CrossMark

Early Smoking, Education, and Labor Market Performance

Ali Palali[1,2]

Abstract This study investigates the effects of early smoking on educational attainment and labor market performance by using mixed ordered and mixed proportional hazard models. The results show that early smoking adversely affects educational attainment and initial labor market performance, but only for males. The probability to finish a scientific degree is 4%-point lower for an early smoker. The effect of early smoking on initial labor market performance is indirect through educational attainment. Once the indirect effect is controlled for there is no direct effect. Moreover, for males only, early smoking has a negative effect on current labor market performance even after conditioning on educational attainment. The probability to have an academic job is 4%-point lower for an early smoker. For females neither education nor labor market performance is affected by early smoking.

✉ Ali Palali
    a.palali@cpb.nl

[1] CPB Netherlands Bureau for Economic Policy Analysis, The Hague, The Netherlands

[2] Department of Economics and Center, Tilburg University, Tilburg, The Netherlands

⚓ Springer

1 Introduction

A vast amount of evidence has piled up about serious negative consequences of smoking since the 1964 Surgeon General's report on the health effects of smoking (Levine et al. 1997). This evidence has lead economists to investigate the potential short term and long term relationship between smoking and various life outcomes such as labor market performance; finding a strong negative association between the two. The literature offers several reasons for this association. The first is the set of causal mechanisms through which smoking adversely affects labor market performance. Some examples of such mechanisms are employer discrimination; health problems, absenteeism and resulting productivity decrease; and hours lost due to smoking breaks (Kristein 1983; Levine et al. 1997; Lee 1999; Halpern et al. 2001; Heineck and Schwarze 2003; Weng et al. 2013).

There is one more causal mechanism through which smoking can affect labor market performance. Smoking negatively affects educational attainment if it is initiated early (Zhao et al. 2012); thus, indirectly deteriorates labor market performance through education. Even though the majority of the documented adverse health effects of smoking is observed in the long term, smoking may have adverse immediate health consequences on young people; if so, early smoking affects education. A report of the Surgeon General in 1994 shows that teenagers who smoke suffer from shortness of breath, increased heart beat and other respiratory problems. Furthermore, they are more vulnerable to the risk of other drug use. Levine et al. (1997) reveal that smoking is associated with decreased physical endurance. In addition, early smoking shows strong association with mental health problems and depression (Andreski and Breslau 1993). Although the nature of this association is yet to be established, there is evidence that mild depression may follow smoking initiation[1] (Steuber and Danner 2006; Goodman and Capitman 2000). Moreover, brain development, cognitive abilities and memory skills of young individuals can also be adversely affected by smoking (Trauth et al. 2000; Jacobsen et al. 2005). Consequently, all these negative effects on health can distort academic achievement.

Health condition of the individuals is not the only way early smoking can affect educational attainment. Since it is forbidden to smoke at schools,[2] smokers need to leave the campus during the breaks and turn back to classrooms after the break. Therefore, they are more likely to be distracted by life outside the school and more likely to return late to classrooms. Moreover, in their seminal paper (Rosenthal and Jacobson 1968) showed that expectations breed the performance. In other words, teachers' expectations about potential performance of students actually affect eventual performance. In the case of early smoking, if teachers form lower expectations about smokers, then early smokers might actually perform worse Furthermore, early smoking can lead students to search for side jobs because they need to finance their new habit. Time

---

[1] However, using instrumental variables approach (Pesko and Baum 2016) find that smoking does not cause an increase or decrease in stress immediately after use.

[2] In the Netherlands where the data used in this study are collected, the students cannot smoke within school premises. This is a part of a general tobacco law passed in 1990.

spent at such jobs eventually reduces the time spent for studying; thus, it subsequently harms students' performance at school.

Admittedly, the negative association between smoking and labor market performance does not have to be causal. It could be the result of a non-causal correlation. Such a correlation between smoking and labor market performance may occur when they are jointly determined by a set of observable and unobservable factors, e.g. parental characteristics, important life events, general attitude towards risk in life, myopic behavior or time preferences. Finally, the third reason is the reverse causality. In other words, labor market performance affects the smoking decision; for example, loss of a job might nudge individuals towards substance use including tobacco.

Keeping such mechanisms in mind, this study analyzes the effects of early smoking on educational attainment and labor market performance. This is not an easy task. The reason is that the aforementioned causal and correlated mechanisms complicate any analysis. The first method to deal with such a complication is to take advantage of instrumental variables. However, most of the instruments used in the literature so far, such as religiosity or parental characteristics, suffer from endogeneity as well. It is hard to assume that such type of individual-level or family-level factors do not have direct effects on educational attainment or labor market performance. French and Zarkin (1995) argue that it is very hard to find reasonable instruments to estimate the effects of alcohol use on wages. Perhaps, the same goes for the effects of smoking. Moreover, several studies discuss the weakness of instruments used for risky health behaviors including smoking and its consequences (French and Popovici 2011; Bound et al. 1995; Conley et al. 2012). Another problem with the IV estimation in the literature is that the negative smoking effect on labor market performance increases in magnitude once the instruments are used (Auld 1998; Zarkin et al. 1998; Van Ours 2004). This finding suggests that unobserved factors that make an individual more likely to smoke also make them perform better in the job market.[3] Although it is technically not possible to refute such a case, the more likely scenario is that unobserved factors that make an individual more likely to smoke, such as ability, time preferences or parental characteristics, make them perform worse in the job market. The same probably goes for educational attainment as well. That means the coefficient for the effect of smoking on labor market performance (or education) should actually decrease in magnitude once the endogeneity is taken into account.

The current study uses a correlated discrete factor approach in order to deal with the endogeneity issue rather than using exclusion restrictions. Heckman and Singer (1984) introduced this approach in order to control for unobserved heterogeneity in hazard rates, and Mroz (1999), for example, used it to estimate the effects of dummy

---

[3] There are other reasons why this can be the case. The first is that maybe there is measurement error in the smoking variables. If so, simple OLS estimations will result in a downward bias. The second is that proposed instruments might have an effect on only a part of the sample, for example occasional users. This is likely to be the case when the instruments are derived from policies such as smoking restrictions, not individual-level or family-level characteristics. However both of these cases suggest that endogeneity of the smoking decision in the wage estimations is not caused by unobserved factors such as ability or time preferences, bur rather mainly caused by measurement errors or type of the instruments.

endogenous variables.[4] First, the dynamics of smoking—accounting for both starting rates and quit rates are analyzed through mixed proportional hazard models. Although the main interest is on the early starting behavior—smoking before the age of 15—the analysis also includes quit rates of smoking to have a complete picture of the unobserved heterogeneity affecting the smoking dynamics. Hazard models provide the best fit to analyze the smoking dynamics as the smoking and quitting decisions are taken in a dynamic setting. Second, educational attainment and labor market performance, are analyzed using mixed ordered probit models with unobserved heterogeneity. Unlike previous studies, this study uses not only hourly wage information but also other indicators to measure labor market performance. I use information on the jobs that respondents have to construct an ordered variable, i.e. job rankings. The advantage of using the job rankings instead of wage information is that the data at hand enable construction of the job ranking variable for both the first job, the first jobs that individuals had, and the current job, the jobs that the individuals had at the time of the survey. However, wage information is available for only current jobs. Finally, smoking dynamics and ordered outcomes (educational attainment and job rankings) are modelled jointly to allow for correlation between unobserved heterogeneity. This controls for unobserved factors that can jointly affect smoking, education and labor market performance. Since reverse causality is not an issue here, because the early smoking behavior occurs before the age of 15, this method corrects for possible endogeneity caused by omitted variables.

The results show that early smoking has a negative effect on educational attainment. Once education is controlled for, the effect of early smoking on the first job rankings vanishes. However, there is still an effect on the current job rankings, which is a finding in line with the existing literature on the wage effects of smoking. An analysis of the probability of moving upward in the job rankings over time supports the aforementioned effects on the current job; showing that those who start smoking early are less likely to move upward. In other words, early smokers not only end up with worse first jobs due the effects through education, but also they are less likely to make a career. Finally, an investigation into the log-hourly-wages shows that reported wage effects of early smoking may be due to the smoking effects on the type of jobs, rather than the wage differentials within the same job.

There are several contributions of this study to the literature on the smoking effects on labor market performance. First, the empirical analysis uses not only the classical hourly wages information to measure labor market performance, but also the initial and the current job rankings. Since the current literature only focuses on the effects on wages, this study provides additional evidence about the effects of smoking on labor market performance. Second, this is the first study which explores the effects of smoking on labor market performance through the effect of educational attainment. It shows that there is an early smoking effect on labor market performance through educational attainment. Third, empirical analysis contrasts the effects of early smoking on the first job rankings and the effects on the current job rankings. Analysis of the

---

[4] This method is used to correct for endogeneity problem in investigations of health and labor market effects of cannabis use and smoking. See Van Ours (2003, 2004, 2006, 2007), and Van Ours and Williams (2009).

first job rankings is interesting as it shows whether early smokers start their job career from a disadvantaged point early in life.

The remainder of this paper is set up as follows. Section 2 introduces the data and briefly presents some stylized facts. Section 3 gives details about the econometric strategy. Section 4 presents and discusses parameter estimates obtained through maximum likelihood estimations, and Sect. 5 concludes.

2 Data and Stylized Facts

2.1 Data

The data used in the empirical part of this study are from the Longitudinal Internet Studies for the Social Sciences (LISS), which comprises detailed data for a representative sample of Dutch population above 16 years old.[5] More specifically, a combined data set—from three specific single-wave collections of information within the LISS data-is used; namely Alcohol and Drugs Study, Work and Schooling Study and Wage Indicator Study. These three surveys, which are explained later in detail, constitute a rich set of information about smoking dynamics, labor market performance, and history of labor market transitions for a sample individuals representing the Dutch population.

Alcohol and Drugs study (2008) is a single wave data set, which is ideal for the purpose of the current study because it contains answers to detailed questions on smoking. Respondents in the LISS panel report whether they have ever used tobacco. If so, they also answer the following question: *At what age, approximately, did you first use tobacco?*. This information allows for the investigation of the determinants of uptake of tobacco, i.e, starting rates of smoking. The respondents who reported ever smoking also report whether they smoked in the last 30 days prior to the survey time. This information is used to estimate the determinants of tobacco cessation, i.e, quit rates of smoking. Analyses of starting and quit rates, then, enables a complete picture of smoking dynamics. This single wave data set consists of 5597 observations in total.

Work and Schooling (2008) and Wage Indicator (2009) are two data sets which focus on the working history of the respondents and their educational attainment. Respondents in the LISS panel answer detailed questions on their educational background (the highest degree of education with a diploma), type of the first and the current job (the job that an individual has at the time of the survey) as well as many other questions on wages, working hours, job satisfaction, etc. Merging these two data sets with Alcohol and Drugs study results in considerable number of missing observations. The resulting merged data set consists of 4030 observations. The reason is that 1567 individuals who participated in Alcohol and Drugs study did not participate in Work and Schooling and Wage Indicator studies. However, in terms of observables, these 1567 respondents

---

[5] LISS panel is a household survey where there can be multiple respondents from the same household, mostly partners. Since empirical estimations are performed separately for males and females, the percentage of same-household respondents is around 5%. Therefore there is no need to control for same-household respondents.

and the remaining 4030 respondents are comparable; therefore, there is no immediate evidence for a selection problem between the data sets.

This paper focuses on three outcome variables: educational attainment, the first job rankings and the current job rankings. Each of the outcomes variables are constructed as ordered variables on a scale from 1 to 9; 1 denotes the lowest educational attainment or the lowest ranked job and 9 denotes the highest educational attainment or the highest ranked job. The exact details of the scales of education and job variables are given in "Appendix 1".[6] As for the job variables, Table 2 displays the details of the rankings. The main idea is that clerical jobs are ranked higher than manual jobs, and non-manual jobs are ranked the highest. Within non-manual jobs, professional ones are ranked higher than managerial ones. As noted earlier the previous literature uses mainly hourly wages to analyze labor market performance. This is neither worse or better than using the job rankings at hand. The advantage of using the job ranking in this study is that this information is available for both the first jobs and the current jobs. Moreover, Table 2 displays the mean hourly wages corresponding to each category in the ordered jobs. Both for males and females, there is a strong positive correlation between hourly wages and job categories, except for the last category. Higher academic jobs pay less than higher supervisory jobs, on average. However, job rankings seem to capture the overall wage differentials, and swapping the last two categories do not cause changes in the empirical findings.

Empirical analysis in the current study uses information only on respondents between 22 and 60 years old. This restriction and the missing observations decrease the sample size to 2174 respondents, 1021 of whom are males. The age restriction is imposed on the sample because most individuals complete their education around the age of 22 and enter the labor market. Moreover, many of them leave the labor market mainly due to early retirement around the age of 60. Not surprisingly, the data at hand also demonstrates this phenomenon. Percentage of those who are in paid employment rises sharply after the age of 22 and drops sharply after the age of 60. Similarly the percentage of those without a job is 24% for under 22 and 57% for above 60, whereas it is only 4% for between 22 and 60.

In addition, there are several sensitivity analyses throughout this study to provide evidence for the robustness of the results. Some of these sensitivity analyses were only possible after merging the data set with other assembled studies within the LISS data. "Appendix 3" briefly discusses data coming from other assembled studies.

