Empir Econ
DOI 10.1007/s00181-014-0842-6

# The effect of education on smoking behavior: new evidence from smoking durations of a sample of twins

**Pierre Koning · Dinand Webbink · Nicholas G. Martin**

Received: 23 November 2012 / Accepted: 18 April 2014


**Abstract** This paper analyzes the effect of education on starting and quitting smoking. We use longitudinal data of Australian twins and estimate duration models for smoking and non-smoking durations. Our approach enables us to take account of the endogeneity of education, censoring of smoking durations, and the timing of starting smoking versus that of completion of education. We find that one additional year of education reduces the duration of smoking with 9 months but has no effect on the decision to start smoking. This finding is robust with respect to different identifying assumptions and seems largely confined to male twins.

**Keywords** Smoking · Duration models · Education

**JEL Classification** C41 · I21

## 1 Introduction

Tobacco smoking is the leading preventable cause of death and disease in many countries. For Australia it has been estimated that 15 % of all deaths were due to tobacco

P. Koning (✉)
Department of Economics, VU University Amsterdam, De Boelelaan 1105, P.O. Box 80510, 1081 HV Amsterdam, The Netherlands
e-mail: p.w.c.koning@vu.nl

D. Webbink
Department of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands

N. G. Martin
Queensland Institute of Medical Research, Brisbane, QLD, Australia

Published online: 20 July 2014

smoking, and many deaths occurred before the age of 65.[1] 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 may yield large returns for public health.

Many studies find better educated individuals to have a better health and a lower risk of mortality (Cutler and Lleras-Muney 2005). Investigating the effect of education on smoking, however, poses various empirical and methodological challenges. First, the level of schooling completed by individuals is typically not randomly assigned but influenced by a multitude of observed and unobserved factors. The endogeneity of education can bias estimates of the effect of education on schooling. This issue has been addressed in several studies that use an instrumental variable approach.[2] Two recent studies exploit variation in educational attainment induced by the Vietnam draft avoidance behavior that increased college attendance in the US (Walque 2007; Grimard and Parent 2007). Both find that education decreases the probability of ever having smoked substantially, but the evidence on quitting smoking is mixed.

Second, smoking decisions have a longitudinal character, with observed smoking durations that typically are incomplete. Estimates that do not take the censoring of smoking durations into account may therefore be inconsistent. In this respect, Douglas and Hariharan (1994) and Douglas (1998) have estimated duration models for the impact of education years on smoking.[3] Using US data from the National Health Interview Survey (NHIS), Douglas and 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.[4]

Third, many individuals already start smoking before completing their schooling. This implies that the effect of schooling on the starting decision of smoking is unclear. In addition, reverse causality might also play a role. In their seminal paper, Farrell and Fuchs (1982) concluded that differences in smoking behavior at age 24 could be fully explained by smoking differences at age 17, when all subjects were still in approximately the same grade. Tenn et al. (2010) elaborate on this idea by exploiting small education differences between similarly selected groups that are one year apart

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

[2] Various recent studies that focus on health outcomes other than smoking also use an instrumental variable approach (Currie and Moretti 2003; Lleras-Muney 2005; Oreopoulos 2006; Kenkel et al. 2006; Lindeboom et al. 2009; Mazumder 2007; Albouy and Lequien 2009). As to the effect on smoking, 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. Kenkel et al. (2006), however, question the validity of parents schooling as instruments.

[3] Walque (2010) also exploits the longitudinal character of smoking data but then focuses on the incidence of smoking. Rather than studying the effects of education, duration models of smoking have been used to estimate the effects of tobacco prices and tobacco regulation (Tauras and Chaloupka 1999; Forster and Jones 2001; Decicca et al. 2007; Boudarbat and Malhotra 2008; Kidd and Hopkins 2004).

[4] Bratti and Miranda (2009) are one of the few studies that take explicit account of endogeneity of smoking decisions by modeling both the decision to enroll in higher education and smoking intensity.

in their life cycle. Similar to Farrell and Fuchs, they conclude that starting smoking is not driven by education, but unobserved "third factors," like time preferences.

