of human capital: evidence from college openings, *Quarterly Journal of Economics*, 118(4), 1495-1532

Cutler, D.M. and A. Lleras-Muney, 2006, Education and health: evaluating theories and evidence, NBER Working Paper 12352.

Decicca, Ph., D. A. Kenkel, A. Mathios, Y-J. Shin and J-Y. Lim, 2007, Youth smoking, cigarette prices and anti-smoking sentiment, *Health Economics*, 17(6), 733-749.

Douglas, S. and G. Hariharan, 1994, The hazard of starting smoking: estimates from a split population duration model, *Journal of Health Economics*, 13(2), 213-230.

Douglas, S., 1998, The duration of the smoking habit, *Economic Inquiry*, 36(1), 49-64.

Fersterer, J. and R. Winter-Ebmer, 2003, Smoking, discount rates, and return to education, *Economics of Education Review*, 22(6), 561-566.

Forster M. and A. M. Jones, 2001, The role of tobacco taxes in starting and quitting smoking: Duration analysis of British data, *Journal of the Royal Statistical Society: Series A*, 164(3), 517-547(31).

Grimard, F. and D. Parent, 2007, Education and smoking: Were Vietnam war draft avoiders also more likely to avoid smoking?, *Journal of Health Economics*, 26(5), 896-926.

Grossman, M., 2005, Education and nonmarket outcomes, NBER Working Paper 11582.

Harris, J.E. and B. Lopez-Valcarel, 2004, Asymmetric Social Interaction in Economics: Cigarette Smoking Among Young People in the United States 1992-1999, NBER Working Paper 10409.

Hill D.J., V.M. White and N.J. Gray, 1991, Australian patterns of tobacco smoking in 1989, *Medical Journal of Australia*, 154, 797-801.

Honoré, B.E., 1993, Identification results for duration models with multiple spells, *Review of Economic Studies*, 60(1), 241-246.

Hougaard, P., B. Harvald and N.V. Holm, 1992, Measuring the similarities between the lifetimes of adult Danish twins born between 1881–1930, *Journal of the American Statistics, Association*, 87, 17–24.

Jones, A.M., 2007, Panel data methods and applications to health economics, in: T.C. Mills and K. Petterson, *Palgrave Handbook of Econometrics, Volume 2: Applied Econometrics*, Palgrave MacMillan.

Kendler K.S, A.C. Heath , N.G. Martin and L.J. Eaves, 1986, Symptoms of anxiety and depression in a volunteer twin population: the etiologic role of genetic and environmental factors, *Archives of General Psychiatry*, 43, 213-221.

Kidd, M.P. and S. Hopkins, 2004, The Hazards of Starting and Quitting Smoking: Some Australian Evidence, *Economic Record*, 80(249), 177-192.

Khwaja, A., D. Silverman and F. Sloan, 2007, Time preference, time discounting, and smoking decisions, *Journal of Health Economics*, 26(5), 927-949.

Kenkel, D., D. Lillard and A. Mathios, 2006, The roles of high school completion and GED receipt in smoking and obesity, *Journal of Labor Economics*, 24(3), 635-660.

Lee, S., 2005, Estimating Panel Data Duration Models with Censored Data, Cemmap working paper CWP 13/03, The Institute for Fiscal Studies Department of Economics.

Lindeboom, M., A. Llena-Nozal and B. Van der Klaauw, 2007, Parental education and child health: evidence form a schooling reform, Free University Amsterdam.

Lleras-Muney, A., 2005, The relationship between education and adult mortality in the United States, *The Review of Economic Studies*, 72(1), 189-221.

Malhotra, N. and B. Boudarbat, 2008, The Hazards of Starting the Cigarette Smoking Habit, *International Journal of Economic Perspectives*, 2(4).

Mazumder, B., 2007, How did schooling laws improve long-term health and lower mortality?, Federal Reserve Bank of Chicago, WP 2006-23 (revised January 24, 2007), Chicago, Il.

Mokdad, A.H., J.S. Marks, D.F. Stroup and J.L. Gerberding, 2004, Actual Causes of Death in the United States, 2000, *Journal of the American Medical Association*, 291, 1238-1245.

Mundlak, Y., 1978, On the Pooling of Time Series and Cross Section Data, *Econometrica*, 46(1), 69-85.

Oreopoulos, P., 2006, Estimating average and local average treatment effects of education when compulsory schooling laws really matter, *American Economic Review,* 96(1), 152-175.

Ridder, G. and I. Tunali, 1999, Stratified partial likelihood estimation, *Journal of Econometrics*, 92(2), 193-232.

Sander, W., 1995, Schooling and quitting smoking, *Review of Economics and Statistics*, 77(1), 191-99.

Spasojevic, J., 2003, Effects of education on adult health in Sweden: results from a natural experiment, Ph. D. Dissertation, City University of New York Graduate Center.

Tauras, J.A. and F.J. Chaloupka, 1999, Determinants of Smoking Cessation: An Analysis of Young Adult Men and Women, NBER working paper 7262.

Waldron, I., 1991, Patterns and causes of gender differences in smoking, *Social Science and Medicine*, 32(9), 989-1005.

Walque, D. de, 2007. Does education affect smoking behaviours? Evidence using the Vietnam draft as an instrument for college education, *Journal of Health Economics*, 26(5), 877-895.

Webbink, D., D. Posthuma, D. Boomsma, J. de Geus and P.M. Visscher, 2008, Do twins have lower cognitive ability than singletons? *Intelligence*, 36(6), 539-547.

Appendix: Factor analysis of discounting variables

Factor analysis of retained discounting variables: factor loadings and uniqueness of variance

Questions	Factor loading	Unique variance
(i) Taking decisions quickly		
"Do you often make decisions in the spur of the moment?" (YES)	0.42	0.82
"Have people said that sometimes you act too rashly?" (YES)	0.57	0.67
"I like to think about things for a long time before I make a decision." (NO)	0.67	0.55
"I usually think about all the facts before I make a decision." (NO)	0.50	0.75
(ii) Making decisions on instinct		
"I nearly always think about all the facts in detail before I make a decision, even when other people demand a quick decision." (NO)	0.40	0.84
"I often do things based on how I feel at the moment, without thinking how they were done in the past." (NO)	0.47	0.78
"I often follow my instincts, hunches, or intuition without thinking through all the details." (YES)	0.34	0.89
(iii) Having debts, no savings		
"Would being in debt worry you?" (NO)	0.21	0.96
"Do you think people spend too much time safeguarding their future with savings and insurances?"(YES)	0.19	0.96
"I am better at saving money than most people" (NO)	0.16	0.98
(iv) Running out of money		
"I often spend money until I run out of cash or get into debt from using too much credit." (YES)	0.70	0.51
"Because I so often spend too much money on impulse, it is hard for me to save money, even for special plans like a holiday." (YES)	0.71	0.50
"I enjoy saving more than spending it on entertainment or thrills." (NO)	0.31	0.90