2.2 Stylized Facts

Figure 1 highlights unconditional dynamics of starting age of smoking in the sample. Panel (a) displays the empirical hazard rates of tobacco uptake for both males and females. The figure shows that starting rates make a peak at the age of 16 and then another—smaller peak at the age of 18 for both males and females. The first peak

---

[6] Note that the percentage of respondents in the first two categories of the education attainment variable is somehow low. However, this does not create identification problems. Lumping the first two categories did not change any of the empirical results and conclusions.

indicates that, conditional on not smoking before, individuals have the highest risk of smoking at the age of 16.[7] Starting rates virtually become zero after the age of 25 for males and 23 for females; indicating that those who do not start smoking until mid 20s are very unlikely to do so afterwords. In other words, individuals mature out of smoking risk in their mid 20s regardless of gender. This finding is replicated by the cumulative starting probability figures in panel (b), where the slope of cumulative probability becomes almost zero around the age of 23 for females and 25 for males. The vertical axis displays the probabilities where the slope of cumulative probability becomes almost zero; indicating that more than 60% of females and 65% of males start using tobacco at some point in time.

"Appendix 2" presents the details and the descriptive statistics of the control variables and the variables of interest. The second row on the right panel shows that around 25% of the individuals, male or female, start using tobacco before the age of 15. The first sub-panel on the right presents the statistics of education variables and shows that, for both males and females, approximately 10% of the respondents have a university degree. Most individuals obtain an applied or a higher vocational degree. Around 4% of the respondents report that their education level is below the compulsory education in the Netherlands (a VMBO degree).[8] The last two sub-panels present the statistics of labor market performance variables. A quick comparison of the figures in the table reveals that there is an upward movement. The percentage of individuals having a lower ranked job is smaller in the current job variable whereas that having a higher ranked job is larger. Since the individuals can move upward in the job rankings after years of experience, this observation is reasonable.

Figure 2 displays the percentage of early smokers in each category of the ordered variables. For males, there is a clear pattern showing that lower ranked categories are mostly filled with early smokers. For females, as can be seen in Fig. 3, there seems to be no obvious pattern in educational attainment and labor market performance. This unconditional and purely descriptive evidence suggests that there is a negative association between early smoking and educational attainment as well as early smoking and labor market performance for males. Whether this association is causal or not is an empirical question.

3 Empirical Model

3.1 Dynamics of Smoking

Two main components of the smoking dynamics are analyzed: starting rates and quit rates. In the starting rates analysis, I assume that individuals become vulnerable to the risk of smoking from age 13 onwards, as only a handful of respondents report a smaller starting age. Specification of the starting rate at time t ($t = 0$ at age 12), conditional

---

[7] As of 2013 the sale of tobacco to people under 18 is illegal in the Netherlands. However, the legal age before 2013 was 16, which explains the peak at the age of 16.

[8] Note that for these cases it is not possible to identify if the early smoking takes place before the education is over. However, assuming that early smoking effect exists for only those with (education > 2) does not change the results as the percentage of those with (education < 3) is very small.

on observed characteristics x and unobserved characteristics u, is

$$\theta_s(t \mid x, u) = \lambda_s(t) \exp\left(x'\beta_s + u\right) \tag{1}$$

where β_s represents the effects of independent variables; $\lambda_s(t)$, individual duration (age) dependence. u denotes Heckman and Singer type discrete unobserved heterogeneity (Heckman and Singer 1984), which is unmeasurable set of differences in individuals' susceptibility to smoking. Duration (age) dependence has a form of flexible step function; $\lambda_s(t) = \exp(\Sigma_k \lambda_k I_k(t))$, where k (= 1,…,9) is a subscript for age categories. $I_k(t)$ presents time-varying dummy variables that are one in subsequent categories, 8 of which are for individual ages (age 13, …, 20) and the last interval is for ages above 20. Given that the model has a constant term in $x'\beta_c$, the first parameter in duration dependence, λ_1, is normalized to 0.

Similar to starting rates of smoking, quit rates are also assessed using a duration model. The LISS panel includes questions on the last month use of tobacco. The specification below assumes that if an individual reports no use of tobacco in the last 30 days, that individual quit smoking in the time period starting from the first use of tobacco until 30 days prior to the survey. Specification of the quit rate at time τ ($\tau = 0$ at the age of initiation), conditional on observed characteristics x_1 and unobserved characteristics v, is

$$\theta_q(\tau \mid x_1, v) = \exp\left(x_1'\beta_q + v\right) \tag{2}$$

Note that this analysis does not contain any duration dependence, because observing the exact time of quitting in terms of respondents' ages is not possible. However, interval censored nature of the data allows for the quit duration analysis (i.e., total duration of use) thanks to the information on the year in which the first use of tobacco takes place. This information gives an interval for quit duration; in other words, even though total duration of smoking is not observed, minimum and maximum values of this duration are known. Explicitly, duration of smoking, denoted by τ, will lie in the interval $[0, \tau_q]$ where τ_q is the difference between age at the time of survey and the age of the first use.

The joint density of completed durations until initiation of smoking and completed durations of smoking is specified as[9]:

$$g_1(t, \tau \mid x, x_1) = \int_v \int_u f_s(t \mid x, u) f_q(\tau \mid x_1, v) dG(u, v) \tag{3}$$

where $G(u, v)$ is the discrete joint mixing distribution of unobserved heterogeneity which allows for the possibility that conditional on the observed characteristics, starting age of smoking and total duration of smoking are correlated through unobserved characteristics. The number of support points in $G(u, v)$ is not predetermined and chosen using the likelihood ratio tests. For example, $G(u, v)$ can have 3 points of support (u_1, v_1), (u_1, v_2), (u_2); with $v_2 = u_2 = -\infty$. The associated probabilities denoted as

---

[9] Details of the econometric specification are given in "Appendix 4".

$\Pr(u_1, v_1) = p_1$, $\Pr(u_1, v_2) = p_2$ and $\Pr(u_2) = p_3$ are assumed to follow a logistic distribution, $p_i = \frac{\exp(\alpha_i)}{\Sigma_{i=1}^{3} \exp(\alpha_i)}$, where α_3 is normalized to zero. This indicates that the model identifies three types of individuals regarding starting and quitting smoking. The first group consists of individuals with a positive starting and positive quit rate. The second group consists of those with a positive starting rate but a zero quit rate. The third group has a zero starting rate, therefore the quit rate does not exist at all.

3.2 Educational Attainment and Labor Market Performance

I use ordered probit models to investigate how early smoking affects educational attainment and labor market performance. First, I assume that the smoking decision is independent from all the unobserved factors than can be correlated with educational attainment, i.e. that the smoking decision is exogenous. Given that such an assumption is, by and large, not plausible, the following section (Sect. 3.3) will present the model that takes account of possible endogeneity.

Educational attainment is measured as an ordinal variable in a scale of 1–9. To exploit the ordinal character of the dependent variable, I use an ordered probit model with discrete unobserved heterogeneity. Such unobserved heterogeneity captures time-invariant person specific unobserved factors that cause systematic differences in educational attainment. The unobserved latent variable in the ordered probit model is

$$y_{ed}^* = x'\beta_{ed} + \rho_{ed}smoke_{15-} + \epsilon_{ed} + e_{ed} \tag{4}$$

where ρ_{ed} represents the effect of early smoking. ϵ_{ed} controls for discrete type of unobserved heterogeneity, which is different from the error term e_{ed}. Furthermore, β_{ed} measures the effect of the control variables. The observed ordered categories and the rest of the specification of the ordered model are given in "Appendix 4".

Similar to educational attainment, labor market performance is also investigated through an ordered probit model. The unobserved latent variable in the analysis of labor market performance is

$$y_j^* = x'\beta_j + \phi_j education + \rho_j smoke_{15-} + \epsilon_j + e_j \tag{5}$$

where ρ_j represents the effect of early smoking. ϕ_j controls for the effect of educational attainment on labor market performance. ϵ_j controls for discrete type of unobserved heterogeneity. β_j measures the effect of the control variables. The rest of the analysis is analogous to the analysis of educational attainment; therefore, the details of the model specifications are omitted. The analysis of labor market performance is the same for the initial and the current job rankings.

3.3 Joint (Correlated) Model

Assuming that smoking is exogenous to educational attainment and labor market performance might be unrealistic. The exogeneity assumption requires that the early smoking decision is orthogonal to any factor that affects educational attainment and labor market performance. It is, however, likely that there are unobserved personal

characteristics that affect all three processes. Some individuals, to exemplify, can exhibit myopic behavior in general by opting for immediate pleasure rather than long term achievement. If such a behavior is formed early in life, then these individuals will be more likely to smoke at an early age and will be less likely to complete higher levels of education and less likely to invest in human capital. Thus, an estimated negative effect of smoking will reflect a correlation rather than causality. Distinguishing causality from correlation by relaxing the exogeneity assumption is crucial to explore the true effects of early smoking.

To distinguish causality from correlation, I adopt a model that controls for correlation between unobserved heterogeneity affecting the smoking decision, educational attainment and labor market performance. To establish a causal effect, all processes are modeled simultaneously such that unobserved factors are allowed to be correlated by using discrete mixing distributions. This correlated discrete factor approach is equivalent to a correlated random effects model. The main idea is that unobserved heterogeneity affecting these three processes can be correlated, i.e. they come from a joint mixing distribution. This is akin to assume that the endogeneity of the smoking decision stems from unobserved time invariant factors affecting early smoking and educational attainment, such as innate ability or rate of time preferences.[10] Since the early smoking decision is taken before the age of 15, reverse causality is not an issue here. Therefore, this assumption fits well in the empirical question that this study investigates. In the absence of reverse causality, this method corrects the endogeneity problem stemming from possible omitted variables.

The joint density function of the completed duration of smoking initiation, duration of smoking, educational attainment and labor market performance—$g_3(t, \tau, y_{ed} = k_{ed}, y_j = k_j \mid x, x_1, x_{2,ed}, x_{2,j})$ is specified as:

$$\int_{\epsilon_{ed}} \int_{\epsilon_j} \int_v \int_u f_s(t \mid x, u) f_q(\tau \mid x_1, v) Prob(y_{ed} = k_{ed} \mid x_{2,ed}, \epsilon_{ed})$$
$$Prob(y_j = k_j \mid x_{2,j}, \epsilon_j) dG(u, v, \epsilon_{ed}, \epsilon_j) \tag{6}$$

where $G(u, v, \epsilon_{ed}, \epsilon_j)$ is a discrete mixing distribution underlying unobserved heterogeneity affecting age of onset of smoking, duration of use, educational attainment and labor market performance.

4 Parameter Estimates

4.1 The Dynamics of Smoking

Table 5 presents the parameter estimates of mixed proportional hazard models for starting and quit rates, for both males and females. The negative coefficient estimates on religiosity and age-cohort dummies show that individuals who were living with religious parents during their adolescence and individuals who belong to older birth

---

[10] Indeed, Kang and Ikeda (2014) show that time preferences related to smoking are interpersonal rather than being intra-personal. Thus, they are persistent over time.

cohorts have smaller hazard rates. In other words, they have a lower probability of initiating smoking. Moreover, males in couples seem to have higher quit rates compared to singles. In addition, those who start smoking at early ages are less likely to quit smoking. The parameter estimates for females display very similar results. Panel (b) in Table 5 presents the estimates for duration (age) dependence parameters. In line with the patterns observed in Fig. 1, smoking initiation makes a peak at the age of 16 and then a smaller peak at the age of 18 for both gender groups.

Panel (c) in Table 5 presents the parameter estimates of unobserved heterogeneity. In all columns, I set the second mass points to minus infinity, thereby allowing for the possibility of zero starting and quit rates. Columns (2) and (3) show that three mass points are identified in the joint mixing distribution. The finding of 3 points of support suggests that there are three types of individuals regarding starting and quitting smoking. The first group consists of individuals with a positive starting and positive quitting rate. The second group consists of those with a positive starting rate but a zero quitting rate; the third group, those with a zero starting rate. For the last group, therefore, quitting rate does not exist at all. The parameter estimates of probabilities associated to these mass points show that 47% of the males and 45% of the females have a positive starting rate and a positive quitting rate; 22% of the males and 18% of the females have a positive starting rate but a zero quitting rate. 31% of the males and 37% of the females have a zero starting rate. Finally, the log-likelihood test statistics presented in the same panel shows that correlation between unobserved heterogeneity affecting starting rates and quit rates is statistically significant.[11] Therefore, it is important to jointly model starting and quit rates to identify the unobserved heterogeneity behind the dynamics of smoking.

4.2 Educational Attainment

Table 6 displays the estimated parameters of individual and correlated ordered probit models. Columns (1) and (3) show that early smoking has a negative association with educational attainment for both males and females. As clarified before, these regressions ignore the possible endogeneity of the smoking decision. Accordingly, the parameter estimates are bound to be inconsistent.

Columns (2) and (4) present the results of the joint models that control for the possible correlation between unobserved heterogeneity affecting education and smoking. As can be seen in both columns, the parameter estimate of the early smoking effect decreases in size. For females early smoking does not have a causal effect on the educational attainment, and previously reported negative effect is purely due to the correlation through unobserved factors. For males, on the other hand, even though the coefficient estimate decreases in the joint model, it remains significant. Therefore, I cannot rule out the possibility that early smoking has a causal effect on educational attainment for males. Finally, the statistics for likelihood ratio tests that appear in panel

---

[11] However, note that a formal LR test is problematic since one of the parameters (α) is not identified under the null hypothesis. This will be the case in the other LR tests reported in this study where the null hypothesis characterizes no unobserved heterogeneity case.

(c) show that correlation between unobserved heterogeneity affecting education and smoking dynamics is statistically significant.[12]

Admittedly, the finding that early smoking has a causal effect on education might seem surprising. One can argue that a few years of smoking would not possibly cause significant health problems that can impair the youth and prevent him or her from completing education. Such an argument has, of course, merit for the current analysis as well because even those who start using tobacco at an early age will not consume it for long years before they finish their education. Undeniably, certain adverse health effects can naturally be observed if individuals use tobacco early. Some of such affects are briefly discussed before. However, for the sake of the argument, I assume that these effects are not observed. If negative health effects are not driving the significant results presented in Table 6, and if none the other mechanisms discussed before is strong enough for an early smoking effect, then what can explain the results?

