Health, Econometrics and Data Group

THE UNIVERSITY of York

HEDG Working Paper 10/02

SCHOOLING AND SMOKING AMONG THE BABY BOOMERS AN EVALUATION OF THE IMPACT OF EDUCATIONAL EXPANSION IN FRANCE.*

Fabrice Etilé
Andrew M. Jones

January 2010

york.ac.uk/res/herc/hedgwp

SCHOOLING AND SMOKING AMONG THE BABY BOOMERS
AN EVALUATION OF THE IMPACT OF EDUCATIONAL EXPANSION IN FRANCE.*

Fabrice Etilé[1¶] and Andrew M. Jones[2]

[1]INRA – ALISS and PSE, Paris, France.
[2]Department of Economics and Related Studies, University of York, United Kingdom.

January 2010.

Abstract

Post-war expansion of education in France transformed the distribution of schooling for the cohorts born between the 1940s and the 1970s. However, throughout this expansion the proportion with the highest levels of qualifications remained stable, providing a natural control group. We evaluate the impact of schooling on smoking, for the beneficiaries of the post-war expansion, by comparing changes in their outcomes across birth cohorts with changes within the control group. We uncover robust evidence that educational expansion contributed to a decline in smoking. Our results also suggest that tobacco control policies have reinforced the schooling-smoking gradient.

Keywords: Smoking, Education, Duration Analysis, France.

* We are grateful to Andrew Clark, Raphael Godefroy and Marc Gurgand for helpful discussion of the paper, and to seminar participants at University of Lausanne, the 26th Journées de Microéconomie Appliquée (Dijon, 2009), University of Sheffield, Paris School of Economics Health Economics seminar. We gratefully acknowledge funding from the Economics and Social Research Council (ESRC) under grant reference RES-060-25-0045.

¶Corresponding author: Fabrice Etilé, INRA-ALISS, 65, Boulevard de Brandebourg, 94205 Ivry-sur-Seine cedex, E-mail: etile@ivry.inra.fr

1 Introduction

It has not always been the case that those with more schooling are less likely to smoke (Nourrisson, 1999; Hughes, 2003). The schooling-smoking gradient, one of the most documented in social epidemiology, has only become negative in most affluent countries over the past forty years or so. For example, Kenkel and Liu (2007) provide evidence that, in the United States, a negative schooling-smoking gradient emerged in the 1960s. This reversal in the gradient coincides with the emergence of firm scientific evidence on the health hazards of smoking, and with the development of tobacco control policies.[1] This paper addresses the question whether there has been a *causal* negative effect of education on smoking in cohorts that grew up in times of educational expansion and stricter tobacco control policies: *the baby-boomer generation* born between 1945 and 1965. We combine ten repeated cross-sectional surveys collected between 1992 and 2003 and analyse several outcomes that characterise smoking: current smoking, transitions into regular smoking and quits. We adopt an instrumental variables strategy that exploits the fact that, throughout the period of educational expansion, the proportion with the highest levels of qualifications remained stable, providing a natural control group. We evaluate the impact of schooling on smoking, for the beneficiaries of the post-war expansion, by comparing changes in their outcomes across birth cohorts with changes within the control group.

---

[1] For instance, between 1980 and 2000, cigarette prices doubled in France, as a consequence of strong tax hikes, and information about the harmful effects of smoking was publicised widely by media campaigns and medical institutions. Over the same period, tobacco expenditures of French households in the three poorest deciles of the income distribution did not vary, while expenditures of the richest households decreased strongly (Godefroy, 2003).

Economists usually consider two main mechanisms through which education may have a causal impact on smoking (Grossman, 2000). First, there is the *educational efficiency hypothesis* which argues that those with more education allocate less resources to smoking (and other health-damaging but pleasurable behaviours) because they are better at obtaining and processing information about health risks. Second, the schooling-smoking gradient may be accounted for by a "value of life" argument. Smoking not only increases the probability of illness, but also the losses incurred, which depend on education through the discounted wage stream one can expect to receive over one's lifecycle. If more educated individuals have higher earnings and know the health risks of smoking, they face higher losses from premature death or disability: education is therefore positively correlated with the opportunity costs of smoking. This is the *opportunity cost* explanation.

These two arguments suggest that educational expansion policies, such as raising the minimum school leaving age or expanding higher education, may decrease smoking and therefore generate positive spillovers for public health. Education may also affect smoking through a complementary pathway: *socialisation by peers*. There is ample evidence of school peer effects influencing the decision to smoke (Gaviria and Raphael, 2001; Kawaguchi, 2004; Powell *et al.*, 2005; Lundborg, 2006; Clark and Lohéac, 2007). This may produce a positive relationship between length of schooling and smoking initiation. However, those individuals who leave school to work at an early age are likely to be influenced by older adults with established smoking habits, especially in occupations where the prevalence of smoking is high. Hence, it is not clear how the peer effects associated with the educational expansion will have affected the schooling-smoking gradient.

The existence of a negative schooling-smoking gradient was challenged by Farrell and Fuchs (1982), who argued that the observed association may be spurious if third factors - for instance a low discount rate or high innate cognitive abilities - have a positive impact on the demand for education and a negative impact on smoking. Since then, assessing the causal impact of education on smoking has become an important empirical issue. Most studies address it by applying instrumental variable (IV) techniques.Our own IV strategy seeks to identify the effect of schooling on smoking by exploiting the educational expansion that took place in France after World War II.

Between 1945 and the 1990s, levels of schooling increased from one generation to another. This expansion was driven by several policy reforms, which raised the compulsory minimum school leaving age, programmed school construction, and lowered financial hurdles and selectivity for access to secondary education tracks and the lower university levels. Developing an idea originally proposed by Gurgand and Maurin (2007), who estimate financial returns to education, we use those individuals with the highest levels of educational attainment – defined here as a Baccalaureat plus 5 years of completed studies - as a "control group". The share of this group in the population remained constant – around 5-10% - for cohorts born between 1945 and 1965, which suggests that selectivity into this level of education was left unaffected by the educational expansion. We therefore assume that the difference between the unobservable characteristics of the control group and those of the rest of the population did not vary across cohorts. However, educational attainment in the rest of the population increased across cohorts, as barriers were progressively lowered for the access to other levels of education (these groups are therefore defined as the "treatment group"). Our IV strategy is to instrument schooling by a set of interaction effects that are constructed by interacting birth cohorts with membership of the treatment group. This set of instruments captures the fact that only individuals in the treatment group were affected by changes in the supply of education, and that the intensity of the treatment varied from one cohort to another.

Estimates of simple linear probability models show negative correlations between years of completed schooling and both current smoking and initiation of regular smoking, and positive correlation with quitting. For both men and women the more educated start later in their lives and smoke for fewer years. The IV estimates have the same signs as standard OLS estimates, but they are higher in magnitude. We interpret this result as evidence that the relationship between education and smoking duration is causal. The effect of educational expansion can be illustrated by comparing smoking prevalence in the 1990s between the 1965 cohort and the 1945 cohort. For these cohorts educational expansion has contributed to a significant reduction of -2.9 percentage points for the average man and -3.2 percentage points for the average woman. It has also increased the hazard of quitting at any given age by 23.0 per cent for women and 12.5 per cent for men, and decreased the hazard of starting by -10.5 per cent for women and -8.5 per cent for men.

A second important finding is that the *opportunity costs* or the *socialisation* arguments seem to be more relevant than the *educational efficiency* argument for three reasons. First, additional IV estimates, that explore heterogeneity in the impact of schooling, reveal that the marginal effect of education does not decrease with the duration of schooling. This is not consistent with the usual notion of decreasing marginal returns that is implicit in the idea of educational efficiency in health production. Second, educational attainment has an impact on the hazard of starting smoking, even though the age of starting generally occurs before people have completed their schooling. Since, for the cohorts born before 1965, most pupils were assigned to a fixed schooling track and, hence, a predictable level of schooling before the age of 14, the estimated effect reflects the impact of expected schooling on smoking, and is therefore more likely to reflect differences in opportunity costs or social norms. Finally, when we also control for interaction effects between innovations in tobacco control policies and educational attainment, we find evidence which suggests that information policies or correlates of these policies have progressively reinforced the schooling-smoking gradient. Recent surveys show that, whatever their level of schooling, most of the French population are aware that smoking is harmful and addictive. Hence, this finding is most likely accounted for by differences in opportunity costs or social norms.

The paper is organized as follows. Section 2 describes changes in the educational system over the 20th Century in France, and explains the identification strategy. Section 3 presents the data. Section 4 estimates the causal effect of education on smoking in the 1990s. Section 5 extends the analysis to the transitions into and out of regular smoking. It also presents results regarding the interaction between education and anti-smoking policies. Section 6 concludes.

2 Educational expansion and the identification strategy

2.1 Education reforms in 20th Century France

There has been a dramatic expansion of educational opportunities for both men and women throughout the 20th Century in France, facilitated by a series of major reforms. Lower level secondary education became free at the turn of the 1930s. The Zay reform, in 1936, extended the compulsory school leaving age (CSLA) from 13 to 14 for cohorts born after 1923. The need for skilled workers and the demand for education increased after World War II. The Berthoin reform of 1959 was designed to encourage access to secondary education for children of all social classes. The CSLA rose from 14 to 16 and new educational tracks were created for individuals born after 1952. As a consequence, secondary school and university attendance, as well as post-Baccalaureat programmes, developed vigorously, especially after Fouchet's reforms in 1963 and 1966, and Faure's law in 1968. A final hurdle for mass education was crossed with the Haby reform, adopted in 1976, which unified educational tracks in the first cycle of secondary education for those born after 1965.

Before 1976, pupils could be assigned into vocational tracks after the second (completed) year of secondary school (the “5ème”), since then this assignment has usually been done after four years (Prost, 1981). In 1984, the Ministry of Education decided that 80% of each cohort should achieve the Baccalaureat level, and a new "vocational" Baccalaureat was created in 1985/1986. Nowadays, about 60 per cent of individuals have the Baccalaureat, whilst this figure was only 20 per cent for those born in the 1950's. Yet, despite this substantial increase in levels of education, the schooling-smoking gradient persists and is even stronger among younger cohorts (Aliaga, 2001).

Figures 1 and 2 provide a detailed picture of the evolution of educational attainment for cohorts born between 1942 and 1976. This is derived from the dataset that is used for estimation (a full description of the data is provided in Section 3). The dataset was constructed by merging individual records from ten cross-sectional surveys conducted by the French national statistics office (INSEE): "Enquête sur les Conditions de Vie des Ménages 1996-2003"(*EPCV1996-EPCV2003*) and "Enquête santé 1992 and 2003" (*ES1992* and *ES2003*). The distribution of the highest level of educational attainment is decomposed by cohort and gender into five levels: five years after the Baccalaureat including the Grandes Ecoles ("BAC+5"), degrees between two and four years after the Baccalaureat ("BAC+2/BAC+3/BAC+4"), the Baccalaureat (*BAC*), short secondary general and vocational/technical or professional education below the Baccalaureat (*BEPC/CAP*), and no secondary qualification or only primary education (*NOQUAL/CEP*). Table A1 in the Appendix defines these education levels more precisely, gives the equivalence between French, UK and US educational programmes according to the ISCED-97 classification, and the equivalence in terms of completed years of schooling (OECD, 1999).