This paper aims to estimate the effect of education on smoking using longitudinal data from Australian twins. To our knowledge, this paper is the first that simultaneously takes into account the three abovementioned issues: (i) the endogeneity of education, (ii) censoring of smoking durations, and (iii) the timing of starting smoking versus the timing of completion of education. We estimate mixed proportional hazard rate models for smoking and non-smoking durations (Abbring and Van den Berg 2003; Van den Berg 2001). 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). We also include age and duration effects and various unique indicators reflecting the discounting behavior of individuals. These variables may affect both the smoking decision and the number of education years (Fersterer and Winter-Ebmer 2003; Khwaja et al. 2007).

The second contribution of this paper is particularly relevant to the starting decision on smoking. By modeling the number of education years as a time variant variable, we avoid biases by individual ability and group behavior factors that also affect the decision to start smoking at young ages. In particular, individuals may decide on education and smoking at early stages of life, when education itself is not completed. We thus build upon Farrell and Fuchs (1982) and Tenn et al. (2010), acknowledging the fact that one should exploit differences in smoking and education across the lifecycle. Our analysis also gives additional insight on the importance of time preferences of individuals—as "third factors"—to explain smoking starting and quitting decisions.

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. Similar to Farrell and Fuchs (1982) and Tenn et al. (2010), we find no effect of education on the decision to start smoking.

## 2 Data description

### 2.1 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.[5] The surveys gathered information on the respondent's family background (parents, siblings, marital status, and children), socioeconomic status (education, employment

[5] See appendix A1 of Webbink et al. (2012) for a discussion on the data collection and the external validity of the Canberra sample.

**Table 1** Summary statistics of covariates selected twins sample ($N = 5{,}378$)

| | Mean | SD |
|---|---|---|
| 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 |

status and income), health behavior (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 for which we observe smoking behavior and educational attainment, measured up to the age of 60. Table 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.[6] In both surveys educational attainment 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 seven categories have been recorded as 5, 9, 11.5, 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).

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 behavior are determined by similar general measures of time preference (Farrell and Fuchs 1982; 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

[6] The education system of Australia is divided into three broad areas: primary school, secondary school, and tertiary education. Tertiary education (or higher education) in Australia is primarily study at university or a technical college in order to receive a qualification or further skills and training (TAFE). TAFE institutions provide a wide range of predominantly vocational tertiary education and generally award qualifications up to the level of advanced diploma, which is below that of Bachelor degree.

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, which are obtained from the survey in 1980, is presented in the appendix to this paper.

Another issue is the external validity of our sample of Australian twins. The first row shows that the sample consists of only 34 % males. It seems that female twins are more likely to participate in these types of surveys. The lower participation of males has also been found for other twin samples (e.g., Le et al. 2005). 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 last rows in Table 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. The lower smoking prevalence in our sample might be attributed to the upward bias in educational attainment and age restrictions used for the estimation sample (below the age of 60).

### 2.2 Smoking durations

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

- Smoking during lifetime: the 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 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.[7] Figure 1 shows that this results in three possible combinations of successive smoking and non-smoking durations that start from the age of 12.[8] 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.

[7] 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.

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

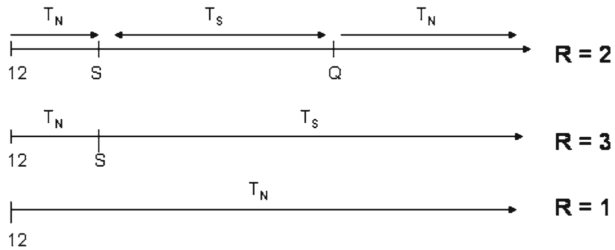


**Fig. 1** Combinations of smoking and non-smoking durations as a function of age, censored and uncensored

**Table 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 presents the sample statistics of the smoking and non-smoking durations, and Figs. 2 and 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 shows that smoking durations mostly start at younger ages, between the age of 12 and 22, which are consistent with other studies (Boudarbat and Malhotra 2008; Kidd and Hopkins 2004). The average starting age is 18 years. Note 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 smoking durations. Figure 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.