An alternative mechanism is the possibility of exogenous time-varying shocks that can simultaneously affect educational attainment and the smoking decision. For example, loss of a friend or a family member or parental divorce can cause frustration and depression; resulting in both lower education and involvement in risky health behaviors including smoking. Panels (a) and (b) of Table 7 attempt to control for some of such possibilities. Panel (a) introduces a dummy variable for early loss of parent(s); panel (b), a dummy variable for early parental divorce. In both cases, "early" means that the mentioned frustrating and depressive event takes place before the age of 15. Under both specifications the smoking effect remains unchanged; therefore, I conclude that the early smoking effect is robust to exogenous childhood shocks.

One can alternatively argue that the proposed joint model is unable to capture unobserved factors affecting the smoking decision and educational attainment. To investigate such a possibility I perform several robustness analyses by introducing control variables that are expected to be highly correlated with the unobserved factors that can affect both processes. One of such unobserved factors could be systematic differences between the rate of time preferences between individuals. If such preferences are formed early in life, then they can explain the negative coefficient estimates of early smoking. Individuals with high rates of time preference will place a higher value on present than on future. Consequently, they will be more likely to enjoy risky health behaviors and less likely to invest in human capital (Levine et al. 1997). In panel (c) of Table 7, I control for certain preference patterns to check the robustness of the smoking effect. If the early smoking effect changes after adding the preference variables, then it means the joint model fails to capture unobserved systematic differences. The preference patterns for which the specification controls are risk aversion, prudence and temperance. "Appendix 3" gives more information on the measurement and the use of these preference variables. For both males and females, parameter estimates in the table show that the results do not change; therefore, such preferences are already captured by the joint model and not driving the main results.

---

[12] Note that in this education estimation late starters, those who start smoking after the age of 14, are in the reference group. Therefore the early smoking effect can be actually a lower bound for the actual effect of smoking on education. Lower panel of Table 6 presents the results after I add another dummy variable for those who start smoking after the age of 14. The early smoking effect does not change.

Another unobserved factor can be the selection into peer groups. For any risky health behavior, peers can have an effect on the individual. If for example, certain students select into certain peer groups, their education and smoking behavior can be affected simultaneously. It should be noted that current econometric model is exactly designed for such unobserved factors. If individuals select into peer groups based on some unobserved factors before smoking initiates, then this will be captured by the model. If selection happens after smoking initiates, based on the smoking behavior, then this is a consequence of smoking rather being a problem. Even though the econometric model is designed to capture such issues, I performed several sensitivity analysis using proxies for peer effects. Panel (d) of Table 7 introduces control variables for parental education. The smoking effect is robust to the inclusion of these variables. Finally, the last panel of Table 7 displays the results of an estimation where I include a dummy variable for the early use of alcohol (before age of 15). There is no data on the peer use of tobacco during childhood. Therefore, the idea here is that the peer effects that can lead an individual to use tobacco can also lead the same individual to involve in other risky health behaviors such as early alcohol use. In other words early alcohol use can be used as an imperfect proxy for the general peer effects. The results show that the early smoking effect is unchanged.[13] All in all, the results of these sensitivity checks show that the joint model successfully controls for correlated unobserved factors that can jointly affect the smoking decision and educational attainment.

Finally, the remaining mechanism possibly explaining the negative smoking effect is that early smoking has indeed a causal effect on education via the various channels mentioned before.[14] Unfortunately the data at hand does not allow for exploration of these mechanisms for there is no information on physical or mental health of individuals at young ages, or on attendance patterns at schools. If there is indeed a causal effect, then those who start smoking at an early age unintentionally enter a different life-labor path. Subsequently, such an early path diversion between smokers and non-

---

[13] In unreported estimations I also included a dummy for the early cannabis use. The coefficient estimate for the smoking effect is -0.26 and -0.09 for males and females, respectively. However, there are only a handful of individuals who start using cannabis before the age of 13. The results are also robust to inclusion of peer use of tobacco at the survey time. If selection into peer groups persists, then controlling for current peer use of tobacco can be used as a proxy for peer use in childhood.

[14] One special aspect of the Dutch education system is that the students are assigned to different types of secondary school after the age of 12 based on their academic success. That means it is possible that certain types of kids get together in different types of schools, which can affect the smoking behavior. In such a case it is hard to assume no-reverse-causality. Even though I try to control for possible peer effects in a reported estimation, I investigate the effects of this early assignment into schools. I performed the baseline joint analysis with a new early smoking dummy which takes a value of 1 if the respondent starts smoking before the age of 13 (the age at which they are in the assigned secondary school). This new analysis cannot suffer from any possible reverse causality because for these starters the provided education is exactly the same. The estimated coefficients for this new early smoking dummy are -0.23 for males, and 0.01 for females. Even though there are only a few individuals starting smoking before the age of 13, reverse causality does not seem to be an issue. The new coefficients are very similar to the baseline results, albeit imprecisely estimated.

In an alternative method, I checked the robustness of the early smoking effect by introducing educational attainment dummies into the starting rates analysis. Even though this estimation is not perfect due to the unavailability of information on the early education tracks of the respondents, it serves well as a sensitivity analysis. The results showed that the negative early smoking effect for males is robust. Full estimation results are available upon request.

smokers can result in serious disadvantages for the former through the accumulation of the effects. The following section will partly shed light on 2 of these disadvantages by presenting the analysis of the first and the current job rankings.

4.3 Labor Market Performance

4.3.1 The First Job Rankings

Table 8 presents the coefficient estimates of the joint models on the first job rankings for males and females. The first column suggests that the early smoking decision has a negative effect on the type of the first job that a male has. Nonetheless, this effect alters once the specification includes educational attainment variables. Considering the discussion about the effects of early smoking on education in the previous section, one can indeed expect a change in the coefficient estimate. However, the change is so manifest that as presented in the second column of the table, keeping the level of education constant, early smoking does not affect the first job. Considering the afore-mentioned mechanisms through which smoking can affect labor market performance, this empirical finding is reasonable. For the first jobs, no employer discrimination is expected as the smoking status of the first time job applicants will not be observable. Moreover, the serious health consequences that can affect the productivity are probably not observed since they take place after long years of use. The remaining mechanism is the educational attainment, which is what the results also indicate.

Additionally, column (3) controls for possible endogeneity of the educational attainment in the first job rankings estimation. The coefficient estimate of early smoking is unchanged; showing that it is robust to the extension of the functional form defining unobserved heterogeneity. The last three columns present the same results for females. For both gender groups, statistics of likelihood ratio tests that appear in panel (d) reveal that the correlation between unobserved heterogeneity affecting smoking dynamics, education and the first job rankings is significant. Therefore, estimations that fail in controlling for this correlation would suffer inconsistency.

Table 9 presents the results of various sensitivity analysis on the joint model for males and females, respectively. The first sensitivity analysis controls for the preference patterns to take account of different rate of time preferences, prudence and patience. Although the sample size decreases substantially, a quick comparison with the columns (3) and (6) of Table 8 reveals that the results remain similar. The second sensitivity analysis controls for the search efforts before an individuals finds his or her first job. Similar to the previous sensitivity analysis, the purpose is to take account of different preferences. This sensitivity analysis also takes account of possible exogenous shocks that can impair abilities to put search effort. The results show that the coefficient estimate of early smoking is robust to this specification for both gender groups. The final sensitivity check controls for possible calendar effects by introducing the year in which an individuals starts his or her first job. The results are also robust to this final specification for both gender groups.

All in all, the results indicate that early smoking has an adverse effect on the first job rankings, but only through education. Therefore, if the academic problems stemming

from early smoking can be prevented, early smokers will be less likely to start their career in the labor market from a disadvantaged point. Admittedly, the first job is only a part of the life-time labor market performance, and the next section will explore the possible long-term effects of early smoking by investigating the effects on the current jobs of individuals.

4.3.2 The Current Job Rankings

Table 10 presents the parameter estimates of the current job ranking regressions.[15]

Column (1) in Table 10 present the results when educational attainment is not controlled for. Column (2) adds the educational attainment dummies into the analysis. Comparing the coefficient estimates of early smoking in columns (1) and (2) shows that, similar to the first job rankings, early smoking has an effect on the current job rankings through education for males.

Column (2) shows the estimates of individual model for males, where early smoking is assumed to be exogenous. There is a negative and significant association between early smoking and the current job rankings. Column (3) presents the results of the correlated model, where the correlation between unobserved heterogeneity is taken into account. The coefficient estimate of early smoking decreases from column (2) to column (3). However, it remains significant.[16] Columns (5) and (6) present the results for females. Neither in the individual model nor in the joint model there is evidence for a negative effect. Likelihood ratio tests show that the joint model is preferred over individual models, which ignore the correlation between unobserved heterogeneity.

Contrasting the results in the previous section with the ones here shows that there is a difference between the first job rankings and the current job rankings in terms of the early smoking effect. Even though there is no evidence for the early smoking effect on the first job rankings (conditional on education), the results in Table 10 show that there is an effect on the current job rankings. Statistically, there can be two main reasons for such a phenomenon. The first is that those who start smoking at an early age are more likely to move downward in the labor market performance rankings, compared to those who do not start smoking early. The second is that early smokers are less likely to move upward. Since there are only a handful of observations where the first job is higher ranked than the current job, the second explanation sounds more plausible.

To check if early smokers are indeed less likely to more upward, I perform a similar ordered probit estimation, where the dependent variable is the difference between the

---

[15] Note that there might be a selection bias in the ranked information on jobs since the ranking exists for individuals who select into employment. In order to check if this is a serious problem in the data at hand, I re-estimated the simple ordered probit model by re-categorizing the job rankings. In the new rankings I set the first category for those who do not have a paid job. The remaining categories are kept the same. The results and interpretation of this estimation remain the same.

[16] One caveat in the current job rankings estimations is that cohort effects can be important. The reason is that since the current job information is collected for everyone in the sample in the same year, young cohorts are less likely to have a current job which is different than their first job. Ideally age categories control for such an effect because they serve as cohort dummies. Moreover, I also included a variable to control for the years elapsed from the moment respondents started working in their first jobs until the survey time. The results did not change, because as expected this new variable is highly correlated with the cohort dummies.

current job rankings and the first job rankings. This variable takes a value of 1 if there is no change or the difference is negative; 2 if the difference is 1; 3 if the difference is 2; and 4 if the difference is more than 2. Table 4 displays the sample statistics of these job transitions. The parameter estimates are reported in Table 11. Since the chances of moving upward in the rankings can depend on the initial position, I also control for the first job type. Coefficient estimate of early smoking shows that males who start using tobacco at an early age are less likely to move upward in the job rankings. This inertia explains why there is no effect on the first job, but there is on the current job. This finding is in line with the expected long term adverse effects of smoking (especially on physical health). Apparently, the accumulation of effects distorts labor market performance and harms individuals' ability to move upward in the job rankings.

Finally, Table 12 succinctly connects the labor market effects of early smoking to the wider literature on the wage effects of tobacco consumption. The dependent variable, in this estimation, is the log hourly wages of the individuals calculated as $\frac{Monthly\ wages}{hours\ worked\ per\ week*4.29}$. For hourly wage estimations, I use information from individuals who report at least 20 h of work per week or at most 60.[17] Unlike the foregoing sections, the dependent variable is not an ordered one in this case. Therefore, "Appendix 4" briefly presents the econometric model that produces the parameter estimates in the table. The main idea behind the estimation is the same; the joint model allows for correlation between unobserved heterogeneity affecting hourly wages, smoking dynamics, education and job rankings.[18]

In order to analyze the effect of early smoking, I initially include the early smoking dummy and the available control variables in the wage equation. The results show that early smoking negatively affect the wages of males. There seems to be no negative effect for females. In order to see the occupation-specific effects of early smoking, I include interactions between the early smoking variable and the job type dummies into the wage equation. Lower panel of Table 12 report the results. For almost all of the job types of males early smoking negatively affects wages. The effect seems to be the highest for lower ranked jobs. For females, on the other hand, there is no evidence for a smoking effect.

Although not reported, I performed several other sensitivity analysis on the current job rankings, wages and the probability of switching to a better job. First, I re-estimated the effect of early smoking on current labor market performance by using information about the current status of smoking. It seems that negative effect on the wages and the current job is higher for those who start smoking early and continue smoking until the time of the survey. However, the coefficient estimate of interaction between early smoking and current smoking is imprecisely estimated. The possible reason is that the vast majority of those who start using tobacco before the age of 15 consists of still-

---

[17] These are natural cutoff points in the data. Only a handful of male respondents report working less than 20 h and more than 60 h. One respondent reports 124 h of working per week.

[18] Similar to the ordered probit model on the current job rankings, there might be a selection bias in the wage estimation because wages exists only for those who select into employment. I controlled for the same bias by estimating a simple Heckman's sample selection model. For both males and females, Mills ratio is found to be insignificant. In short, selection bias due to sample selection is not an issue in the wage estimations.

smokers. Second, I controlled for risk attitudes in the analysis of current labor market performance. The reason is that preferences and risk attitudes can affect incentives to invest further in human capital, e.g. through on the job trainings. The results in Tables 11 and 12 are found to be robust to the inclusion of the risk attitude and preference variables. Furthermore, using ordered logit models instead of probit ones did not also change the results, only rescaled the coefficient estimates.

4.4 Magnitude of the Effects

Figure 4 briefly displays the magnitude of the early smoking effect on educational attainment and the current job rankings for only males. The numbers in the figure are obtained as follows. First, I simulated the probabilities of belonging to the each category in the ordered choices for those who smoked before the age of 15 using the estimates presented in column 2 of Table 6 and column 2 of Table 10. Then, the same probabilities are simulated for those who did not smoke before the age of 15. The numbers in the figure reflect the differences between these probabilities for the each category. For all of the other control variables, sample means are used in the simulations.