Figure 1. Distribution of qualifications by cohorts, females (N=24,605)

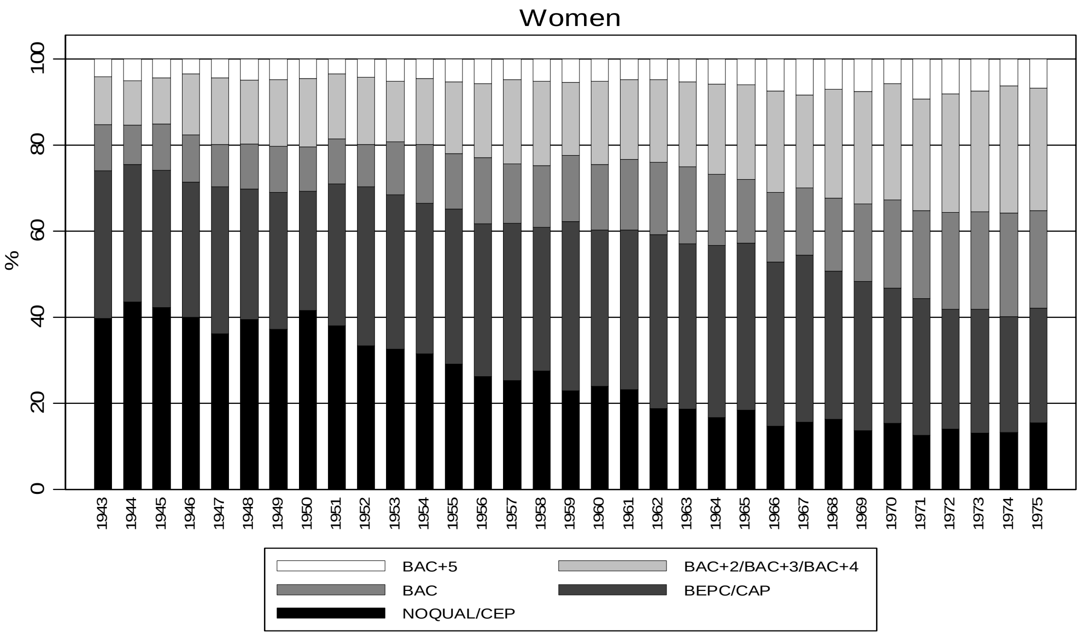
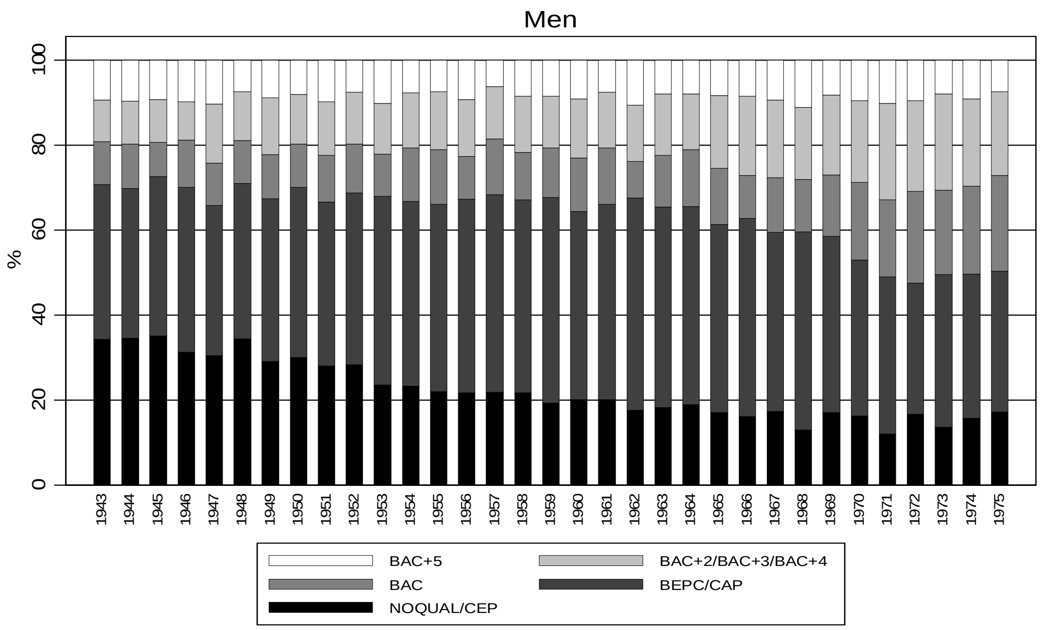


Figure 2. Distribution of qualifications by cohorts, males (N=26,249)



Note: these distributions of qualifications by gender and cohort were computed using **individual** data from the “Enquête sur les Conditions de Vie des Ménages” (EPCV1996-2003) and “Enquête santé” (ES1992 and ES2003).

Figures 1 and 2 also highlight an important empirical fact: the share of those individuals with a BAC+5 level did not increase until the cohorts born in the mid-1960s. If this group of individuals were not affected by systematic changes in unobservable factors over the cohorts born between the 1940s and the mid-1960s, they can provide a legitimate "control group", in the sense that the composition of the group was not affected by the reforms in the supply of education[2].

2.2 Identification strategy

Consider the following simple econometric model for the relationship between smoking and schooling:

$$Y_{i,a} = \alpha E_i + \beta X_{i,a} + \varepsilon_{i,a} \qquad \textbf{(1)}$$

where $Y_{i,a}$ is a measure of smoking behaviour for individual i at age a, E_i is the number of completed years of schooling. $X_{i,a}$ is a set of control variables (including age and cohort effects). $\varepsilon_{i,a}$ represents the impact of unobserved factors. The following additive decomposition is assumed:

$$\varepsilon_{i,a} = u_i + \eta_{i,a} \text{ with } E\left(\eta_{i,a} \middle| X_{i,a}, E_i\right) = 0 \qquad \textbf{(2)}$$

where u is some unobserved trait that reflects third factors such as cognitive ability and/or time preferences, and $\eta_{i,a}$ is an i.i.d. random shock.

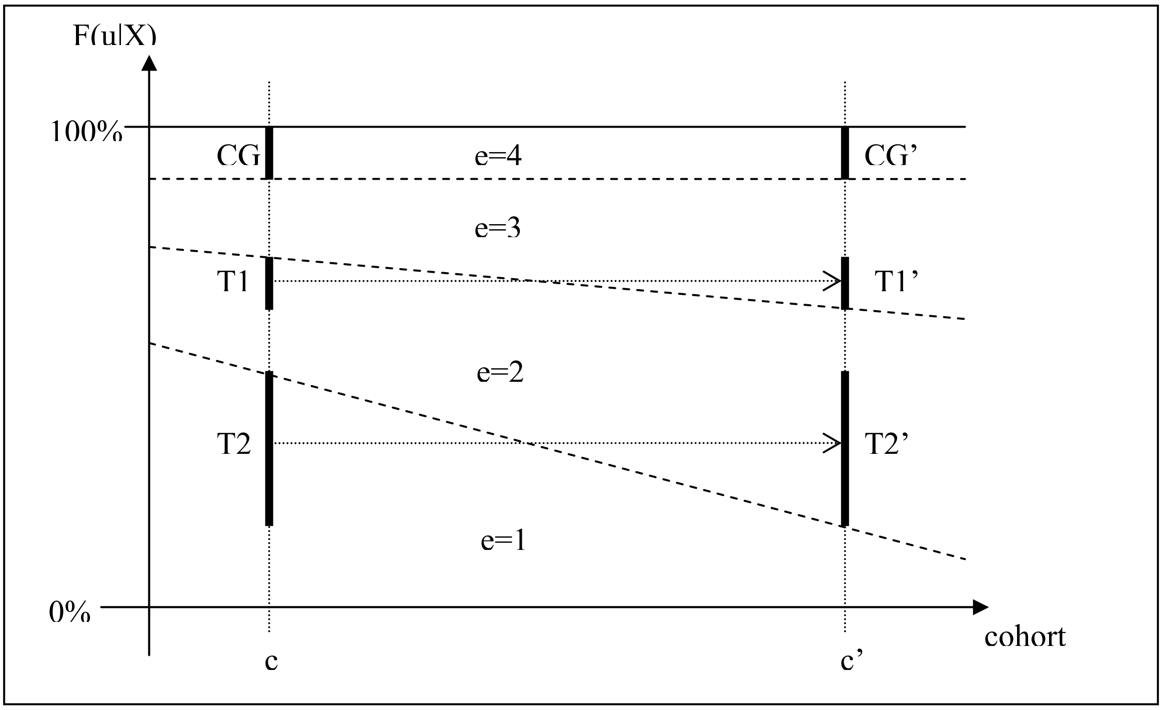
[2] Albouy and Wanecq (2003) and Albouy and Tavan (2007) find evidence that social inequalities in the access to the BAC+5 level are fairly stable for these cohorts. Making a distinction between the BAC+5 university degrees and the Grandes Ecoles degrees reveals that inequalities slightly decreased for the former and slightly increased for the latter.

Our IV strategy is to instrument E_i by a set of dummies constructed by interacting membership of the "treatment" group (those with levels of education below "BAC+5") and cohort dummies. The intuition is that these interaction effects are correlated with changes in education policies (such as school construction or institutional reforms of rules of access to education levels), but that these changes did not affect the control group, *i.e.* those individuals who would have anyway managed to get the highest schooling attainments given their observable and unobservable characteristics. As there are several cohorts and only one endogenous variable, over-identifying restrictions are available to test the validity of the instrument set.

Figure 3 illustrates the basis for our identification strategy. Using a stylised version of Figures 1 and 2, it shows the evolution of shares of participants across four levels of schooling by birth cohorts. The fourth level (e=4) constitutes the control group, as its share (at the top of the figure) remains constant. We suppose that the effect of observable exogenous variables X, other than cohort effects, have been controlled for. Then, the vertical axis gives the cumulated percentages of shares, which can be interpreted as a measure of the unobserved heterogeneity that affects selection into different levels of schooling, conditional on X (given by, $F(u|X)$, where $F(.|.)$ is the conditional distribution function of u).

In this set-up, the IV estimation strategy implicitly compares the smoking gap at a given age between individuals in CG (the "control group") and individuals in T1 or T2, to the smoking gap between CG' and T1' or T2'. This is similar to a difference-in-differences (*DiD*) strategy. As such, its validity rests upon the implicit assumption that the difference in the unobserved traits between the control group and the rest of the population remains constant across cohorts. The over-identifying restrictions provide a test of this assumption.

Figure 3. Identification strategy.



The identification of the causal effect of education relies on those individuals whose education level would be different, given their unobserved characteristics, if they belonged to a different cohort. For example, those in T1, with e=2, would have been in T1', with e=3, had they belonged to cohort c'. This identification strategy will capture the impact of schooling at the levels where changes across cohorts have been most influential. For example, in Figure 3, the size of T2 is bigger than the size of T1. Following Angrist and Imbens (1995), the estimator identifies a weighted average of marginal effects of education for each schooling level. This approach gives more weight to schooling levels whose share changed most across cohorts. As shown in Figures 1 and 2, for cohorts born before 1970, educational expansion mostly affected the lower part of the schooling distribution, where returns are expected to be higher. The results will thus provide information about educational expansion policies targeting primary and lower secondary schooling levels.

Equation (1) is estimated for several dependent variables. Section 4 focuses on outcomes observed at the time of the interview: an indicator for current smoking status and, among those who smoked at some point in their life, an indicator for whether they had quit smoking at the time of the interview. Section 5 examines transitions into and out of smoking, by using retrospective information about the age of onset and the age of quitting.

