*Moiz Bhai**

The Earning Losses of Smokers

Abstract

Using within-family variation from twins and siblings, I find that smokers earn approximately 16% less than nonsmokers. Possible explanations for this earning difference are addiction-related productivity declines and earning reductions from higher health insurance costs. To investigate further, I use variation in the provision of employer-supplied health insurance (ESHI) to examine the mechanism of whether the addiction or insurance component has a larger influence on earnings. While I generally observe a larger earning penalty for smokers with ESHI than smokers without ESHI, the earning difference is statistically indistinguishable from zero.

Current version: January 22, 2020
Keywords: twins, compensating differentials, smoking, incidence of smoking, employer-supplied health insurance
JEL-codes: I12, I13, J31
Corresponding author: Moiz Bhai
mxbhai@ualr.edu

Department of Economics, University of Arkansas at Little Rock, 2801 S. University Ave, Little Rock, AR 72204.
Email: mxbhai@ualr.edu

© The Author(s). 2020 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.
Cite as: Bhai. *IZA Journal of Labor Economics* (2020) 9:1.
https://doi.org/10.2478/izajole-2020-0001

1 Introduction

Since the publication of the Surgeon General's report (U.S. Department of Health, Education, and Welfare, 1964) on the harmful effects of cigarette consumption, a vast literature has emerged exploring the economic consequences of smoking. Early analyses on the economic impacts of smoking anticipated that smoking would have adverse impacts in the labor market. Such studies predicted that poor health would adversely affect productivity and raise the cost of providing health care (Luce and Schweitzer, 1978; Oster et al., 1984), thereby lowering earnings. However, precisely quantifying the effect of smoking on earnings has proven to be challenging. Comparisons between smokers and nonsmokers often reveal significant differences on observable characteristics, which raises the possibility that significant differences may also exist on unobservable characteristics. As a result, credible evidence on the effects of smoking on earnings has been elusive[1].

To address the concerns from differential selection into smoking, studies investigating the effect of smoking on earnings have often invoked strong econometric assumptions. Many of these studies have used distinct approaches such as covariance restrictions (Auld, 2005), instrumental variables (Anger and Kvasnicka , 2010; Van Ours, 2004), longitudinal records (Grafova and Stafford, 2009; Levine et al., 1997), and more recently twin siblings (Lång and Nystedt, 2018) to overcome issues of selection. On the whole, these studies estimate the total effect of smoking on earnings, and they find that smoking reduces earnings between 8% and 24%. Despite the large earning penalty, it remains unclear whether the differences in earnings occur from the diminished productivity of smokers, their use of healthcare services, skill differences, or some combination of each explanation.

To address concerns of skill differences between smokers and nonsmokers on observed and unobserved characteristics, I exploit within-family variation in smoking from twin and singleton siblings using data from the first three waves of the National Survey of Midlife Development (MIDUS). Since individuals in families have similar backgrounds and face similar environments, differences on unobservable characteristics are smaller within families than between two random individuals. Thus, I contrast the within-family estimates with the more traditional approaches using a representative sample of Americans. Estimates from the representative sample of Americans in MIDUS show that smokers tend to earn approximately 16% less than nonsmokers. The estimates from the within-family models show a penalty of similar magnitudes of approximately 16% reduction in earnings for smokers. Overall, the estimates for the earning reduction for smokers from within-family models are statistically similar to the estimates from traditional models. At the minimum, these findings suggest that genetic differences do not influence the earning gap between smokers and nonsmokers.

Next, I disentangle the hypothesized mechanisms that contribute to the earning differences – namely, healthcare costs and addiction-related productivity declines – by exploiting the provision of employer-supplied health insurance (ESHI), which is offered at the firm level. I first compare the earnings of smokers with ESHI to the earnings of nonsmokers with ESHI. Analysis by ESHI status reveals that smokers with ESHI experience an economically and statistically significant reduction in earnings compared to nonsmokers. I compare the earnings

1 The study of Chaloupka and Warner (2000) contains a thorough selection of early studies evaluating the effects of smoking on wages.

of smokers with ESHI to those of nonsmokers without ESHI to examine whether healthcare costs or addiction-related productivity is the primary factor that influences the earning disparity. Since in the latter comparison, smokers bear the costs of their own healthcare, any earning differences should reflect addiction-related productivity losses. Although the earning effect for smokers without ESHI is large and negative compared to that for nonsmokers without ESHI, the difference is statistically indistinguishable from zero. Consequently, the empirical facts from this difference-in-differences (DiD) framework suggests that healthcare costs for smokers is a primary driver of their reduction in earnings.