As to the pattern of quitting hazards in Fig. 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

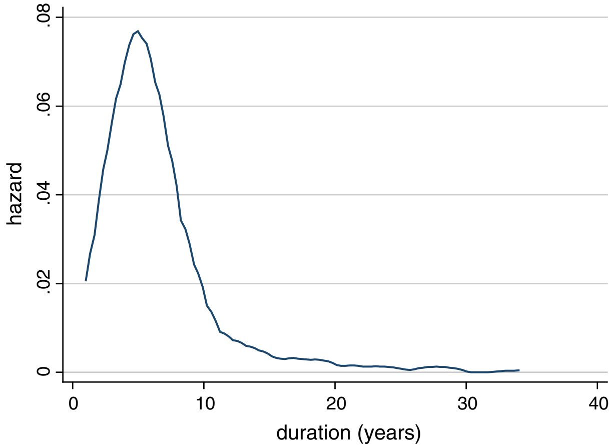


**Fig. 2** Observed hazard rates of starting smoking

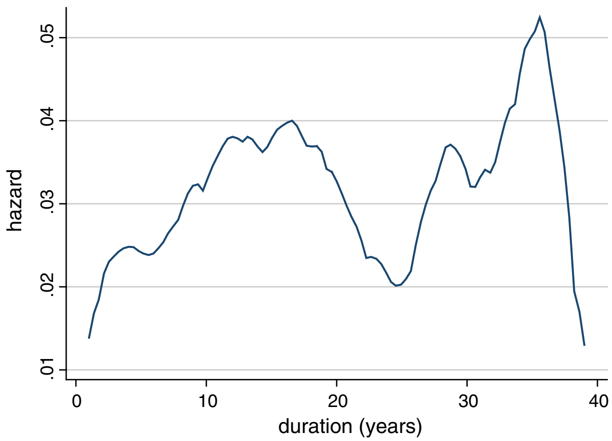


**Fig. 3** Observed hazard rates of quitting smoking

and Hopkins (2004) and Douglas (1998). Essentially, this hump-shaped pattern may result from three sources: habit formation, selection effects, and age effects. When modeling the quitting hazard, we therefore allow for all these effects in the MPH specification.

Figures 4 and 5 show Kaplan–Meier estimates for the survival functions of smoking and non-smoking durations for respondents that have not completed high school (education = 0), which only have completed high school (education = 1) and those who have received further schooling after high school completion (education = 2). The Kaplan-Meier estimates provide evidence for the probability of stopping smoking to increase with education, while the probability to quit smoking decreases with education.

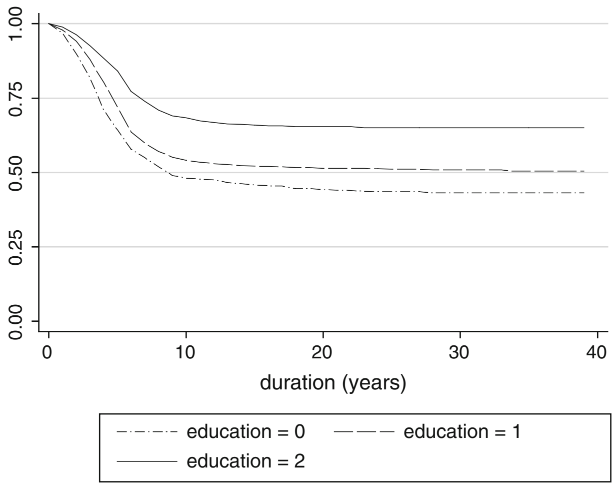


**Fig. 4** Kaplan–Meier estimates of non-smoking durations by education level

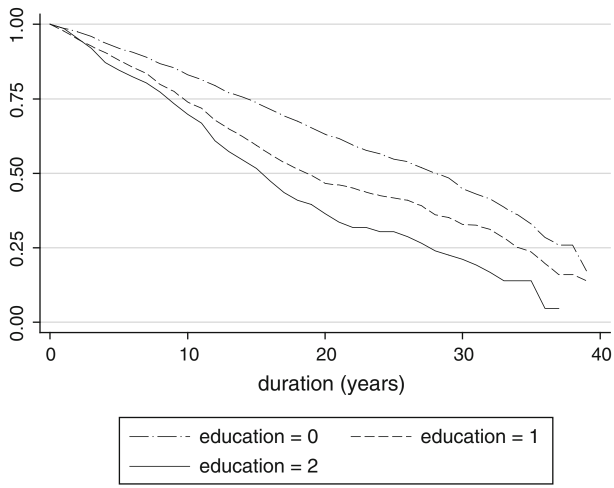


**Fig. 5** Kaplan–Meier estimates of smoking durations for education levels

## 3 Empirical strategy

### 3.1 The MPH model

Our research strategy entails the use hazard rate models to examine the impact of education on smoking and non-smoking spells. Within the current context, 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 $T$.