2.3 Comparison with previous studies

A number of studies have used IV methods to identify the causal effect of schooling on smoking (or, more generally, health). These estimation strategies can often be interpreted in a DiD framework. For instance, Grimard and Parent (2007) instrument education by interactions between gender and birth cohort for the cohorts at risk of induction into the Vietnam war. While de Walque (2007) uses a continuous measure of the risk of induction.[3] These papers compare trends in smoking across cohorts between men and women, the latter being the control group. Both studies present evidence that education has a negative impact on smoking, and increases the likelihood of quitting for smokers.

The literature also uses instruments related to educational policies, , arguing that variations in the supply of education or in the institutional constraints on schooling do not affect smoking other than through their indirect effect on the demand for education. Kenkel *et al.* (2006), using US data, and Jürges *et al.* (2009), using German data, instrument schooling by state-specific educational policy variables, such as compulsory schooling laws or the provision of secondary schools. As the timing of educational expansion differed from one state to another, schooling decisions vary by birth cohorts and states. Hence, the comparison

[3] As college attendance was a way to avoid the draft, cohorts of US men that were at risk of induction had good reason to enroll in college after high school even if it would not have been their choice otherwise.

of smoking prevalence and schooling levels between states and between cohorts identifies the impact of education on smoking. These papers provide evidence of negative effects of education on current and lifetime smoking.

Schooling reforms have also been used to infer the effect of education on smoking or health using only variation across cohorts. Using French data, Albouy and Lequien (2008) exploit the two reforms that extended the compulsory school leaving age (CSLA) from 13 to 14 and 14 to 16 for those cohorts born between 1923 and 1952 and after 1952 respectively, to estimate the impact of the school leaving age (their measure of education) on survival at age 80. The instrument is merely a dummy variable that indicates whether the individual belongs to a birth cohort that was affected by the reform. Identification relies on a regression discontinuity design (RDD) that relates differences in smoking to differences in schooling between the first birth cohorts that were affected by the reform and cohorts born one to three years earlier. This strategy uses the clear and discontinuous increase in the proportion of individuals leaving school after age 16 in cohorts born just after 1952. The standard errors of their estimates are such that neither the exogeneity of schooling, nor the null that schooling has no effect can be rejected.

Albouy and Lequien (2008) use the number of years spent at school, which may not reflect actual achievements. Instead we use the level of schooling, which is measured on a convenient continuous metric: the number of completed years of schooling according to the ISCED classification.[4] Figures 1 and 2 also display evidence that the post-war educational

[4] Indeed, years of schooling *per se* are not an appropriate indicator of health knowledge. As stated by Chevalier et al. (2004), "knowledge may […] come in indivisible "lumps" and it make sense for these to be associated with credentials".

expansion did not induce any significant discontinuity in the distribution of schooling levels between cohorts born before and after 1952. As a consequence, it is not possible to use a cohort-based discontinuity design.

3 Data and methods

This paper studies smoking behaviour through current smoking status, and the decisions to start and quit smoking. This section presents the datasets and the main outcome variables and explains how we select the birth cohorts in order to ensure the validity of the main identifying assumption.

3.1 Data sets and outcome variables

A smoker or a former smoker is defined as someone who smokes or used to smoke at least one cigarette a day.

Section 4 uses data on smoking status at the time of the interview that are drawn from ten pooled cross-sectional data sets: eight repeated cross-sections of the *Enquête Permanente sur les Conditions de Vie des Ménages 1996-2003* [*EPCV1996-EPCV2003*: Permanent Survey on the Conditions of Living of French Households]; and two cross-sections of the *Enquête Santé* 1992 and 2003 [*ES1992* and *ES2003*: Health Surveys].[5] These are nationally representative surveys of households and individuals conducted by the French National Institute for

[5] In the EPCV surveys, individuals who smoke only a pipe or cigars (about 0.5% of the sample) are regarded as non smokers. 0.1% of the sample report smoking a pipe or cigars as well as cigarettes and they are dropped from the estimation sample. ES2003 defines a third category: occasional smokers who did not smoke every day at the time of the survey (about 5% of the sample). They are excluded from the empirical analysis.

Statistics and Economic Studies (INSEE) using a common sampling methodology.[6] More precise data on lifetime smoking are available in ES1992, EPCV2001 and ES2003, and these are used to analyse smoking status among those individuals who had ever smoked, *i.e.* whether they had quit smoking or not at the time of the interview. In all surveys, a core section gives detailed information on individual characteristics. Education is measured using a very precise nomenclature, in particular for post-Baccalaureat education levels. In the *EPCV* surveys, one in three individuals aged over 15 in each household are randomly drawn to answer a special health section. In the *ES* surveys, health and socio-demographic information were collected for all household members. The main characteristics of the survey respondents are summarised in Tables A2 and A3 in the Appendix.

3.2 Information about smoking transitions

EPCV2001 and ES2003 provide retrospective information about smoking transitions: including the age of starting and the age of quitting for former smokers. Section 5 uses this information to analyse the decisions to initiate and to quit smoking (see e.g., Douglas and Hariharan (1994), Forster and Jones (2001) or Lopez-Nicolas (2002)).

EPCV2001 asks current and former smokers "How old were you when you started to smoke daily?", which was asked to current and former smokers. In ES2003, age of starting for current smokers can be constructed by using answers to the question “for how many years have you smoked?”. For former smokers, answers to two questions have to be used: “How long ago did you quit (in months or years)?” and “for how many years have you smoked?”. Given the date of birth, the date of interview, and information about quitting, it is possible to

[6] Details are available on http://www.insee.fr/fr/methodes/sources/pdf/Methodologie_EPCV_fixe.pdf and http://www.insee.fr/fr/methodes/default.asp?page=sources/ope-enq-sante.htm .

bound the age of onset within intervals of one to two years, with lower and upper limits that are not necessarily integer-valued. We chose to measure age of onset by the lowest integer age in the interval (choosing the highest one does not alter the results). An important feature of the dataset is that many individuals have not started smoking at the time of the survey. For them, age of onset is right-censored at their age at the time of the survey, minus one year since they could start smoking before their next birthday.

We assume that individuals are at risk of starting to smoke when they reach 10-years old. A pseudo-panel is then constructed by expanding each individual's data by the number of years at risk, *i.e.* age of onset minus 10 for those who smoked at some point in their life and age at the time of the interview minus 10 for never smokers. The dependent variable $Y_{i,a}$ of equation **(1)** is then a dummy, which equals 1 only if the individual decided to start at a (the age of onset) for smokers and former smokers, and takes the value 0 for other years and for never smokers. Since individuals who have already started are no longer at risk and are removed from the sample, the linear probability model **(1)** is equivalent to assuming a discrete time hazard model, and working with the hazard of starting smoking at age a.

To analyse the decision to quit smoking, we use the sub-sample of smokers and former smokers for whom we know the time since the most recent quit. The dependent variable is a dummy variable which indicates, for each year since age of smoking onset, if the individual quit during that year. Smoking duration is computed as the date at the time of the survey minus the date of starting minus, for former smokers only, the time since quitting. It can be bounded within intervals of one to two years. Using information about the date of birth and the date of interview, we can then bound the age of quitting in intervals of one year (for current smokers and most former smokers) to three years (for a few former smokers). We use the lower integer age in the intervals to define age of quitting. We then construct a panel by expanding each individual observation by age of quitting minus age of starting, *i.e.* smoking duration. For individuals who still smoke at the time of the interview, $Y_{i,,a}$ takes the value 0 at every age. For former smokers, $Y_{i,,a}$ equals 1 if quitting takes place at age a (it is then the last observation for an individual), and 0 otherwise. This sample is used to model the decision to quit in a discrete time hazard framework.[7]

3.3 Sample selection

Figure A1 in the Appendix shows the prevalence of lifetime smoking by year of birth for all individuals born between 1930 and 1975 (using ES1992, EPCV2001 and ES2003 only). The proportion of lifetime smokers is concave in the year of birth, with a maximum for cohorts born around 1970. This pattern reflects a mortality bias: lifetime smokers are less likely to be represented in older cohorts, since they tend to die earlier. The mortality bias may have important consequences for the estimation of education and price effects, because the elderly respondents surveyed in 2001 are, on average, less educated and less prone to have

---

[7] Since the data set has been constructed on the basis of retrospective information about smoking, voluntary or involuntary recall errors may threaten the validity of our results. This would be the case in particular, if errors are correlated with the instruments, *i.e.* if they vary both with cohorts and education. To check whether recall errors are a potential threat, we have compared smoking prevalence by gender and education computed from our pseudo-panel for 1980/1981 and 1991/1992, with the prevalence provided by the "Enquête Santé 1980-1981" (ES1981) and "Enquête Santé 1991-1992" (ES1992). The prevalence of smoking computed from the pseudo-panel and the prevalence given by ES1992 does not exhibit systematic differences by either birth cohort or education level for both men and women. There are more differences with prevalence computed for 1980/1981. Smoking status in 1980/1981 is likely to be under-reported 20 years later, and that underreporting increases with education, especially in younger cohorts (with a difference of about 20 percentage points for well-educated men and women born after 1959). This might be explained by a selection-by-mortality effect, or stronger anti-smoking norms among the well-educated. Using IV techniques should help to address problems of errors in variables.

started smoking.[8] For this reason, and following other authors (*e.g.* de Walque, 2007), we restrict the sample to individuals aged under 60 at the time of the survey. We also leave out those individuals who were still studying or were under 24 years-old, to avoid measurement errors on education, except for those who had already reached the "Grandes Ecoles" level.

The main assumption underlying our identification strategy is that the share of those individuals with the top level of education (the control group) remained constant across cohorts. Hence, we have to define the top level, and to choose a cohort-window over which its share remains constant. We use all individuals with at least five completed years of schooling after the Baccalaureat (BAC+5) as a control group.[9] Figures 1 and 2 show that the share of individuals with a BAC+5 degree increased for women born after 1960, and for menborn after the end of the 1960s. This rise in the proportion of BAC+5 graduates is the consequence of the Haby reform, which favoured access to higher education for cohorts born after 1965. Estimates of the share of BAC+5 in each cohort reveals that, in cohorts born between 1945 and 1965, about 4.30% of women and 9.24% achieved a BAC+5 degree, and no cohort significantly departed from this average (see Figures A2 and A3 in the Appendix). Hence, we focus on these 21 cohorts, the "baby-boomers". This leaves us with 20,335 women and

[8] A well-known epidemiological study that monitored male British doctors over 50 years shows that survival probabilities of smokers and non-smokers begin to diverge around age 50 (Doll *et al.*, 2004). Such selection bias can go either way, depending on whether smokers' mortality risk increases or not with education. We find no education-smoking gradient in these data for the elderly, which could be interpreted as evidence that the more educated smokers have a lower risk of mortality.

[9] Gurgand and Maurin (2007) use men who attended the *Grandes Ecoles* (GE) as their control group. In our sample there are not many individuals with a GE diploma in each cohort and the control group would be very

18,785 men for whom current smoking status is available. The corresponding sample sizes for quitting, age of smoking onset and age of quitting are 3,421, 5,030 and 1,685 for women and 4,830, 4,282 and 2,285 for men.

4 The impact of education on current smoking and quitting

4.1 Main results

Tables 1 and 2 present the estimates of the impact of education attainment on current smoking status and smoking cessation for those who have ever smoked, for men and women. We use linear probability models, and the coefficients show the effect, in percentage points, of one more year of schooling.[10] Each table presents OLS and IV estimates of equation (1), using two different set of controls. The OLS results are reported in the odd columns and the IV estimates appear in the even columns. The IV regressions do not use the full set of interactions between cohort memberships and membership of the treatment group. This is because the right-hand-side variables would be highly multicollinear in the first step regressions. We reduced the set of instruments to four dummies that were constructed by interacting membership of the control group with membership of cohorts born between 1953 and 1956, 1956 and 1959, 1960 and 1962, 1962 and 1965. This choice was guided by two factors. First, cohorts born before 1953 did not benefit from the Berthoin reform. Second, we carefully examined the difference in years of schooling between cohorts and between the

---

small. Choosing only GE as the control group would also rule out the possibility of studying smoking among women, because the share of women attending the GE has never been stable.