The results show that early smoking decreases the probability of finishing a high level of education. The effect is the largest on the probability of finishing a scientific degree. The probability of completing scientific education is 4%-point lower for someone who smoked before the age of 15. Similarly, early smoking decreases the probability of having an academic job by almost 4%-point in the long run.

5 Conclusion

There is a small literature studying the causal effects of smoking on labor market performance. The majority of the studies within this literature focuses on earnings or hourly wages. The literature on the relationship between smoking and educational attainment is even smaller. A handful of studies explore the association between smoking behavior and education without establishing causal effects.

This study focuses on the effects of early initiation of smoking on educational attainment and labor market performance. It uses not only hourly wage information but also ranking of jobs to measure labor market performance as this information is available for both the first jobs and the current jobs. Since educational attainment and job ranking variables have ordinal character, ordered probit models are used in estimations. The results indicate that there is a strong negative association between smoking and education as well as smoking and labor market performance for both the first job and the current job rankings. However, it is possible that smoking and education, and smoking and labor market performance are jointly determined by a set of unobserved factors. To tackle this endogeneity problem, the current study uses a correlated discrete factor approach, which is equivalent to a correlated random effects model in which the main idea is that unobserved personal characteristics affecting smoking, education and labor market performance can be correlated. In the absence of reverse causality, this method yields causal effects as it controls for the endogeneity problem stemming from omitted variables.

The results show that early smoking has a negative effect on educational attainment for males only. This negative effect is robust to several sensitivity checks such as controlling for risk preferences and depressive childhood events. Apparently, smoking does not only have long term negative consequences, but also can start affecting one's life earlier on. Once education is controlled for, there is no evidence for the early smoking effect on the first jobs. The only effect seems to be through education. This effect suggests that early smokers start their labor market career from a disadvantaged point, and this disadvantage is due to that early smokers perform worse in schools. No causal effect is found for females.

Unlike the first job rankings, educational attainment is not the only channel through which early smoking affects the current job rankings of males. The results show that there is still an effect on the current job conditional on educational attainment; which is a finding in line with the existing literature on the wage effects of smoking. Proposed mechanisms for the wage effects of smoking are mainly discrimination, serious health consequences and smoking breaks at workplaces. All these mechanisms might work after years in the labor market. All in all, it seems that the adverse effects of early smoking accumulates over time and early smokers who start their career with low ranked jobs become stuck in those jobs or become less likely to make a career. Finally, an analysis of the log-hourly-wages shows that reported wage effects of early smoking may be due to the possible effects on the type of jobs.

The reported effects of early smoking on educational attainment of males suggest that policies against smoking, targeting especially youth, can be indeed effective. It is not only the case that early smoking affects solely education, but it also affects other important life outcomes through education. Therefore fight against the adverse effects of smoking needs to begin very early in schools for once the smoking effect on educational attainment materializes, there might be several long term consequences. If the negative effect of early smoking on education can be prevented, the indirect adverse effects can also be prevented. The easiest way to do so is, of course, to prevent young individuals from initiation into smoking. This requires a much more detailed analysis of the determinants of tobacco uptake. Only then it is possible to identify the more vulnerable individuals and fight against the negative aspects of tobacco uptake. Furthermore, the difference between males and females in terms of the early smoking effect indicates that there is need for further analysis of the mechanisms through which smoking affects education and labor market performance. Apparently, some of the proposed mechanisms work only for males.

Finally, the existence of early smoking effect on the current job conditional on educational attainment suggests that early smoking affects labor market performance through other channels in the long run. Apparently, the problems related to early initiation of smoking accumulate, and it is not only the education that matters in the long run. This further calls for preventive measures.

Compliance with Ethical Standards

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

Appendix 1: Ordered Variables

Tables 1 and 2 below show the ordered categories of the education and the job ranking variables. In the Netherlands, currently, the compulsory education ends at the age of 16; corresponding to a VMBO degree. Any of the other degrees are based on voluntary

Table 1 Ordered categories of education variable

$Education_i$	Explanation	Dutch abbreviation	Expected age at graduation
1	Primary or no education		16−
2	Special or other education		16−
3	Preparatory middle-level applied education (vocational)	VMBO	16
4	Preparatory middle-level applied education (technical, combined)	VMBO	16
5	Higher general continued education	HAVO	17
6	Preparatory scholarly education	VWO	18
7	Middle-level applied education	MBO	19
8	Higher vocational education	HBO	21
9	Scientific education (bachelor and higher)	WO	21+

Table 2 Ordered categories of the job variables

Job_i	Explanation	Hourly wages	
		Males	Females
1	Agrarian job (e.g., farm worker)	12.5	9.2
2	Unskilled and trained manual work (e.g., cleaner)	13.3	12.2
3	Semi-skilled manual work (e.g., driver, factory worker)	14.6	13.8
4	Skilled manual work (e.g., car mechanic, electrician)	15.4	13.9
5	Other clerical work (e.g., administrative assistant)	17.7	14.6
6	Intermediate supervisory or commercial work (e.g., department manager)	18.4	15.2
7	Intermediate academic or independent job (e.g., teacher, nurse)	19.7	17.0
8	Higher supervisory job (e.g, manager, director)	28.5	24.4
9	Higher academic or independent job (e.g., physician, scholar)	25.4	20.7

education. Due to the changes in the education system, name of the degrees obtained by the older part of the sample may be different. For those cases, I used the current categories that correspond to the old ones, according to CBS (Statistics Netherlands).

The information on jobs come jointly from Working and Schooling and Wage Indicator data sets. The first job and the current job have the same categories, disregarding side jobs and holiday jobs. The columns on the right display the average hourly wages (in Euros) for each category. There are other variables which might possibly be used in the ordering of the jobs. However, for none of these variables such as satisfaction with job, satisfaction with wages or working hours there is enough variation between job categories. It seems that almost all of the respondents have similar levels of satisfaction with their jobs, wages or working hours regardless of the job that they have.

Appendix 2: Descriptive Statistics

Definition of the variables used throughout this study and their summary statistics are given below (Tables 3 and 4):

- Smoke: Dummy variable with a value of 1 if individual ever used tobacco; 0 otherwise.

- $Smoking_{15-}$: Dummy variable with a value of 1 if individual used tobacco before age of 15; 0 otherwise.

- Background variables.
  - Religious: Amount of the times that the parents of the respondent visited the church in a week when the respondent was 15 years old.
  - Migrant: Dummy variable with a value of 1 if individual is migrant; 0 otherwise.

- Cohort effects

  - Cohort (30−) (Reference): Dummy variable; 1 if aged below 30; 0 otherwise.
  - Cohort (30–36): Dummy variable; 1 if aged between 30 and 36; 0 otherwise.
  - Cohort (37–43): Dummy variable; 1 if aged between 37 and 43; 0 otherwise.
  - Cohort (44–49): Dummy variable; 1 if aged between 44 and 49; 0 otherwise.
  - Cohort (50+):Dummy variable; 1 if aged above 50; 0 otherwise.

- Urbanization level

  - Very urban: Dummy variable with a value of 1 if the municipality of residence is very urban (population density per km^2 is above 2500); 0 otherwise.
  - Urban: Dummy variable with a value of 1 if the municipality of residence is urban (population density per km^2 is between 1000 and 2500); 0 otherwise.
  - Rural: Dummy variable with a value of 1 if the municipality of residence is rural (population density per km^2 is between 500 and 1000); 0 otherwise.
  - Very rural (Reference): Dummy variable with a value of 1 if the municipality of residence is very rural (population density per km^2 is below 500); 0 otherwise.

- Domestic situation

Table 3 Descriptive statistics of the variables

	Males		Females	
	Mean	SD	Mean	SD
Variables of interest				
Smoke	0.69	0.46	0.64	0.48
Smoke<15	0.26	0.45	0.24	0.43
Educational attainment				
Primary or no	0.02	0.15	0.02	0.12
Special or other	0.02	0.15	0.03	0.16
Prep. vocational	0.11	0.32	0.1	0.29
Prep. technical	0.11	0.32	0.13	0.34
General continued	0.05	0.22	0.06	0.24
Prep. scholarly	0.05	0.21	0.03	0.17
Middle applied	0.26	0.44	0.29	0.45
Higher vocational	0.24	0.43	0.26	0.44
Scientific	0.13	0.34	0.09	0.28
The first job rankings				
Agrarian	0.05	0.21	0.01	0.1
Unskilled manual	0.07	0.26	0.08	0.27
Semi-skilled manual	0.16	0.37	0.05	0.21
Skilled manual	0.16	0.37	0.01	0.12
Other clerical	0.2	0.4	0.42	0.49
Int. supervisory	0.12	0.33	0.07	0.25
Int. academic/independent	0.13	0.34	0.31	0.46
Higher supervisory	0.02	0.14	0.02	0.12
Higher academic/independent	0.08	0.27	0.03	0.18
The current job rankings				
Agrarian	0.03	0.17	0.01	0.1
Unskilled manual	0.02	0.13	0.08	0.26
Semi-skilled manual	0.1	0.3	0.04	0.2
Skilled manual	0.13	0.34	0.02	0.12
Other clerical	0.16	0.37	0.32	0.47
Int. supervisory	0.18	0.38	0.1	0.3
Int. academic/independent	0.17	0.38	0.34	0.47
Higher supervisory	0.12	0.32	0.05	0.22
Higher academic/independent	0.1	0.3	0.05	0.22
Control variables				
Religious	2.14	1.87	2.15	1.89
Migrant	0.14	0.34	0.11	0.32
Urbanization dummies				
Very urban	0.11	0.32	0.12	0.33
Urban	0.51	0.44	0.49	0.43

Table 3 continued

	Males		Females	
	Mean	SD	Mean	SD
Rural	0.22	0.41	0.21	0.41
Very rural	0.16	0.37	0.18	0.38
Domestic situation				
Single wo. children	0.14	0.35	0.13	0.34
Married wo. children	0.3	0.46	0.29	0.46
Married w. children	0.52	0.5	0.5	0.5
Single w. children	0.03	0.16	0.07	0.25
Cohort dummies				
Cohort (30−)	0.07	0.12	0.07	0.15
Cohort (30–36)	0.18	0.38	0.21	0.41
Cohort (37–43)	0.29	0.46	0.28	0.46
Cohort (44–49)	0.31	0.47	0.32	0.47
Cohort (50+)	0.15	0.4	0.11	0.35
Risk attitudes				
Risk averse	0.45	0.5	0.6	0.49
Prudence	0.57	0.49	0.57	0.5
Temperance	0.43	0.5	0.49	0.5
Parents' education				
Mother edc	4.36	2.88	4.27	2.53
Father edc	5.27	3.1	5.35	3.19
Childhood events				
Parents death	0.04	0.18	0.03	0.18
Parent divorced	0.04	0.19	0.05	0.21
Search efforts (first job)				
No search	0.46	0.5	0.42	0.49
Short search	0.33	0.47	0.39	0.49
Moderate search	0.11	0.31	0.09	0.29
Intensive search	0.1	0.31	0.11	0.31
Moderate search	0.11	0.31	0.09	0.29
Intensive search	0.1	0.31	0.11	0.31

- Single wo. children (Reference): Dummy variable; 1 if individual is single without children; 0 otherwise.
- Married wo. children: Dummy variable with; 1 if individual is married without children; 0 otherwise.
- Married w. children: Dummy variable with; 1 if individual is married with children; 0 otherwise.
- Single w. children: Dummy variable with; 1 if individual is single with children; 0 otherwise.

Table 4 Job transitions (from the initial job rankings to the current job rankings)

	Males		Females	
	Mean	SD	Mean	SD
The current job < the first job	0.03	0.18	0.09	0.29
The current job = the first job	0.54	0.50	0.66	0.47
The current job > the first job	0.43	0.50	0.25	0.43
The current job = the first job +1	0.13	0.34	0.08	0.28
The current job = the first job +2	0.14	0.34	0.09	0.29
The current job > the first job +2	0.16	0.37	0.08	0.26

Table 4 shows the percentage of individuals in each category based on the difference between the first job rankings and the current job rankings

- Risk attitudes

    - Risk Aversion: Dummy variable; 1 if the individual displays risk aversion; 0 otherwise.
    - Prudence: Dummy variable; 1 if the individual displays prudence; 0 otherwise.
    - Temperance: Dummy variable; 1 if the individual displays temperance; 0 otherwise.

- Parental characteristics

    - Mother's education: Highest degree the mother of the individual obtained; 1–9.
    - Father's education: Highest degree the father of the individual obtained; 1–9.
    - Parental loss: Dummy variable; 1 if the individual lost at least one of his or her parents before the age of 15.
    - Parental divorce: Dummy variable; 1 the parents of the individuals divorced before the age of 15.

- Search efforts before the first job

    - No search: Dummy variable; 1 if the individual reports no search effort for the first job.
    - Short search: Dummy variable; 1 if individual spent less than 1 month to find his or her first job.
    - Moderate search: Dummy variable; 1 if individual spent 1-3 months to find his or her first job.
    - Intensive search (Reference): Dummy variable; 1 if individual spent more than 3 months to find his or her first job.

Appendix 3: Additional Data Sets from the LISS Panel

In addition to the three main single wave studies of the LISS panel—Alcohol and Drugs, Working and Schooling, and Wage Indicator—the following two data sets are used in some sensitivity analysis.

Life History Questionnaire

Life history questionnaire is a single wave study consisting of several parts that concern family histories of the respondents. In each part of the survey the respondents answer different questions regarding their family situation during childhood. The questionnaire was made available in 2012, and in total 5231 individuals completed the questionnaire.

Measuring Higher Order Risk Attitudes of the General Population

This single wave questionnaire concerns the measurement of the degree of risk aversion, prudence and temperance of respondents by recording answers given to several choices between lotteries. The respondents were presented lottery choices after being assigned to different situations by means of random dice throws. Their choices between lotteries are used to measure the risk attitudes.

