MAX-PLANCK-INSTITUT FÜR SOZIALRECHT UND SOZIALPOLITIK
MAX PLANCK INSTITUTE FOR SOCIAL LAW AND SOCIAL POLICY

mea Munich Center for the Economics of Aging

Educational differences in smoking: selection versus causation

Hendrik Jürges, Sophie-Charlotte Meyer

02-2017

MEA DISCUSSION PAPERS

Educational differences in smoking: selection versus causation

Hendrik Jürges, Sophie-Charlotte Meyer

Zusammenfassung:

Wir untersuchen die Ursachen von Bildungsunterschieden im Rauchverhalten mit Daten des deutschen Mikrozensus. Auf Basis der retrospektiven Informationen über das Alter bei Rauchbeginn vergleichen wir altersspezifische Hazardraten dafür, mit dem Rauchen anzufangen, von (künftigen) niedrig und hoch gebildeten Personen. Wir finden, dass bis zu 90% der Bildungsunterschiede im Rauchen bis zum Alter von 16 Jahren, d. h. vor Ende der Schulpflicht entstehen. Diese Unterschiede bleiben bis ins Erwachsenenalter konstant. Darüber hinaus untersuchen wir die Rolle gesundheitsbezogenen Wissens (gemessen durch Tätigkeit in Gesundheitsberufen). Dieses trägt praktisch nicht zur Erklärung des Rauchverhaltens bei. Unsere Ergebnisse deuten darauf hin, dass hauptsächlich (unbeobachtete) Faktoren, die sowohl mit dem Rauchen als auch mit Bildungsentscheidungen zusammenhängen, für Bildungsunterschiede im Rauchen verantwortlich sind. Nur kleine Anteile der Bildungslücke scheinen auf allgemeine oder gesundheitsbezogene Bildung zurückzuführen zu sein. Die Wirksamkeit bildungspolitischer Maßnahmen zur Bekämpfung des Rauchens ist daher wahrscheinlich begrenzt.

Abstract:

We investigate sources of educational differences in smoking. Using a large German data set containing retrospective information on the age at smoking onset, we compare age-specific hazard rates of starting smoking between (future) low and high educated individuals. We find that up to 90% of the educational differences in smoking develop before the age of 16, i.e. before compulsory schooling is completed. This education gap persists into adulthood. Further, we examine the role of health-related knowledge (proxied by working in health-related occupations) and find it hardly explains smoking decisions. Our findings suggest that (unobserved) factors determining both the selection into smoking and education are almost exclusively responsible for educational differences in smoking. Only small parts of the education gap seem to be caused by general or health-specific education. The effectiveness of education policy to combat smoking is thus likely limited.

Keywords:
education; smoking initiation; health-related knowledge

JEL Classification:
I12, J22, J13

Educational differences in smoking: selection versus causation

Hendrik Jürges, Sophie-Charlotte Meyer*

Schumpeter School of Business and Economics, University of Wuppertal

21 December 2016

Abstract

We investigate sources of educational differences in smoking. Using a large German data set containing retrospective information on the age at smoking onset, we compare age-specific hazard rates of starting smoking between (future) low and high educated individuals. We find that up to 90% of the educational differences in smoking develop before the age of 16, i.e. before compulsory schooling is completed. This education gap persists into adulthood. Further, we examine the role of health-related knowledge (proxied by working in health-related occupations) and find it hardly explains smoking decisions. Our findings suggest that (unobserved) factors determining both the selection into smoking and education are almost exclusively responsible for educational differences in smoking. Only small parts of the education gap seem to be caused by general or health-specific education. The effectiveness of education policy to combat smoking is thus likely limited.

Keywords: education; smoking initiation; health-related knowledge
JEL classification: I12, J22, J13

*Address for correspondence: Rainer-Gruenter-Str.21 (FN), 42119 Wuppertal, Germany, E-Mail: *juerges@uni-wuppertal.de*.

1 Introduction

Educational differences in smoking, with lower educated individuals being more likely to smoke than higher educated individuals, have been widely documented (e.g. Kenkel et al., 2006; de Walque, 2007; Cutler and Lleras-Muney, 2010; Jürges et al., 2011; Maralani, 2013). The economic literature discusses at least three partly related reasons why high educated individuals smoke less (de Walque, 2010). First, education is an investment raising future income, which increases the marginal return to health capital and leads to a higher optimal health stock (Grossman, 1972; Becker and Mulligan, 1997; Grossman, 2006; Becker, 2007). Second, education can change the inputs into health production itself (Rosenzweig and Schultz, 1981; Kenkel, 1991). For instance, higher educated individuals may be more likely aware of the harmful effects of smoking or better able to process health information, such as following medical advice. This highlights the role of health-related knowledge. Whereas these two explanations claim that at least part of the education-smoking link is causal, a third explanation stresses selection due to individual differences in relevant characteristics, such as time preferences or willpower, which lead to higher education and healthier behavior simultaneously (Fuchs, 1982; Farrell and Fuchs, 1982).

Researchers have recently tried to estimate the causal effect of education on smoking – with mixed findings (see Grossman, 2015, for an overview). In order to tackle the endogeneity of education, one exploits presumably exogenous variation in education, such as changes in compulsory schooling, the avoidance of the Vietnam War draft due to college enrollment, distance to college, abolition of school fees, school construction programs, and so on. Whereas some studies find a strong protecting effect (Kenkel et al., 2006; de Walque, 2007; Jürges et al., 2011), others find no evidence that education affects smoking behavior (Park and Kang, 2008; Kemptner et al., 2011; Clark and Royer, 2013; Lundborg, 2013).

In the present paper, we follow a different approach based on a simple argument originally raised by Farrell and Fuchs (1982): to the extent that educational differences in smoking appear already in adolescence, and thus before differences in education emerge, the education-smoking link is driven by selection rather than causation. Put differently, if formal education affected smoking behavior, educational differences should surface after formal education is completed, not before, as the outcome (smoking) must follow the cause (education). Farrell and Fuchs clearly show that *future* educational attainment explains smoking patters at age 17 as much as actual attainment explains smoking patterns at age 24 and conclude that "additional years of schooling is not causally related to smoking" (p. 229). This simple but effective empirical strategy to separate selection and causation has found surprisingly little resonance in the literature. Only fairly recently, de Walque (2010) has replicated and expanded the analysis by Farrell and Fuchs, in addition stressing the effect of college education on smoking cessation. Based on smoking biographies, he finds

that the correlation between education and smoking tends to increase past age 25 also after controlling for individual fixed effects which capture time-invariant confounders like time preferences. This finding suggests a causal effect of college education – acquired after starting to smoke – on quitting smoking.

In our analysis, we first analyze if educational disparities in smoking appear before or after compulsory schooling is completed. We use the German Microcensus, an administrative data set with more than one million individuals containing retrospective information on the age of smoking onset. As the previous literature, we find a strong negative association between education and smoking. However, our analyses suggest that educational differences in smoking onset develop while individuals are in school. Comparing age-specific hazard rates of starting to smoke of future high and low educated individuals from age 10 to age 25, we find that hazard rate ratios are largest at the earliest ages and become smaller as individuals get older. Some 90% of the education gap in ever smoking are determined before compulsory education in Germany is completed. These findings contradict the notion that the education-smoking link is largely due to a causal effect of education on smoking. Rather, differences in characteristics determining both the selection into smoking and education are more likely responsible for educational differences in smoking.

Our data does not contain information on the year respondents have stopped smoking, therefore it is not possible to construct complete individual smoking biographies. However, we have repeated cross-sections, which allows us to construct pseudo-panel data on smoking behavior following birth cohorts over time. Based on this pseudo-panel, we study the development of the education gradient in current smoking up to age 50. Our results are again in line with the selection hypothesis: educational differences in smoking remain virtually constant as cohorts get older. Thus a causal effect of (higher) education on stopping smoking cannot be the main explanation for these differences.

The second part of our analysis focuses on the potential role of health information or health-related education in explaining educational differences in smoking. Here we address the theoretical argument that education affects smoking because it provides health-related knowledge, specifically knowledge about the health effects of smoking. Previous studies have shown that smoking was more prevalent among higher than lower educated in cohorts born before 1930 because smoking was viewed as part of a sophisticated life style. Gradually lower educated individuals copied this behavior and the smoking rates of low and high educated aligned. (e.g. Brenner, 1993; Pampel, 2005; Schulze and Mons, 2006; Piontek et al., 2010; de Walque, 2010; Vedøy, 2014; Bricard et al., 2015; Pampel et al., 2015). Only after the adverse effects of smoking on health became more widely known, the high educated began to reject smoking, and smoking rates among higher educated fell below those of the lower educated. Analyzing smoking histories from 1940–2000 in the US, de Walque (2010) shows that smoking declined since the harmful consequences of

tobacco consumption have become publicized in the beginning of the 1960s. This decline started earlier and was steeper among higher educated individuals, eventually leading to the inversion of the educational gradient. De Walque concludes that education facilitates the access to health-related information and/or increases the ability to process this information. Studies using direct measures of health knowledge do not support this argument. For instance, Kenkel (1991) and Mocan and Altindag (2012) find that individuals with good knowledge about the health consequences of smoking generally smoke less, but this only explains a small part of the observed relationship between general education and smoking. Johnston et al. (2015) use UK health survey data to construct an index of health knowledge. While their OLS model suggest that education is significantly related to better health knowledge, IV estimates based on changes in compulsory schooling indicate that there is unlikely a causal effect of education on health knowledge.

In our analysis we use indirect information on health knowledge. Specifically, we compare smoking initiation and cessation rates between individuals working in health-related occupations, such as nurses, physicians, or pharmacists, and individuals working in other occupations with equivalent levels of formal education. If health-related knowledge affects smoking behavior, individuals working in the health sector should be less likely to start smoking and more likely to quit smoking *after* they have completed their occupational education and training.[1] Our results indicate that health-related knowledge, if at all, can have only minor effects on an individual's decision to quit smoking. In contrast, individuals working in the health sector, especially doctors and pharmacists, are less likely to smoke in adolescence, i.e. before they take up their health-related education. This finding lends further support that selection rather than causation (running through health knowledge) largely accounts for differences in smoking behavior.[2]

2 Post-compulsory Schooling and Smoking

2.1 Data and Measurement

To investigate educational differences in smoking initiation we use five waves of the German Microcensus 1989, 1999, 2003, 2005, and 2009. The Microcensus is an annual official survey of 1% of German households covering approximately 800,000 individuals per wave.[3]

[1] It is perhaps ironic that one of first landmark studies that demonstrated the hazards of cigarette smoking used data on UK physicians (Doll and Hill, 1954).

[2] An interesting study in this context is Han et al. (2011), who show that Chinese medical students have much better knowledge regarding the many dangers of smoking but are not less likely to smoke.

[3] Participation in the Microcensus is mandatory, but answering health-related questions is voluntary. Before 2005, health-related questions were asked of a random subsample of 50% of respondents in 1989 and 45% of respondents in 1999 and 2003. The data were provided by the Research Data Centers of the Federal Statistical Office and the Statistical Offices of the Länder in Düsseldorf, Germany, analyzed on-site (further information: http://www.forschungsdatenzentrum.de/en/).

We restrict the sample to respondents born between 1930 and 1989 living in West Germany with valid information on all variables.[4] The analytical sample contains more than 1,000,000 individuals. Sample statistics of the relevant variables used for the whole study population, as well as for the subpopulations of never and ever smoking individuals are presented in Table 1.

Table 1: Sample statistics; proportions and averages

Variable		All	Never smoker	Ever smoker
Demographics				
Male		0.492	0.406	0.584
Age		45.6	46.3	44.8
Cohort				
1930-1954		0.471	0.423	0.448
1955-1967		0.253	0.327	0.288
1968-1989		0.276	0.250	0.264
Education				
High education: $\geq$ (Fach-) Abitur		0.243	0.275	0.208
High education: $>$ Q.75		0.248	0.286	0.207
Occupation (| same educational level)[a]				
Physicians/pharmacists (academic)		0.060	0.064	0.055
Health-related (intermediate)		0.073	0.075	0.071
	%		51.9	48.1
	N	1,036,321	538,324	497,997

Note: [a]The sample size for occupation differs as the sample is restricted to individuals with equal educational attainment: academic: 129,311; intermediate: 261,413; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Smoking Behavior

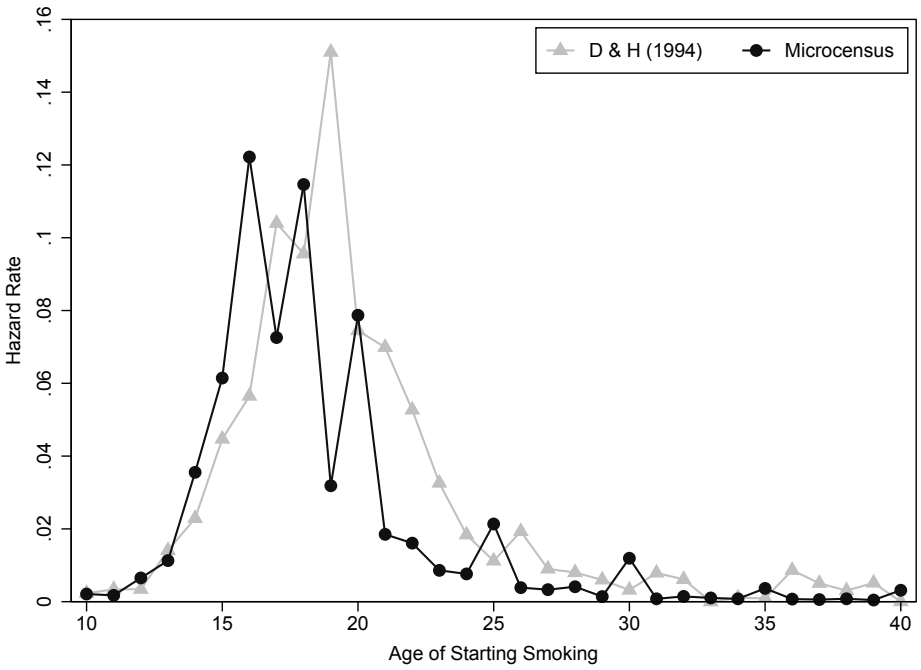
Respondents were asked whether they smoke currently and, if not, whether they ever smoked. Combining both questions indicates whether an individual ever smoked, which happens to be the case for 48% of the sample. Our key variable is age at smoking uptake. We use this information to compute hazard rates of smoking onset, shown in Figure 1 for the entire sample. The probability of starting smoking is largest between the age of 15 and 20. At each age, roughly 10% of previous non-smokers have started smoking. The peaks at age 16, 18, 20 might be explained by recall error (e.g. due to rounding).[5]

[4]We focus on West German respondents because of differences in education systems across East and West. However, analyses based only the East German population are quite similar and lead to the same conclusions (see Figure A.2 in the Appendix). The lower year of birth bound ensures a sufficient number of highly educated respondents for each year of birth, especially among women, in the older cohorts, while the upper bound ensures a sufficient number of respondents who have actually completed school by the time of the survey.