[10] Recent papers on education and smoking also use linear specifications (Grimard and Parent, 2007; de Walque, 2007).

treatment and the control groups, and the trend in this difference-in-difference displays kinks for cohorts born in 1957, 1960 and 1962. The bottom part of the Tables (lines 6 and 7) presents two specification tests to assess the robustness of the IV procedure. Line 6 reports the F-statistics associated with the first stage model. The instruments are relevant if the F-statistics is higher than 10, otherwise, they are said to be weak, and the IV estimates are likely to be strongly biased (Bound, Jaeger and Baker, 1995). Line 7 shows the p-value for a Hansen test of the over-identifying restrictions. The instruments can be safely excluded from the smoking equation if they are orthogonal to the residuals, which is indicated by a p-value higher than 0.10.

Columns 1 and 2 (for women) and Columns 5 and 6 (for men) show estimates of the effect of education with controls for a quadratic cohort trend, dummies for the year of interview, dummies for the region of residence, dummies for the type of residential area (set of control variables (1)). Regarding current smoking prevalence, for women, the OLS estimates of the returns to one more year of schooling is -0.35 percentage points, while for men it is -1.27 (Table 1, Columns 1 and 5). The IV estimates are higher, respectively -1.51 for women and -2.75 percentage points for men. The F-statistics are high (645.2 and 1287.9) and the p-value of the Hansen over-identification test is higher than 0.10. All coefficients are statistically significant at the level of 1%.

As shown in Table 2 (Columns 1 and 5), schooling also has a positive and significant impact on quitting smoking for those who have ever smoked. Once again, the IV estimates are larger than the OLS estimates. Additional computations reveal that one more year of schooling increases the prevalence of quits (the ratio of former to ever smokers) by 1.65 percentage points for women, and 2.16 percentage points for men. These estimates can be used to illustrate the impact of the post-war educational expansion on levels of smoking observed several decades later. As average schooling increased by 2.09 years for women and

1.06 years for men between the oldest and the youngest cohorts in the estimation sample, educational expansion policies have on average contributed to a significant reduction of the smoking prevalence in the 1965 cohort as compared to the 1945 cohort: -2.94 percentage points for men (1.06 × 2.771) and -3.16 percentage points for women (2.09 × 1.514).[11]

Since selection into social classes is heavily influenced by schooling, one may argue that these estimates reflect differences in social norms toward smoking between the top and the bottom of the social ladder, rather than the effect of education through opportunity costs or information on health risks. Smoking is now widely considered as a stigma rather a distinction, and is associated with negative traits such as a lack of self-control (Hughes, 2003). To test this argument, we introduce in the regressions a control variable (SES+) for occupations that include executives, the self-employed, professionals, employees in contact with the public and white collar workers.[12] We also control for real household income and interaction terms between the cohort trend and SES+, the cohort trend and the time trend (see the descriptive statistics in Table A3 in the Appendix A). This does not change either the OLS or the IV results (Columns 3 and 4, and 7 and 8). Hence, if differences in social norms and socialisation drive the current schooling-smoking gradient, they seem to be related to the education of individuals rather than to their occupation.

---

[11] We recognise that these predictions must be treated cautiously. One implicit assumption is that the age-period-cohort effects are fully taken into account by the model: predictions outside the -support of the data must be valid to make possible the *ceteris paribus* comparison between the oldest and youngest cohorts.

[12] This includes "intermediary professions" in the public sector and administrative occupations of the private sector. The dichotomy here follows closely Bourdieu (1979, chap. 3).

Table 1. Smoking status at the time of the interview - cohorts born between 1945 and 1965

This Table presents **the average effect, in % points, of one more year of schooling** on the probability to smoke at the time of the interview.

Gender	**Women**				**Men**			
N	20,335 individual observations				18,785 individual observations			
Technique	***OLS***	***IV***	***OLS***	***IV***	***OLS***	***IV***	***OLS***	***IV***
Years of schooling	-0.349***	-1.514***	-0.325***	-1.886***	-1.273***	-2.771***	-1.012***	-2.745***
	(5.92)	(5.80)	(4.88)	(4.85)	(18.92)	(13.65)	(13.10)	(8.47)
Set of control variables (1)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Set of control variables (2)	No	No	Yes	Yes	No	No	Yes	Yes
F-statistics, excluded instruments first stage		645.22		283.93		1287.94		502.69
Hansen p-value		0.195		0.176		0.863		0.732

Table 2. Smoking cessation at the time of the interview - cohorts born between 1945 and 1965

This Table presents **the average effect, in % points, of educational expansion** on the probability to have quit smoking for those individuals who have ever smoked.

Gender	**Women**				**Men**			
Sample characteristics	3,421 individual observations				4,830 individual observations			
Technique	***OLS***	***IV***	***OLS***	***IV***	***OLS***	***IV***	***OLS***	***IV***
Years of schooling	1.257***	1.651**	1.175***	1.838*	0.956***	2.155***	0.844***	2.138***
	(7.59)	(2.03)	(6.59)	(1.68)	(7.14)	(3.84)	(5.76)	(2.69)
Set of control variables (1)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Set of control variables (2)	No	No	Yes	Yes	No	No	Yes	Yes
F-statistics, excluded instruments first stage		82.67		40.87		172.91		90.43
Hansen p-value		0.843		0.882		0.747		0.700

Notes for Tables 1 and 2: Absolute t statistics in parenthesis; *= significant at the level of 10%; ** = significant at the level of 5%; *** = significant at the level of 1%; Set of controls (1): a quadratic cohort trend, dummies for the year of interview, dummies for the region of residence, dummies for the type of residential area; Set of controls (2): real household income per unit of consumption (OECD scale), SES+, interaction between a cohort trend and a time trend, interaction between SES+ and a cohort trend, SES+_missing, interaction between SES+_missing and a cohort trend.

4.2 Heterogeneity in the impact of schooling

Grossman (2004) notes that IV estimates of returns to education are generally higher than the OLS estimates. This is because identification of the IV estimates relies on exogenous shocks that affect mainly the bottom end of the distribution of education levels where, *a priori*, the marginal health returns to education should be higher. To test for the heterogeneity of returns across the distribution of schooling, we estimate alternative specifications wherein years of schooling is replaced by indicators of whether the individual has at least x years of completed schooling, with $x = 5, 9, 11, 12, 13, 14$ or 16.

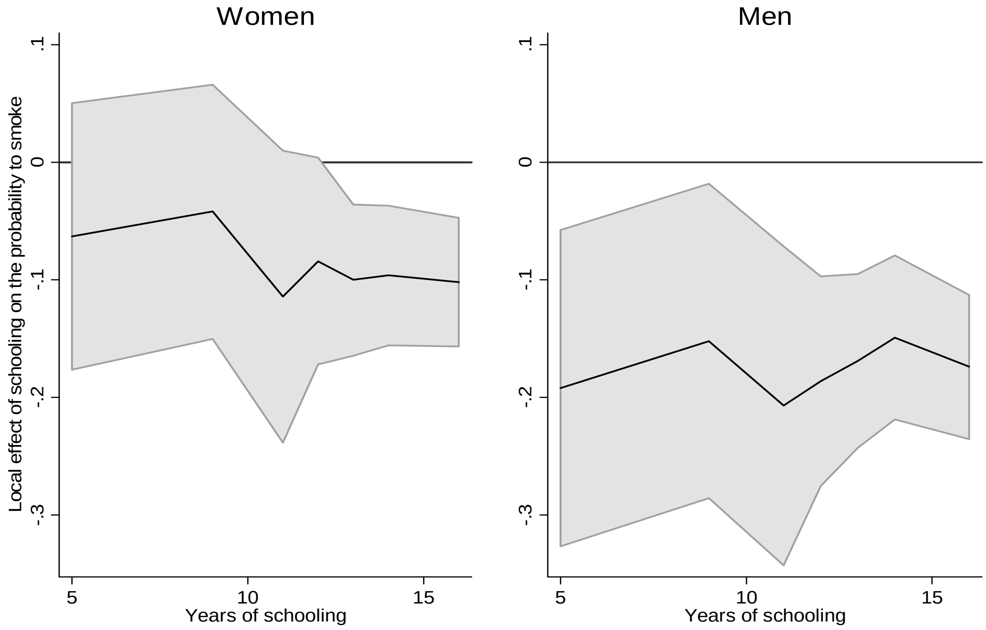
As the measures of smoking and schooling are both dichotomous they are modelled jointly using bivariate probit models. Smoking is modelled as a function of the schooling dummy and the control variables that have been used previously. In addition, to allow explicitly for the possibility of pre-policy differences in smoking and heterogeneity in returns between the control and the treatment group, we introduce a fixed effect specific to the control group (the dummy for BAC+5). Schooling is modelled as a function of the same control variables, plus the instruments that were used in the IV regressions. BAC+5 is excluded as it predicts the outcome perfectly (for all BAC+5 graduates, the schooling dummy equals 1). These regressions produce local average treatment effects: they show whether being induced by the educational expansion to attain at least x years of education produces benefits in terms of a lower propensity to smoke. Statistically significant estimates of the schooling coefficient means that the marginal returns to smoking are significant at x and that educational expansion produced significant changes between cohorts in the proportion of individuals having at least x years of schooling.[13]

Figure 4 shows evidence that, for men, the marginal impact of schooling on current smoking does not vary significantly with years of schooling. For instance, having at least 9 years of education decreases the probability of smoking by 15.2 percentage points, while having at least 14 years of schooling decreases it by 14.9 percentage points (see also Table A3, Men, Column 2 *vs.* Column 6). For men, educational expansion has had the same impact on smoking for all those individuals who benefited from it. This is also the case for women, although the impact on current smoking is smaller and not statistically significant in the lower part of the distribution of schooling.

Figure 5 shows that the picture is a little different for smoking cessation. For men, the policy impact is more substantial for secondary qualifications (BEPC, CAP and BAC) *i.e.* between 9 and 12 years of schooling. In particular, without the educational expansion, the probability of quitting among men who would not have had a CAP degree would have been much lower (by 42.8 percentage points, as shown in Table A4, Men, Column 3). For women, the benefits of educational expansion in terms of quitting are concentrated in the lower part of the schooling distribution (between 33.8 and 39.3 percentage points). Having at least a Baccalaureat does not produce additional benefits as compared to having at least a primary or a lower secondary qualification (BEPC or CAP).