The findings from this paper contribute to the economics of smoking by providing new estimates on the effect of smoking on earnings that potentially address issues of skill differences between smokers and nonsmokers. Furthermore, the findings have labor market implications, as the results suggest that firms adjust compensation on the full dimensions of worker quality, incorporating workers' health investments such as smoking. Legislation at both the national level, such as the Health Insurance Portability and Accountability Act, and at the state level, such as smoker protection laws, impose restrictions that prohibit differential insurance prices beyond a certain threshold for smokers. The prominence of smoking-related insurance legislation makes understanding the earning dynamics between health insurance and health investments such as smoking especially salient. Recent legislation such as the Affordable Care Act also imposes restrictions on the pricing of insurance to smokers and separates health insurance from employment for a segment of the population. The expansion of public insurance and the decoupling of insurance with employment raise important implications for the incidence of health behaviors such as smoking.

Additionally, this paper contributes to the discussion in public policy regarding the incidence of smoking. Scholars have often disagreed about the social costs of smoking (Manning et al., 1989; Chaloupka and Warner, 2000), as well as to what degree smokers impose negative externalities on the society. The results presented in this paper suggest that smokers do pay for at least some portion of their healthcare costs through reduced earnings. This finding is relevant because as of 2015 approximately 17% of Americans continue to smoke (CDC, 2015), and discussions on tobacco control policy remain prominent in the public sphere both in the US and internationally. The relationship between earnings and smoking might be more pertinent in the international context given the comparatively high rates of smoking in the developing world.

2 Conceptual framework

Following the framework on ESHI from Currie and Madrian (1999) and Gruber (2000), firms compensate workers based on their marginal product. Workers can then choose compensation offers where they trade off monetary compensation in exchange for fringe benefits such as health insurance. As shown in Equation 1, the total earnings denoted by E are based on the marginal product for individual worker i and contain monetary compensation denoted by W and nonmonetary compensation denoted by C. Health insurance forms a significant portion of the nonmonetary component. Because of various transaction costs and legislative restrictions on differential pricing of insurance, firms do not individually price health insurance,[2] and it

2 There is some variation in pricing based on smoking status and through differences in health plan choices.

is offered at the firm level. Therefore, firms are unable to completely adjust compensation for fringe benefits, but they can still adjust total compensation by reducing monetary earnings.

$$E_i = MPL_i = W_i - C_i \tag{1}$$

The economic framework presented in equation 1 provides a pathway denoting how addiction and healthcare costs can serve as mechanisms that explain the earning differences between smokers and nonsmokers. If addiction results in productivity decline, and healthcare costs are fixed, then a reduction in monetary earnings will occur to reflect diminished worker productivity. Even though firms may not directly observe individual smoking habits, they can observe other attributes of addiction such as smoking breaks, sick days, and nicotine withdrawal. Nicotine withdrawal occurs immediately upon cessation of smoking, which influences the physiological state of smokers. In addition to declines in productivity at work, smokers also miss more days of work (Lundborg, 2007; Halpern et al., 2007), which is observable to firms. The next pathway involves usage of health benefits. If healthcare costs differ between smokers and non-smokers for the same level of productivity, while health plans remain fixed at the firm level, then the differential healthcare costs can also be reflected in the reduced earnings of smokers.

Table 1 contains disaggregated estimates from Berman et al. (2014) on various addiction and health insurance costs of smoker. Table 1 provides bounds for the various economic costs of smoking. Addiction-related productivity declines include smoking breaks, missed days of work, and nicotine withdrawal. The health insurance component includes the disproportionate cost of providing health benefits for smokers. The bounds for health insurance and addiction-related costs overlap in terms of whether healthcare costs versus productivity have larger impacts.[3] The notable exception in Table 1 concerns pension costs, as smokers tend to have lower life expectancies than nonsmokers and thereby require pension payments for a shorter duration. Nevertheless, the pension savings from smoking are trivial with respect to health costs and addiction-related productivity declines.