[5]Retrospective information may be prone to response bias if respondents do not remember the exact age of smoking initiation. Studies have found that in comparison with longitudinal records, recalled

Alternatively, the peaks at age 16 and 18 could also be due to the fact that in Germany, 16 was the legal smoking age (until 2007) and 18 is the legal age (since 1975). In any case, Figure 1 shows that most of the action in terms of starting smoking happens up to age 20. For comparison, Figure 1 also shows the hazard rates reported by Douglas and Hariharan (1994) for the U.S. (data were from the 1978 and 1979 NHIS and cover birth cohorts 1939 to 1953). At their maximum, hazard rates are of similar size as in our data. However, the U.S. pattern appears to be shifted somewhat to the right, i.e. people started smoking later on average. Especially in their early twenties, non-smokers were more likely to take up smoking in the U.S. than in our data.

Figure 1: Empirical hazard rates of starting smoking



Source: German Microcensus 1989, 1999, 2003, 2005, 2009; Douglas and Hariharan, 1994, Figure 1.

Formal Education

We measure formal education as the highest attained school leaving certificate. As described elsewhere in more detail (e.g. Kemptner et al., 2011), students in German secondary school visit one of three different tracks (basic, intermediate, academic). Track choice, made at age 10, is largely based on performance in primary school although parental background is also an important predictor (e.g. Jürges and Schneider, 2011; Lehmann and Peek, 1997). Each of the three tracks leads to a different leaving certi-

information on smoking status was fairly accurate (Krall et al., 1989). There is limited evidence that individuals tend to overestimate the age at onset (Bright and Soulakova, 2014), but there is no evidence on educational differences in this recall bias. The implications for our research are thus unclear.

ficate. Compulsory schooling ends upon completion of either of the two lower tracks. Students then continue receiving vocational training. Only those completing academic track (German: *Fachabitur* or *Abitur*), which takes 2 to 3 years longer than the other two tracks, are allowed to enter university. In the following, we consider as high educated everyone who has acquired a university entrance qualification and thus received 2 to 3 years of post-compulsory schooling.[6] In our data, about 24% of the individuals are coded as high educated.

One important concern when analyzing educational differences in smoking over long periods of time is that selection into higher education has changed in recent decades (de Walque, 2010). In many countries including Germany tertiary education has been made accessible to increasing proportions of the population, thereby changing the inherent ability distribution within each education segment. For instance, whereas about 10% of individuals born in 1940 finished academic track, more than 30% of those born 1980 did (see Figure 1 in Jürges et al., 2011). This inevitably changed the nature of a university entrance examination. To see if our results are robust to these changes we also used a relative education measure. This measure labels as high educated everyone in the top quartile of the education distribution, where education is measured by years of full-time education. Our results are astonishingly robust to using this measure and we relegate all results to Appendix C. We further note that the fact that absolute and relative measures of education yield very similar results in terms of educational differences in education can be interpreted as a sign that the content of education does not matter, but leave a deeper analysis of this issue to future research.

"Information" Cohorts

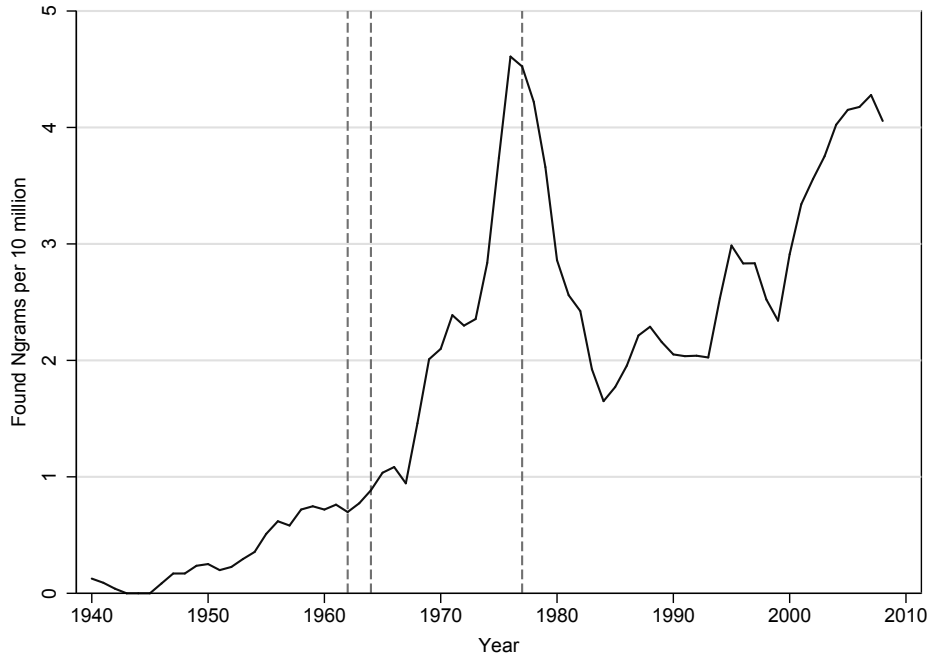
Similar to Farrell and Fuchs (1982), we define groups of year of birth cohorts according to the relevant historical events that made the harmful consequences of smoking known to the public (such as the publication of the Surgeon General report in the U.S.). Since there is no general consensus about these events in Germany, we tried to identify pivotal years by analyzing the development of the proportion of Ngrams in German publications published between 1950 and 2010 that relate to the hazards of smoking (see Figure 2).[7]

Our selected Ngrams appear first during the 1950s. The relative number has increased sharply in late 1960s, peaked in 1976, then declined again until the 2000s. This development suggests that in Germany the public debate on the harmful effects of tobacco consumption followed the reports in the UK (1962) and the US (1964) with a lag of a few years. However, in 1964 the popular German news magazine *DER SPIEGEL* published a

[6] We use the terms post-compulsory education vs. compulsory education and higher vs. lower education interchangeably.

[7] Google Ngram search terms: (Rauchen ist ungesund) + (Rauchen und Gesundheit) + Raucherlunge + Raucherbein + (Rauchen schädlich) + (Rauchen ist schädlich) + (Rauchen und Lungenkrebs) + (Rauchen Lungenkrebs) + (Rauchen tödlich) + (Rauchen ist tödlich).

Figure 2: Proportion of selected Ngrams in German publications that relate to the hazards of smoking



Note: 1962: Royal College of Physicians Report (UK) | 1964: Surgeon's General Report (US) | 1977: ban on tobacco advertising on German TV; *Source:* Google Ngram 2016.

special report on smoking (Der Spiegel, 20.1.1964) immediately following the publication of the US report. This indicates that information on the dangers of smoking reached the broader German public already in the mid 1960s. The debate finally resulted in policy action in 1977 when tobacco advertising on German radio and TV was banned.[8] Based on these two events, we classify our cohorts according to the "available" health information at age 10: born until 1954, born 1955–1967, and born 1968–1989.

Most members of the oldest cohort were likely not aware of the health-damaging consequences of smoking when they grew up, because the debate was mainly confined to the medical literature. Members of the second cohort (born 1955–1967) grew up after this knowledge reached the broader public by the publication of the first US Surgeon's General Report in 1964 as well as the report in the news magazine *DER SPIEGEL*. The US also marks the beginning of awareness campaigns in the US (de Walque, 2010; Kenkel and Sindelar, 2011). Individuals growing up during this time therefore could have known about the harmful effects of tobacco consumption, possibly the information reached the high educated earlier than the low educated. In contrast, all members of the third cohort (born 1968 to 1989), who reached adolescence after the passing of the ban on tobacco

[8]Notably, health warnings on cigarette packages were introduced only in 2003, which coincides with another rise in the number of relevant Ngrams.

advertising on German TV in 1977 should have been aware of the harmful consequences of smoking. Thus arguably both low and high educated could have been fully informed.

2.2 Results

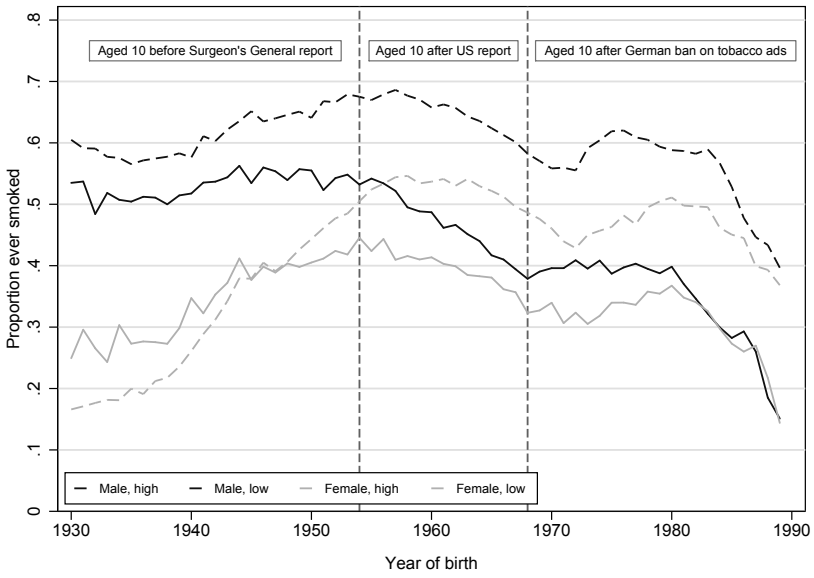
2.2.1 Educational Differences in Ever Smoking Across Cohorts

Figure 3 (a) shows long-term trends in ever smoking by year of birth, respectively for both sexes and education levels. Smoking prevalences generally increase until the 1950s cohorts, but more so among women than among men. The share of ever smokers peaks among individuals born in the 1950s and generally declines from then on. This pattern suggests that the publication of the US Surgeon's General Report (and the subsequent media coverage in Germany) might have been crucial for Germany as well. The sharp decline in the prevalence of smoking for men and women born after 1985 likely arises due to a composition effect. Individuals born in the late 1980s are relatively young (aged 16–20) compared to earlier cohorts when we observe them and still might take up smoking. While earlier-born men tend to have smoked a lot more frequently than women, these gender differences have nearly vanished (conditional on education level). In contrast to the gender gap, the education gap has widened. While there are hardly any differences by education in ever smoking among men from older birth cohorts, prevalences are diverging across cohorts with higher smoking rates among low-educated men. High educated women born up to 1945 have smoked more often than low educated, but the educational differences have flipped sign and are nearly as strong now as they are for men. In both sexes, the disparities appear to be most pronounced for individuals born in the most recent years. In general, this cohort pattern is similar to earlier findings based on German data (Brenner, 1993; Piontek et al., 2010).

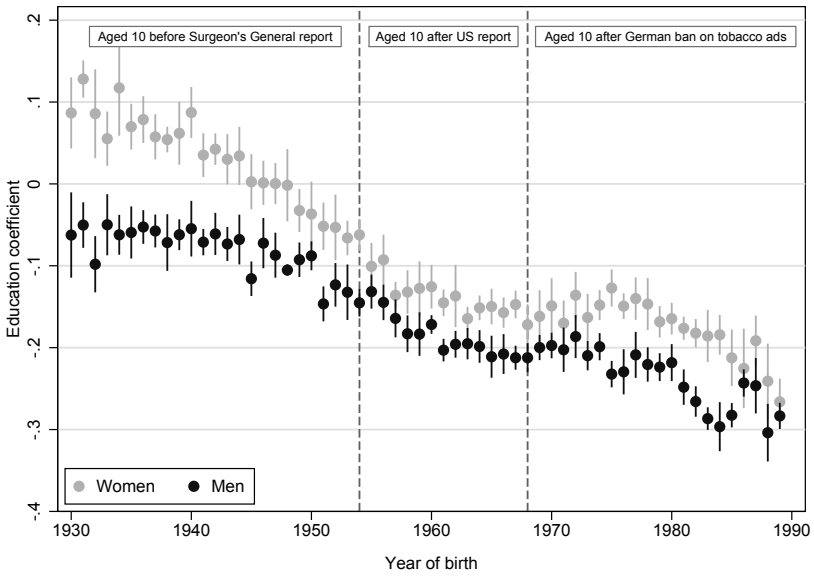
Figure 3 (b) plots education differences by cohort directly, estimated from linear probability models, controlling for German nationality, state fixed effects and a fourth order polynomial in age. Educational differences have clearly increased in absolute value across birth cohorts and are larger for men than for women. While high educated men born in 1930 have a 6 percentage point lower probability to ever smoke, their counterparts born in the 1980s have a 25 percentage point lower probability. The development is almost parallel for women, with high educated women born between 1930–1945 being even more likely to ever smoke than lower educated women of this generation. That estimates become more negative — younger birth cohorts showing a larger education gap — is in line with previous findings from other countries (e.g. de Walque, 2010).

Figure 3: Cohort trends in the proportion of respondents who ever smoked, by gender and education, and education differences in smoking, by gender

(a) Proportion of respondents who ever smoked



(b) OLS estimates of education differences in smoking



Note: Coefficients and the corresponding 95% confidence intervals were obtained from separate OLS regressions of ever smoking on high education by year of birth; Control variables included: German nationality, fixed effects for states of residence, fourth order polynomial in age; Robust standard errors clustered at region*cohort level in parentheses; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

2.2.2 Educational Differences in Age at Smoking Onset

The fact that the youngest cohort, which had the largest amount of information on the dangers of smoking, also has the largest education gap suggests education might be important in using that information. We now concentrate on this cohort and study educational differences in age at smoking initiation.[9] The upper panel of Figure 4 shows the estimated (log) hazard rates of taking up smoking at ages 10 to 25 and the corresponding 95% confidence intervals – separately for the low and high educated, and for men and women. The estimated hazard rates follow a similar age pattern: they increase steadily from age 10, peak at age 16 and then decline.

Turning to education differences, let us first look at the results for men. Here we find significantly higher hazard rates to start smoking in low compared to high educated men from age 10 to age 19. Remarkably, the education differences (shown as log hazard rate ratios, presented in the lower panel) become smaller the longer individuals are in school and thus the more education they have acquired. After school has ended also for the high educated (from age 20 onwards), the hazard rates of the low and high educated are remarkably similar and the differences become insignificant. A causal effect of post-compulsory education would suggest the exact opposite. Thus starting smoking in adulthood is hardly related to schooling. Results for women are very similar. To summarize, educational differences in smoking uptake are largest before education is completed and even before the minimum school leaving age of 16 is reached. After age 20, when high educated individuals already have acquired their university entrance qualification, the differences between the two educational groups become negligible.