I use the same strategy in Noussair et al. (2014) to measure the incidence of prudence, temperance, and risk aversion. I measure risk aversion with the number of safe choices an individual makes, out of the five decisions involving a sure payoff and a risky lottery. Similarly, I measure prudence with the number of prudent choices, and temperance with the number of temperate choices that an individual makes. Then, I assume that an individual is risk averse (prudent, temperate) if the number of choices is greater than 3. Since the data set used in Noussair et al. (2014) is the same, a more detailed explanation of the questions and measurement strategy can be found in their study.

Appendix 4: ML of the Starting Rates, Quit Rates, Ordered Probit Models and the Wage Equation

Starting Rates

The specification of the hazard rate for starting rates (following Eq. 1) yields the following functional form for the conditional density function of the completed durations until the uptake of tobacco;

$$f_s(t \mid x, u) = \theta_s(t \mid x, u) \exp\left(-\int_0^t \theta_s(\mu \mid x, u)d\mu\right) \tag{7}$$

Integrating out the unobserved heterogeneity in the conditional density in Eq. 7 gives the following density function for the duration until tobacco uptake (t) conditional on x, but unconditional on u:

$$f_s(t \mid x) = \int_u f_s(t \mid x, u)dG(u) \tag{8}$$

where $G(u)$ is a discrete mixing distribution where the number of support points are chosen based on empirical tests. In case of 2 support points, two types of individuals exist regarding the hazard rate for tobacco uptake: those who are more likely to smoke and those who are less likely to smoke. Each individual has a probability of belonging to one of these types, and the probabilities, being the same for everyone, are denoted as follows: $\Pr(u = u_a) = r$ and $\Pr(u = u_b + u_a) = 1 - r$. r has a logit specification; $r = \frac{\exp(\alpha)}{1+\exp(\alpha)}$, where α is the parameter defining the probabilities and to be estimated by the model. The mixed proportional hazard framework assumes that α does not depend on any observables, including calendar or age effects.

The log-likelihood that accounts for the discrete nature of the observations on the onset age of smoking is

$$\sum_{i=1}^{n} d_{s,i} \log \left[F_s(t_i - 1) - F_s(t_i) \right] + \left(1 - d_{s,i} \right) \log \left[1 - F_s(t_{s,i}) \right] \tag{9}$$

where i is an index for individual, n is the number of individuals in the sample and $d_{s,i}$ is a dummy variable that is equal to 1 if an individual started using tobacco and equal to 0 otherwise.

Quit Rates

The conditional density function for the completed durations until the last use (following Eq. 2) is

$$f_q(\tau \mid x_1, v) = \theta_q(x_1, v) \exp(-\theta_q(x_1, v) \cdot \tau) \tag{10}$$

where x_1 is the set of control variables. v is the unobserved heterogeneity in the quit rates. Integrating out the conditional density function over the interval of maximum and minimum values that the duration can take on, yields the distribution function, F_q

Using the distribution function, F_q, I specify the following log-likelihood for the analysis of quit rates

$$\sum_{i=1}^{m} d_{q,i} \log \left[F_q(\tau_{q,i}) \right] + \left(1 - d_{q,i} \right) \log \left[1 - F_q(\tau_{q,i}) \right] \tag{11}$$

where m is the number of individuals that ever used tobacco and $d_{q,i}$ is a dummy variable that has a value of 1 if the individual stopped using tobacco and a value of 0 if the individual did not stop using tobacco. Individuals who report using tobacco in the last 30 days are right censored, i.e. they are assumed to be non-quitters. I perform the quitting analysis using only those who ever use tobacco; otherwise quit rates do not exist. Similar to starting rates I assume that there are 2 unobserved heterogeneity groups.

Ordered Probit Models

The observed ordered categories in the data are

$$
y_{ed} = \begin{cases}
1 & \text{if } y_{ed}^* \leq \mu_1 \\
2 & \text{if } \mu_1 < y_{ed}^* \leq \mu_2 \\
. \\
. \\
k_{ed} - 1 & \text{if } \mu_{k_{ed}-2} < y_{ed}^* \leq \mu_{k_{ed}-1} \\
k_{ed} & \text{if } \mu_{k_{ed}-1} < y_{ed}^*
\end{cases} \tag{12}
$$

where μ's are to be estimated threshold parameters in the ordered probit model and k_{ed} is the number of alternatives in the ordered choice; 9. Assuming that the error term e_{ed} has a standard normal distribution, one can write the following probabilities for the ordered probit model, conditional on observable and unobservable individual heterogeneity[19]:

$$
\Pr(y_{ed} = k_{ed}|x_{2,ed}, \epsilon_{ed}) = \Phi(\mu_{k_{ed}} - x_{2,ed}'\beta_{ed} - \epsilon_{ed}) - \Phi(\mu_{k_{ed}-1} - x_{2,ed}'\beta_{ed} - \epsilon_{ed}). \tag{13}
$$

where $\Phi(.)$ is standard normal cdf. with $\Phi(\mu_0 - x_{2,ed}'\beta_{ed} - \epsilon_{ed}) = 0$ and $\Phi(\mu_9 - x_{2,ed}'\beta_{ed} - \epsilon_{ed}) = 1$. Since there are heterogeneity specific constants in the model, I set the first threshold parameter μ_1 to zero. The other threshold parameters are modeled in the following way so as to ensure that the probabilities are positive and thresholds are ordered: $\mu_2 = \gamma_1^2, \mu_3 = \mu_2 + \gamma_2^2, \ldots,$ and $\mu_{k_{ed}-1} = \mu_{k_{ed}-2} + \gamma_{k_{ed}-2}^2$. Analogous to starting and quit rates, integrating out the unobserved heterogeneity in the conditional probabilities, given in Eq. 13, yields the unconditional ones. Explicitly;

$$
\Pr(y_{ed} = k_{ed}|x_2) = \int_{\epsilon_{ed}} Prob(y_{ed} = k_{ed}|x_2, \epsilon_{ed})dG(\epsilon_{ed}) \tag{14}
$$

where $k \in \{1, 2, 3, \ldots, 9\}$, denoting ordered choices. I assume that $G(\epsilon_{ed})$ is a discrete mixing distribution. In case that there are 2 support points, conditional on observed characteristics, there are 2 types of individuals in the ordered choices on educational attainment: high education types and low education types. The associated probabilities are: $\Pr(\epsilon_{ed} = \epsilon_{a,ed}) = p$ and $\Pr(\epsilon_{ed} = \epsilon_{b,ed} + \epsilon_{a,ed}) = 1 - p$, where p is modeled using a logit specification, $p = \frac{\exp(\alpha)}{1+\exp(\alpha)}$. Finally, the likelihood function of the ordered probit models is $\prod_N Prob(y_{ed} = k_{ed}|x_2)$.

---

[19] For simplicity I write $x_{2,ed}'\beta_{ed} = x'\beta_{ed} + \rho_{ed}smoke_{15-}$

Wage Equation

In order to investigate the wage effects of smoking, I first estimate a simple linear model where log hourly wage is a function of early smoking for males and females separately. To take account of unobserved heterogeneity in the wage equation, I use a discrete factor approach. The wage equation is specified as

$$w_i = \beta_0 + \beta_1 x_{4,i} + \beta_2 t_i + \beta_3 dt_i + \omega_i + e_i \tag{15}$$

where w_i is log hourly wage, x_i represents personal characteristics, t_i is a dummy indicator of early smoking and dt_i is the interaction of early smoking dummy with the first job ranking dummies. ω_i is the unobserved heterogeneity component of the wage equation for hourly wages. e_i is the error term. This discrete factor approach makes the probability density function conditional on unobserved heterogeneity, which can be integrated out once we assume a functional form. The resulting log likelihood of this linear model is

$$L = \sum_{i=1}^{N} log\left(\int_{\omega} f(w|x_4, \omega)dG(\omega)\right) \tag{16}$$

where $f(.)$ is the probability density function of the normal distribution and $G(\omega)$ is a discrete mixing distribution as in the previous cases.

The joint model—for starting rates, quit rates, education, the current job and wages-is specified as:

$$\int_{\omega}\int_{\epsilon_{ed}}\int_{\epsilon_j}\int_v\int_u f_s(t \mid x, u)f_q(\tau \mid x_1, v)Prob(y_{ed} = k_{ed} \mid x_{2,ed}, \epsilon_{ed})Prob(y_j$$
$$= k_j \mid x_{2,j}, \epsilon_j)f_c(w \mid x_4, \omega)dG(u, v, \epsilon_{ed}, \epsilon_j, \omega) \tag{17}$$

where $dG(u, v, \epsilon_{ed}, \epsilon_j, \omega)$ is the mixing distribution.

Appendix 5: Estimates

See Tables 5, 6, 7, 8, 9, 10, 11 and 12.

Table 5 Parameter estimates of the mixed proportional hazard models on the starting and quit rates of smoking

	Males		Females	
	(1)	(2)	(3)	(4)
a. Control variables				
Religious	-0.32 (7.3)**	0.01 (0.2)	-0.05 (2.3)**	0.06 (0.8)
Migrant	-0.13 (1.2)	0.12 (0.8)	-0.13 (1.0)	0.40 (1.0)
Very urban	0.57 (3.8)**	-0.11 (0.5)	0.33 (2.5)	-0.21 (0.5)
Urban	0.27 (2.2)**	0.12 (0.8)	0.21 (1.7)*	-0.30 (0.8)
Rural	0.30 (2.2)**	0.10 (0.6)	0.23 (1.6)	-0.31 (0.7)
Cohort (30–36)	-0.30 (0.8)	-0.15 (0.4)	-0.27 (2.2)**	1.21 (0.5)
Cohort (37–43)	-0.22 (0.7)	-0.85 (2.3)**	-0.56 (3.6)**	-0.58 (0.3)
Cohort (44–49)	-0.37 (1.1)	-1.24 (3.3)**	-0.32 (1.1)	-1.47 (0.9)
Cohort (50+)	-0.10 (0.3)	-1.27 (3.4)**	-0.39 (1.9)*	-1.12 (0.7)
Married wo. children		0.57 (3.1)**		0.73 (2.0)**
Married w. children		0.73 (4.2)**		0.96 (2.6)**
Single w. children		0.25 (0.6)		-0.50 (1.2)
Starting age		0.66 (3.8)**		2.23 (4.8)**
b. Age (duration) dependence				
λ_{20+}	0.22 (1.3)		0.11 (0.4)	
λ_{20}	0.31 (0.9)		1.30 (4.8)**	
λ_{19}	0.74 (2.9)**		1.00 (3.9)**	
λ_{18}	1.30 (7.9)**		1.60 (10.4)**	
λ_{17}	0.35 (1.8)*		0.66 (3.9)**	
λ_{16}	1.22 (9.1)**		1.14 (8.8)**	
λ_{15}	0.66 (5.3)**		0.66 (5.1)**	
λ_{14}	0.01 (0.0)		0.40 (3.2)**	

Table 5 continued

	Males		Females	
	(1)	(2)	(3)	(4)
c. Unobserved heterogeneity				
u_a, v_a	−1.20 (3.6)**	−4.02 (8.4)**	−2.12 (11.0)**	−5.38 (3.0)**
u_b, v_b	−∞	−∞	−∞	−∞
α_1	0.39	(4.9)**	0.23	(3.0)**
α_2	−0.36	(3.3)**	−0.70	(5.5)**
−LogLikelihood	2366.8		2580.6	
LR-Test	40.5**		34.2**	
Observations	1071		1103	

LR-Test presents LR statistics for the significance of correlation between unobserved heterogeneity affecting starting and quit rates. Null hypothesis assumes no unobserved heterogeneity in quit rates

Identified correlation in the mixing distribution indicates 3 types of individuals. The first group consists of individuals with a positive starting and positive quit rate. The second group consists of those with a positive starting rate but a zero quit rate. The third group has a zero starting rate, therefore the quit rate does not exist at all. The associated probabilities are 47, 22, 31% for males, and 46, 18, 36% for females

Absolute *t*-statistics in parentheses. *,** Statistical significance at 10 and 5%, respectively

Table 6 Parameter estimates of the ordered probit models on educational attainment

	Males		Females	
	(1)	(2)	(3)	(4)
$Smoking_{15-}$	−0.27 (3.5)**	−0.20 (2.6)**	−0.12 (1.7)*	−0.04 (0.5)
a. Control variables				
Religious	0.04 (2.0)**	0.04 (2.1)**	0.06 (3.7)**	0.06 (3.5)**
Migrant	0.06 (0.7)	0.06 (0.6)	0.15 (1.6)	0.15 (1.5)
Very urban	0.43 (3.3)**	0.47 (3.6)**	0.28 (2.2)**	0.28 (2.1)**
Urban	0.09 (0.9)	0.09 (0.9)	−0.04 (0.4)	−0.06 (0.6)
Rural	−0.02 (0.2)	−0.02 (0.1)	−0.12 (1.1)	−0.14 (1.3)
Cohort (30–36)	0.14 (0.3)	0.14 (0.3)	−0.04 (0.1)	−0.01 (0.0)
Cohort (37–43)	−0.28 (0.7)	−0.28 (0.7)	−0.34 (0.8)	−0.30 (0.7)
Cohort (44–49)	−0.26 (0.6)	−0.25 (0.6)	−0.54 (1.3)	−0.49 (1.2)
Cohort (50+)	−0.34 (1.8)*	−0.33 (1.8)*	−0.87 (2.1)**	−0.83 (2.0)**
b. Ordered probit thresholds				
γ_1	0.97 (21.5)**	0.98 (20.9)**	0.99 (20.8)**	1.00 (20.7)**
γ_2	0.66 (22.7)**	0.66 (22.6)**	0.75 (26.1)**	0.75 (25.8)**
γ_3	0.39 (14.5)**	0.40 (14.3)**	0.44 (16.8)**	0.44 (16.5)**
γ_4	0.37 (14.1)**	0.37 (13.8)**	0.30 (11.8)**	0.30 (11.8)**
γ_5	1.16 (52.6)**	1.17 (51.7)**	1.24 (58.6)**	1.24 (58.0)**
γ_6	0.31 (9.8)**	0.31 (9.6)**	0.40 (12.0)**	0.40 (11.9)**
γ_7	1.02 (24.3)**	1.02 (24.0)**	0.88 (20.7)**	0.88 (20.3)**