Empirical evaluations to examine addiction versus health costs have been limited. A notable study by Cowan and Schwab (2011) extends the theory of compensating differentials to health behavior such as smoking and examines how the provision of ESHI influences earnings. They use pooled person-year data from the National Longitudinal Survey of Youth 1979 to use

Table 1 The annual costs of a smoker (in 2010 dollars)

	Best estimate	High range	Low range
Excess absenteeism	$517	$576	$179
Presenteeism	462	1,848	462
Smoking breaks	3,077	4,103	1,641
Excess healthcare costs	2,056	3,598	899
Pension benefits	–296	0	–296
Total costs	5,816	10,125	2,885

Source: Berman et al. (2014).

Notes: The table shows the differences in costs of employing a smoking employee versus a nonsmoking employee. "Presenteeism" refers to the costs arising from nicotine withdrawal.

3 The estimates in Table 1 should be interpreted with caution because they include some studies that do not fully address the issue of selection into smoking.

a DiD research design that exploits variation from within-smokers who switch between ESHI and non-ESHI employment. Their approach addresses concerns of unobserved skill differences since they exploit variation in ESHI, which tends to be offered at the firm level. Their research design only generates credibly estimates for smokers' health insurance costs and not the total costs of smoking on earnings. Consistent with the theory of compensating differentials, they find that smokers with ESHI employment tend to earn less.

3 The National Survey of Midlife Development

This study uses data from the National Survey of Midlife Development or MIDUS (Brim et al., 2011; Ryff et al., 2012; Ryff et al., 2013-2014). Data collection for MIDUS occurred in 1996, 2006, and 2014. MIDUS contains a nationally representative sample of Americans, a subsample of twins and singleton siblings, and a large oversample of urban underrepresented individuals. The MIDUS sampling for twins and siblings involves a "snowballing" component. During selection for the representative sample, interviewers inquired whether the respondent had a twin or a sibling, and then sought to contact this twin or sibling for inclusion in the family subsample. MIDUS surveyed all individuals from both the family and representative samples on questions about their health, employment, retrospective family background, human capital, and labor market participation. This allowed for the family sample to be compared to the nationally representative sample.

To construct my data, I transform the annual earning data, which are arranged on a categorical scale, to a continuous measure of earnings by using the midpoint values of the respective category.[4] For the lowest category of earnings, I use 1/3 the lowest value of earnings, and for the highest category of earnings, I use 3/2 the maximum value of earnings.[5] I then use the consumer price index from the Bureau of Economic Analysis to deflate earnings to constant 2006 dollars. I define individuals as having ESHI if they responded that their insurance is provided through their employer. I use responses on past smoking history status to construct variables for former smokers and response on current smoking status to identify current smokers. For education, I use indicator variables for high school completion, college attendance, and college completion. Lastly, I use labor market participation during the survey years to construct my labor supply outcomes.

Table 2 contains the descriptive statistics for the national, twin, singleton sibling, and the full sample. The average age of study participants is approximately 48 years, and one-fifth of the sample currently smokes, with nearly half of the sample having smoked at one point in their lives. Notable earning differences are evident for the sibling sample, which earns more than the full sample, which are also statistically significant at the 1% significance level. Besides earnings, the sibling and twin sample appear similar to the full sample on most characteristics. This outcome is consistent with MIDUS' random sampling approach in the search for twins and siblings. Nevertheless, to assess the representative nature of the family sample, I test for the similarity between the family sample and the representative sample in Table 3

4 A strength of MIDUS is that the annual earnings categories are spaced unevenly with more categories for lower income levels, thereby improving precision. There are 34 bins for income categories, with bins at the lower points of the income distribution spaced in increments of $1,000, with the highest increment bins separated by $5,000.