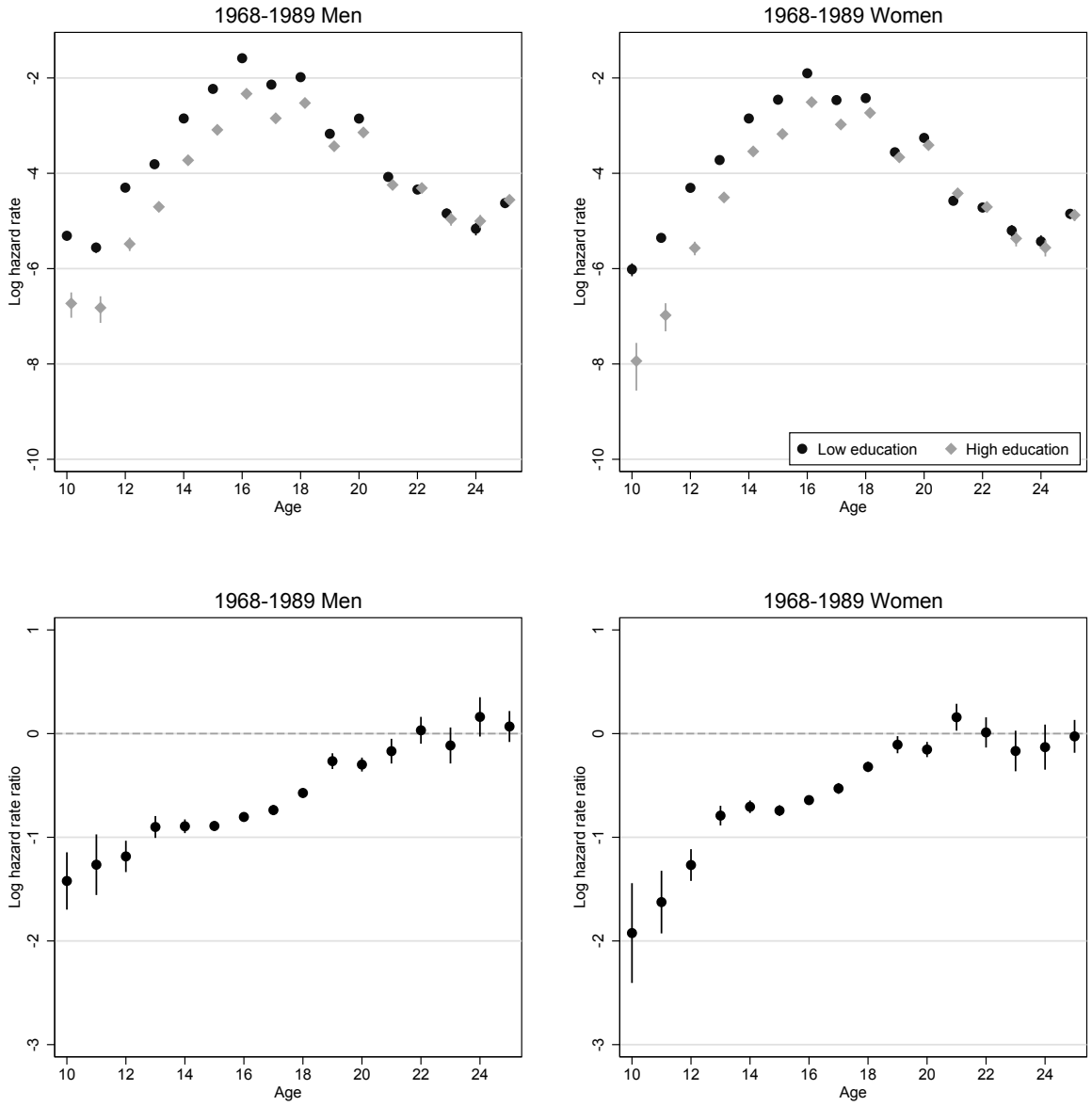
Figure 5 shows this in yet another way, namely by the *cumulative* proportion of men and women who ever smoked before a given age. [10] The share of individuals who ever smoked rises more sharply in age among the low than among the high educated. The education differences increase up to about age 18 but hardly change thereafter. This pattern is similar for men and women, but again more pronounced for men.

At the age of 25 the education difference in the proportion of respondents who ever smoked amounts to 20 percentage points for men and 14 percentage points for women. Note that the educational difference in ever smoking until the age of 16 is nearly as high: 17 percentage points for men and 13 percentage points for women. In other words, educational differences in smoking at the age of 16 account for 85% and 93% of the total difference in ever smoking at the age of 25 among men and women, respectively. Following the discussion in Section **??**, we argue that if at all, at most 15% (7%) of the differences in

[9]Results for the two older cohorts are shown in Figures B.1 and B.2 in the Appendix. By concentrating on the youngest cohorts, we also avoid potential bias due to selective mortality of smokers. In the analysis of the oldest cohort, we have addressed this issue by excluding respondents aged 60 and over from the analysis, see Appendix E.

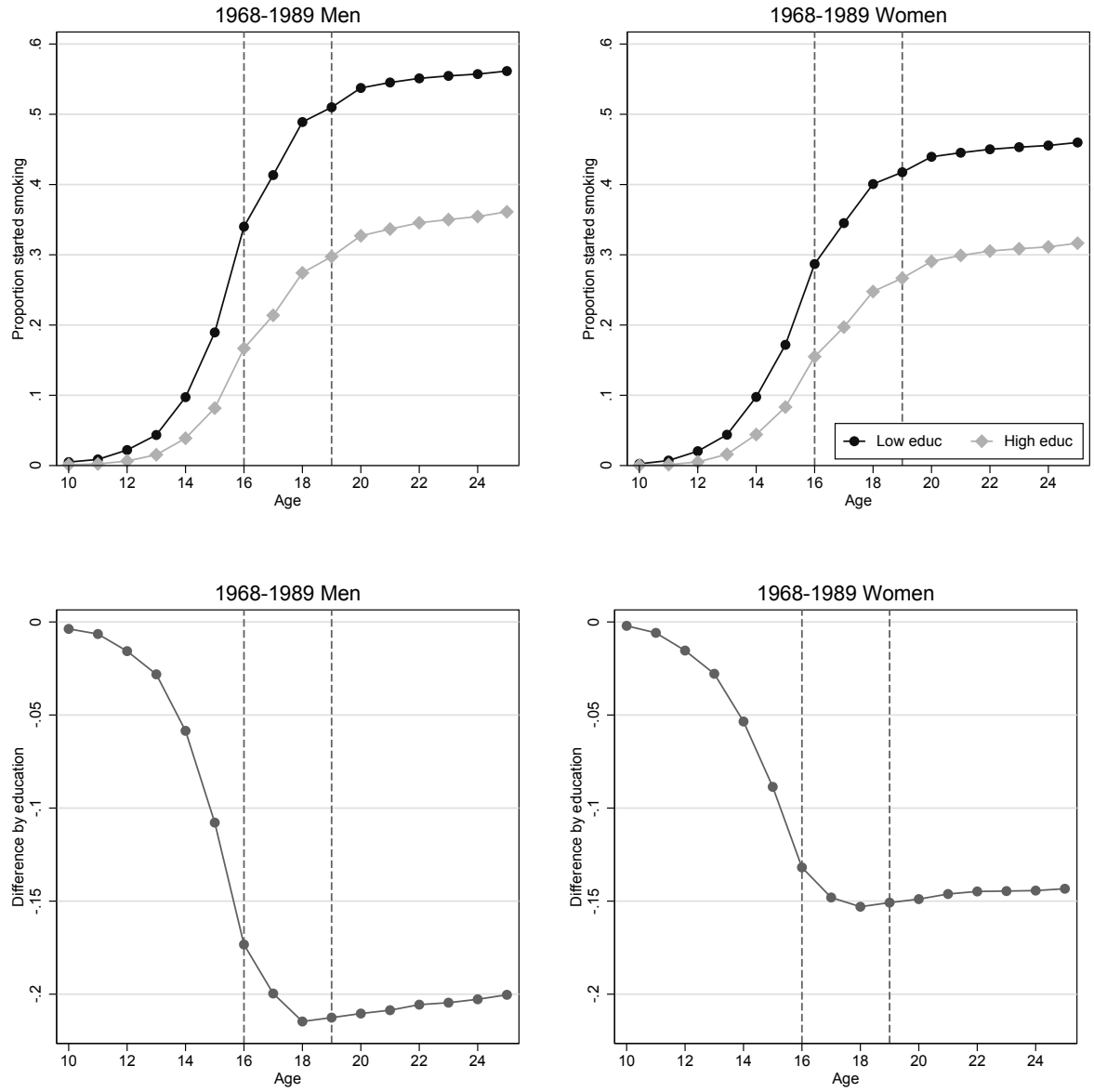
[10]We again focus on individuals of the most recent cohort born between 1968 and 1989. See Figure B.3 and B.4 in the Appendix for the results pertaining to older cohorts.

Figure 4: Hazard rates: Smoking initiation by completed formal education



Note: Upper panel: each circle/diamond presents the age-specific log-hazard rate and the corresponding 95% confidence interval obtained from a discrete time event history model taking smoking initiation as failure event; Lower panel: presents the calculated hazard rate ratios and the corresponding 95% confidence interval; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure 5: Educational differences in smoking initiation until a given age



Note: Figures in the upper panel show the education-specific distributions while the figures in the lower panel display its calculated differences (higher educated - lower educated); *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

ever smoking between high and low educated individuals might be attributed to a causal effect of post-compulsory education.

2.3 Smoking behavior in adulthood

As noted before, the Microcensus does not allow constructing complete individual smoking biographies. However, due to our data being repeated cross-sections containing information on current smoking, we are able to track birth cohorts' smoking behavior over time, at least partially. We know, for each adult respondent, the smoking status at two points in time: at age 20 and at the age of the survey.[11] For instance, for everyone who was interviewed at the age of, e.g., 40 (born 1949, 1959, 1963, 1965, or 1969), we know whether they smoked at 20 and whether they still smoke at 40. We use this information to examine if the cross-sectional education gradient in smoking in these cohorts changes in adulthood. If the gradient is found to increase, education potentially had an influence on smoking behavior in adulthood, particularly on stopping smoking. If the education gradient remains unchanged, this is evidence against such influence.

We estimate education gradients as before by the coefficients of OLS regressions of current smoking on education and basic covariates such as survey year, region, German nationality. In order to increase the number of observations, we have not only looked at respondents who were exactly 40 years old, for instance, when they were interviewed, but all respondents aged 38 to 42 (i.e. within a five year age band around the pivotal age). Age trends within these bands were accounted for by controlling for age at interview relative to the pivotal age. In the following, we only show and discuss results for our most recent cohort (born 1968 to 1989) whom we observe up to age 40. Results for older cohorts are qualitatively similar and can again be found in the Appendix (Tables B.1 and B.2). In addition to our earlier analyses we also show education gradients when education is measured by having completed college (as in de Walque, 2010), which typically happens after age 20 but before age 30, and years of full-time education.

The results are shown in Table 2. Each number shows the percentage point difference in smoking at different ages between high and low educated in the above sense (Panel A), between college graduates and others (Panel B), and per year of full-time education (Panel C). The overall picture is clear: Education gradients hardly change as respondents get older. Interpreted in the same way as before, this means smoking differences by (future) college education are almost exclusively due to selection into education rather than a causal effect of college education on smoking behavior. This can also be seen

[11] We assume that everyone who started smoking before age 20 still smokes at age 20. To justify this assumption, we undertook some complementary analyses based on the German Socio-economic Panel (SOEP), which suggested that 2% of men and 4% of women who report to have started smoking in adolescence stopped before the age of 20. See Wagner et al. (2007) for information on the Socio-economic Panel study. Data for years 1984-2013, version 30, SOEP, 2015, doi:10.5684/soep.v30.

if one compares coefficients across Panels A and B. At age 20 the gradient is larger regarding college education (Panel B) than regarding post-compulsory schooling (Panel A). This suggest that college graduates are a slightly stronger selection in terms of smoking behavior than those who acquired the necessary entrance qualification.

Table 2: Educational differences in current smoking at age 20, 30 and 40 (percentage point differences estimated by OLS). Education differences at age 20 are derived from retrospective information

	Men		Women	
Cohort observed at...	Age 30	Age 40	Age 30	Age 40
Panel A: High education				
Smoked at age 20	-0.2036***	-0.1938***	-0.1484***	-0.1578***
	(0.0055)	(0.0097)	(0.0053)	(0.0094)
Smoked at age 30	-0.2184***		-0.1622***	
	(0.0054)		(0.0049)	
Smoked at age 40		-0.2036***		-0.1611***
		(0.0090)		(0.0083)
N	32,993	10,781	34,091	11,090
Panel B: College education				
Smoked at age 20	-0.2273***	-0.2108***	-0.1656***	-0.1751***
	(0.0067)	(0.0109)	(0.0065)	(0.0111)
Smoked at age 30	-0.2419***		-0.1637***	
	(0.0062)		(0.0056)	
Smoked at age 40		-0.2065***		-0.1691***
		(0.0098)		(0.0092)
N	27,901	9,743	28,371	9,704
Panel C: Years of education				
Smoked at age 20	-0.0371***	-0.0337***	-0.0279***	-0.0295***
	(0.0009)	(0.0014)	(0.0008)	(0.0014)
Smoked at age 30	-0.0413***		-0.0313***	
	(0.0008)		(0.0008)	
Smoked at age 40		-0.0359***		-0.0324***
		(0.0013)		(0.0013)
N	32,993	10,781	34,091	11,090

Note: Control variables included: German nationality, fixed effects for states of residence, dummies for the survey year, age (age at interview minus target age); Robust standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

3 Post-Schooling Health Education

3.1 Data and Empirical Approach

The previous analyses suggest that general education is unlikely the driving force behind observed differences in smoking initiation. Individuals start smoking before schooling is

competed – leading to a reversed order of cause and effect. In this section, we complement our analysis by studying whether health-related knowledge (in contrast to general education) can explain an individual's smoking decisions.

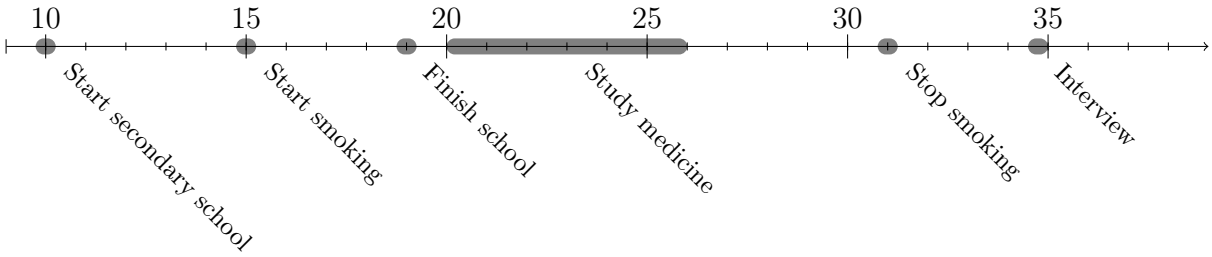
We focus on health education acquired post-schooling, because too little is known about health-knowledge taught in secondary school.[12] But we know more about health-knowledge acquired in college or during vocational training. Licensure ensures that health professionals have completed health-related education. This holds not only for academics (doctors and pharmacists) but also at an intermediate level of general education (nurses, midwifes, etc.). Thus we define respondents working in health-related occupations according to the German classification of occupations (KldB 1992) as individuals who acquired health-related knowledge. We first compare the smoking behavior of physicians and pharmacists on the one hand and other academics on the other hand. Both groups have enjoyed the same overall level of education (13 years in school plus 5 to 6 years in college) but with decidedly different content. As shown in Table 1, 6% of the college graduates in our data are working as physician or pharmacist. Second, we compare individuals with intermediate schooling (German: *Realschulabschluss*) working in a health-related occupation, for instance as a nurse, with a comparable individual in another occupation (10 years of schooling plus 3 years of vocational training). Health related occupations are those with KldB code 85: nurse, physiotherapist, masseur, midwife, nutrition consultant, alternative practitioner, physician's assistant, medical technical assistant, pharmaceutical technician, and speech therapist. 7% of the intermediately educated individuals pursue a profession within the health sector (see Table 1).