$$\theta = \Pr(T < t < T + \mathrm{d}t \,|t \geq T) \tag{1}$$

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):

$$\theta^{d}_{ijt,\tau} = \lambda^{d}_{0}(t) \exp\left(\alpha^{d} educ_{ij,\tau} + \boldsymbol{X}_{ij\tau}\boldsymbol{\beta}^{d}\right) \psi^{d}(\tau)\, \boldsymbol{\upsilon}^{d}_{j} \tag{2}$$

where $i$ indicates the individual ($i = 1..I$), $j$ indicates the twin pair ($j = 1..\frac{I}{2}$), $t$ is the elapsed duration, and $\tau$ indicates calendar time. Equation (2) shows that the MPH specification consists of four parts, representing the genuine duration dependence $\lambda_0$, variation in hazards due to observed individual and twin specific characteristics $\boldsymbol{X}$, education years (*educ*), calendar time effects $\psi$, and unobserved twin pair specific characteristics $\boldsymbol{\upsilon}$, respectively.

Duration dependence in the (non-)smoking decision is specified by the baseline hazard, $\lambda^{d}_{0}(t)$. A sufficiently flexible baseline specification is needed to take account of habit formation in the quitting hazard. Accordingly, we model genuine duration dependence in the quitting hazard as a (semi-parametric) polynomial function of the elapsed duration. With 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.

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 $\tau$. 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 1955), gender, birth weight, and the four proxies for the discounting behavior of the respondents. The age of respondents varies with calendar time $\tau$.

Finally, calendar time effects itself are modeled as dummies affecting all respondents equally at the same time intervals. We distinguish between three periods: prior to 1970, 1970 to 1980, and the years thereafter. By including calendar time effects, we control for general time trends in tobacco prices and tobacco taxes and other economic variables that have been found to affect the starting and quitting decisions of smoking (Forster and Jones 2001; Kidd and Hopkins 2004; Boudarbat and Malhotra 2008).

### 3.2 Identification

For both the starting and quitting hazards unobserved twin effects are taken into account by the time-invariant random effects $\boldsymbol{\upsilon}^d$. 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). 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. We thus specify the twin specific effects in the following auxiliary regression:

$$\ln \upsilon_j^d = \frac{\gamma^d}{2} \sum\nolimits_{i=1,2} \max_t \left(educ_{ij}\right) + \ln \zeta_j^d \tag{3}$$

with $d = Q, S$. In Eq. (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 $\zeta_j^d$ 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. Adding average values of education years as a controls for unobserved heterogeneity allows us to disentangle the well-known "within" from the "between" estimators of both coefficients. Thus the coefficient estimate of education years, $\alpha$, is identified from variation "within" twin pairs—both in completed education years and variation in education years per individual over time.

### 3.3 Maximum likelihood estimation

To estimate the model in Eqs. (2) and (3), we need to make closing assumptions on the distribution of the twin random effects $\zeta_j^d (d = S, Q)$. We do this in a non-parametrical fashion, assuming $K$ mass points for $\zeta_j^d$ with probability weights $P_1,\ P_2, \ldots, 1-P_1-\ldots-P_{K-1}$, respectively (Heckman and Singer 1984). Thus, the unknown distribution of $\boldsymbol{\zeta}$ is represented by a distribution with a finite number of points of support, where the first point of support is normalized to $\{0, 0, 0, 0\}$. This specification acknowledges the fact that some individuals may have very low smoking starting hazards and thus are very likely to never start smoking at all (this group constitutes 57 % of our sample). More specifically, with one point of support equal to zero, the MPH model is equivalent to the more conventional split population model, where the probability of not starting smoking at all is estimated separately (Douglas and Hariharan 1994; Douglas 1998; Kidd and Hopkins 2004; Boudarbat and Malhotra 2008).