5 Alternative approaches to top coding and bottom coding of earnings do not influence the results in a meaningful way.

Table 2 Descriptive statistics for The National Survey of Midlife Development

	Random	Sibling	Twin	All
Age	45.58 (10.53)	47.33 (9.60)	45.23 (10.01)	45.67 (10.30)
Female	0.50 (0.50)	0.52 (0.50)	0.53 (0.50)	0.51 (0.50)
Employer insurance (ESHI)	0.58 (0.49)	0.57 (0.50)	0.582 (0.49)	0.583 (0.49)
Schooling	14.244 (2.48)	14.717 (2.37)	14.1 (2.40)	14.351 (2.46)
High school	0.94 (0.23)	0.97 (0.16)	0.94 (0.24)	0.95 (0.22)
Some college	0.67 (0.47)	0.76 (0.43)	0.65 (0.48)	0.69 (0.46)
College graduate	0.41 (0.49)	0.48 (0.50)	0.39 (0.49)	0.43 (0.50)
Non-white	0.13 (0.33)	0.05 (0.22)	0.07 (0.26)	0.10 (0.30)
Earnings	51,830 (43,073)	59,041 (46,269)	52,508 (42,374)	54,510 (44,322)
Log earnings	10.44 (1.08)	10.58 (1.11)	10.47 (1.08)	10.50 (1.08)
Smoke	0.22 (0.41)	0.19 (0.39)	0.21 (0.41)	0.21 (0.41)
Ever smoke	0.52 (0.50)	0.45 (0.50)	0.46 (0.50)	0.49 (0.50)
n	5,615	5,681	4,078	11,306

Source: National Survey of Midlife Development (1996, 2006, and 2014).

Notes: Standard deviations are under the mean values of the variables. Sample size represents person-year observations.

Abbreviation: ESHI, employer-supplied health insurance.

Table 3 Testing for sample selection: singletons vs. twins

Question	Singletons	Twins	Difference	*P*-value
Mother's education				
Has less than high school	0.360	0.343	−0.017	0.29
Graduated high school	0.402	0.403	0.009	0.59
Attended some college	0.129	0.130	0.009	0.45
College graduate	0.108	0.108	0.0002	0.98
Schooling (years)	11.20	11.44	0.242**	0.03
Father's education				
Has less than high school	0.409	0.408	0.007	0.68
Graduated high school	0.325	0.293	−0.044**	0.04
Attended some college	0.089	0.100	−0.010	0.28
College graduate	0.177	0.191	0.014	0.33
Schooling (years)	11.04	11.11	−0.076	0.60
n	5,615	4,078		

Source: National Survey of Midlife Development (1996, 2006, and 2014).

Notes: Sample size is in person-years. *P*-values are from Two sample *t*-test for equality of mean values between singletons and twins. Statistical significance is denoted by the following: $^{**}P < 0.05$.

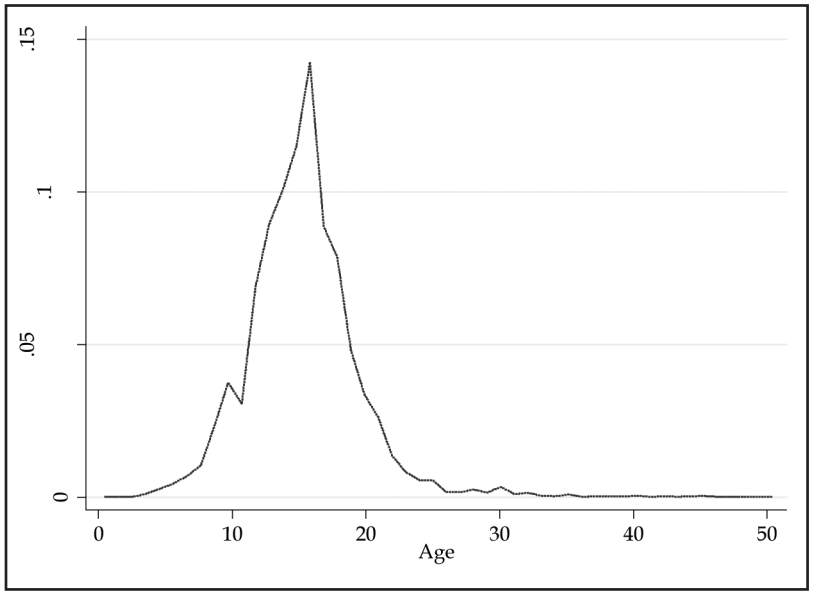
by comparing family background variables. The twin sample is statistically similar to the representative sample on all variables except for father's high school completion.

Next, I examine smoking history. Figure 1 shows the distribution of ages at smoking initiation. Most individuals begin smoking in their teen years with a small portion of individuals initiating as preteens and a small portion initiating in early adulthood. The age of initiation is

important because it shows that individuals start smoking when they are young, possibly suggesting that within-family estimates can be useful since there are more commonalities within-families when twins and siblings are young.