To illustrate our empirical approach, consider the fictitious life course of Doctor Bob – shown in Figure 6. Bob, who attends academic secondary track, starts smoking at the age of 15. He finishes post-compulsory schooling at the age of 19. At 20, Bob takes up his studies of medicine (for six years) and acquires health-specific knowledge. At 31 he decides to stop smoking and a few years later he enters our sample. Obviously, Bob's decision to start smoking is made before he receives health education and the former cannot be affected by the latter. Still, there might be a strong negative correlation between smoking in adolescence and planning to become a doctor, for instance because the parents were doctors, too, and warned against the dangers of smoking, or because of future-oriented time preferences.

[12]The Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany (KMK) publishes recommendations on health education at schools, also in respect of addiction prevention. There exist three different circulations of these recommendations: "Gesundheitserziehung und Schule" published 01.06.1979, "Sucht- und Drogenprävention" 03.07.1990 and "Empfehlung zur Gesundheitsförderung und Prävention in der Schule" published 15.11.2012. See https://www.kmk.org for details. Rather than providing specific guidance these recommendations are worded in general terms by primarily describing core competencies the students should posses, similar to the school subject curricula.

We study the strength of this "selection effect" into studying medicine rather than economics, say, by comparing smoking rates of *future* academics at the age of 20, i.e. when they usually take up their studies. By restricting the comparison to academics, we aim to eliminate the more general selection effect into high versus low education. One limitation of our data is that they lack information on the age when individuals actually start studying. We assume this is to be age 20 for everyone because in Germany, this is the average age at which men and women start studying at university – independent of the field (Feuerstein, 2008). When analyzing quit rates, we compare only academics who smoked at the age of 20. If health-related education has a negative effect on smoking behavior, we should find higher quit rates among doctors than among other academics.

Figure 6: Fictional life course (academic): Acquiring health education



When analyzing selection into health education and the effect of health education on quit rates for intermediately educated individuals, e.g. nurses or physiotherapists, we follow the same approach. The only differences is that the pivotal age in this case is 16, not 20. This is when compulsory schooling ends and vocational training starts.

3.2 Results

Table 3 shows raw gender- and occupation-specific incidences to start and stop smoking.[13] We first focus on university graduates, i.e. we compare physicians/pharmacists with other academics as presented in Panel a). Columns 1 and 3 compare the cumulative proportions of men and women in medical and non-medical occupations who started smoking until they were 20. Individuals who become physicians or pharmacists later in life were already less likely to start smoking before the age of 20, i.e. before receiving health education. Among male physicians and pharmacists (Column 1), 27% started to smoke until age 19, whereas 31% of academic men in other occupations did. In other words, men becoming physicians or pharmacists later in life have a 4.2 percentage point smaller probability to take up smoking before they begin receiving health education. While the overall proportion of academic women taking up smoking is smaller than the proportion of men (Column 3), the difference between women with medical and non-

[13]Estimates obtained from OLS models controlling for German nationality, state fixed effects and a 4th order polynomial in age lead to very similar results (see A.3).

Table 3: Proportions starting and stopping smoking by occupation

	Men		Women	
	Start smoking (1)	Stop smoking (2)	Start smoking (3)	Stop smoking (4)
a) Conditional on academic education				
Physicians/pharmacists	0.2708	0.5995	0.1894	0.5893
	(0.0067)	(0.0141)	(0.0068)	(0.0195)
N	4,435	1,201	3,368	638
Other academics	0.3125	0.5866	0.2305	0.5740
	(0.0017)	(0.0032)	(0.0019)	(0.0047)
N	74,162	23,171	47,346	10,905
Difference	-0.0417***	0.0129	-0.0411***	0.0154
	(0.0069)	(0.0145)	(0.0071)	(0.0201)
b) Conditional on secondary education				
Health professionals	0.1560	0.3588	0.1032	0.4022
	(0.0089)	(0.0297)	(0.0023)	(0.0116)
N	1,679	262	17,362	1,790
Other occupations	0.1366	0.3734	0.0992	0.3593
	(0.0010)	(0.0039)	(0.0008)	(0.0041)
N	107,280	14,635	135,092	13,380
Difference	0.0194**	-0.0146	0.0040*	0.0429***
	(0.0098)	(0.0300)	(0.0024)	(0.0123)

Note: Proportions in start smoking are defined before the age of 20/16, i.e. before health education is received; Proportions in stop smoking are conditional on taking up smoking before the age of 20/16; Standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

medical education is similar (4.1 percentage points). To summarize, individuals who study medicine or pharmacy and become doctors or pharmacists are already different in terms of their smoking behavior before acquiring the health-specific knowledge that is taught in medical school.

Quit rates among academics who smoked before the age of 20 are close to 60%, independent of sex and field (Panel a, Columns 2 and 4). For both men and women, we find marginally higher quit rates (1.3 and 1.5 percentage points, respectively) among those working in a medical profession. These differences are not statistically significant, however. Overall, our results suggest that 6 years of medical education at university have at best a very small effect on the likelihood of stopping smoking.

Our findings for non-academics in Panel b) are somewhat different. The proportion of individuals taking up smoking before the age of 16, when they begin their vocational training, is even slightly higher (1.9 percentage points for men and 0.4 percentage points for women) for individuals working as health professionals. Regarding smoking cessation, men working in health-related occupations who have started smoking before the age

of 16 are even 1.5 percentage points *less* likely to quit smoking. However, only few intermediate-educated men work in the health sector, so that the standard error of the proportion of those who quit is fairly large and thus also the estimated difference between men in medical and non-medical occupations. In contrast to men, women who received occupation-specific health education have a 4.3 percentage point higher probability to stop smoking. This difference is statistically significant and might be attributed to the health-specific knowledge acquired during vocational training. In fact, intermediately educated women seem to be the only group for which health knowledge has a sizable affect. Why this is the case is unclear. Many women who smoke stop when they become pregnant – in order not to compromise the health of their unborn child. One might speculate that health education at the intermediate level also teaches about those health risks, so that knowledgeable women who become pregnant are more likely to quit.

Our analysis on health-related knowledge raises a few concerns. First, as already mentioned, we assume that individuals who start smoking until age 20 (16) do not quit smoking before they are 20 (16), because we do not know the age at which individuals stop smoking. This might bias our results if individuals who choose health-related occupations are more likely to stop smoking *before* they take up their medical studies or vocational training. Supplementary analyses using the German SOEP indicate that the proportion quitting before age 20 is only about 3%. It remains possible that we overstate the proportion of individuals who smoke at age 20 and study medicine or pharmacy. In this case, however, we underestimate the selection effect and overestimate the effect of health-knowledge. Second, some individuals might have abandoned their medicine studies before receiving a degree but completed another (non-medical) study. This might bias downwards our estimated effect of health knowledge, as those individuals received some health education and might thus be more likely to quit smoking. However, we believe that this bias is negligible as in Germany, the dropout rate for medicine is below 10% (Heublein et al., 2012). Third, physicians are a highly selective group of individuals even among academics because the admission to medical school in Germany is highly competitive and generally favors individuals with the best school grades. Thus physicians could also be positively selected on unobserved "non-cognitive" skills or time preferences and not be comparable with other academics, even if both groups are restricted to individuals who smoke at age 20. In fact, our findings support the interpretation that unobserved characteristics likely determine the choice of occupation, i.e. post-schooling health education, as well as the decision to start smoking. Neglecting these unobserved characteristics in our comparison of quit rates, however, will lead to an overestimate of the health-knowledge effect. To conclude, none of these concerns seem to be strong enough to cast serious doubt on our results.

4 Summary and Conclusion

Recent empirical studies exploiting credible exogenous variation in schooling have yielded mixed evidence – even within one country and using the same data – regarding the question whether the link between education and smoking is causal. For instance, Jürges et al. (2011) find an effect of education on smoking exploiting academic track openings in Germany, whereas Kemptner et al. (2011) find little evidence on a causal effect by exploiting changes in compulsory schooling. In this paper we complement the earlier instrumental variables studies and follow a simple but effective approach suggested in Farrell and Fuchs (1982) in order to explore *how much* of the relationship between education and smoking can possibly be causal. This approach exploits a specific characteristic of smoking that probably sets it apart from other health behaviors. Almost all smokers start smoking in adolescence or early adulthood, that is, at a time before formal education is completed. The finding that (future) education explains smoking has led Farrell and Fuchs (1982) to claim that education cannot be a major determinant of smoking initiation as the cause must precede the effect. We apply this reasoning analyzing a large German data set of more than one million observations which contains retrospective data on smoking initiation. We find that about 90% of the differences in smoking between low and high educated individuals are already present at age 16, before compulsory education is completed. Whether an individual ever smokes is thus predominantly determined at an age before education differences are likely to be effective.

One of the few papers that have taken a similar approach is de Walque (2010), who argues, however, that one should also consider possible differences in smoking cessation in adulthood that could be due to education. Our data, which consist of repeated cross-sections, allow us to partly study the development of current smoking during adulthood, for instance before and after completing college. We find that educational differences in smoking hardly change in adulthood. We interpret this result as support for our main hypothesis: selection rather than causation drives most of the association between education and smoking.

Further, we examine the role of health-related knowledge. General education has only little medical content, so it is not clear how much of the general education-smoking gradient is due to health knowledge that is acquired in school. We approach this question by comparing individuals of identical formal education levels but with different academic or vocational training. Specifically, we compare the smoking behavior of physicians and pharmacists with other academics before and after attending university or college. The results are again largely in line with our main hypothesis. Doctors and pharmacists are less likely to smoke, but this is already before they start studying. Afterwards they are *not* more likely to stop smoking. On a lower level of general education, we compare health workers such as nurses with otherwise similar respondents. Here, we find that

(female) nurses are similarly likely to smoke before they start their vocational training but indeed more likely to quit smoking (conditional on having smoked already at age 16). This speaks in favor of some effect of health related knowledge.

Overall, however, our results cast doubt on the external validity of studies finding a strong protective effect of education on smoking by exploiting exogenous variation in schooling. We show that at best, education can have a very small impact on smoking. This is in line with recent studies that find limited evidence for causal effects of education on smoking behavior or health in general. Future research should focus on the family and school context to understand selection into smoking and why it is so strongly correlated with education decisions in order to design successful primary prevention programs.

References

Becker, G. S. (2007). Health as human capital: synthesis and extensions, *Oxford Economic Papers* **59**(3): 379–410.

Becker, G. S. and Mulligan, C. B. (1997). The endogenous determination of time preference, *The Quarterly Journal of Economics* **112**(3): 729–758.

Brenner, H. (1993). A birth cohort analysis of the smoking epidemic in West Germany., *Journal of Epidemiology and Community Health* **47**(1): 54–58.

Bricard, D., Jusot, F., Beck, F., Khlat, M. and Legleye, S. (2015). Educational inequalities in smoking over the life cycle: an analysis by cohort and gender, *International Journal of Public Health* **61**(1): 101–109.

Bright, B. C. and Soulakova, J. N. (2014). Evidence of telescoping in regular smoking onset age, *Nicotine & Tobacco Research: Official Journal of the Society for Research on Nicotine and Tobacco* **16**(6): 717–724.

Clark, D. and Royer, H. (2013). The effect of education on adult mortality and health: Evidence from Britain, *The American Economic Review* **103**(6): 2087–2120.

Cutler, D. M. and Lleras-Muney, A. (2010). Understanding differences in health behaviors by education, *Journal of Health Economics* **29**(1): 1–28.

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

de Walque, D. (2010). Education, information, and smoking decisions: Evidence from smoking histories in the United States, 1940–2000, *Journal of Human Resources* **45**(3): 682–717.

Der Spiegel (20.1.1964). Rauchen (cover topic), **1964**(4).

Doll, R. and Hill, A. B. (1954). The mortality of doctors in relation to their smoking habits, *British Medical Journal* (4877): 1451–1455.

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

Farrell, P. and Fuchs, V. R. (1982). Schooling and health, *Journal of Health Economics* **1**(3): 217–230.

Feuerstein, T. (2008). Entwicklung des Durchschnittsalters von Studierenden und Absolventen an deutschen Hochschulen seit 2000, *Wirtschaft und Statistik* (7): 603–608.

Fuchs, V. R. (1982). Time preferences and health: An exploratory study, *in* V. R. Fuchs (ed.), *Economic Aspects of Health*, Chicago Press, pp. 93–120.

Grossman, M. (2006). Chapter 10: Education and nonmarket outcomes, *in* E. Hanushek and F. Welch (ed.), *Handbook of the Economics of Education*, Elsevier, pp. 577–633.

Grossman, M. (1972). On the concept of health capital and the demand for health, *Journal of Political Economy* **80**(2): 223–255.

Grossman, M. (2015). The relationship between health and schooling: What's new?, *Nordic Journal of Health Economics* **3**(1): 7–17.

Han, M.-Y., Chen, W.-Q. and Chen, X. (2011). Do smoking knowledge, attitudes and behaviors change with years of schooling? a comparison of medical with non-medical students in china, *Journal of Community Health* **36**(6): 966–974.

Heublein, U., R., J., Schmelzer, R. and Sommer, D. (2012). Die Entwicklung der Schwund- und Studienabbruchquoten an den deutschen Hochschulen. Statistische Berechnungen auf der Basis des Absolventenjahrgangs 2010, *Forum Hochschule 3/2012*, Hochschul-Informations-System.

Johnston, D. W., Lordan, G., Shields, M. A. and Suziedelyte, A. (2015). Education and health knowledge: Evidence from UK compulsory schooling reform, *Social Science & Medicine* **127**: 92–100.

Jürges, H., Reinhold, S. and Salm, M. (2011). Does schooling affect health behavior? Evidence from the educational expansion in Western Germany, *Economics of Education Review* **30**(5): 862–872.

Jürges, H. and Schneider, K. (2011). Why young boys stumble: Early tracking, age and gender bias in the german school system, *German Economic Review* **12**(4): 371–394.

Kemptner, D., Jürges, H. and Reinhold, S. (2011). Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany, *Journal of Health Economics* **30**(2): 340–354.

Kenkel, D. S. (1991). Health behavior, health knowledge, and schooling, *Journal of Political Economy* **99**(2): 287–305.

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

Kenkel, D. S. and Sindelar, J. (2011). Economics of health behaviors and addictions: Contemporary issues and policy implications, *in* S. Glied and P. C. Smith (eds), *The Oxford Handbook of Health Economics*, Oxford University Press, New York, pp. 206–231.

Krall, E. A., Valadian, I., Dwyer, J. T. and Gardner, J. (1989). Accuracy of recalled smoking data., *American Journal of Public Health* **79**(2): 200–202.