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. All these parameters are estimated with Maximum Likelihood. Conditional upon the points of support $\zeta^d$ $(d = S, Q)$ and for respondent type $R$, there are three possible outcomes for the individual log likelihood contribution $\Lambda$:

$$\Lambda_{ij}\left(T_{N1}, T_{N2}, T_S, R|\zeta^S, \zeta^Q\right) = L_{ij}\left(T_{N1}|\zeta^S\right) \\ \times L_{ij}\left(T_s|\zeta^Q\right)^{I(R\neq 3)} L_{ij}\left(T_{N2}|\zeta^s\right)^{I(R=1)} \tag{4}$$

where $T_{N1}$, $T_{N2}$, and $T_s$ indicate the (two) non-smoking and smoking durations, and $I$ is a 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:

$$\boldsymbol{\Lambda} = \Pi_j\left[\sum_{k=1}^{K} P_k\left\{\Lambda_{1j}\left(.|\zeta_k^S, \zeta_k^Q\right) + \Lambda_{2j}(.|\zeta_k^S, \zeta_k^Q)\right\}\right] \tag{5}$$

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 3 shows the Maximum Likelihood estimation results of Eqs. (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 first impose the restriction that there is no correlation between the two hazard rates. It turns out that 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.

The coefficient estimates in Table 3 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 %, but our result is in line with Tenn et al. (2010). It is likely that our result can be explained by our estimation method, that exploits the "within-twin" variance, rather than cross-sectional

**Table 3** 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 | |

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. Especially at young ages the influence of peers on smoking (where we control for) is substantial (Harris and Lopez-Valcarel 2008). Similarly, if decisions regarding education are made at young ages and are correlated with the smoking decision at the same time, this yields an upward bias to our coefficient estimates as well. Using our method, such biasing effects cannot be picked up by our education variable and are controlled for, as in the relevant time period education differences are only small.

We also re-estimated the model for 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.

Various covariates have an effect on the decision to start smoking. Smoking durations are more likely to start at young ages (with the biggest peak at 18 years of age), and for younger cohorts, women and individuals with high birth weight.[9] 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 hazard rate that is close to zero (with a probability weight of 29 %).

Regarding the quitting decision, we do find a significant effect of education. 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] The psychological literature suggests that traditional sex roles can explain gender differences in smoking (Waldron 1991).

As to the other covariates, quitting smoking is less 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

[9] As shown in Table 3 we have included missing dummies for three explanatory variables (i.e., father's education, mother's education, and birth weight). The interpretation of the dummy coefficients is cumbersome, since our model is non-linear, and the education variables can take a range of values. In light of our current analysis, however, the costs of dropping missing observations are higher. In particular, for 25 % of the observations we miss at least one of the relevant three variables. Dropping observations with missing variables would thus harm the efficiency of our estimates.

[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.

**Table 4** 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) Completed education as (time invariant) variable | −0.120 | 0.095 |
| | (0.013)*** | (0.021)*** |
| (iv) Discounting variables excluded as controls | −0.012 | 0.102 |
| | (0.019) | (0.022)*** |

quitting decreases strongly with the smoking duration. At the same time, the likelihood of ongoing smoking durations decreases as a result of aging. 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, Table 4 presents the estimated effect of years of schooling on quitting and starting smoking for various specifications, with the attention predominantly focussed on the identification assumptions on the twin effects.[11]

### 5.1 Unobserved heterogeneity

To start with, 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. Related to this, we assume unobserved effects in the hazard rates only to vary among twin pairs. As a first robustness check, we relaxed this second assumption by modeling the unobserved heterogeneity distribution as individual effects, implying

[11] We tested the sensitivity to measurement errors in reported smoking and completed education years. 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 to similar estimation results. Another robustness test entailed the estimation of non-linear education effects, but this did not change our estimation results either.

that there is no correlation of individual effects within twin couples. As Table 4 shows, this reduces the number of repeated spells per stratus, causing the efficiency of the estimated distribution of unobserved effects to reduce.[12] At the same time, the fit of the smoking duration model to the data increases 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—when twins show larger differences through the effect of differential environments. For both the non-smoking and smoking durations, however, this model variant does not yield different coefficient estimates of the education variables.