Table 6 continued

	Males		Females	
	(1)	(2)	(3)	(4)
c. Unobserved heterogeneity				
$\epsilon_{ed,a}$	2.18 (5.1)**	2.28 (5.2)**	2.56 (6.0)**	2.61 (6.1)**
$\epsilon_{ed,b}$		−0.39 (4.7)**		−0.33 (3.7)**
α_1		0.33 (4.2)**		0.16 (2.1)**
α_2		−0.22 (2.3)**		−0.57 (5.3)**
−LogLikelihood	4199.4	4188.1	4363.3	4356.6
LR-Test		22.6**		6.7**
Observations	1017	1017	1103	1103
Controlling for the late smoking behavior				
$Smoking15-$	−0.36 (3.9)**	−0.22 (2.2)**	−0.15 (1.9)*	−0.04 (0.4)
$Smoking14+$	−0.14 (1.7)*	0.00 (0.0)	−0.18 (2.4)**	−0.10 (1.2)
−LogLikelihood	4198.0	4187.5	4362.1	4356.1
Observations	1017	1017	1103	1103

LR-Test presents LR statistics for the significance of correlation between unobserved heterogeneity affecting smoking dynamics and education. Null hypothesis assumes no unobserved heterogeneity in education

For the joint models only the coefficient estimates for educational attainment are presented

Identified correlation in the mixing distribution indicates 3 types of individuals with perfect correlation between starting behavior and educational attainment. The first group consists of individuals with a positive starting, positive quit rate and low educational attainment. The second group consists of those with a positive starting rate, a zero quit rate and low educational attainment. The third group has a zero starting rate and high educational attainment. The associated probabilities are 45, 25, 30% for males, and 43, 20, 37% for females

Absolute *t*-statistics in parentheses. *,** Statistical significance at 10 and 5%, respectively

Table 7 Parameter estimates of sensitivity analysis on education

	Males (1)		Females (2)	
a. Early parental divorce				
$Smoking_{15-}$	−0.22	(2.5)**	−0.01	(0.1)
Parental divorce	−0.13	(0.7)	−0.14	(1.0)
−LogLikelihood	3089.6		3271.5	
Observation	758		825	
b. Early parental loss				
$Smoking_{15-}$	−0.22	(2.5)**	−0.01	(0.1)
Parent loss	−0.06	(0.3)	−0.09	(1.5)
−LogLikelihood	3089.8		3270.7	
Observation	758		825	
c. Risk attitudes				
$Smoking_{15-}$	−0.19	(2.8)**	0.05	(1.0)
Risk aversion	−0.09	(0.9)	−0.09	(0.9)
Prudence	0.18	(1.8)*	0.21	(2.0)**
Temperance	0.11	(1.1)	0.11	(1.1)
−LogLikelihood	2426.7		2267.1	
Observation	585		579	
d. Educational attainment of parents				
$Smoking_{15-}$	−0.21	(2.4)**	−0.03	(0.3)
Mother's education	0.01	(0.7)	0.04	(2.6)**
Father's education	0.03	(1.7)*	−0.01	(0.6)
−LogLikelihood	3010.2		3198.8	
Observation	737		810	
e. Use of alcohol				
$Smoking_{15-}$	−0.22	(2.9)**	−0.01	(0.9)
Early use of alcohol	0.12	(0.3)	0.34	(0.8)
−LogLikelihood	4187.9		4356.1	
Observation	1017		1103	

Absolute *t*-statistics in parentheses. *· ** Statistical significance at 10 and 5%, respectively

Table 8 Parameter estimates of the ordered probit models on the first job rankings

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
Smoking15−	−0.16 (2.0)**	−0.07 (0.9)	−0.06 (1.2)	−0.15 (1.8)*	−0.10 (1.4)	−0.07 (1.1)
a. Control variables						
Religious	0.01 (0.8)	0.00 (0.0)	0.02 (1.0)	0.05 (3.0)**	0.03 (1.3)	0.05 (3.5)*
Migrant	0.02 (0.2)	0.05 (0.5)	0.08 (0.9)	0.17 (1.7)*	0.09 (0.8)	0.12 (1.8)
Very urban	0.61 (5.0)**	0.37 (2.8)**	0.55 (3.5)**	0.27 (2.1)**	0.15 (1.7)*	0.17 (1.7)*
Urban	0.32 (3.4)**	0.26 (2.7)**	0.34 (2.9)**	0.08 (0.9)	0.10 (1.1)	0.12 (1.3)
Rural	0.18 (1.7)*	0.16 (1.5)	0.20 (1.4)	0.04 (0.4)	0.09 (0.8)	0.10 (0.9)
Cohort (30–36)	0.49 (2.2)**	0.52 (2.5)**	0.66 (2.1)**	0.18 (0.9)	0.27 (1.4)	0.25 (1.3)
Cohort (37–43)	0.17 (0.8)	0.40 (2.0)**	0.41 (1.4)	−0.06 (0.3)	0.15 (0.8)	0.14 (0.7)
Cohort (44–49)	0.47 (2.1)**	0.71 (3.5)**	0.67 (2.2)**	0.13 (1.6)*	0.49 (2.6)**	0.48 (2.5)**
Cohort (50+)	0.53 (2.3)**	0.83 (3.9)**	0.81 (2.6)**	−0.06 (1.3)	0.44 (2.1)**	0.43 (2.0)**
b. Educational attainment						
Special or other		0.71 (2.7)**	−0.81 (2.1)**		1.67 (5.8)**	1.84 (5.6)**
Prep. vocational		0.19 (0.8)	0.18 (0.7)		0.35 (1.4)	0.34 (1.4)
Prep. technical		0.54 (2.3)**	0.53 (2.1)**		0.91 (3.7)**	0.91 (3.7)**
General continued		1.02 (3.7)**	1.06 (3.8)**		1.05 (3.9)**	1.06 (4.0)**
Prep. scholarly		0.92 (3.6)**	0.86 (3.1)**		0.88 (2.4)**	0.89 (2.4)**
Middle applied		1.25 (6.1)**	0.94 (3.9)**		1.61 (7.2)**	1.63 (7.2)**
Higher vocational		1.85 (6.5)**	0.92 (2.4)**		1.94 (6.4)**	1.95 (6.3)**
Scientific		2.19 (10.0)**	1.26 (4.3)**		2.44 (10.0)**	2.48 (10.1)**

Table 8 continued

	Males						Females					
	(1)		(2)		(3)		(4)		(5)		(6)	
c. Ordered probit thresholds												
γ_1	0.71	(16.6)**	0.72	(16.3)**	0.74	(15.9)**	1.00	(17.6)**	1.10	(17.8)**	1.12	(16.6)**
γ_2	0.79	(27.4)**	0.82	(28.4)**	0.85	(26.5)**	0.51	(14.7)**	0.55	(14.6)**	0.55	(14.7)**
γ_3	0.67	(27.4)**	0.72	(27.1)**	0.76	(23.7)**	0.25	(8.1)**	0.28	(8.0)**	0.28	(7.9)**
γ_4	0.74	(31.3)**	0.80	(31.2)**	0.88	(21.3)**	1.11	(49.1)**	1.18	(47.7)**	1.19	(47.3)**
γ_5	0.61	(23.3)**	0.66	(23.1)**	0.76	(15.1)**	0.43	(17.8)**	0.46	(17.8)**	0.46	(17.6)**
γ_6	0.75	(24.1)**	0.81	(24.5)**	0.94	(15.2)**	1.15	(39.3)**	1.22	(38.4)**	1.23	(37.5)**
γ_7	0.36	(9.1)**	0.39	(8.9)**	0.44	(8.0)**	0.43	(8.1)**	0.46	(7.9)**	0.46	(7.8)**
d. Unobserved heterogeneity												
$\epsilon_{j,a}$	1.13	(4.9)**	0.01	(0.0)	1.77	(3.4)**	2.17	(9.2)**	−2.65	(1.7)*	3.12	(2.8)**
$\epsilon_{j,b}$	−0.23	(2.8)**	−0.05	(0.6)	−1.75	(6.0)**	−0.19	(1.9)*	0.95	(3.0)**	−2.18	(1.9)*
α_1	0.35	(4.4)**	0.35	(4.4)**	−0.48	(1.9)**	0.17	(2.2)**	0.17	(2.3)**	−3.94	(5.6)**
α_2	−0.22	(2.2)**	−0.22	(2.2)**	−1.52	(4.5)**	−0.60	(5.4)**	−0.63	(5.4)**	0.20	(2.5)**
α_3					−0.80	(3.0)**					−0.71	(4.9)**
α_4					0.32	(3.1)**						
α_5					0.01	(0.4)						

Table 8 continued

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
−LogLikelihood	4573.4	4442.0	6091.1	4289.5	4172.7	5876.7
LR-Test-1	26.4**	26.2**	24.1**	30.1**	30.0**	22.2**
LR-Test-2			14.4**			0.6
Observations	1017	1017	1017	1103	1103	1103
Columns (1) and (4) when the endogeneity of the smoking decision is ignored						
Smoking15−	−0.19 (2.3)**			−0.20 (2.5)**		

LR-Test-1 presents LR statistics for the significance of correlation between unobserved heterogeneity affecting smoking dynamics and the first job (for the last column, also education). Null hypothesis assumes no unobserved heterogeneity in the first job

LR-Test-2 presents LR statistics for the significance of correlation between unobserved heterogeneity affecting education and the first job

For the joint models only the coefficient estimates for job rankings are presented

Identified correlation in the mixing distribution indicates 6 types of individuals for males. (1) positive starting rate, positive quit rate, low educational attainment, low job rankings. (2) positive starting rate, zero quit rate, low educational attainment, low job rankings. (3) zero starting rate, high educational attainment, low job rankings. (4) positive starting rate, positive quit rate, low educational attainment, high job rankings. (5) positive starting rate, zero quit rate, low educational attainment, high job rankings. (6) zero starting rate, high educational attainment, high job rankings. The associated probabilities are 13, 5, 10, 29, 22, 21%. For females 4 types are identified (1), (2), (6) as in males and the last being the merged group of (3), (4) and (5). The associated probabilities are 45, 18, 36, 2%

Absolute t-statistics in parentheses. *, ** Statistical significance at 10 and 5%, respectively

Table 9 Parameter estimates of sensitivity analysis on the first job rankings

	Males (1)		Females (2)	
a. Controlling for the risk attitude				
$Smoking_{15-}$	−0.05	(0.4)	−0.06	(0.7)
Risk aversion	0.07	(0.7)	0.02	(0.5)
Prudence	0.19	(1.8)*	0.06	(0.5)
Temperance	−0.05	(0.5)	−0.10	(1.3)
−LogLikelihood	3466.2		3004.6	
Observation	582		573	
b. Time spent to search for the first job				
$Smoking_{15-}$	−0.04	(0.5)	−0.09	(1.0)
Short search	−0.01	(0.1)	0.10	(1.6)
Moderate search	0.07	(0.5)	0.19	(1.1)
Intensive search	0.47	(2.4)**	0.33	(1.3)
−LogLikelihood	3397.5		2867.4	
Observation	570		546	
c. Year at the time of first job				
$Smoking_{15-}$	−0.05	(0.6)	−0.08	(1.1)
Year at the time of first job	0.01	(2.0)**	0.00	(0.7)
−LogLikelihood	5922.3		5582.3	
Observation	990		1054	

Absolute *t*-statistics in parentheses

*, ** Statistical significance at 10 and 5%, respectively

Table 10 Parameter estimates of the ordered probit models on the current job rankings

	Males (1)	(2)	(3)	Females (4)	(5)	(6)
Smoking15–	−0.31 (4.2)**	−0.23 (2.7)**	−0.18 (2.9)**	−0.10 (1.4)	−0.03 (0.4)	−0.04 (0.5)
a. Control variables						
Religious	0.02 (0.9)	0.00 (0.0)	0.12 (0.2)	0.04 (2.2)**	0.00 (0.1)	−0.01 (0.6)
Migrant	−0.13 (1.4)	−0.13 (1.4)	−0.14 (0.1)	0.08 (0.8)	−0.02 (0.2)	−0.09 (0.6)
Very urban	0.71 (5.9)**	0.47 (3.6)**	0.51 (3.7)**	0.51 (4.3)**	0.38 (3.0)**	0.39 (3.1)**
Urban	0.33 (3.5)**	0.29 (3.1)**	0.32 (3.3)**	0.29 (3.4)**	0.34 (3.9)**	0.35 (3.9)**
Rural	0.24 (2.2)**	0.26 (2.4)**	0.28 (2.6)**	0.20 (2.0)**	0.28 (2.8)**	0.29 (2.8)**
Cohort (30–36)	−0.78 (3.0)**	−0.12 (3.9)	0.12 (3.9)	−0.10 (0.4)	−0.06 (0.3)	−0.07 (0.3)
Cohort (37–43)	−1.01 (3.9)**	0.05 (3.7)	0.05 (3.7)	−0.41 (1.7)*	−0.12 (1.1)	−0.13 (1.0)
Cohort (44–49)	−0.70 (2.7)**	0.28 (2.2)*	0.28 (2.2)*	−0.34 (1.4)	0.02 (0.1)	0.00 (0.0)
Cohort (50+)	−0.82 (3.2)**	0.18 (2.4)*	0.18 (2.4)*	−0.58 (2.3)**	0.07 (0.3)	0.11 (0.4)
b. Educational attainment						
Special or no		1.02 (4.1)**	0.06 (0.2)		1.91 (5.4)**	0.81 (5.2)**
Prep. vocational		0.17 (0.8)	0.19 (0.8)		0.47 (1.4)	0.45 (1.4)
Prep. technical		0.47 (2.1)**	0.50 (2.1)**		1.00 (3.0)**	0.99 (3.0)**
General continued		1.33 (5.1)**	1.36 (5.1)**		1.12 (3.2)**	1.12 (3.1)**
Prep.scholarly		1.36 (5.7)**	1.40 (5.7)**		1.02 (2.6)**	1.00 (2.5)**
Middle applied		1.59 (7.8)**	1.61 (7.8)**		1.77 (5.5)**	1.76 (5.4)**
Higher vocational		2.15 (3.5)**	2.19 (3.6)**		2.03 (5.6)**	2.02 (5.5)**
Scientific		2.78 (12.7)**	2.81 (12.7)**		3.00 (8.8)**	3.00 (8.6)**
c. Ordered probit thresholds						
γ_1	0.48 (8.0)**	0.49 (8.5)**	0.50 (8.0)**	1.05 (18.5)**	1.05 (18.5)**	1.05 (17.5)**
γ_2	0.88 (21.7)**	0.88 (21.7)**	0.89 (20.7)**	0.53 (13.7)**	0.53 (13.7)**	0.53 (13.4)**