Lehmann, R. H. and Peek, R. (1997). *Aspekte der Lernausgangslage von Schülerinnen und Schülern der fünften Klassen an Hamburger Schulen: Bericht über die Untersuchung im September 1996*, Behörde für Schule, Jugend und Berufsbildung, Amt für Schule.

Lundborg, P. (2013). The health returns to schooling - What can we learn from twins?, *Journal of Population Economics* **26**(2): 673–701.

Maralani, V. (2013). Educational inequalities in smoking: The role of initiation versus quitting, *Social Science & Medicine* **84**: 129–137.

Mocan, N. H. and Altindag, D. T. (2012). Education, cognition, health knowledge, and health behavior, *Working Paper 17949*, National Bureau of Economic Research.

Pampel, F. C. (2005). Diffusion, cohort change, and social patterns of smoking, *Social Science Research* **34**(1): 117–139.

Pampel, F. C., Legleye, S., Goffette, C., Piontek, D., Kraus, L. and Khlat, M. (2015). Cohort changes in educational disparities in smoking: France, Germany and the United States, *Social Science & Medicine* **127**: 41–50.

Park, C. and Kang, C. (2008). Does education induce healthy lifestyle?, *Journal of Health Economics* **27**(6): 1516–1531.

Piontek, D., Kraus, L., Müller, S. and Pabst, A. (2010). To what extent do age, period, and cohort patterns account for time trends and social inequalities in smoking?, *SUCHT* **56**(5): 361–371.

Rosenzweig, M. R. and Schultz, T. P. (1981). *Education and Household Production of Child Health*, American Statistical Association, Washington, Proceedings of the American Statistical Association, Social Statistical Section.

Schulze, A. and Mons, U. (2006). The evolution of educational inequalities in smoking: A changing relationship and a cross-over effect among German birth cohorts of 1921–70, *Addiction* **101**(7): 1051–1056.

Vedøy, T. F. (2014). Tracing the cigarette epidemic: An age-period-cohort study of education, gender and smoking using a pseudo-panel approach, *Social Science Research* **48**(0): 35–47.

Wagner, G., Frick, J. and Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP): Scope, Evolution and Enhancements, *Schmollers Jahrbuch: Zeitschrift für Wirtschafts- und Sozialwissenschaften* **127**(1): 139–169.

A Supplementary tables and figures

Table A.1: OLS estimates: Educational differences (high vs low) in the proportion of respondents who ever smoked, by information cohort

	Men		Women	
Cohort	Ever smoked	N	Ever smoked	N
1930–1954	-0.0866***	224,805	0.0098	239,439
	(0.0070)		(0.0119)	
1955–1967	-0.1867***	149,394	-0.1381***	149,468
	(0.0057)		(0.0060)	
1968–1989	-0.2266***	135,440	-0.1667***	137,775
	(0.0070)		(0.0066)	

Note: Controls included: 4th order polynomials of age, i.region*cohort, German citizenship, year; Robust standard errors clustered at region*cohort level in parantheses; * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Table A.2: OLS estimates: Educational differences (high vs low) in the proportion of respondents who stopped smoking, by information cohort

	Men		Women	
Cohort	Stopped smoking	N	Stopped smoking	N
1930–1954	0.0902***	134,539	0.1294***	76,074
	(0.0073)		(0.0082)	
1955–1967	0.1227***	88,928	0.1461***	74,032
	(0.0047)		(0.0034)	
1968–1989	0.1266****	67,463	0.1310***	56,961
	(0.0045)		(0.0091)	

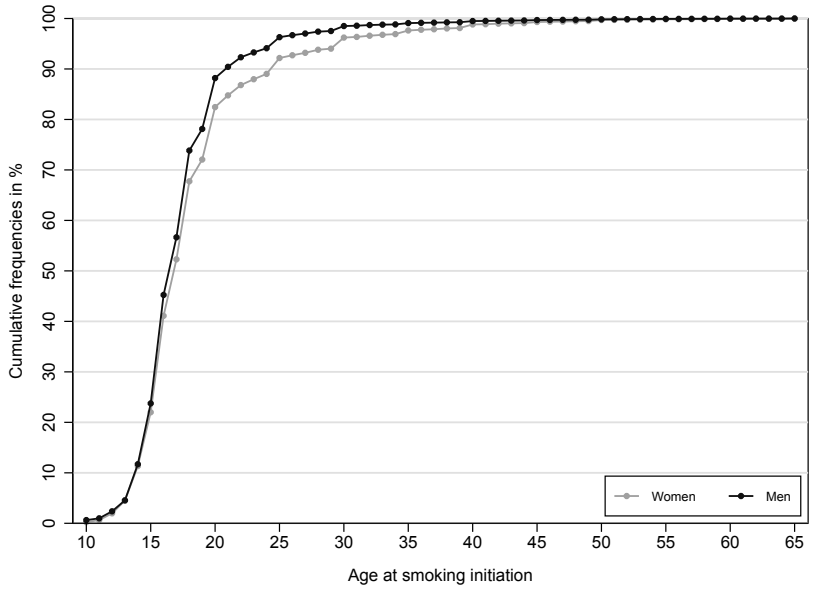
Note: Controls included: 4th order polynomials of age, i.region*cohort, German citizenship, year; Robust standard errors clustered at region*cohort level in parantheses; * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Table A.3: OLS estimates: Occupational differences (medical versus other) in the proportion of respondents who started smoking before they started their occupational training and in the proportion who stopped smoking

	Men		Women	
	Start < 20/16	Stop smoking	Start < 20/16	Stop smoking
a) Conditional on academic education				
Physicians/pharmacists	-0.0560***	0.0193**	-0.0535***	0.0158
	(0.0084)	(0.0096)	(0.0071)	(0.0203)
N	78,597	24,372	50,714	11,543
b) Conditional on intermediate education				
Health occupations	0.0079	0.0175	-0.0207***	0.0385***
	(0.0096)	(0.0133)	(0.0028)	(0.0053)
N	108,959	14,897	152,454	15,170

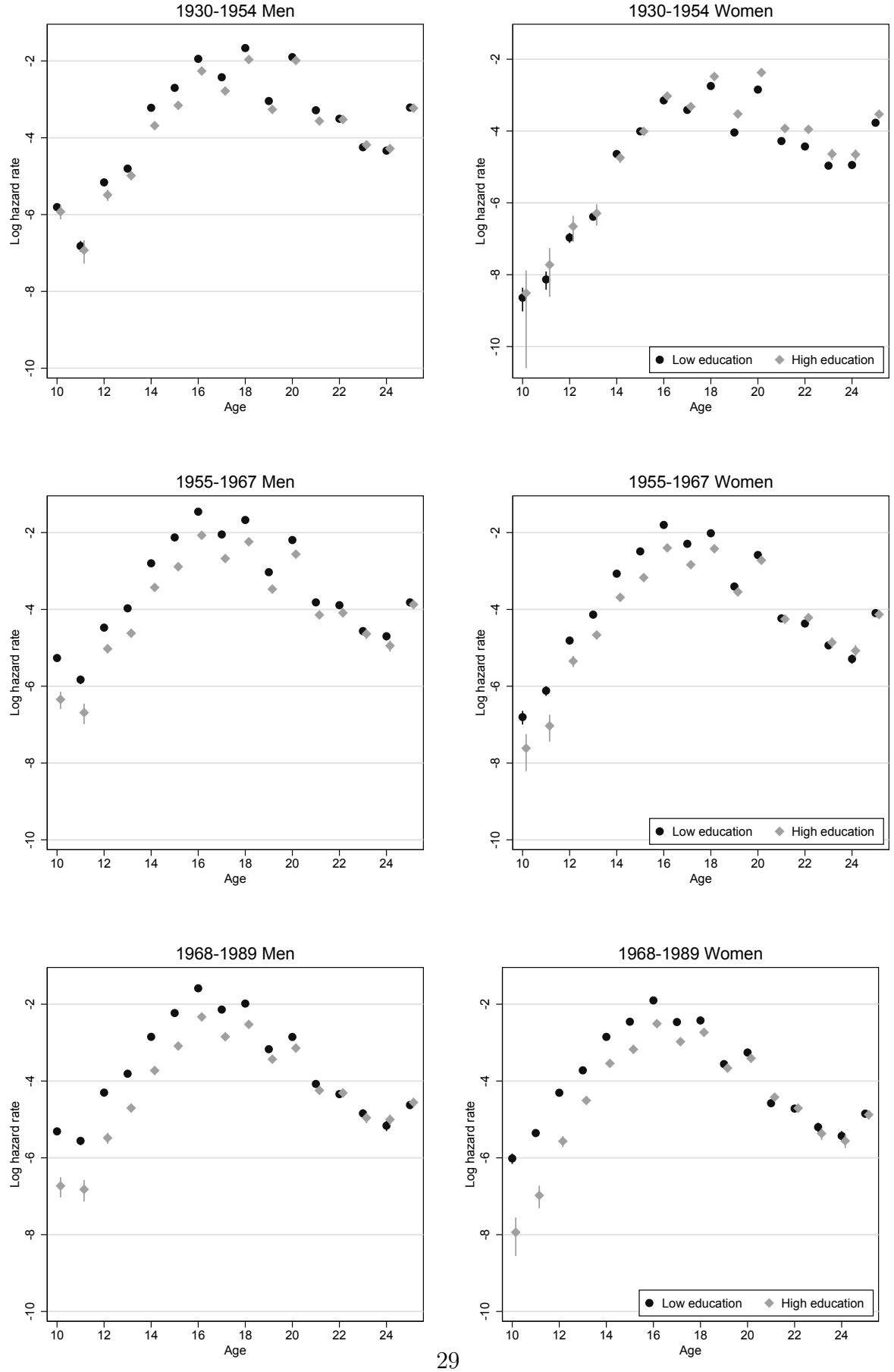
Note: Results for stop smoking are conditional on taking up smoking before the age of 20/16; Control variables included: German nationality, fixed effects for states of residence, fourth order polynomial in age; Robust standard errors clustered at region*cohort level in parantheses; * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure A.1: Cumulative distribution of age at smoking initiation (Microcensus data)



Source: German Microcensus 1989, 1999, 2003, 2005, 2009.

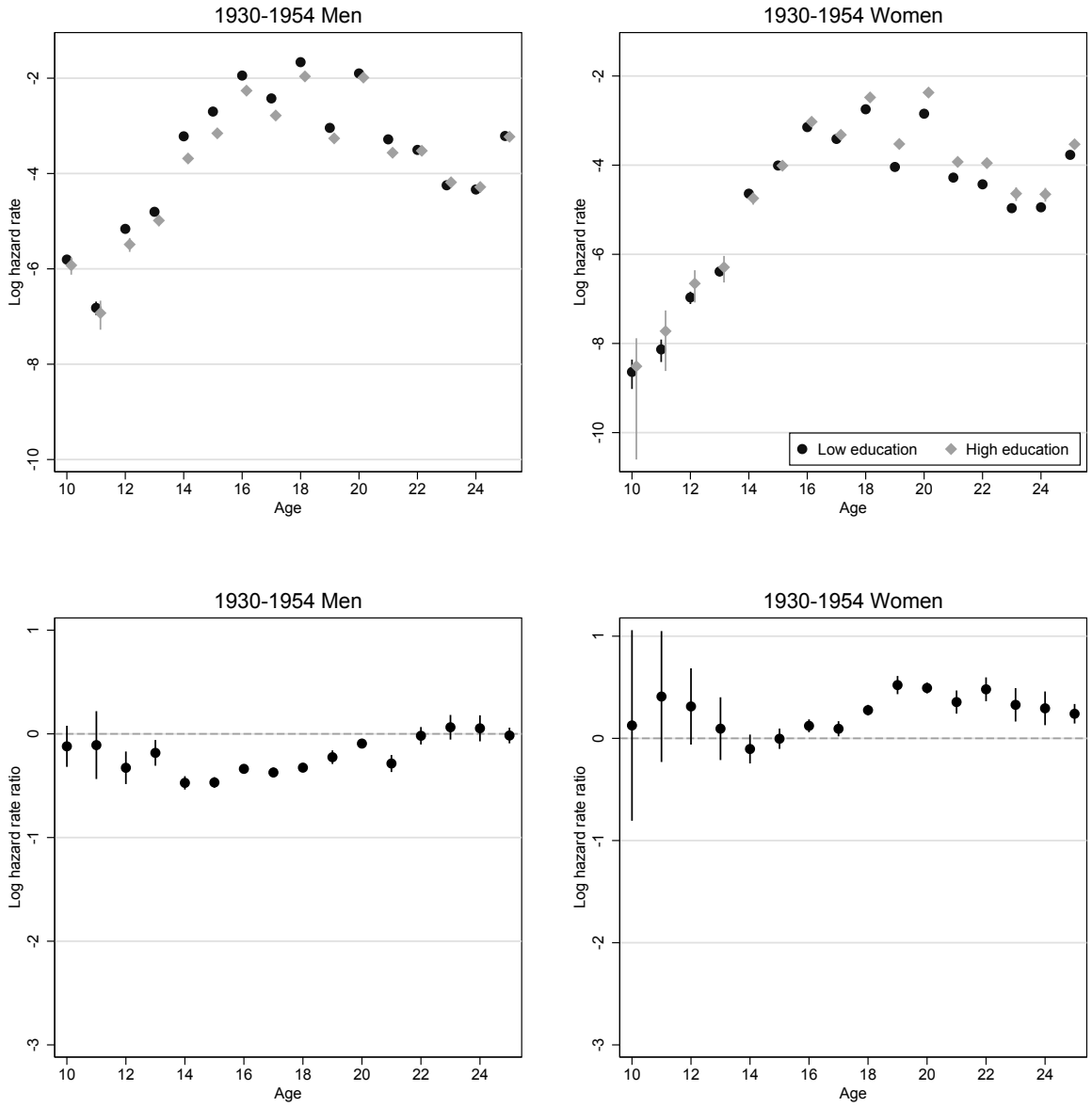
Figure A.2: Log hazard rates: smoking initiation by completed formal education, East Germany



Source: German Microcensus 1989, 1999, 2003, 2005, 2009.