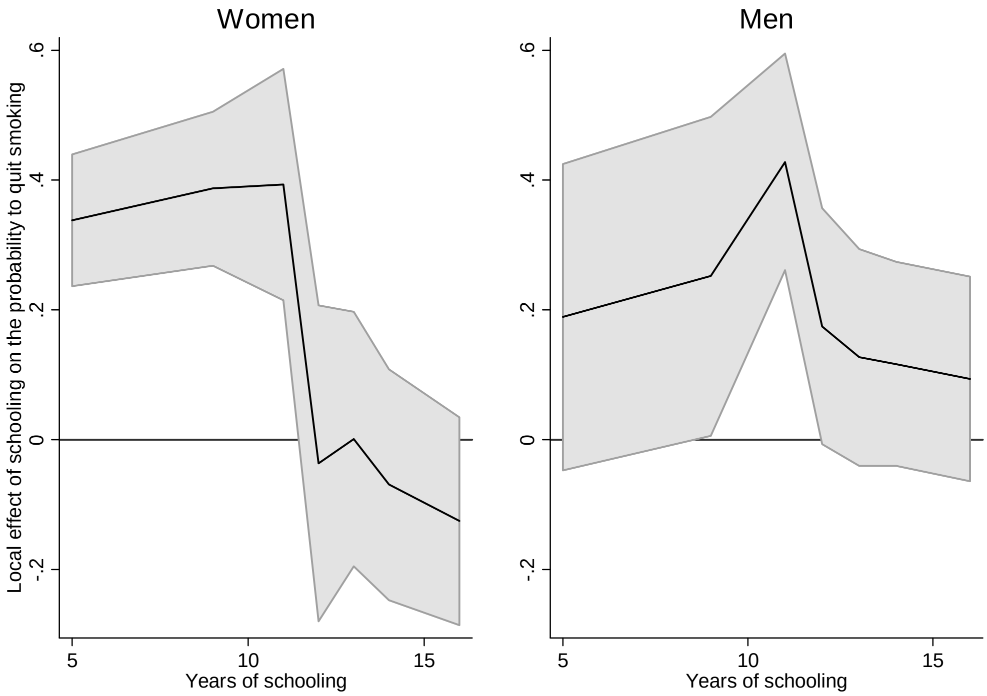
[13] The strength of the exclusion restrictions is tested by a χ^2 test which is displayed at the bottom of Tables A4 and A5.

Figure 4. Local effect of schooling on the probability to smoke .



Note: this figure shows for women and men the effect of having more than *x* years of education on the probability of smoking at the time of the interview. The grey area represents the 95% confidence interval, and the black line the estimated effect.

Figure 5. Local effect of schooling on the probability of smoking cessation .



Note: this figure shows for women and men the effect of having more than *x* years of education on the probability of having quit at the time of the interview. The grey area represents the 95% confidence interval, and the black line the estimated effect.

As shown by Figures 1 and 2, educational expansion had most impact on those at the bottom end of the distribution of schooling. If the marginal returns to schooling on health were decreasing, we would thus expect an upward (respectively downward) sloping relationship between years of schooling and local returns in Figure 4 (respectively Figure 5).[14] This only occurs for women's probability to quit. As a consequence, here, the hypothesis that the marginal returns are decreasing is likely not to hold. As emphasised in the introduction, the marginal returns have three components: the information effect, the opportunity cost effect and the socialisation effect. The marginal returns in terms of capability to process information – the information effect - are likely to be decreasing, as individuals with basic literacy and numeracy skills understand the major dangers from smoking: estimates suggest that almost 90% of the French population and 85% of smokers recognize that smoking is addictive and harmful (HCSP, 1998). Hence, the marginal returns are likely to reflect education-related differences in opportunity costs and/or socialisation rather than differences in the capability to understand health warnings.

The returns that are identified in this section are observed several decades after the educational expansion. The next section tests whether these results also hold when one examine transitions into and exits from regular smoking.

5 The effect of education on smoking transitions

The analysis of individual transitions into and exits from smoking is carried out using retrospective information on the age of onset and time since quitting drawn from *EPCV2001*

---

[14] Note that the marginal effects cannot be summed to produce for instance, the effect of having 16 years of schooling instead of 5: we here estimate local effects which are not suitable for such extrapolations.

and *ES2003*. The identification strategy relies on changes in education policies between cohorts and education levels, but the cohorts also differ in the length of their exposure to anti-smoking policies. Hence, in the previous regressions, the latter are a potential confounder of the education effect. By using longitudinal information on smoking transitions, we are able to disentangle the age/cohort and the time effects of public policies. We also provide some suggestive evidence that returns to education have changed over time and have interacted with tobacco control policies.

5.1 Empirical model and specifications

As described in Section 3.2. we constructed two pseudo-panels that indicate for each individual *i,* at each age *a*, if there is a transition into or out of smoking for those individuals who are at risk for such transitions. In such a discrete-time framework, a linear probability model of these transitions can be used to estimate the hazards of starting and quitting smoking. The panels can further be matched with aggregate time-series data to control for changes in public health policies.

In France, laws to ban smoking in public places and pro-smoking advertising, and to set up prevention policies were adopted in 1976 (the Veil law) and 1991 (the Evin law). Figure A4 in the Appendix shows that yearly tobacco sales per capita continuously increased from 1949 until Veil law, with a parallel drop in the real relative price of cigarettes. Then both prices and aggregate sales remained stable for almost 15 years. From the early 1990s, there have been substantial price increases. Indeed, while the Veil law marked a clear shift in information policies, with restrictions on advertising and the development of information campaigns, the main innovation of the Evin law was to allow the Government to remove tobacco prices from the computation of the Consumer Price Index, as used in the process of European Monetary Union. Then, large tax increases were coupled with even more vigorous information-based anti-smoking campaigns (Nourrisson, 1999, Berlivet, 2000). The variable, VEIL, which equals 1 if calendar year at age *a* is higher than 1976, controls for the advent of anti-smoking campaigns. Figure A4 also shows two breaks in price trends, in 1976 and 1991, even if there is no discontinuity *per se* in the price series. These continuous changes in prices are controlled by using the logarithm of price, LOGPRICE.

Figures A5 and A6 in the Appendix display non-parametric estimates of the hazards for starting and quitting. The more educated ("BAC and more"), start less and quit more at any given age or smoking duration, with the exception of well-educated women born between 1945 and 1955 who have higher hazards for starting than less educated women. The hazard of starting is increasing, peaks between age 18 and 20, and then decreases. It is almost zero after age 35. The hazard of quitting is globally increasing but is irregular. This is due to rounding or heaping effects, whereby individuals tend to declare smoking durations that are multiples of 5. An appropriate set of dummies are therefore used to control for duration dependence in a flexible manner.[15] We also introduce a linear time trend and a linear cohort trend. It is possible to include the latter, because of the panel nature of the dataset and, in models for starting, because time effects are controlled by a linear trend and the smoking policy variables rather than a fully saturated model.

---

[15] Duration dependence is an important issue in models of smoking duration (Forster and Jones, 2001). In the model for starting, we introduce 16 dummies for each year from 10-years old to 25-years old, all longer duration being in the reference group. In our analysis of quitting, we introduce 8 dummies only. Given that there are heaping and rounding effects in the self-reporting of smoking duration, the dummies indicate whether the duration is in one of the following intervals [0,3], [4,7][8,12],[13,17],[18,22],[23,27],[28,32],[33,100]. Allowing for one dummy per year has a negligible impact on the results.

5.2 The impact of education on starting and quitting

Columns 1 and 2 of Tables 3 and 4 display OLS and IV estimates of the effect of education on the hazard of starting and quitting for women. Columns 5 and 6 reports results from the same models for men. The coefficients are to be interpreted as variations in percentage points in the hazard of starting or quitting.

These results show that omission of controls for anti-smoking policies does not undermine the results presented in Section 4. An increase of one year in completed years of schooling decreases significantly the hazard of starting by 0.05 percentage points for women, and by 0.20 percentage points for men (Columns 2 and 6 in Table 3). One would *a priori* attribute this correlation to the existence of unobserved heterogeneity, as for most individuals the education level we observe is higher than the one they had at the age of onset. However, before the Haby law (*i.e.* for cohorts born before 1966), selection into the three main education tracks was decided, for most individuals, between 11 and 13 (Prost, 1981). Most people were able to anticipate their eventual level of schooling when they decided whether to start or not. The negative education-starting correlation in men may thus reflect the *causal* effect of opportunity costs related to an *expected* schooling level that will be realised later in life. Another explanation is that pro-smoking norms in the workplaces of those individuals who dropped out of school early were stronger than within-school peer effects for those who stayed on.

Regarding the decision to quit, the results are quite similar: one more year of schooling increases the hazard of quitting by 0.15 percentage points for women and 0.22 percentage points for men. These effects may seem modest in size, but the average hazard of quitting is 1.4 per cent for women and 1.6 per cent for men. Increasing the probability of quitting by 0.15 and 0.22 percentage points corresponds to an increase in the hazard of quitting by about 11% for women and 12% for men. Regarding the decision to start, the corresponding decreases in the hazard of starting are smaller: -5 per cent for women and -8 per cent for men. Overall these estimates imply that educational expansion has contributed to the decrease in smoking prevalence by increasing the likelihood of quitting and decreasing the likelihood of starting in younger generations as compared to older generations.

Regarding the policy variables, it should be noted that price increases have the expected effect: they reduce the likelihood of starting and increase the likelihood of quitting. The policy variable VEIL has a positive impact on the decision to quit but it is also positively associated with starting. This suggests that information campaigns that were set-up from 1976 onwards may not have had the expected effect. This may not be surprising, as information campaigns (especially in the 1980s and the 1990s) emphasised the long-term costs of smoking, which may be of less relevance for young people (O'Donoghue and Rabin, 2001).

Table 3. Decision to start smoking - cohorts born between 1945 and 1965

Gender	**Women**				**Men**			
N	5,030 individuals;145,467 age-individual observations				4,282 individuals; 95,765 age-individual observations			
Technique	***OLS***	***IV***	***IV***	***IV***	***OLS***	***IV***	***IV***	***IV***
Years of schooling	0.000	-0.054*	-0.393***	-0.371**	-0.064***	-0.201***	-0.212***	-0.511***
	(0.07)	(1.91)	(2.85)	(2.13)	(6.39)	(7.97)	(5.20)	(5.02)
LOGPRICE*Years of Schooling			-1.442**				-0.043	
			(2.52)				(0.26)	
VEIL*Years of Schooling				0.415*				0.403***
				(1.89)				(3.40)
LOGPRICE	-0.894***	-0.873***	13.290**	-1.280***	0.225	0.287	0.743	0.000
	(6.05)	(5.90)	(2.36)	(4.51)	(0.88)	(1.12)	(0.42)	(0.00)
VEIL	0.617***	0.623***	0.496***	-3.479	0.604**	0.618**	0.615**	-3.673***
	(4.01)	(4.05)	(3.09)	(1.57)	(2.37)	(2.43)	(2.40)	(2.72)
Cragg-Donald or F-statistics, excluded instruments first stage		59.84	2.09	6.56		171.10	11.34	21.75
Hansen p-value		0.513	0.582	0.001		0.145	0.000	0.439

Table 4. Decision to quit smoking - cohorts born between 1945 and 1965

Gender	**Women**				**Men**			
N	1,685 individuals; 37,183 age-individual observations				2,285 individuals; 54,870 age-individual observations			
Technique	***OLS***	***IV***	***IV***	***IV***	***OLS***	***IV***	***IV***	***IV***
Years of schooling	0.066***	0.152**	0.279**	-0.141	0.068***	0.222***	0.256*	-0.978**
	(5.28)	(2.30)	(1.97)	(0.92)	(6.40)	(3.23)	(1.78)	(2.53)
LOGPRICE*Years of Schooling			0.616				0.158	
			(1.24)				(0.31)	
VEIL*Years of Schooling				0.310*				1.264***
				(1.77)				(3.11)
LOGPRICE	3.620***	3.644**	-2.379	3.594**	3.372***	3.373***	1.855	2.834***
	(7.74)	(7.78)	(0.49)	(7.72)	(8.88)	(8.85)	(0.38)	(6.15)
VEIL	0.471*	0.477*	0.492**	-2.599	0.427**	0.443**	0.447**	-11.875***
	(1.93)	(1.95)	(2.01)	(1.58)	(2.06)	(2.12)	(2.15)	(3.04)
Cragg-Donald or F-statistics, excluded instruments first stage		24.79	4.03	2.58		21.89	4.56	1.82
Hansen p-value		0.206	0.003	0.112		0.124	0.039	0.716

Notes for Tables 3 and 4: These Tables present **the effect of one more year of education, in % points,** on the hazard of starting and quitting smoking. Absolute t statistics, corrected for clustering on individuals, are reported in parentheses; *= significant at the level of 10%; ** = significant at the level of 5%; *** = significant at the level of 1%; Set of controls: a linear cohort trend, a linear time trend, age dummies to control for duration dependence.