Table 10 continued

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
γ_3	0.79 (25.1)**	0.79 (25.1)**	0.79 (24.6)**	0.30 (8.3)**	0.30 (8.3)**	0.30 (8.1)**
γ_4	0.76 (37.4)**	0.76 (27.4)**	0.76 (26.5)**	1.09 (43.0)**	1.09 (43.0)**	1.09 (41.9)**
γ_5	0.77 (28.9)**	0.77 (28.9)**	0.77 (28.4)**	0.54 (21.7)**	0.54 (21.7)**	0.54 (21.1)**
γ_6	0.81 (28.9)**	0.81 (28.9)**	0.81 (27.3)**	1.16 (43.7)**	1.16 (43.7)**	1.16 (41.6)**
γ_7	0.81 (23.3)**	0.81 (23.3)**	0.81 (22.9)**	0.68 (15.4)**	0.68 (15.4)**	0.68 (14.9)**
d. Unobserved heterogeneity						
$\epsilon_{j,a}$	2.56 (9.3)**	1.56 (4.7)**	3.71 (4.7)**	2.41 (9.1)**	1.01 (2.5)**	1.04 (2.5)**
$\epsilon_{j,b}$			−2.19 (3.1)**			−0.18 (0.7)
α_1			−4.11 (7.9)**			−0.20 (1.4)**
α_2			−4.18 (7.9)**			−0.10 (0.7)
α_3			0.30 (3.9)**			
α_4			−0.18 (2.1)**			
−LogLikelihood	2047.4	1832.5	6005.3	1851.5	1684.2	5988.0
LR-Test-1			60.4**			50.8**
LR-Test-2			5.2**			1.8
Observations	1017	1017	1017	1103	1103	1103

LR-Test-1 presents LR statistics for the significance of correlation between unobserved heterogeneity affecting smoking dynamics and the current job (for the last column, also education). Null hypothesis assumes no unobserved heterogeneity in the current job

LR-Test-2 presents LR statistics for the significance of correlation between unobserved heterogeneity affecting education and the current job

For the joint models only the coefficient estimates for job rankings are presented

Identified correlation in the mixing distribution indicates 5 types of individuals for males. (1) positive starting rate, positive quit rate, low educational attainment, low job rankings. (2) positive starting rate, zero quit rate, low educational attainment, high job rankings. (3) zero starting rate, high educational attainment, low job rankings. (4) positive starting rate, positive quit rate, low educational attainment, high job rankings. (5) zero starting rate, high educational attainment, high job rankings. The associated probabilities are 42, 26, 1, 1, 31%. For females, only perfect correlation is identified between educational attainment and job rankings for females

Absolute *t*-statistics in parentheses. *, ** Statistical significance at 10 and 5%, respectively

Table 11 Parameter estimates of the ordered probit models on the probability of having a higher ranked job

	Males (1)		Females (2)	
Smoking$_{15-}$	−0.27	(2.7)**	0.01	(0.0)
a. Control variables				
Religious	0.01	(0.4)	−0.02	(0.9)
Migrant	−0.13	(1.1)	−0.12	(0.8)
Very urban	0.34	(2.1)**	0.34	(2.0)**
Urban	0.18	(1.5)	0.25	(2.0)**
Rural	0.31	(2.4)**	0.16	(1.1)
Cohort (30–36)	0.07	(3.3)	−0.19	(0.6)
Cohort (37–43)	0.41	(2.3)**	0.02	(0.1)
Cohort (44–49)	0.64	(3.5)**	0.03	(0.1)
Cohort (50+)	0.56	(3.1)**	0.06	(0.2)
b. Educational attainment				
Special or other	1.13	(3.4)**	1.82	(4.2)**
Prep. vocational	0.35	(1.3)	0.56	(1.5)
prep. technical	0.45	(1.7)*	0.89	(2.5)**
General continued	1.42	(4.7)**	1.02	(2.6)**
Prep. scholarly	1.35	(4.6)**	1.08	(2.5)**
Middle applied	1.46	(5.7)**	1.43	(4.1)**
Higher vocational	1.34	(3.1)**	1.42	(2.8)**
Scientific	1.83	(6.5)**	2.32	(6.1)**
c. The first job types				
Unskilled manual	0.78	(3.7)**	0.38	(1.1)
Semi-skilled manual	−0.01	(0.1)	−0.73	(2.0)**
Skilled manual	−0.64	(3.7)**	−0.71	(1.8)*
Other clerical	−0.67	(3.9)**	−1.77	(5.3)**
Int. supervisory	−1.73	(10.1)**	−2.74	(7.9)**
d. Ordered probit thresholds				
γ_1	0.67	(24.3)**	0.61	(19.1)**
γ_2	0.79	(23.4)**	0.82	(18.7)**
e. Unobserved heterogeneity				
ϵ_a	0.15	(1.5)	0.27	(2.6)**
ϵ_b	−0.13	(0.3)	−0.36	(0.6)

Table 11 continued

	Males (1)		Females (2)	
α_1	0.35	(4.4)**	0.18	(0.1)
α_2	−0.22	(2.2)**	−0.61	(0.1)
−LogLikelihood		3513.6		3377.9
Observations		1017		1103

For the joint models only the coefficient estimates for the probability of having a higher ranked job are presented

Identified correlation in the mixing distribution indicates 3 types of individuals with perfect correlation between starting behavior, educational attainment and probability o having a higher ranked job. The first group consists of individuals with a positive starting, positive quit rate, low educational attainment, low probability to have a higher ranked job. The second group consists of those with a positive starting rate, a zero quit rate, low educational attainment, low probability to have a higher ranked job. The third group has a zero starting rate, high educational attainment, high probability to have a higher ranked job. The associated probabilities are 38, 29, 33% for males, and 44, 20, 36% for females

Absolute t-statistics in parentheses. $*,**$ Statistical significance at 10 and 5%, respectively

Table 12 Parameter estimates of the joint model on log-hourly wages

	Males (1)		Females (2)	
$Smoking_{15-}$	−0.08	(1.9)*	−0.04	(1.1)
a. Control variables				
Religious	0.00	(0.5)	0.00	(0.3)
Migrant	0.01	(0.1)	−0.10	(2.2)**
Very urban	−0.02	(0.9)	0.14	(2.0)**
Urban	0.04	(0.1)	0.06	(1.9)
Rural	0.03	(1.0)	0.01	(0.1)
Cohort (30–36)	0.28	(3.8)**	0.38	(3.9)**
Cohort (37–43)	0.41	(4.2)**	0.46	(3.5)**
Cohort (44–49)	0.44	(5.2)**	0.52	(4.2)**
Cohort (50+)	0.48	(8.2)**	0.33	(4.8)**
σ	0.31	(53.6)**	0.49	(62.9)**
b. Educational attainment				
Special or other	0.20	(1.5)	0.01	(0.3)
Prep. vocational	0.31	(2.3)**	0.02	(0.2)
Prep. technical	0.43	(3.1)**	0.01	(0.1)
General continued	0.53	(4.2)**	0.10	(0.4)
Prep. scholarly	0.48	(3.8)**	0.14	(1.1)
Middle applied	0.57	(5.1)**	0.12	(0.8)
Higher vocational	0.70	(2.1)**	0.24	(1.1)
Scientific	0.30	(2.0)**	0.13	(0.5)

Table 12 continued

	Males (1)		Females (2)	
c. Unobserved heterogeneity				
ω_1	2.26	(14.1)**	1.54	(3.1)**
ω_2	−0.29	(1.9)*	0.11	(0.4)
α_1	0.34	(3.2)**	−0.12	(0.1)
α_2	−0.18	(1.1)	−0.03	(0.1)
−LogLikelihood	5464.6		4135.9	
Observations	897		661	
Introducing the interaction between the job types and the smoking indicator				
$Smoking_{15-}$	−0.40	(12.9)**	−0.04	(1.1)
× Unskilled manual	0.08	(1.5)	0.01	(0.3)
× Semi-skilled manual	0.22	(1.9)*	0.02	(0.2)
× Skilled manual	0.31	(1.9)*	0.01	(0.1)
× Other clerical	0.35	(2.7)**	0.10	(0.4)
× Int. supervisory	0.37	(2.4)**	0.14	(1.1)
× Int. academic/independent	0.41	(2.8)**	0.12	(0.8)
× Higher supervisory	0.62	(3.9)**	0.24	(1.1)
× higher academic/independent	0.22	(1.2)	0.13	(0.5)
−LogLikelihood	5474.3		4132.0	
Observations	897		661	

For the joint models only the coefficient estimates for wages are presented

Identified correlation in the mixing distribution indicates 3 types of individuals with perfect correlation between starting behavior, educational attainment and hourly wages. The first group consists of individuals with a positive starting, positive quit rate, low educational attainment and low wages. The second group consists of those with a positive starting rate, a zero quit rate, low educational attainment and low wages. The third group has a zero starting rate, high educational attainment and high wages. The associated probabilities are 43, 26, 31 for males, and 31, 34, 35% for females

Absolute t-statistics in parentheses. *,** Statistical significance at 10 and 5%, respectively

Appendix 6: Figures

See Figs.