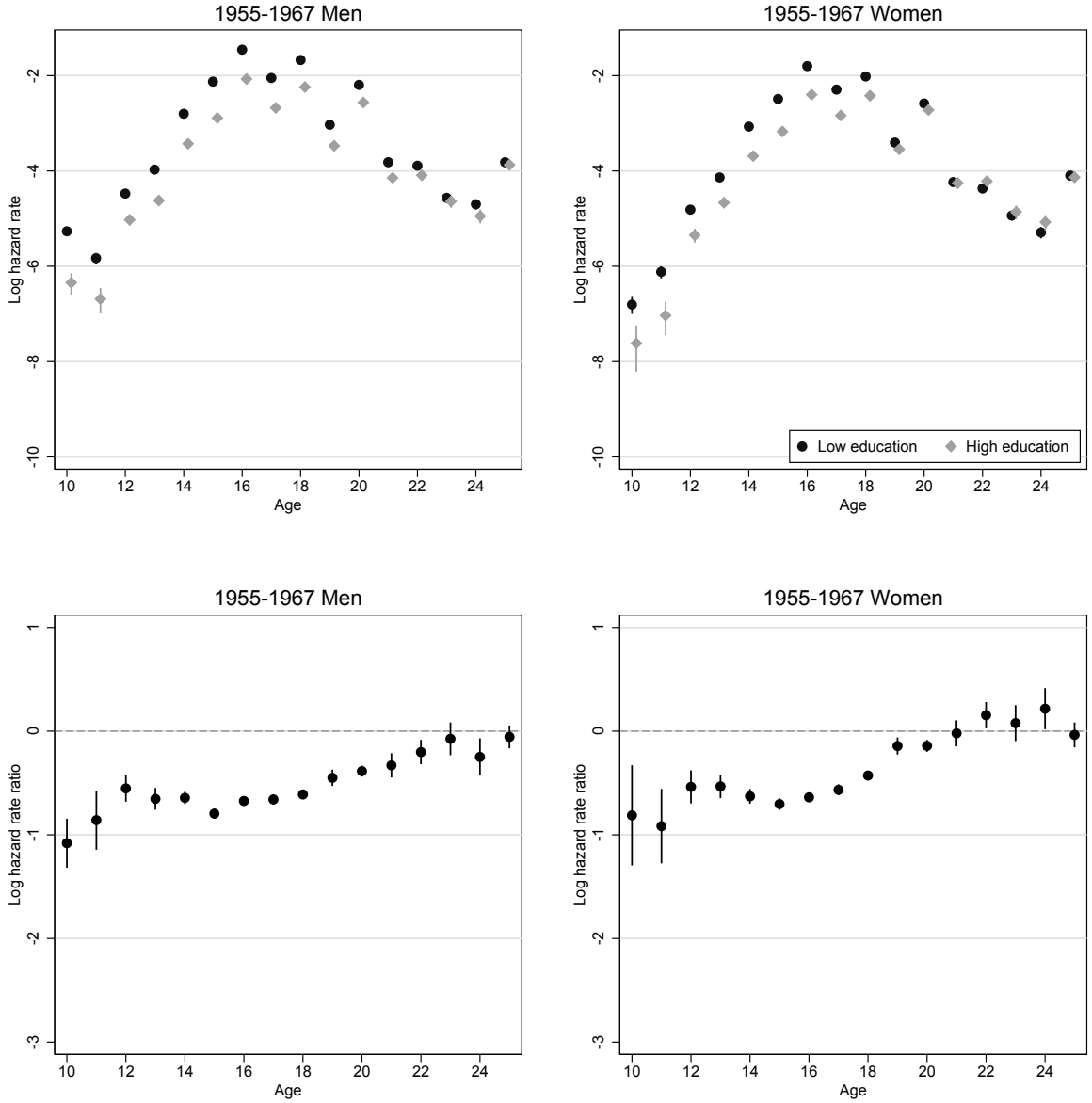
B Results for 1930-54 and 1955-67 cohorts

Figure B.1: Hazard rates: Smoking initiation by completed formal education: 1930-1954



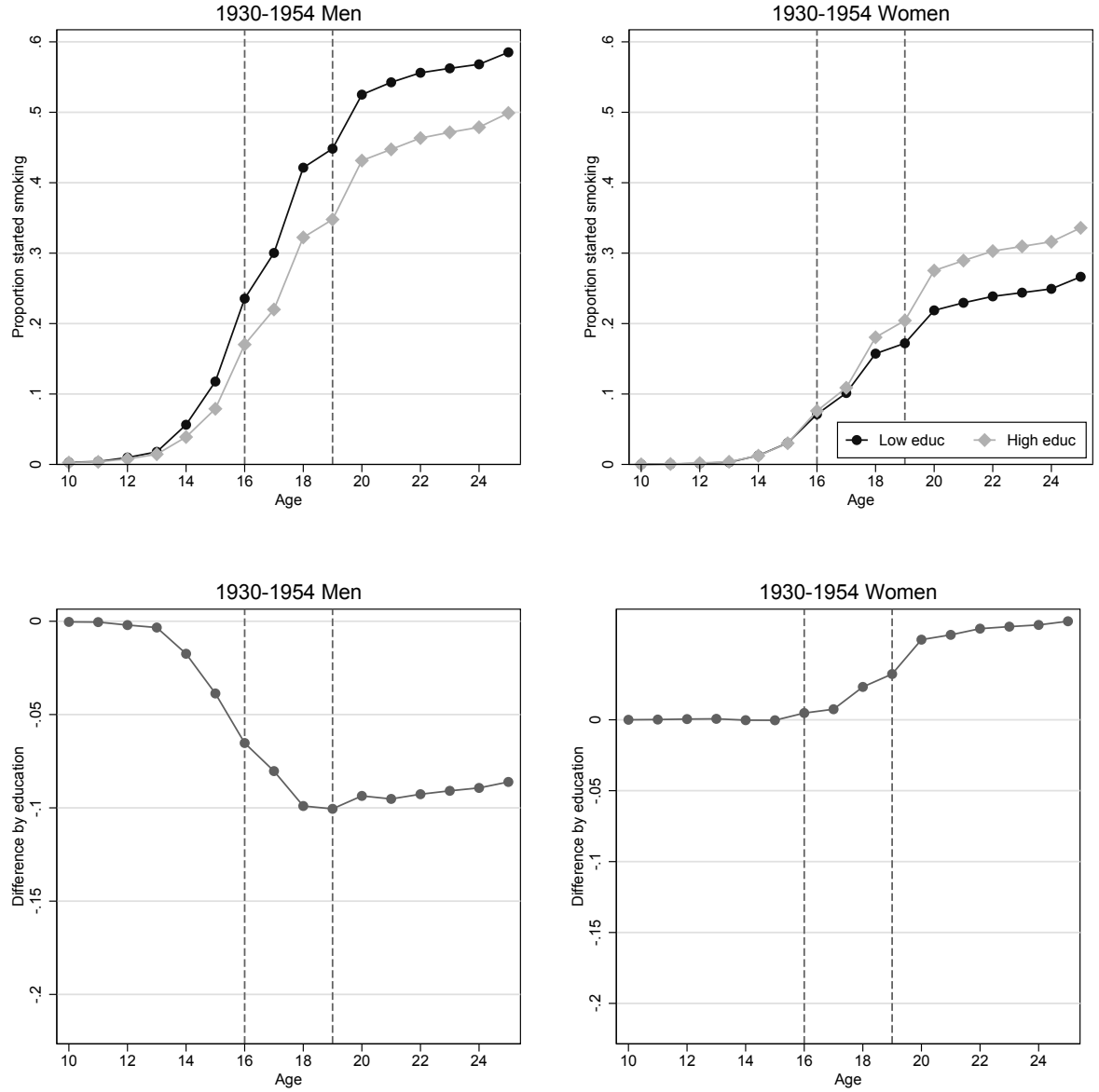
Note: Upper panel: each circle/diamond presents the age-specific log-hazard rate and the corresponding 95% confidence interval obtained from a discrete time event history model taking smoking initiation as failure event; Lower panel: presents the calculated hazard rate ratios and the corresponding 95% confidence interval; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure B.2: Hazard rates: Smoking initiation by completed formal education: 1955-1967



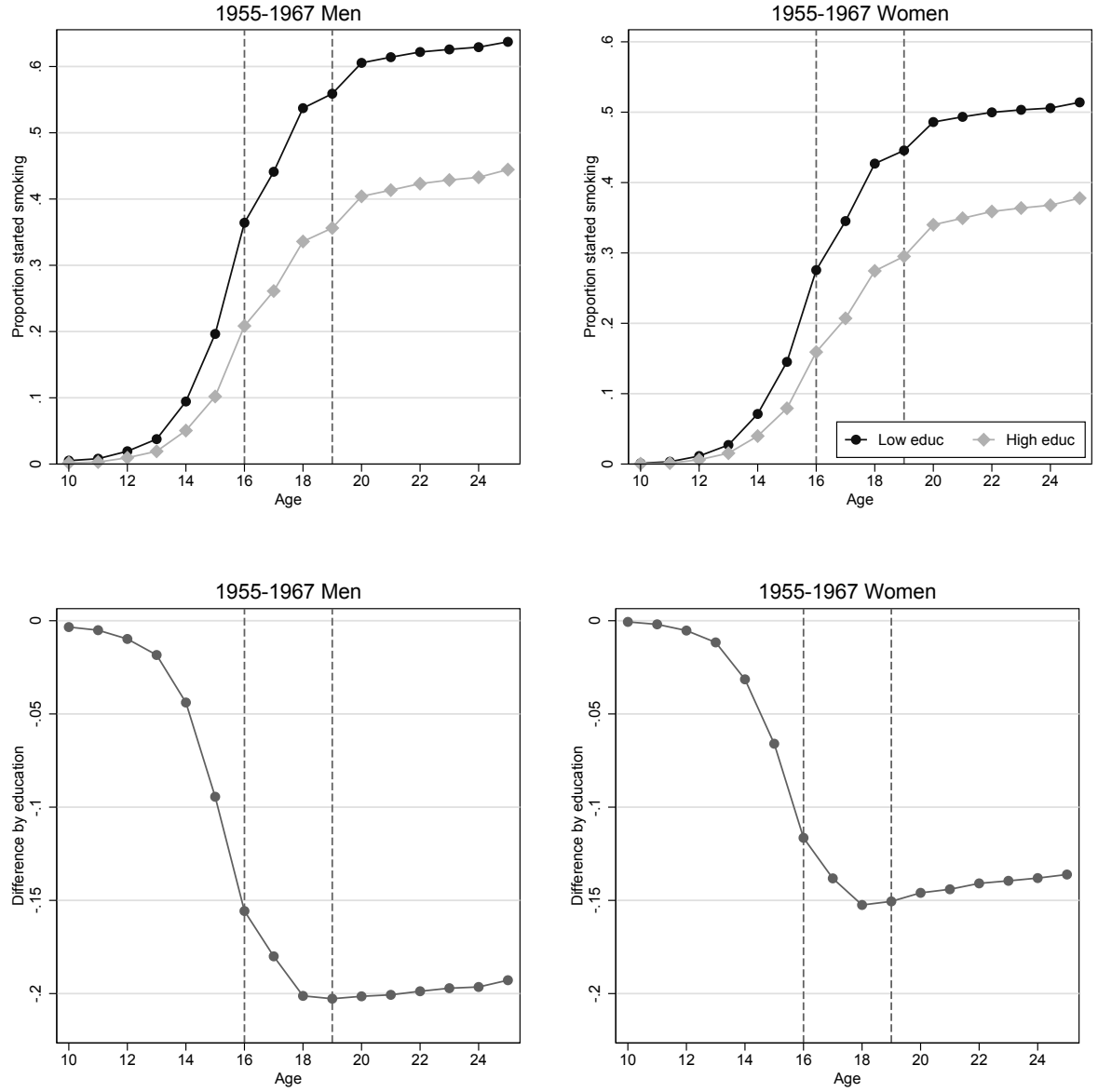
Note: Upper panel: each circle/diamond presents the age-specific log-hazard rate and the corresponding 95% confidence interval obtained from a discrete time event history model taking smoking initiation as failure event; Lower panel: presents the calculated hazard rate ratios and the corresponding 95% confidence interval; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure B.3: Educational differences in smoking initiation until a given age, cohort 1930–1954



Note: Figures in the upper panel show the education-specific distributions while the figures in the lower panel display its calculated differences (higher educated - lower educated); *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure B.4: Educational differences in smoking initiation until a given age, cohort 1955–1967



Note: Figures in the upper panel show the education-specific distributions while the figures in the lower panel display its calculated differences (higher educated - lower educated); *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Table B.1: Educational differences in current smoking at age 20, 30, 40 and 50 (percentage point differences estimated by OLS): Cohort 1955–1967. Education differences at age 20 are derived from retrospective information

	Men			Women		
Cohort observed at:	Age 30	Age 40	Age 50	Age 30	Age 40	Age 50
Panel A: High education [0;1]						
Smoked at age 20	-0.2136***	-0.1943***	-0.1603***	-0.1827***	-0.1441***	-0.1093***
	(0.0095)	(0.0054)	(0.0103)	(0.0103)	(0.0055)	(0.0072)
Smoked at age 30	-0.2001***			-0.1924***		
	(0.0094)			(0.0096)		
Smoked at age 40		-0.1966***			-0.1614***	
		(0.0051)			(0.0049)	
Smoked at age 50			-0.1606***			-0.1370***
			(0.0064)			(0.0063)
N	13,059	38,776	24,206	12,333	38,750	24,702
Panel B: College education [0;1]						
Smoked at age 20	-0.2279***	-0.2211***	-0.1862***	-0.2045***	-0.1582***	-0.1369***
	(0.0118)	(0.0062)	(0.0081)	(0.0134)	(0.0070)	(0.0089)
Smoked at age 30	-0.2356***			-0.2032***		
	(0.0112)			(0.0119)		
Smoked at age 40		-0.2161***			-0.1673***	
		(0.0056)			(0.0059)	
Smoked at age 50			-0.1753***			-0.1454***
			(0.0071)			(0.0075)
N	12,522	34,619	21,753	11,708	32,827	20,617
Panel C: Years of education						
Smoked at age 20	-0.0359***	-0.0332***	-0.0265***	-0.0321***	-0.0265***	-0.0202***
	(0.0014)	(0.0008)	(0.0010)	(0.0015)	(0.0008)	(0.0010)
Smoked at age 30	-0.0360***			-0.0355***		
	(0.0014)			(0.0014)		
Smoked at age 40		-0.0343***			-0.0317***	
		(0.0007)			(0.0008)	
Smoked at age 50			-0.0276***			-0.0256***
			(0.0009)			(0.0009)
N	13,059	38,776	24,206	12,333	38,750	24,702

Note: Control variables included: German nationality, fixed effects for states of residence, dummies for the survey year, age (age – the age of interest); Robust standard errors in parantheses; * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Table B.2: Educational differences in current smoking at age 20, 40 and 50 (percentage point differences estimated by OLS): Cohort 1930–1954. Education differences at age 20 are derived from retrospective information