5.3 Tobacco control policies and changes in the schooling-smoking gradient

As noted in the introduction, the schooling-smoking gradient has not always been negative. Kenkel and Liu (2007) provide evidence that, in the US, the negative gradient appeared in the 1960s, coinciding with the emergence of a positive gradient between schooling and knowledge about smoking risks. But in the 1980s, the latter became flat, as information percolated more widely, while the schooling-smoking gradient still persists. Kenkel (1991) found that the negative correlation between health knowledge and smoking increases in magnitude with schooling: health knowledge matters more when education is higher. Hence, smoking policies may also affect the slope of the relationship between education and smoking.

One way to look at interactions between smoking and education policies is to interact LOGPRICE and/or VEIL with the number of completed years of schooling. These new interaction variables can be instrumented by interactions between cohort dummies, LOGPRICE and/or VEIL, and an indicator for not being a BAC+5 graduate. The explanatory power of the excluded instruments is tested via the Cragg-Donald F statistic, although it may not be applicable in the context of the linear probability model (Stock and Yogo, 2005).

The results are displayed in Tables 3 and 4 wherein Columns 3 and 7 show, for women and men respectively, estimates from a specification with LOGPRICE interacted with schooling. Regarding the coefficients on schooling, the changes are generally minor, except for women’s decision to start. This is the only case for which schooling interacted with price has a statistically significant impact. The interaction effect is negative: increases in cigarette prices are associated with an increase in the negative effect of education. The direct effect of education is still negative, but the direct price effect becomes positive. However, given the magnitude and the sign of the interaction effect, the average price effect remains negative. The Cragg-Donald statistics is quite low, around 2.1, which indicates that the IV estimator does not perform much better than the OLS estimator. Regarding men's decisions to start and quit, and women's decision to quit, the over-identifying restrictions are rejected.

Columns 4 and 8 display results from a specification that interactsVEIL with schooling. Here, there is a positive and significant interaction effect in men's decision to start smoking (+0.40 percentage points per year of schooling), while the direct effect of VEIL is negative (-3.67 percentage points): the total effect of VEIL is positive for a man with the average level of education. Hence, most men were more likely to start smoking after the Veil law and, perhaps surprisingly, this before-after contrast increases with education.

Regarding the decision to quit smoking, the direct effect of education is now negative for both women and men (-0.14 and -0.98 percentage points respectively). However, the interaction effect is positive and fairly high. This suggests that the education-quit gradient was flat or slightly negative before the Veil law, and became positive after the Veil law. This evidence must be taken with caution, as the Cragg-Donald statistics are quite low (although the two F-statistics associated with the first-stage regressions for schooling and schooling*VEIL are well above the usual threshold of 10).

In the end, the empirical evidence presented here provides suggestive evidence that tobacco control policies, or correlates of these policies such as changes in social norms, have reinforced the education-smoking gradient through their impact on the decision to quit smoking.[16] However, they are also associated with an unexpected reduction in the education gradient in smoking onset for men.

[16] Berlivet (2000) provides very convincing evidence that it is the changes of social norms in the upper classes that caused the adoption of the Veil law, and not the converse.

6 Conclusion

We have used a cohort-based strategy to identify the causal effect of educational attainment on smoking in France. We find clear and robust evidence that the post-war expansion of educational opportunities has contributed to a decrease in the likelihood of smoking for those individuals who benefited from the expansion.

Our results also suggest that the current schooling-smoking gradient is explained mainly by education-related differences in opportunity costs, and perhaps socialisation during youth and adulthood, rather than by differences in information about health risks. First, the marginal effect of education is not monotonically decreasing, as should be the case if it acted only by enhancing the efficiency of health production. Second, almost all individuals are now informed about the dangers of smoking and the education gradient in the decision to quit still persists.[17] Finally, we find that education has an impact on the decision to start smoking, which often occurs before schooling is completed. If the social gradient in health behaviour is mainly explained by differences in opportunity costs, it would be interesting to test whether redistributive and labour market policies have an indirect effect on social heterogeneity in health behaviours, by reducing the variation in opportunity costs. This is left for future research.

---

[17] Following this line of argument, Fuchs (2004) notes that "To explain the education–health connection, some researchers have proposed that those with more schooling are quicker to act on new health information or take advantage of improvements in medical technology. This seems reasonable, but is it important? The persistence of the negative gradient between education and cigarette smoking many decades after information about the harmful effects of smoking became widespread raises questions about the robustness of this explanation". Kenkel (1991) also notes that, in the US, differences in health knowledge explain no more than 20% of the education-smoking gradient, which may favour the opportunity costs argument.

REFERENCES

Albouy, V. and Lequien, L. (2008), “Does compulsory education lower mortality?”, *Journal of Health Economics*, 28, 155-168.

Albouy, V. and Tavan, C. (2007), “Accès à l’enseignement supérieur en France: une democratisation réelle mais de faible ampleur”, *Economie et Statistique*, 410, 3-22.

Albouy, V. and Wanecq, T. (2003), “Les Inégalités sociales d’accès aux grandes écoles”, *Economie et Statistique*, 361, 27-52.

Aliaga, C. (2001), "Le tabac : vingt ans d'usage et de consommation", *INSEE Première*, 808, 1-4.

Angrist J. and Imbens G. (1995), “Two-stage least squares estimation of average causal effects models with variable treatment intensity”, *Journal of the American Statistical Association*, 90, 431-442.

Berlivet, L. (2000), *Une santé à risques. L’action publique de lutte contre l’alcoolisme et le tabagisme en France (1954-1999)*. Thèse de doctorat en Droit et Science Politique, Université de Rennes 1.

Bound, J., Jaeger D.A., Baker, R.M. (1995), “Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak”, *Journal of the American Statistical Association*, 90, 443-450 .

Bourdieu, P. (1979), *La Distinction*, Paris : Editions de Minuit.

Chevalier, A., Harmon, C., Walker, I. and Zhu, Y. (2004), "Does education raise productivity, or just reflect it?", *Economic Journal*, 114, F499-F517.

Clark, A.E. and Loheac, Y. (2007), “'It wasn't me, it was them!' Social influence in risky behavior by adolescents”, *Journal of Health Economics*, 26, 763-784.

Clark, A.E. and D’Angelo, E. (2010), “Upward Social Mobility, Well-Being and Political Preferences: Evidence from the BHPS”, Paris School of Economics, *mimeo*.

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

Doll, R., Peto, R., Boreham, J. and Sutherland, I. (2004), "Mortality in relation to smoking: 50 years' observations on male british doctors", *British Medical Journal*, doi:10.1136/bmj.38142.554479.AE (published 22 june 2004), 1-9.

Douglas, S. and Hariharan, G. (1994), "The Hazard of Starting Smoking : Estimates from a Split Population Duration Model", *Journal of Health Economics*, 13, 213-230.

Farrell, P. and Fuchs, V.R. (1982) "Schooling and health: The cigarette connection", *Journal of Health Economics*, 1, 217-230.

Forster, M. and Jones, A.M. (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, 517-547.

Fuchs, V.R. (2004), "Reflections on the socio-economic correlates of health", *Journal of Health Economics*, 23, 653-662.

Gaviria, A. and Raphael, S. (2001), “School-based peer effects and juvenile behavior”, *Review of Economics and Statistics*, 83, 257-268.

Godefroy, R. (2003), “Les taxes sur les cigarettes sont-elles régressives”, *Economie Publique*, 3-28.

Grossman, M. (2004), “The demand for health, 30 years later: a very personal retrospective and prospective reflection”, *Journal of Health Economics*, 23, 629-636.

Grossman, M. (2000), “The human capital model”, in *Handbook of Health Economics* (Culyer, A.J. and Newhouse, J.P. eds), Amsterdam : Elsevier, vol. 1A, chap. 7, 347-408.

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

Gurgand, M. and Maurin, E. (2007), “A large scale experiment: Wages and educational expansion in France”, PSE Working Paper n°2007-21, available at http://www.pse.ens.fr/document/

HCSP (1998), *La Santé en France : 1994-1998*, Paris : La Documentation Française / Haut Comité de la Santé Publique.

Hughes, J. (2003), *Learning to Smoke: Tobacco Use in the West*, Chicago: The University of Chicago Press.

Jürges, H., Steffen, R. and Salm, M. (2009), “Does schooling affect health behavior? Evidence from the educational expansion in Western Germany”, *IZA Discussion Paper* n°4330.

Kawaguchi D. (2004), “Peer effects on substance use among American teenagers”, *Journal of Population Economics*, 17, 351-367.

Kenkel, D.S. (1991), "Health behavior, health knowledge, and schooling", *Journal of Political Economy*, 99, 287-305.

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

Kenkel, D.S. and Liu F. (2007), “The evolution of the schooling-smoking gradient”, Working Paper, Department of Policy Analysis and Management, Cornell University.

Lundborg P. (2006), “Having the wrong friends? Peer effects in adolescent substance use”, *Journal of Health Economics*, 25, 214-233.

López-Nicolás, A. (2002), "How important are tobacco prices in the propensity to start and quit smoking? An analysis of smoking histories from the Spanish National Health Survey", *Health Economics*, 11, 521-535.

Nourrisson, D. (1999), *Histoire Sociale du Tabac*, Paris : Editions Christian.

O’Donoghue, T. and Rabin, M. (2001), “Risky Behavior among Youths: Some Issues from Behavioral Economics”, in *Risky Behavior among Youths: An Economic Analysis* (J. Gruber, ed.), Chicago: University of Chicago Press and NBER, 29-68.

OECD (1999), "Classifying Educational Programmes – Manuel for ISCED-97 Implementation in OECD Countries – 1999 Edition", online document available at http://www.oecd.org

Powell L.M., Tauras J.A. and Ross H. (2005), “The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behaviour”, *Journal of Health Economics*, 24, 2005, 950-968

Prost, A. (1981), *Histoire de l'enseignement et de l'éducation. Tome IV. Depuis 1930*, Paris : Perrin.

Senik, C. (2008), “Ambition and jealousy: Income interactions in the ‘old’ Europe versus the ‘new’ Europe and the United States”, *Economica*, 75, 495-513.

Stock, J.H. and Yogo, M. (2005), “Testing for weak instruments in IV regression”, in *Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas Rothenberg* (Donald W. K. Andrews and James H. Stock, eds.), Cambridge: Cambridge University Press, 80-108.

APPENDIX: DESCRIPTIVE STATISTICS AND SOME ADDITIONAL RESULTS.