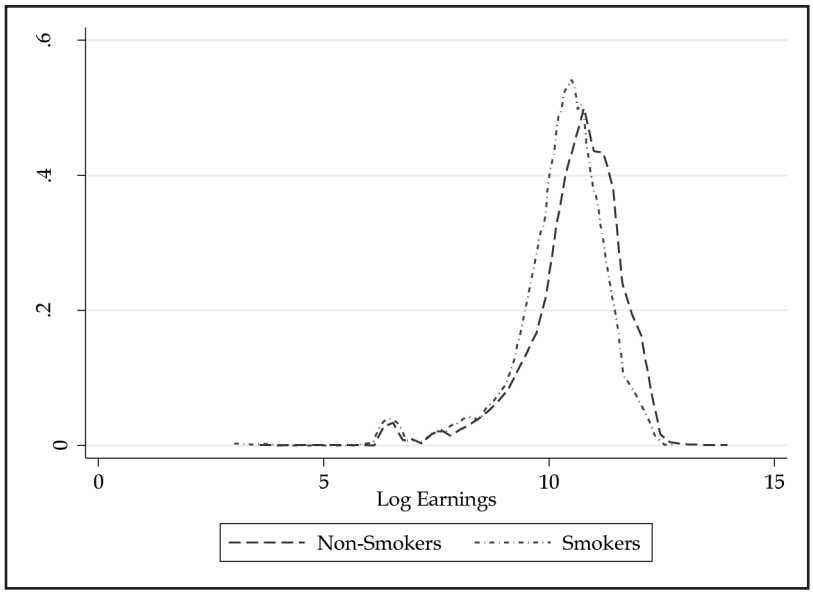
In Figure 2, I show the distribution of log earnings between smokers and nonsmokers. Figure 2 reveals that smokers tend to earn less than nonsmokers. I test this difference using a Kolmogorov–Smirnov test for equality of distributions and find that the distributional differences are statistically significant at the 1% significance level. To evaluate whether demographic characteristics could possibly influence this difference, I compare whether smokers

Figure 1 When do smokers initiate?



Source: The National Survey of Midlife Development (1996).

Figure 2 Log earnings of smokers versus nonsmokers.



Source: The National Survey of Midlife Development (1996, 2006, and 2014).

and nonsmokers are different on demographic and ability variables in Table 4. Notable and statistically significant differences are detected for age and gender, as smokers tend to be younger than nonsmokers and are more likely to be male. Besides the demographic differences, smokers show a 1-year difference in completed schooling and tend to be thinner than non-smokers as measured by the body mass index (BMI).

4 Estimation strategy

4.1 The full sample

Equation 2 provides a detailed framework for the econometric relationship between smoking and earnings. Ideally, the outcome variable log earnings (in real 2006 dollars) denoted by E for individual i at time t is regressed upon smoking status denoted by S and demographic characteristics X, which includes age, race, and gender, along with characteristics such as ability, tastes, and preferences denoted by Z. Both demographic and nondemographic characteristics influence earnings and the decision to smoke. Attempting to control for these attributes, especially for tastes and preferences, is often challenging. Consequently, estimating the effect of smoking on earnings poses empirical challenges: if smoking is correlated with unobserved attributes such as productivity that positively affect earnings, then the failure to account for these omitted variables biases the smoking coefficient in typical cross-section models. The direction of the bias is unclear and depends on which unobserved mechanisms have bigger impacts. For example, the omission of ability from the regression estimation would produce a downward bias because it is positively correlated with earnings and negatively correlated with smoking. The omission of other attributes such as tastes and preferences that could be positively correlated with smoking and earnings would produce an upward bias on the effect of smoking.

Table 4 Are smokers different than nonsmokers?

	Non-smoker		Smoker		Difference	*P*-value
	Mean	STD	Mean	STD		
Age	48.11	12.10	45.24	10.91	2.87	<0.01***
Female	0.50	0.50	0.51	0.50	0.01	0.72
Employer insurance	0.57	0.50	0.53	0.50	0.04	<0.01***
Schooling	14.60	2.46	13.19	2.24	1.41	<0.01***
High school	0.96	0.20	0.89	0.32	0.07	<0.01***
Some college	0.72	0.45	0.52	0.50	0.20	<0.01***
College graduate	0.48	0.50	0.21	0.41	0.27	<0.01***
Non-white	0.10	0.30	0.10	0.31	0.00	0.98
Earnings	55,221	45,972	41,744	35,082	13,477	<0.01***
Log earnings	10.47	1.14	10.24	1.05	0.23	<0.01***
Body mass index	30.38	16.54	28.99	16.65	1.39	0.01**
n	8,998		2,308			

Source: National Survey of Midlife Development (1996, 2006, and 2014).