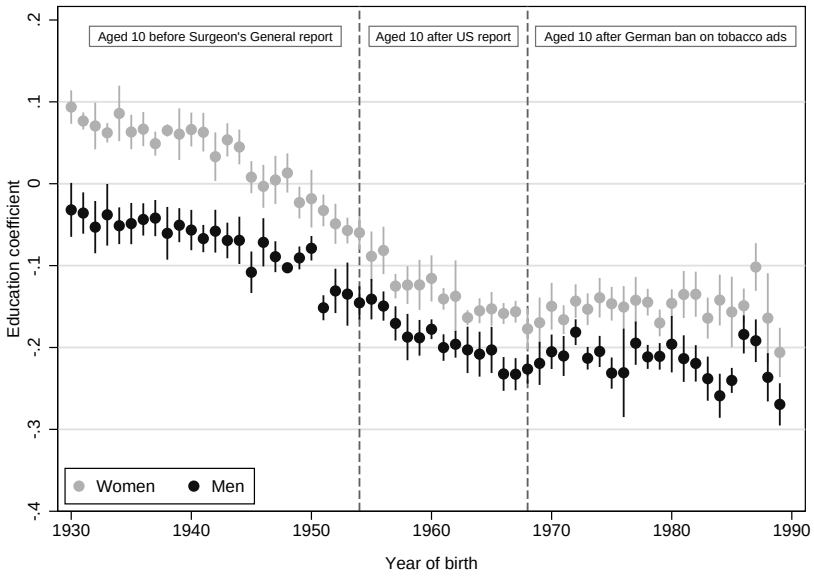
	Men		Women	
Cohort observed at:	Age 40	Age 50	Age 40	Age 50
Panel A: High education [0;1]				
Smoked at age 20	-0.1120***	-0.1275***	0.0134	-0.0266***
	(0.0130)	(0.0070)	(0.0165)	(0.0080)
Smoked at age 40	-0.1474***		-0.0900	
	(0.0125)		(0.0152)	
Smoked at age 50		-0.1339***		-0.0755***
		(0.0065)		(0.0071)
N	8,955	30,249	8,106	30,042
Panel B: College education [0;1]				
Smoked at age 20	-0.1341***	-0.1381***	-0.0009	-0.0320***
	(0.0143)	(0.0078)	(0.0190)	(0.0097)
Smoked at age 40	-0.1977***		-0.0913	
	(0.0133)		(0.0175)	
Smoked at age 50		-0.1395***		-0.0921***
		(0.0070)		(0.0083)
N	8,887	28,037	8,018	25,312
Panel C: Years of education				
Smoked at age 20	-0.0182***	-0.0191***	0.0021	-0.0009
	(0.0017)	(0.0009)	(0.0020)	(0.0010)
Smoked at age 40	-0.0248***		-0.0130	
	(0.0016)		(0.0019)	
Smoked at age 50		-0.0213***		-0.0124***
		(0.0009)		(0.0009)
N	8,955	30,249	8,106	30,042

Note: Control variables included: German nationality, fixed effects for states of residence, dummies for the survey year, age (age – the age of interest); Robust standard errors in parantheses; * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

C Alternative education measure: relative education

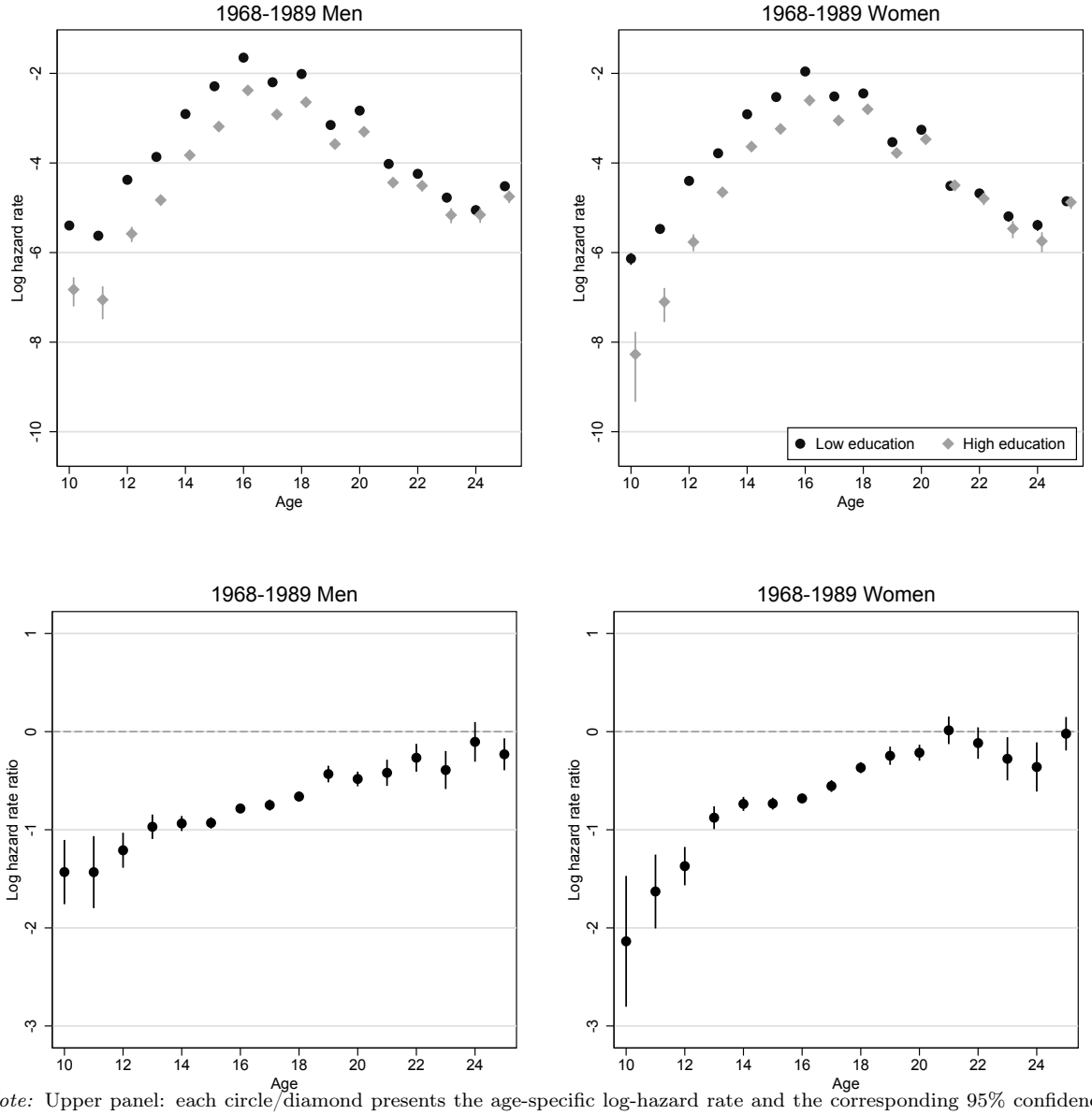
Figure C.1: Cohort trends in ever smoking by gender and relative education (being in the top quartile of one's cohort)

(a) OLS estimates: percentage point differences between high and low educated



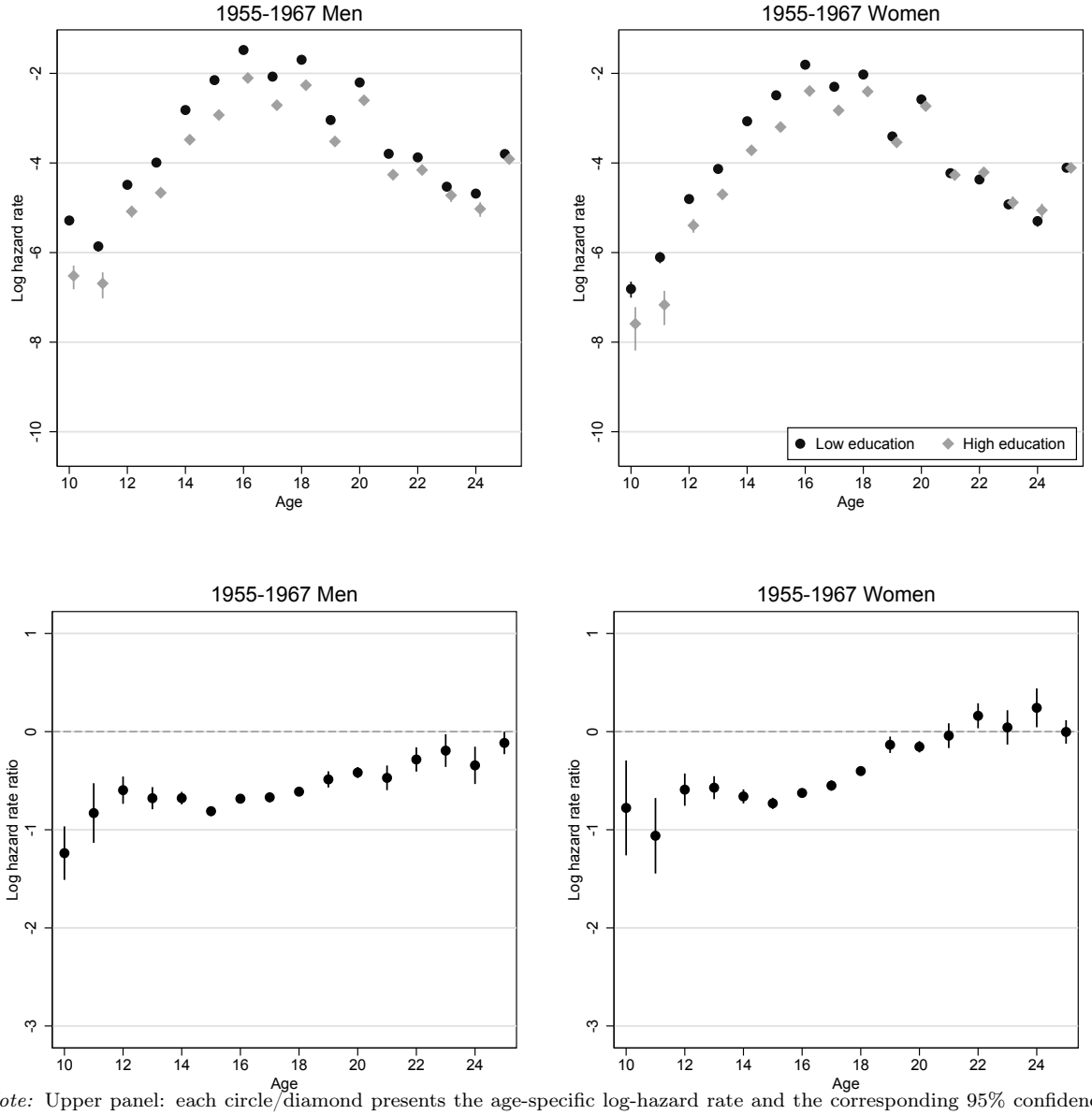
Note: Coefficients and the corresponding 95% confidence intervals were obtained from separate OLS regressions of ever smoking on high education (upper quartile) by year of birth; Control variables included: German nationality, fixed effects for states of residence, fourth order polynomial in age; Robust standard errors clustered at region*cohort level in parentheses; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure C.2: Hazard rates: Smoking initiation by completed relative education (being in the top quartile of one's cohort): Cohort 1968–1989



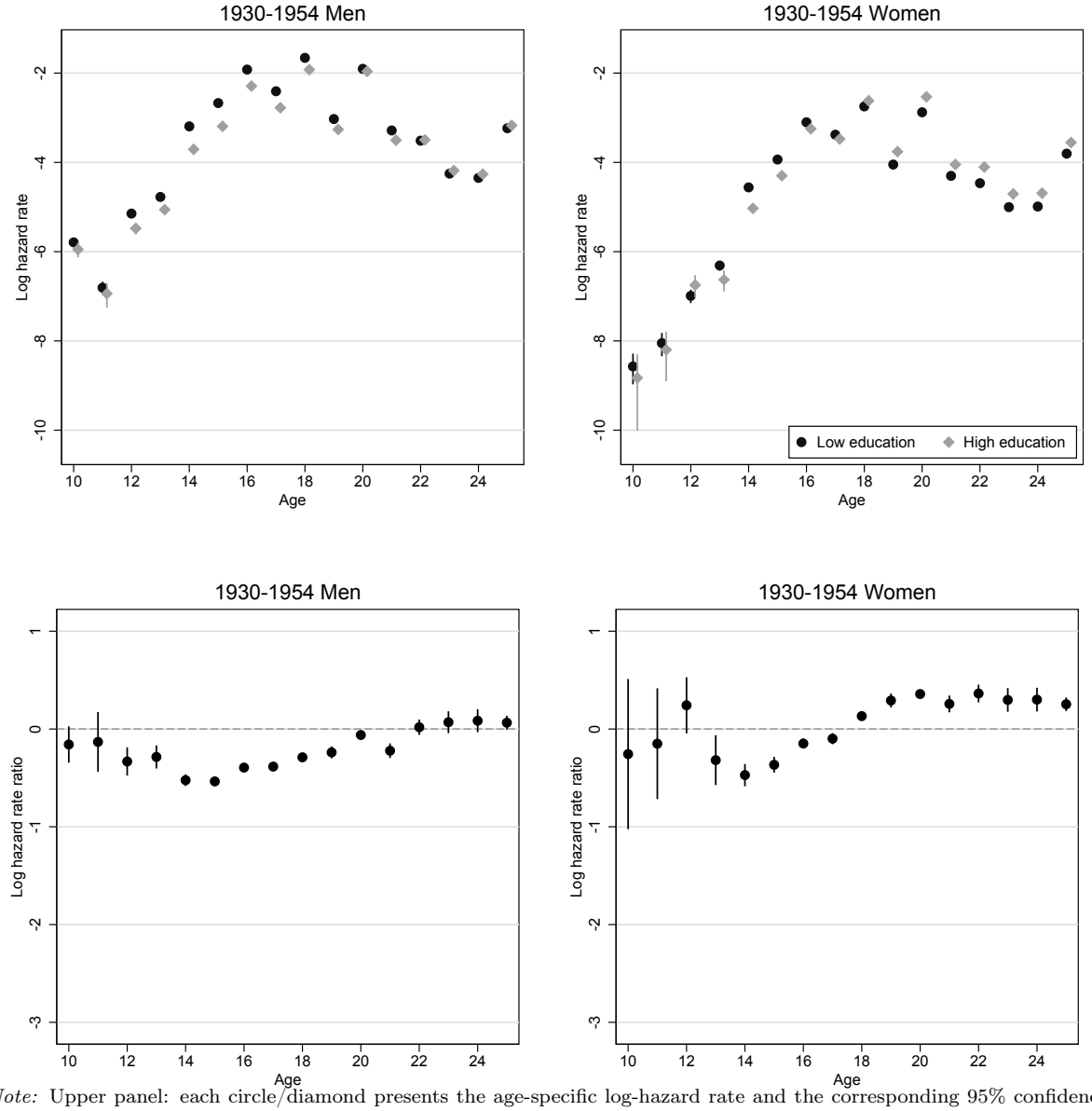
Note: Upper panel: each circle/diamond presents the age-specific log-hazard rate and the corresponding 95% confidence interval obtained from a discrete time event history model taking smoking initiation as failure event; Lower panel: presents the calculated hazard rate ratios and the corresponding 95% confidence interval; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure C.3: Hazard rates: Smoking initiation by completed relative education (being in the top quartile of one's cohort): Cohort 1955–1967



Note: Upper panel: each circle/diamond presents the age-specific log-hazard rate and the corresponding 95% confidence interval obtained from a discrete time event history model taking smoking initiation as failure event; Lower panel: presents the calculated hazard rate ratios and the corresponding 95% confidence interval; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