Table A1. Definition of qualifications with international equivalents

Variable's name	Educational programmes (French name)	Description with some U.K. and/or U.S. equivalents	ISCED-97 level	Equivalence in formal years of schooling
NOQUAL	No qualifications	No qualifications	0	0
CEP	Certificat d'Etudes Elémentaires (C.E.P.), Diplôme de fin d'étude obligatoire.	No equivalent certification for adult literacy and numeracy	1	5
BEPC	Brevet des collèges (B.E.P.C.), Brevet d'Enseignement Primaire Supérieur, Brevet Elémentaire.	Certification for having completed the first stage of secondary school: *cf.* grade 9 in the U.S.A., Certificates of Secondary Education grades 1-5 in the U.K.…	2A	9
CAP	Certificat d'Aptitude Professionnelle (CAP), Brevet d'Etudes professionnelles.	Vocational Qualifications: *cf.* GNVQ Foudation and Intermediate levels in U.K.	3C	11
BAC	Baccalauréat (general or technical), Baccalauréat (first part), Certificat de fin d'études secondaires, Brevets professionnels, Brevet supérieur.	National Diplomas which certify High School vocational, professional or general studies: *cf.* GCE A– and S-level or GNVQ A-level in U.K., High School Diploma in the U.S.A..	3A, 3B, (3C)	12, except 11 for those with only the first part of the baccalaureat and 13 for those with a vocational baccalaureat
BAC+2	Bac+1 and Bac+2	Programmes in 1 or 2 years after the Baccalaureat: *cf.* Vocational Certificate or Academic Associates's Degree Programme in the U.S.A., Higher National Diploma etc. in U.K..	4A, 4C, 5B, (5A with first two years of university successfully completed).	14
BAC+3 *BAC+4*	Bac+3 and Bac+4	Programmes in 3 or 4 years after the baccalaureat	5A with at least three years at the university completed, first cycle only	16
BAC+5	DEA, DESS, 3ème cycle, "Grandes Ecoles"	Master	5A, second cycle	18
GE	"Grandes Ecoles"	Unknown		18

Table A2. Brief description of the datasets

Survey	**Enquête Permanente sur les Conditions de Vie des Ménages** *EPCV* (Permanent Survey on Household Living Conditions)								**Enquête Santé** *ES* (Health Survey)	
Year of the survey	1996	1997	1998	1999	2000	2001	2002	2003	1992	2003
N	14,847	14,248	13,667	14,319	13,441	12,653	13,991	13,740	21,555	25,828
Number of individuals with non missing information on the following outcomes – all birth cohorts.										
Current smoking status	4,114	10,945	10,450	10,965	10,287	5,194	10,733	5,625	17,158	20,943
Lifetime smoking						5,194			17,058	20,617
Age of smoking onset						5,187				20,138
Age of smoking cessation						2,390				7,712

Table A3. Descriptive statistics – estimation samples

	Estimation sample	
Variables (mean, standard error if continuous)	*Current smoking status (N=39120)*	*Decision to quit (N = 8251)*
Smoker	31.0%	61.8%
Lifetime smoker	49.4%	100%
Former smoker (for lifetime smokers only)	38.3%	38.3%
Year of birth	1955.2 (6.0)	1955.5 (5.86)
Years of schooling	9.68 (5.17)	9.65 (5.03)
BAC+5	6.7%	5.6%
ES1992	16.3%	43.5%
EPCV1996	4.2%	0.0%
EPCV1997	10.3%	0.0%
EPCV1998	9.7%	0.0%
EPCV1999	10.3%	0.0%
EPCV2000	9.7%	0.0%
EPCV2001	4.5%	11.7%
EPCV2002	9.7%	0.0%
ES2003	18.6%	44.8%
Region 1	17.3%	15.9%
Region 2	21.3%	24.4%
Region3	7.6%	8.3%
Region 4	9.4%	9.1%
Region 5	12.8%	11.7%
Region 6	9.8%	9.7%
Region 7	11.0%	10.1%
Region 8	10.8%	10.8%
Lives in a rural area	26.0%	27.5%
Lives in a small city	30.2%	28.8%
Lives in a metropolitan area	43.8%	43.7%
Logarithm of the household real equivalenced income	9.60 (0.66)	9.68 (0.72)
CSP+ : membership of upper social classes (recoded as 0 if missing)	31%	23.8%
CSmissing : social class missing	17.8%	30.3%

Table A4. Smoking status at the time of the interview - cohorts born between 1945 and 1965

This Table presents **the effect of having at least *x* years of schooling on the probability to smoke** at the time of the interview.

Gender	**Women**							**Men**						
N	20,335 individual observations							18, 785 individual observations						
***x* (in years) =**	**5**	**9**	**11**	**12**	**13**	**14**	**16**	**5**	**9**	**11**	**12**	**13**	**14**	**16**
Schooling $\geq x$	-0.063	-0.042	-0.114*	-0.084*	-0.100***	-0.096***	-0.102***	-0.192***	-0.152**	-0.207***	-0.186***	-0.169***	-0.149***	-0.174***
	(-1.09)	(-0.76)	(-1.80)	(-1.87)	(-3.05)	(-3.17)	(-3.65)	(-2.80)	(-2.23)	(-2.99)	(-4.09)	(-4.48)	(-4.19)	(-5.57)
BAC+5	-0.056***	-0.055***	-0.048***	-0.036**	-0.018	-0.015	0.015	-0.104***	-0.101***	-0.090***	-0.050***	-0.026	-0.031	0.022
	(-4.08)	(-4.06)	(-3.38)	(-2.22)	(-0.96)	(-0.79)	(0.59)	(-7.87)	(-7.60)	(-6.59)	(-2.92)	(-1.31)	(-1.49)	(0.77)
Correlation	0.085	0.055	0.153	0.099	0.122	0.106*	0.102*	0.161	0.109	0.209*	0.149*	0.112*	0.100*	0.123**
	(0.92)	(0.56)	(1.31)	(1.09)	(1.78)	(1.67)	(1.80)	(1.61)	(1.05)	(1.87)	(1.88)	(1.72)	(1.69)	(2.52)
$\chi^2(4)$ statistics instruments	47.47	38.83	95.56	222.14	426.64	464.59	834.87	48.57	49.19	120.19	412.86	633.55	715.12	1002.61

Table A5. Smoking cessation at the time of the interview - cohorts born between 1945 and 1965

This Table presents **the effect of having at least *x* years of schooling on the probability of smoking cessation** at the time of the interview..

Gender	**Women**							**Men**						
N	3,421 individual observations							4,830 individual observations						
***x* (in years) =**	**5**	**9**	**11**	**12**	**13**	**14**	**16**	**5**	**9**	**11**	**12**	**13**	**14**	**16**
Schooling $\geq x$	0.338***	0.387***	0.393***	-0.036	0.001	-0.069	-0.125	0.189	0.252**	0.428***	0.175*	0.127	0.117	0.094
	(6.54)	(6.38)	(4.32)	(-0.29)	(0.01)	(-0.76)	(-1.53)	(1.57)	(2.01)	(5.03)	(1.89)	(1.49)	(1.46)	(1.17)
BAC+5	0.104**	0.0981**	0.0650	0.0923*	0.0889	0.109**	0.138**	0.077**	0.070**	0.039	0.026	0.023	0.031	0.015
	(2.37)	(2.28)	(1.50)	(1.85)	(1.58)	(1.98)	(2.05)	(2.46)	(2.25)	(1.28)	(0.71)	(0.55)	(0.73)	(0.28)
Correlation	-0.541***	-0.596**	-0.648**	0.217	0.108	0.230	0.329**	-0.193	-0.300	-0.690**	-0.156	-0.088	-0.113	-0.050
	(2.64)	(2.55)	(2.08)	(1.04)	(0.68)	(1.54)	(2.37)	(0.88)	(1.13)	(2.06)	(1.13)	(0.74)	(1.03)	(0.53)
$\chi^2(4)$ statistics instruments	13.46	4.04	16.45	53.95	102.52	110.33	154.07	7.68	5.88	20.28	101.92	149.57	185.23	241.23

Note: These estimates are produced using a **bivariate probit model** wherein smoking and "having at least *x* years of schooling" are the two dependent variables, and the former is a function of the latter, the sets of control variables (1) and (2), and BAC+5. Schooling is modelled as a function of the same set of control variables, and the set of instruments used to produce the results in Tables 1 and 2. Absolute t statistics in parenthesis: *= significant at the level of 10%; ** = significant at the level of 5%; *** = significant at the level of 1%; All regressions control for the sets of control (1) and (2): a quadratic cohort trend, dummies for the year of interview, dummies for the region of residence, dummies for the type of residential area, real household income per unit of consumption (OECD scale), SES+, interaction between a cohort trend and a time trend, interaction between SES+ and a cohort trent, SES+_missing, interaction between SES+_missing and a cohort trend.

Figure A1. Sample proportion of lifetime smokers by year of birth (ES1992, EPCV2001, ES2003).