a

b
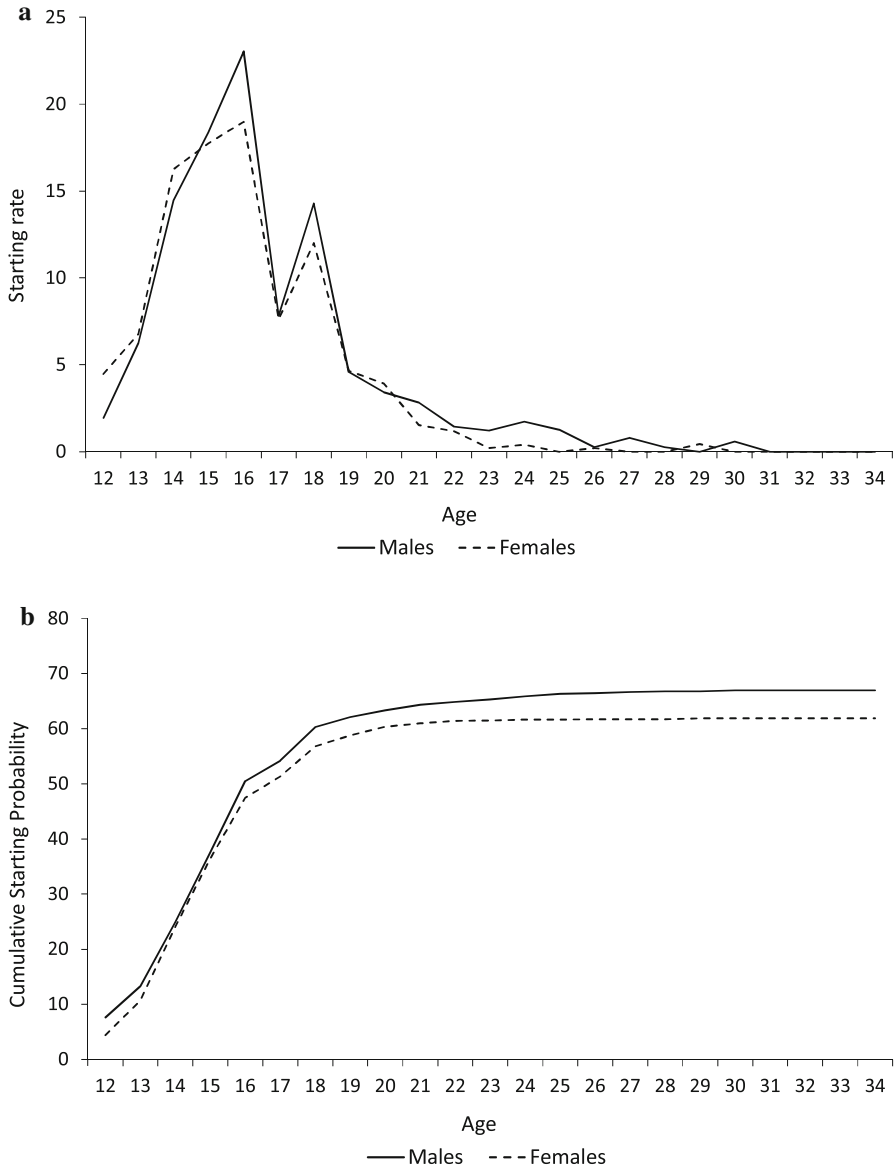

Fig. 1 Starting rates and cumulative starting probabilities of smoking (in %). **a** Starting rates, **b** cumulative starting probabilities
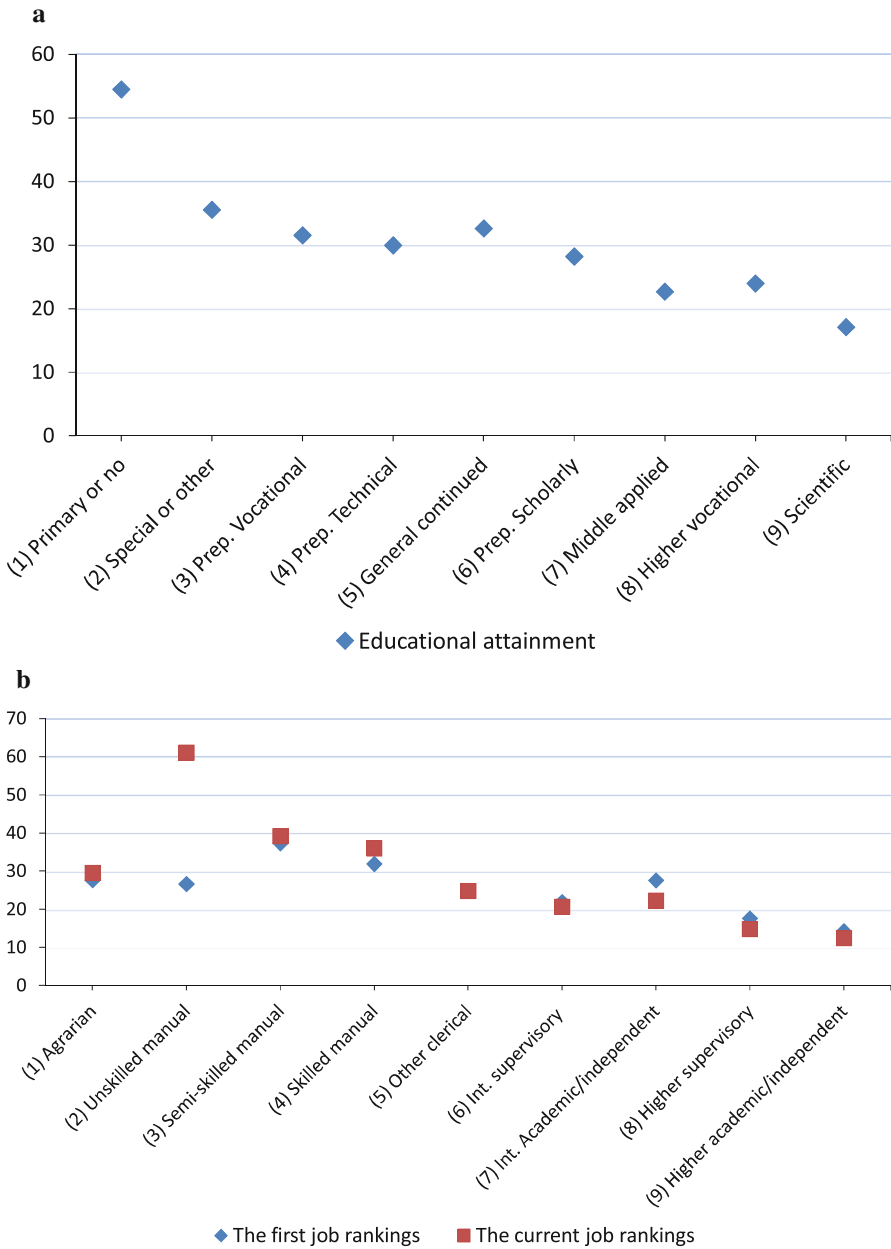
a

b

Fig. 2 Males: proportion of early smokers in each category of the ordered variables (in %). **a** Educational attainment, **b** labor market performance
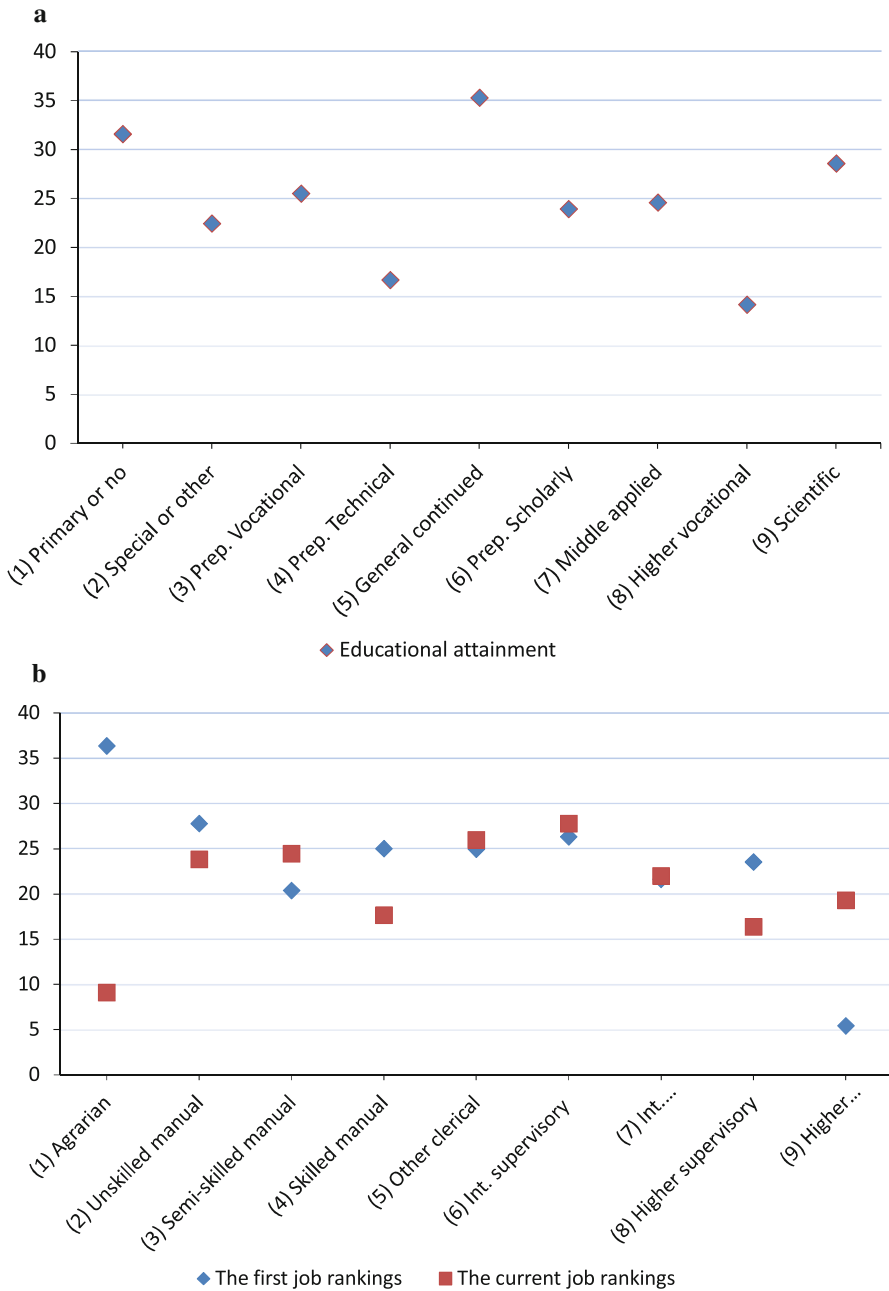
a

b

Fig. 3 Females: proportion of early smokers in each category of the ordered variables (in %). **a** Educational attainment, **b** labor market performance
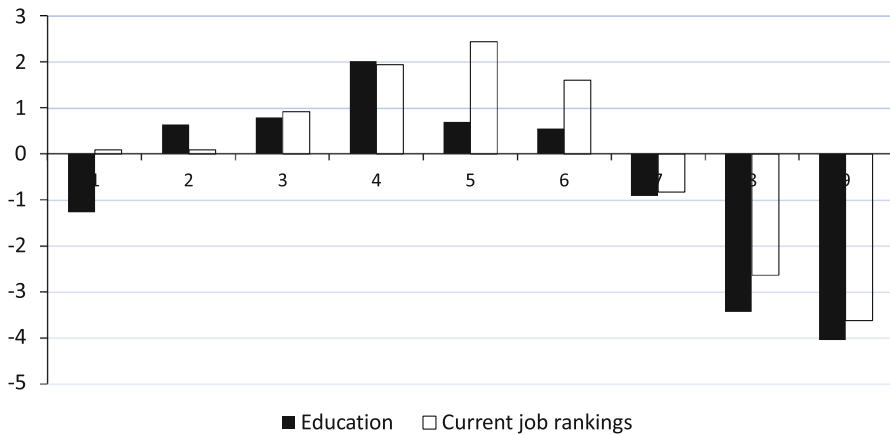
Fig. 4 Males: differences in the estimated probabilities of belonging to one of the ordered categories for those who start smoking before the age of 15 and those who do not (in %). The figure above is obtained as follows: First, I simulated the probabilities of belonging to the each category in the ordered choices for someone who smoked before the age of 15. Then, I simulated the same probabilities for someone who did not smoke before the age of 15. The levels above reflect the difference between these probabilities for the each category. For all of the other control variables, I used the sample means in simulations

References

Andreski, P., & Breslau, N. (1993). Smoking and nicotine dependence in young adults: Differences between blacks and whites. *Drug and Alcohol Dependence*, *32*(2), 119–125.

Auld, M. C. (1998). Wages, alcohol use and smoking: Simultaneous estimates. Discussion paper, no: 98/08. Department of Economics, University of Calgary.

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of American Statistical Association*, *90*(430), 443–450.

Conley, T. G., Hansen, C. B., & Rossi, P. E. (2012). Plausibly exogenous. *The Review of Economics and Statistics*, *94*(1), 260–272.

French, M. T., & Popovici, I. (2011). That instrument is lousy! In search of agreement when using instrumental variables estimation in substance use research. *Health Economics*, *20*, 127–146.

French, M. T., & Zarkin, G. (1995). Is moderate alcohol use related to wages? Evidence from four worksites. *Journal of Health Economics*, *14*, 319–344.

Goodman, E., & Capitman, J. (2000). Depressive symptoms and cigarette smoking among teens. *Pediatrics*, *106*, 748–755.

Halpern, M. T., Shikiar, R., Rentz, A. M., & Khan, Z. M. (2001). Impact of smoking status on workplace absenteeism and productivity. *Tobacco Control*, *10*, 233–238.

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

Heineck, G., & Schwarze, J. (2003). Substance use and earnings: The case of smokers in Germany. Discussion paper, no: 743, IZA.

Jacobsen, L. K., Krystal, J. H., Mencl, W. E., Westerveld, M., Frost, S. J., & Pugh, K. R. (2005). Effects of smoking and smoking abstinence on cognition in adolescent tobacco smokers. *Biological Psychiatry*, *57*(1), 56–66.

Kang, M.-I., & Ikeda, S. (2014). Time discounting, present biases, and health-related behaviors: Evidence from japan. *Health Economics*, *23*, 1443–1464.

Kristein, M. M. (1983). How much can business expect to profit from smoking cessation? *Preventive Medicine*, *12*(2), 358–381.

Lee, Y. L. (1999). Wage effects of drinking and smoking: An analysis using Australian twins data. Discussion paper 99/22. Department of Economics, The University of Western Australia.

Levine, P. B., Gustafson, T. A., & Velenchik, A. D. (1997). More bad news for smokers? The effects of cigarette smoking on wages. *Industrial and Labor Relations Review*, *50*(3), 493–509.

Mroz, T. A. (1999). Discrete factor approximations in simultaneous equation model: Estimating the impact of a dummy endogenous variable on a continious outcome. *Journal of Econometrics*, *92*, 233–274.

Noussair, C., Trautmann, S., & van de Kuilen, G. (2014). Higher order risk attitudes, demographics, and financial decisions. *Review of Economic Studies*, *81*, 325–355.

Pesko, M., & Baum, C. (2016). The self-medication hypothesis: Evidence from terrorism and cigarette accessibility. *Economics and Human Biology*, *22*, 94–102.

Rosenthal, R., & Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review*, *3*, 16–20.

Steuber, T., & Danner, F. (2006). Adolescent smoking and depression: Which comes first? *Addictive Behaviors*, *31*, 133–136.

Trauth, J. A., McCook, E. C., Seidler, F. J., & Slotkin, T. A. (2000). Modeling adolescent nicotine exposure: Effects on cholinergic systems in rat brain regions. *Brain Research*, *873*(1), 18–25.

Van Ours, J. C. (2003). Is cannabis a stepping stone for cocaine? *Journal of Health Economics*, *22*, 539–554.

Van Ours, J. C. (2004). A pint a day raises a man's pay; but smoking blows that gain away. *Journal of Health Economics*, *23*, 863–886.

Van Ours, J. C. (2006). Cannabis, cocaine and jobs. *Journal of Applied Econometrics*, *21*, 897–917.

Van Ours, J. C. (2007). The effect of cannabis use on wages of prime age males. *Oxford Bulletin of Economics and Statistics*, *69*, 619–634.

Van Ours, J. C., & Williams, J. (2009). Why parents worry: Initiation into cannabis use by youth and their educational attainment. *Journal of Health Economics*, *28*, 132–142.

Weng, S. F., Ali, S., & Leonardi-Bee, J. (2013). Smoking and absence from work: Systematic review and meta-analysis of occupational studies. *Addiction*, *108*(2), 307–319.

Zarkin, G. A., Mroz, T. A., Bray, J. W., & French, M. T. (1998). The relationship between drug use and labour supply for young men. *Labour Economics*, *5*(4), 385–409.

Zhao, M., Konishi, Y., & Glewwe, P. (2012). Does smoking affect schooling? Evidence from teenagers in rural china. *Journal of Health Economics*, *31*, 584–598.