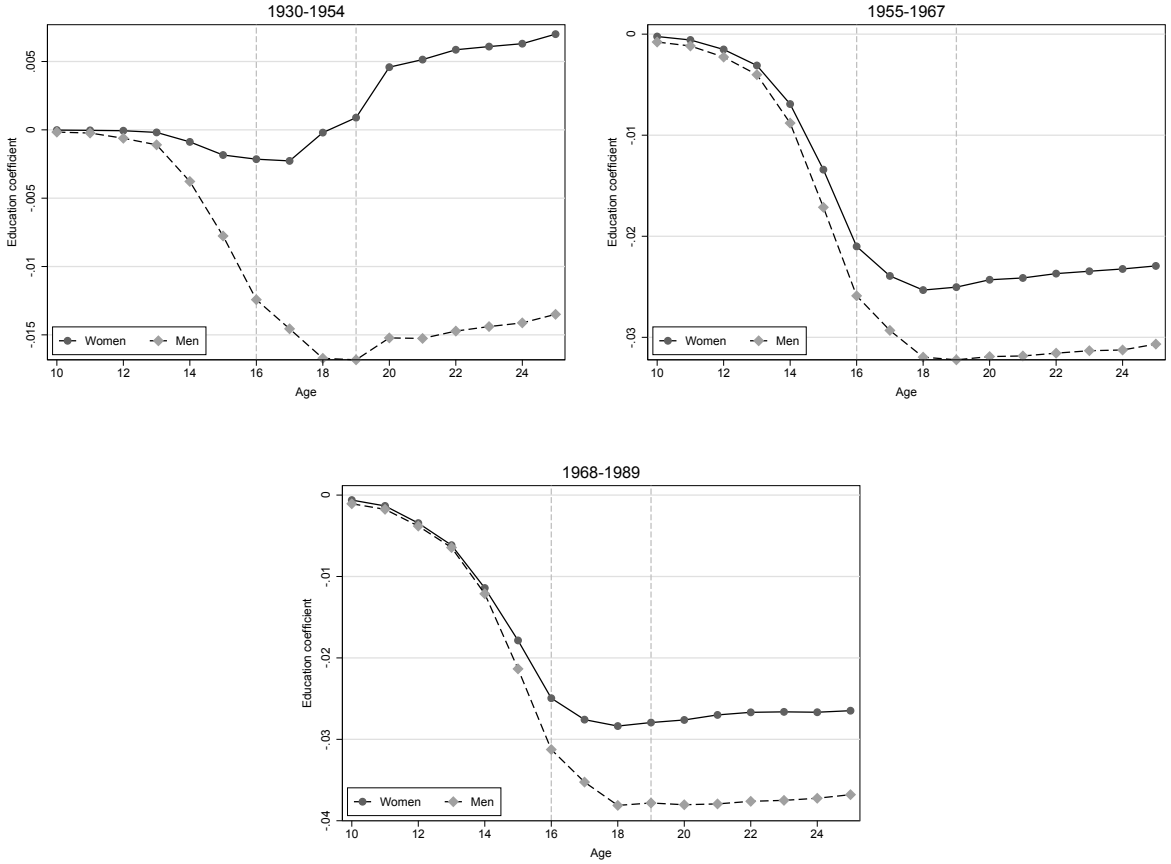
Figure C.4: Hazard rates: Smoking initiation by completed relative education (being in the top quartile of one's cohort): Cohort 1930–1954



Note: Upper panel: each circle/diamond presents the age-specific log-hazard rate and the corresponding 95% confidence interval obtained from a discrete time event history model taking smoking initiation as failure event; Lower panel: presents the calculated hazard rate ratios and the corresponding 95% confidence interval; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

D Alternative education measure: years of education

Figure D.1: Educational differences in smoking initiation until a given age: OLS coefficient on years of education

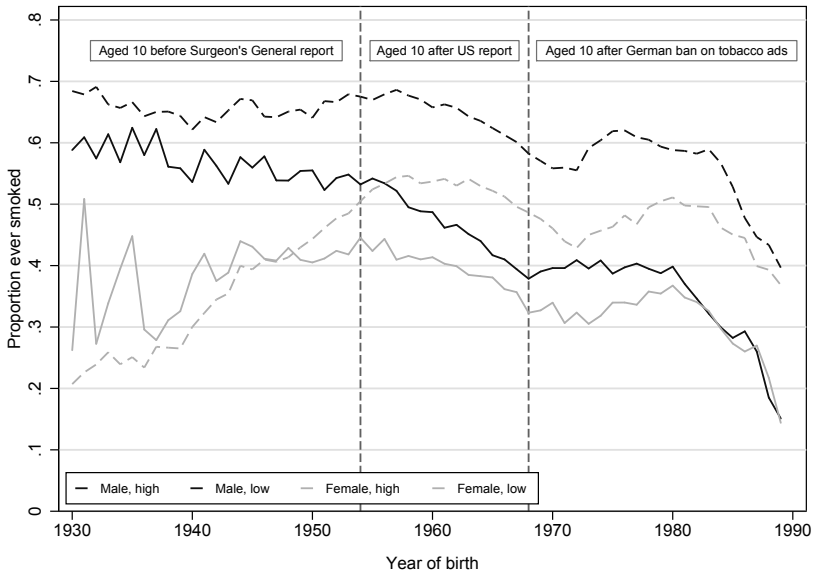


Note: Figures show the coefficient for years of education on smoking at a given age; Sample restricted to individuals aged 25 and over; mean years of education are 10.7 for women and 11.2 for men; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

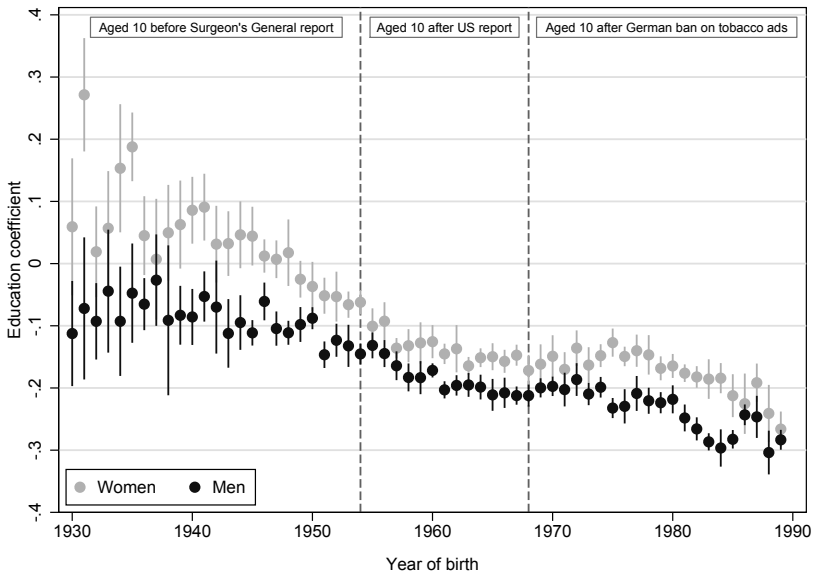
E Results excluding respondents aged 60 and older (to mitigate selective mortality)

Figure E.1: Cohort trends in ever smoking by gender and education, excluding respondents aged 60 and older

(a) Prevalences

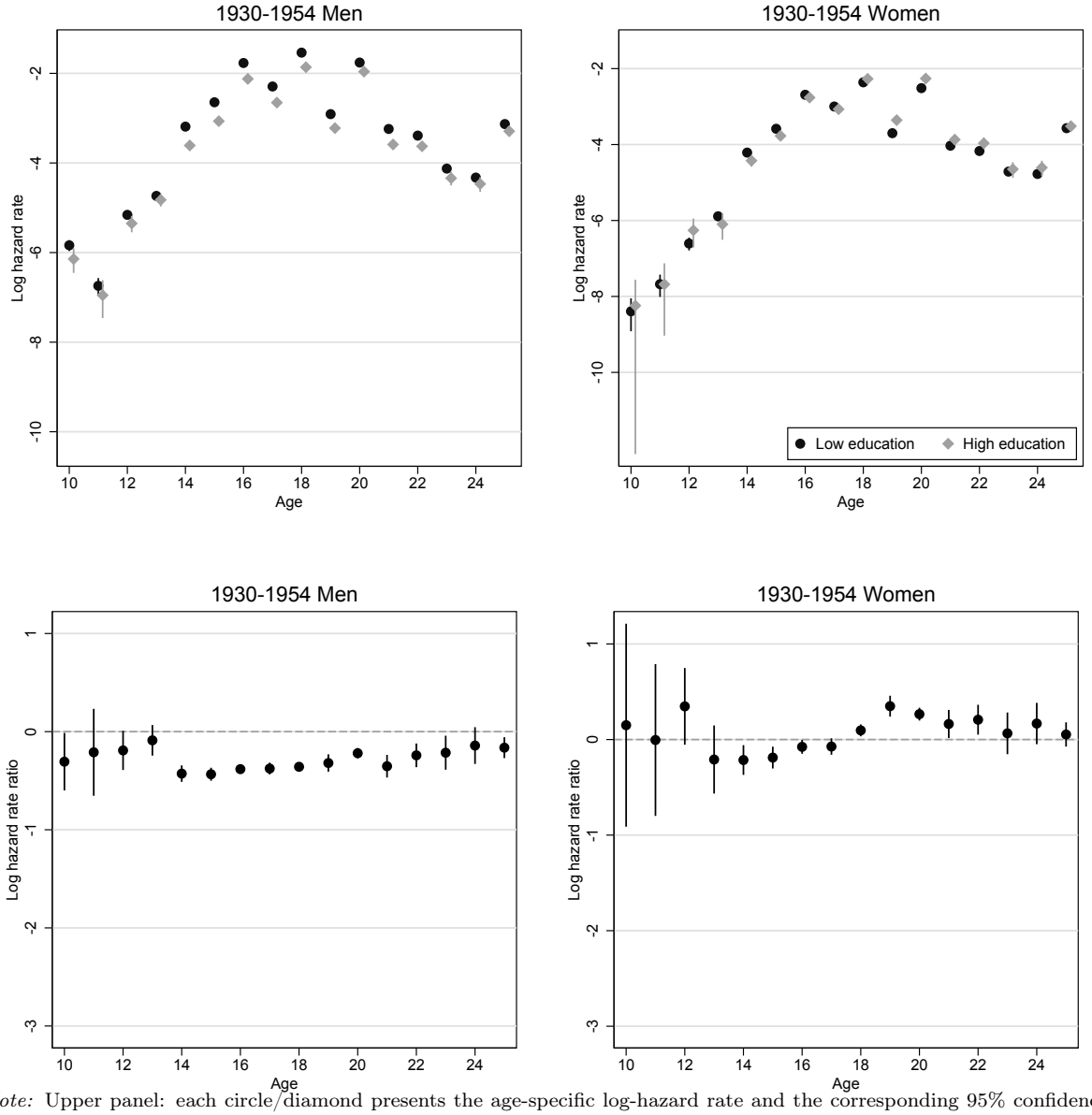


(b) OLS estimates



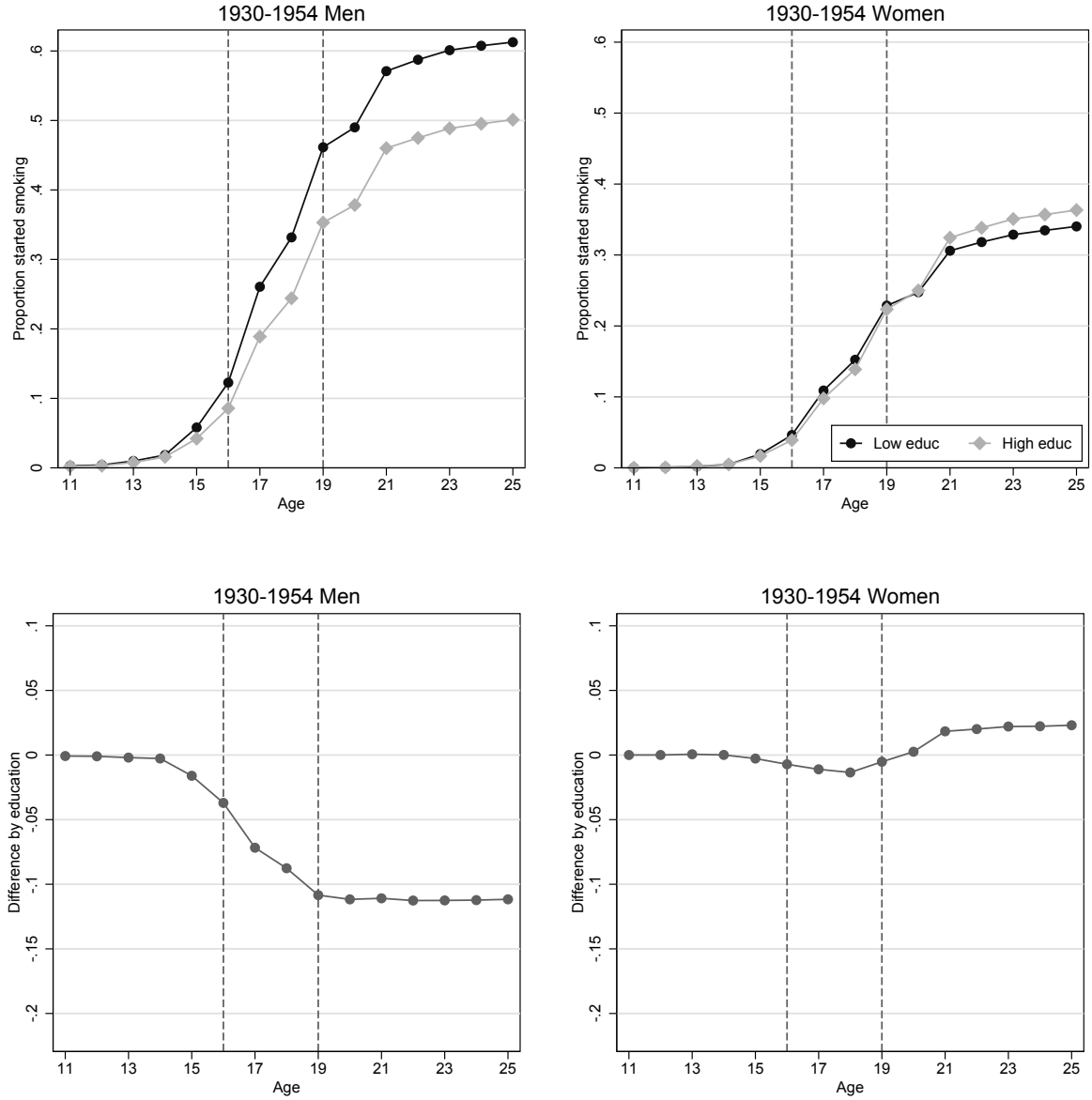
Note: Coefficients and the corresponding 95% confidence intervals were obtained from separate OLS regressions of ever smoking on high education by year of birth; Control variables included: German nationality, fixed effects for states of residence, fourth order polynomial in age; Robust standard errors clustered at region*cohort level in parentheses; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure E.2: Hazard rates: Smoking initiation by completed formal education: Cohort 1930–1954, excluding respondents aged 60 and older



Note: Upper panel: each circle/diamond presents the age-specific log-hazard rate and the corresponding 95% confidence interval obtained from a discrete time event history model taking smoking initiation as failure event; Lower panel: presents the calculated hazard rate ratios and the corresponding 95% confidence interval; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure E.3: Educational differences in smoking initiation until a given age: Cohort 1930–1954, excluding respondents aged 60 and older



Note: Figures in the upper panel show the education-specific distributions while the figures in the lower panel display its calculated differences (higher educated - lower educated); *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.