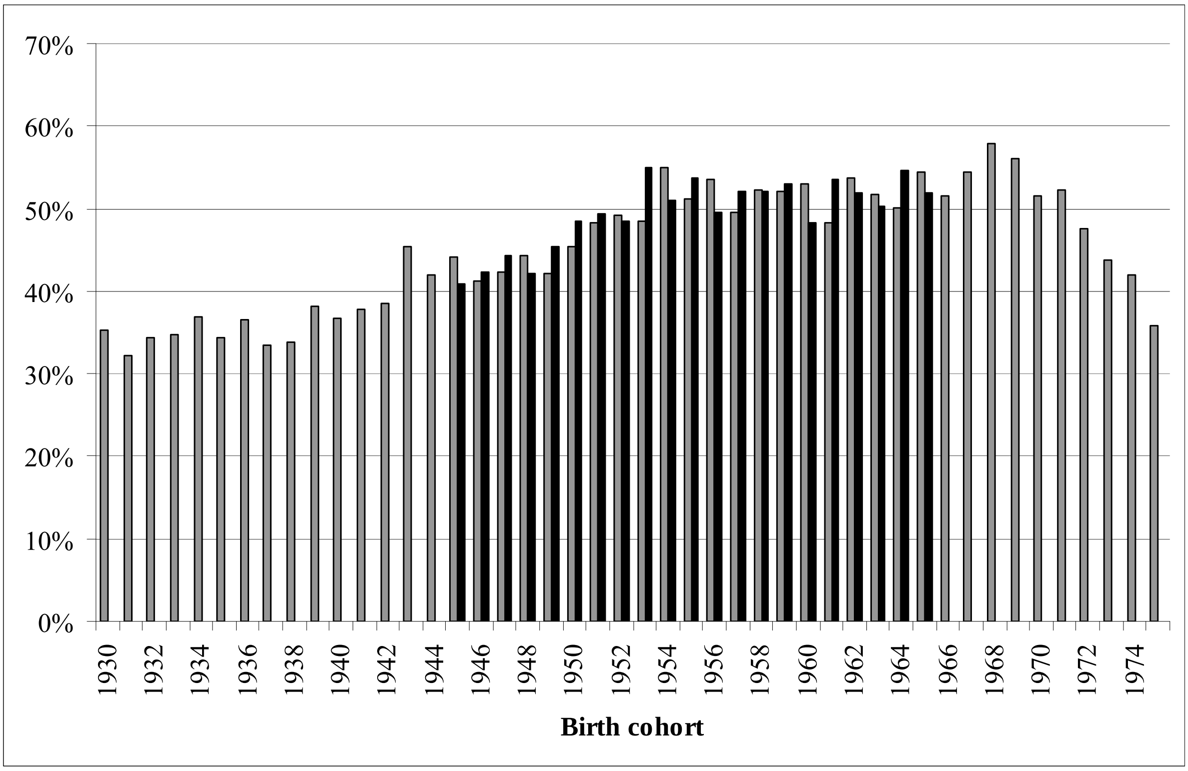
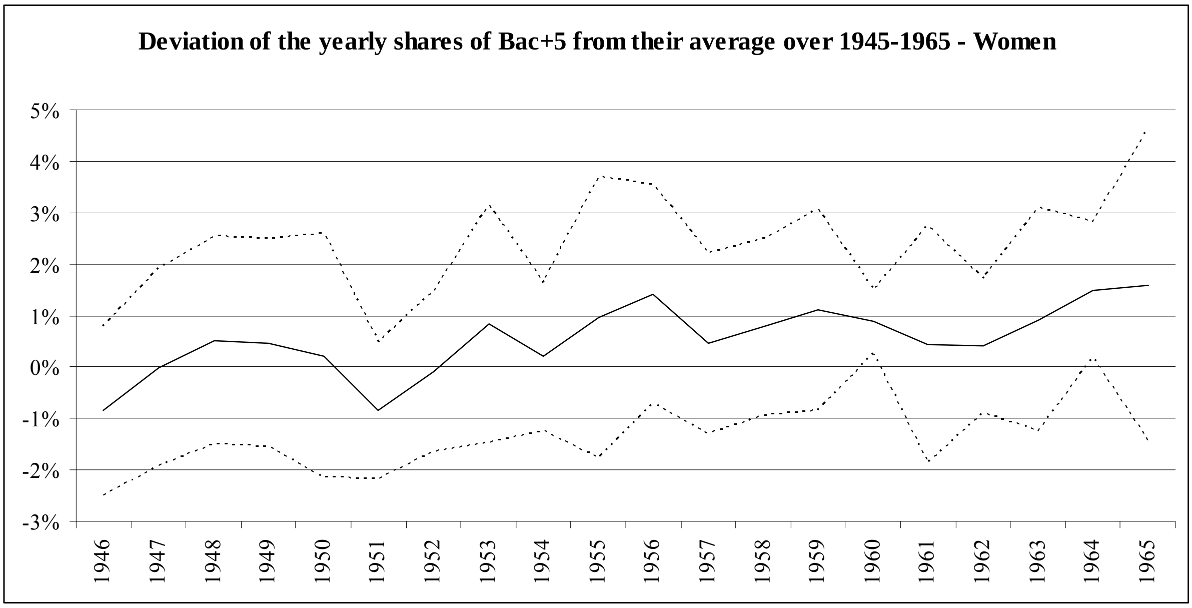
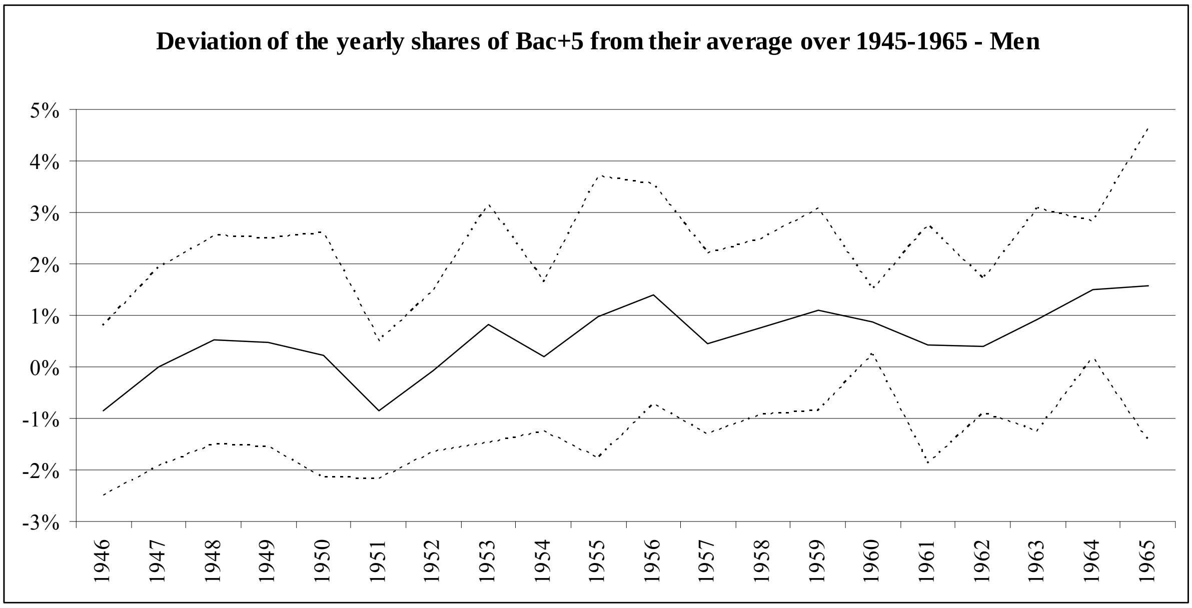


Figure A2. Test of "stability of the share of BAC+5" hypothesis - Women.



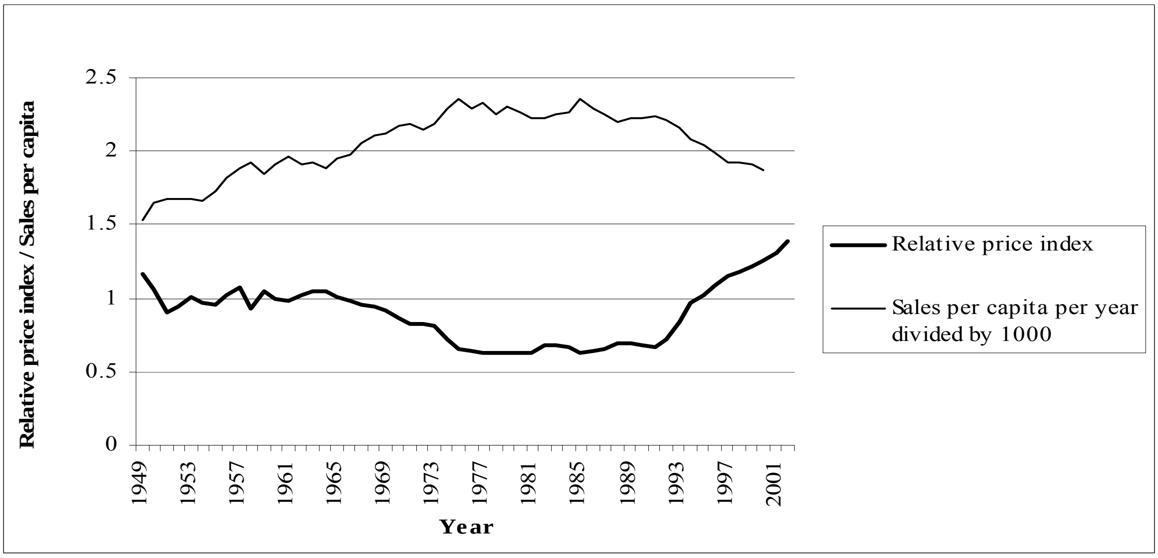
Note: the black continuous line shows on the y-axis the difference between the proportion of BAC+5 graduates in a cohort *c* of the x-axis and the average proportion of BAC+5 graduates for all cohorts born between 1945 and 1965. For instance, the cohort born in 1965 has about 1.5% more BAC+5 graduates than the average. The doted black lines represent the 95% confidence intervals.

Figure A3. Test of "stability of the share of BAC+5" hypothesis - Men.



Note: same as in Figure A2.

Figure A4. Trends in price and sales.



Note: Yearly sales data were obtained from the tobacco industry documentation centre (Centre de Documentation et d'Information sur le Tabac). We aggregated cigarette and loose tobacco (for hand-made cigarettes, pipes and chewing) sales, with a conversion rate of 1g per cigarette. The relative price index was constructed using INSEE data for the period (1949-2002). To obtain sales per capita, sales were divided by INSEE yearly figures for the total population aged over 15 in France.

Figure A5. Empirical hazard of starting smoking by gender, education and cohorts.

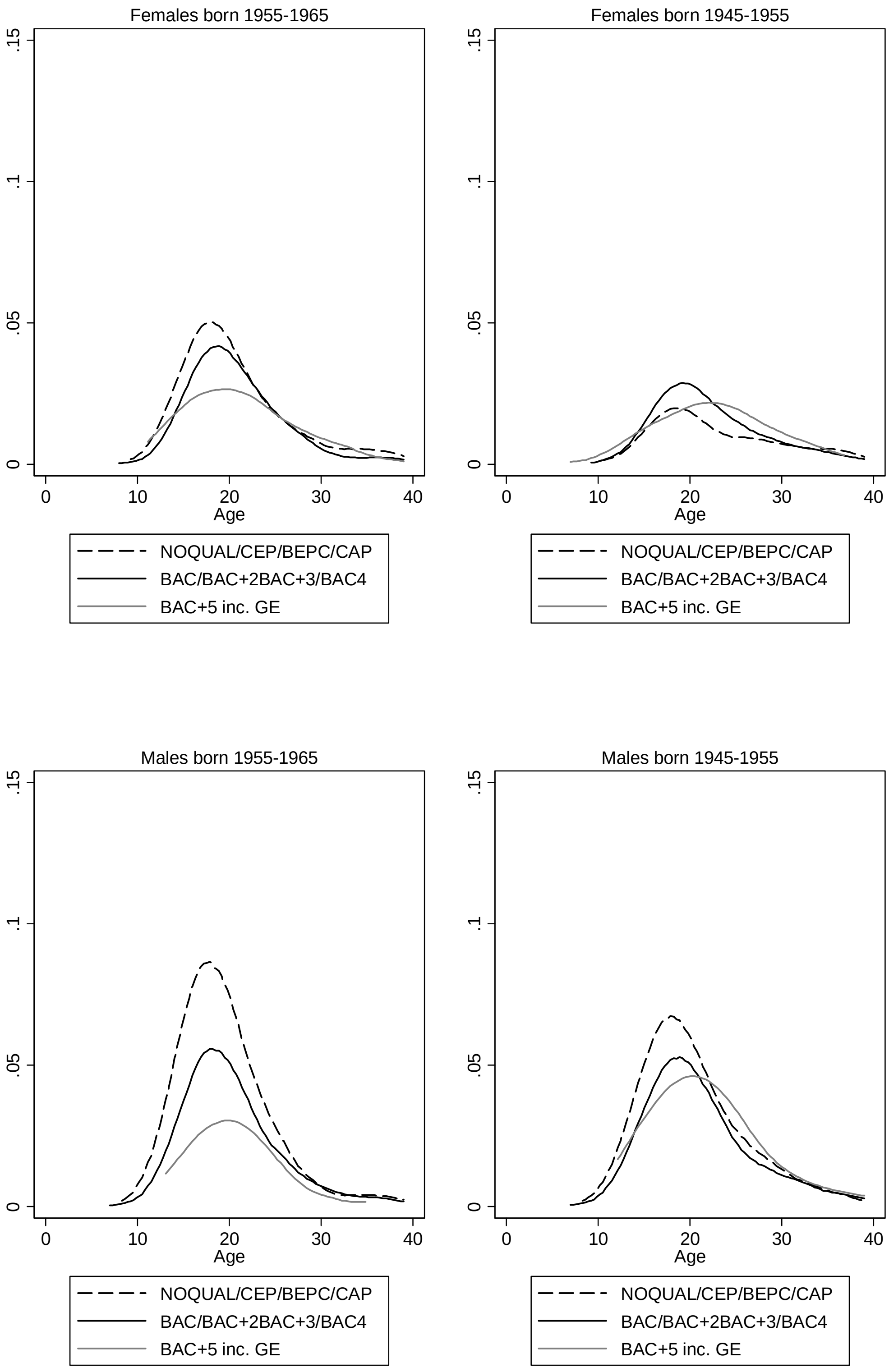


Figure A6. Empirical hazard of quitting smoking by gender, education and cohorts.