Second, we tested the robustness of our results by zooming into the sub sample of identical ("monozygotic") twins in our data. The assumption that unobserved effects are equal per twin pair may be less restrictive for this sub-sample. Again, this did not result in significant or substantial differences from the outcomes of the benchmark model.

Third, we restricted the observed variation in education years to cross-sectional variation ("between twin pairs") in *completed education years* only. 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 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 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.

Finally, we re-estimated our model without the discounting variables as controls. In particular, if discounting behavior 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. However, we find coefficient estimates not to change as a result of this, suggesting that the effects of time preferences are largely absorbed by the twin specific effects.

## 5.2 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

[12] With modeling 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.

linear probability model (LPM) for smoking 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.[13]

The implied effect on smoking incidence we find is substantially smaller than those of studies following the instrumental variables approach, like Grimard and Parent (2007) who find one year of education to reduce the incidence of smoking with approximately 8 %-points.[14] This estimate should, however, be interpreted as a local average treatment effect 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 group, as most individuals start smoking between the age of 12 and 18.

Our results are more similar to those obtained by Tenn et al. (2010) who find no effect on smoking behavior. This does not come as a surprise, as they also estimate average treatment effects, while constructing suitable control groups' framework to control for selection effects (instead of twin effects as in our case). Still, one important difference with their analysis is that we use information of the complete smoking histories of individuals, rather than young individuals only. This allows us to estimate the effect of education on smoking cessation (at higher ages) as well, and may explain why we do find a significant, albeit small, impact.

## 6 Conclusions

We conclude 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 studies that using an instrumental variables approach, we find no effect of education on the decision to start smoking. This difference in findings can be explained by the modeling 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.

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—at least not in the time period under consideration in our analysis. 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.

Our results are robust with respect to a variety of sensitivity checks. At the same time, however, care should be taken in the transferability of these results to other

[13] It should be noted here that the LPM estimates are identified from 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.

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

populations. In this respect, 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. Moreover, our sample of twins contains significantly more females than males. 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; Webbink et al. 2008)—it is possible that our results might therefore not be fully transferable to the population at large. In addition, it should be stressed that respondents in our sample decided to start smoking prior to 1990, which is a period where tobacco control did not appear to be a priority in Australia (Treasury Australia 2012). It may well be that the health risks associated with smoking did not receive much attention in the curriculum of schools, particularly prior to the 1990s.

## Appendix: Factor analysis of discounting variables

| 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 |

## References


Abbring J, Van den Berg G (2003) The non-parametric identification of treatment effects in duration models. Econometrica 71:1491–1517

Albouy V, Lequien L (2009) Does compulsory education lower mortality? J Health Econ 28(1):155–168

Baker L, Treloar S, Reynolds C, Heath A, Martin N (1996) Genetics of educational attainment in Australian Twins: sex differences and secular changes. Behav Genet 26(2):89–102

Bauer T, Göhlman S, Sinning M (2007) Gender differences in smoking behavior. Health Econ 16(9):895–909

Black S, Devereux P, Salvanes K (2007) From the cradle to the labour market? The effect of birth weight on adult outcomes. Q J Econ 122(1): 409–439

Boudarbat B, Malhotra M (2008) The hazards of starting the cigarette smoking habit. Int J Econ Perspect 2(4):93–106

Bratti M, Miranda A (2009) Non-pecuniary returns to higher education: the effect of smoking intensity in the UK. Health Econ 19(8):906–920

Calvin C, Fernandes C, Smith P, Visscher P, Deary I (2009) Is there still a cognitive cost of being a twin in the UK? Intelligence 37:243–248

Chamberlain G (1982) Multivariate regression models for panel data. J Econometrics 18:5–46

Currie J, Moretti E (2003) Mother's education and the intergenerational transmission of human capital: evidence from college openings. Q J Econ 118(4):1495–1532