Notes: The columns containing STDs are next to the mean values. *P*-values are from Two sample *t*-test for equality of means between smokers and nonsmokers. Statistical significance is denoted by the following: $^{**}P < 0.05$ and $^{***}P < 0.01$.

Abbreviation: STD, standard deviation.

$$E_{it} = \alpha + \tau_t + \beta S_{it} + \gamma X_{it} + \pi Z_{it} + \varepsilon_{it} \quad (2)$$

To overcome the empirical challenges, I begin with Equation 3 as the baseline specification. First, I pool data across all three waves of the survey and I regress earnings on smoking status and demographic characteristics. The key parameter of interest b is a binary indicator for whether the respondent smokes. Essentially, this regression compares the earnings of smokers to nonsmokers conditional on covariates. Alternatively, I reestimate the model given in equation 3 for former smokers. By using former-smokers, I compare the earnings of former-smokers to never-smokers. In all specifications, I include survey year fixed effects and a vector of *X* that includes covariates for age, age squared, gender, and whether the respondent is non-white. In additional specifications, I control for differences in education.[6]

The use of log earnings parametrizes the coefficient on smoking or the smoking penalty as a percent of earnings. For all regressions involving log earnings as a dependent variable with a binary variable as a regressor, I use the approach of Kennedy (1981) to approximate the effects of smoking on earnings. I cluster standard errors at the individual level to reflect repeated person-level observations and to address heteroscedasticity and autocorrelation over time.

$$E_{it} = \alpha + \tau_t + \beta S_{it} + \gamma X_{it} + \varepsilon_{it} \quad (3)$$

Nevertheless, equation 3 does not include characteristics such as ability, preferences, and tastes, which might jointly determine smoking status and earnings. To address concerns of omitted variables, I exploit the longitudinal component of MIDUS to estimate an individual fixed-effects model given by equation 4. Estimates from the fixed-effects model compare within-person changes of smoking status and its effect on earnings. They address the concern of unobserved time-invariant characteristics that are omitted such as tastes and preferences. While the estimates from the individual fixed-effects model are causal under an assumption of time-invariant unobservables, they capture the local average treatment effect (LATE) of quitters because during this study's time period, most of the smokers are quitting and not initiating as would be expected based on smoking patterns in midlife. The LATE for quitters is an interesting parameter but does not capture the effects of smoking on earnings. Since a non-negligible number of individuals report transitioning from nonsmokers to smokers, I reestimate the fixed-effects model comparing former-smokers to never-smokers as an alternative measure of the effect of quitting smoking on earnings.

$$E_{it} = \alpha_i + \tau_t + \beta S_{it} + \gamma X_{it} + \varepsilon_{it} \quad (4)$$

4.2 **Sibling by year identification strategy**

The sibling and twin analysis is motivated by the fact that individuals within families share similar genetics, environments, and possibly the same factors that influence smoking decisions. Because individuals raised in the same family share similar environments and genetics that influence the production of hard skills like ability and soft skills such as personality, the use of within-family models would mitigate the bias from traditional comparisons of smokers to nonsmokers. This research design is useful especially since a recent work on smoking initiation

6 In Table A1, I show that smoking does not influence labor supply, thereby excluding one possible pathway through which smoking might influence labor market outcomes.

suggests that peers might play a role (Nikaj, 2017). More specifically, the use of within-family variation addresses these concerns of omitted variable bias because unobservable attributes are smaller within families than outside of families. Thus, a family fixed-effects analysis can produce improved estimates over traditional cross-section Ordinary Least Squares (OLS) (Card, 2001). The correlates on smoking behavior, namely, why people initiate and continue to smoke, are factors ranging from family attitudes, peer pressure, sociodemographic factors, personality/social skills, stress, and availability (Center for Substance Abuse, 1997). Since many of these causal factors that influence smoking decisions occur when individuals are young, this provides auxiliary evidence that within-family estimates might ameliorate the bias in estimating the effect of smoking on earnings.

MIDUS contains both monozygotic and dizygotic twins, and I pool both types to form the twin sample.[7] Monozygotic (identical) twins share the same genetic makeup, whereas dizygotic (fraternal) twins share half the genetic makeup and are siblings born at the same time. Analysis of the twin siblings imposes stricter restrictions compared to singleton siblings because twins are more similar than non-twin siblings, as they are raised together and share a greater proportion of genetic material. More specifically, the twin design reduces threats of potential confounders from the sibling design, such as differences in spacing, birth order, and family size, in addition to differential parental endowments because twins have parents who are of the same age.

In the family design, I identify the effect of smoking by comparing the earnings of a sibling (or twin) smoker to the earnings of a sibling (or twin) nonsmoker. Equation 5 shows the baseline specification for the within-family models. Again, I examine how earnings for individuals are influenced by smoking status. Here, I replace the individual fixed effects with twin-by-year, sibling-by-year, or family fixed effects where *f* indexes family. In subsequent specifications, I introduce controls for schooling to examine how schooling differences influence the effect of smoking on earnings.

$$E_{ift} = \alpha_f + \tau_t + \beta S_{ift} + \gamma X_{ift} + \varepsilon_{ift} \tag{5}$$

Using variation from within-family models has some of its own limitations. The first concern arises with the disparate smoking decisions within twin and sibling sets. If the same unobserved factors that induce people to smoke in the traditional cross-sectional models also cause disparate smoking decisions within siblings and twins, then estimates from within-family models suffer from the same bias. If unobservable attributes are smaller within families, a plausible assumption because individuals within families tend to be more similar than individuals between families, then the within-family estimates are still useful because they provide a bound on the upward bias compared to the traditional estimates on the effect of smoking on earnings. Second, the within-family estimates are likely to exacerbate measurement error in response on smoking status, and this would introduce a downward bias and attenuate the coefficient on the smoking if the measurement error is classical.

Other concern deals with generalizability of results from twin-based analyses to the general population. Since twinning rates vary with maternal age and different ethnicities, this

7 Table A2 contains the effect of smoking on earnings by zygosity. The estimates for the effects of smoking are statistically similar across zygosity, which implies that the genetic influences on the effect of smoking on earnings are small. In fact, the results are very much in line with Lång and Nystedt (2018), who used Swedish twins in finding negative effects of smoking on earnings, but the estimates for smoking are statistically similar across zygosity.

might mean that twins might be different from singletons. I conservatively interpret estimates from the twin by year fixed-effects models as LATEs, because twin-sibling dynamics might differ from singleton siblings. For estimates from the sibling by year fixed-effects models, I interpret them as a form of average treatment effect (ATE) that could be generalizable to the representative population.

Finally, to attempt to evaluate the mechanisms that influence the earning differences between smokers and nonsmokers, I estimate DiD models given by equation 6 for the family samples of twins and siblings.

$$E_{ift} = \alpha_f + \tau_t + \theta(S^* \text{ESHI})_{ift} + \beta S_{ift} + \sigma \text{ESHI}_{ift} + \gamma X_{ift} + \varepsilon_{ift} \quad (6)$$

I use variation from differences in smoking uptake in sibling sets (first difference) and the difference between ESHI in sibling sets (second difference). The interaction between smoking status with ESHI is the coefficient of interest denoted by θ, which is the DiD estimator. Estimates from equation 6 provide bounds on how much health insurance-related costs adversely affect earnings for smokers.

5 Results

5.1 Main results

Table 5 contains results for the full sample. In the first column of Panel A, I compare the earnings of smokers to nonsmokers without including schooling, which serves as a proxy for ability. I find that smoking produces an economically and statistically significant reduction in earnings of approximately 24%. Controlling for schooling in the second column reduces the earning penalty to less than 17%, a relative decline of 25%. The large sensitivity of the smoking coefficient to the schooling control suggests that smokers and nonsmokers differ on latent ability. Next, in the third and fourth columns, I introduce individual fixed effects to examine

Table 5 The earnings impact of smoking for the full sample

	Full sample			
	1	2	3	4
Panel A				
Smoker				
n = 8,975	−0.235*** (0.022)	−0.168*** (0.024)	0.063 (0.089)	0.064 (0.090)
Panel B				
Former-smoker				
n = 7,280	−0.061** (0.025)	−0.019 (0.025)	−0.041 (0.072)	−0.041 (0.072)
Covariates				
Education	Yes	No	Yes	Yes
Individual	No	Yes	No	Yes

Source: National Survey of Midlife Development (1996, 2006, and 2014).