Cutler D, Lleras-Muney A (2005) Education and health: evaluating theories and evidence. NBER Working Paper 12352

de Walque D (2007) Does education affect smoking behaviors? Evidence using the Vietnam draft as an instrument for college education. J Health Econ 26(5):877–895

de Walque D (2010) Education, information, and smoking decisions. evidence from smoking histories in the United States, 1940–2000. J Human Resour 45(3):682–717

Decicca P, Kenkel D, Mathios A, Shin Y-J, Lim J-Y (2007) Youth smoking, cigarette prices and anti-smoking sentiment. Health Econ 17(6):733–749

Douglas S, Hariharan G (1994) The hazard of starting smoking: estimates from a split population duration model. J Health Econ 13(2):213–230

Douglas S (1998) The duration of the smoking habit. Econ Inq 36(1):49–64

Farrell P, Fuchs V (1982) Schooling and health: the cigarette connection. J Health Econ 1(3):217–230

Fersterer J, Winter-Ebmer R (2003) Smoking, discount rates, and return to education. Econ Educ Rev 22(6):561–566

Forster M, Jones A (2001) The role of tobacco taxes in starting and quitting smoking: duration analysis of British data. J R Stat Soc 164(3):517–547

Grimard F, Parent D (2007) Education and smoking: were Vietnam war draft avoiders also more likely to avoid smoking? J Health Econ 26(5):896–926

Harris J, Lopez-Valcarel B (2008) Asymmetric social interaction in economics: cigarette smoking among young people in the United States 1992–1999. J Health Econ 27:249–264

Heckman J, Singer B (1984) A method for minimizing the impact of distributional assumptions in econometric models for duration data. Econometrica 52(2):271–320

Hougaard P, Harvald B, Holm N (1992) Measuring the similarities between the lifetimes of adult Danish twins born between 1881 and 1930. J Am Stat Assoc 87:17–24

Kidd M, Hopkins S (2004) The hazards of starting and quitting smoking: some Australian evidence. Econ Rec 80(249):177–192

Khwaja A, Silverman D, Sloand F (2007) Time preference, time discounting, and smoking decisions. J Health Econ 26(5):927–949

Kenkel D, Lillard D, Mathios A (2006) The roles of high school completion and GED receipt in smoking and obesity. J Labor Econ 24(3):635–660

Le A, Miller P, Heath A, Martin N (2005) Early childhood behaviours, schooling, and labour market outcomes: estimates from a sample of twins. Econ Educ Rev 24(1):1–17

Lindeboom M, Llena-Nozal A, Van der Klaauw B (2009) Parental education and child health: evidence form a schooling reform. J Health Econ 28:109–131

Lleras-Muney A (2005) The relationship between education and adult mortality in the United States. Rev of Econ Stud 72(1):189–221

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

Miller P, Mulvey C, Martin N (1995) What do twin studies reveal about the economic returns to education? A comparison of Australian and U.S. findings. Am Econ Rev 85(3):586–599

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. Am Econ Rev 96(1):152–175

Sander W (1995) Schooling and quitting smoking. Rev Econ Stat 77(1):191–199

Tauras J, Chaloupka F (1999). Determinants of smoking cessation: an analysis of young adult men and women. NBER working paper 7262

Treasury Australia (2012) Issues in tobacco taxation, Treasury, Australia. 2012–01-03. :http://www.treasury.gov.au/documents/2086/PDF/Document_57.pdf

Tenn S, Herman D, Wendling B (2010) The role of education in the production of health: n empirical analysis of smoking behavior. J Health Econ 29(3):404–417

Van den Berg G (2001) Duration models: specification, identification, and multiple duration. In: Heckman JJ, Leamer E (eds) Handbook of Econometrics, vol 5. North-Holland, Amsterdam

Waldron I (1991) Patterns and causes of gender differences in smoking. Soc Sci Med 32(9):989–1005

Webbink D, Postuma D, Boomsma D, de Geus E, Visscher P (2008) Do twins have lower cognitive ability than singletons? Intelligence 36(6):539–547

Webbink D, Vujic S, Koning P, Martin N (2012) The effect of childhood conduct disorder on human capital. Health Econ 21(8):928–945