Notes: Huber–White clustered standard errors are in the parentheses. All individuals are between 25 and 66 years, and all regressions include controls for race, gender, and age. Panel A compares the earnings of smokers to nonsmokers, whereas Panel B compares ever smokers or former-smokers to never-smokers. Statistical significance denoted by the following: **P < 0.05 and ***P < 0.01.

the influence of (changing) smoking status on earnings. Both columns reveal a statistically indistinguishable effect. The individual fixed-effects models capture the effect of quitting on earnings because smoking initiation rates in adulthood are very low but quit rates are high. Since a small portion of the sample switches smoking status between three waves and with the reductions in sample size for later waves, it is likely that the study is underpowered to detect the within-person effects of changing smoking status on earnings.

In Panel B of Table 5, for former-smokers, I compare the earnings of former-smokers to never-smokers. In the first column, without controls for schooling, former-smokers earn approximately 6% less than individuals who are never-smokers. Upon inclusion of controls for schooling, the earnings of former-smokers are statistically indistinguishable from the earnings of never-smokers. This indicates that the reduction in earnings reflects differential selection as opposed to the long-lasting effects of smoking. The pattern is consistent with the general results on the earnings of former-smokers. Approximately 2% of the sample returns to smoking regularly between the survey waves. The comparison between former-smokers and never-smokers measures the effect of quitting on earnings. Similar to the earlier specification that uses individual fixed effects, the findings show no effect of being a former-smoker on earnings. Next, in the analyses comparing former-smokers to non-smokers, I again find statistically indistinguishable effects in the individual fixed-effects specifications.

In Table 6, I present within-family estimates from siblings and twins. Beginning with column 1 of Panel A in the pooled OLS sibling sample, I find large reductions in earnings of sibling smokers of approximately 25%. By including controls for schooling, the reduction in earnings declines to approximately 25%. In the third and fourth columns, I present estimates from more rigorous models that include sibling by year fixed effects that likely minimize omitted variable biases compared to the pooled regressions. In the third column, I find smokers earn approximately 16% less, and controlling for schooling in the fourth column does not affect

Table 6 The earnings impact of smoking for the family sample

	Family level			
	1	**2**	**3**	**4**
Panel A: Siblings (n = 4,080)				
Smoker	−0.270*** (0.033)	−0.199*** (0.037)	−0.165** (0.072)	−0.156** (0.073)
Former-smoker	−0.023 (0.038)	−0.024 (0.038)	−0.079 (0.101)	−0.062 (0.103)
Panel B: Twins (n = 2,774)				
Smoker	−0.287*** (0.038)	−0.182*** (0.045)	−0.195** (0.078)	−0.163** (0.080)
Former-smoker	0.014 (0.046)	0.045 (0.047)	−0.120 (0.107)	−0.112 (0.112)
Covariates				
Education	No	Yes	No	Yes
Family	No	No	Yes	Yes

Source: National Survey of Midlife Development (1996, 2006, and 2014).

Notes: Huber–White clustered standard errors are given in parentheses. All individuals are between 25 and 66 years, and all regressions include controls for race, gender, and age. The first two columns compare across siblings/twins, and the last two columns measure smoking within siblings/twin sets. The coefficient on smoke compares the earnings of smokers to nonsmokers, whereas the coefficient on former-smoker compares ever-smokers or former-smokers to never-smokers. Statistical significance is denoted by the following: **$P < 0.05$ and ***$P < 0.01$.

the size or significance of the coefficient, and I continue to find a 16% reduction. The robustness of the effect size after controlling for schooling suggests that family fixed effects appear to adequately handle differences in the coefficient arising from measures of observed ability. Overall, in all sibling models with and without family fixed effects, the effect of smoking remains large, statistically significant, and negative, and former-smokers do not have statistical differences in their earnings from never-smokers.

In Panel B of Table 6, I examine the effect of smoking on earnings both across and within twin pairs. Beginning with the first column that contains pooled OLS for twins, I observe that the effect of smoking has a large negative effect on earnings of approximately 28%, which is significant at the conventional levels of significance. The second column includes controls for schooling, which reduces the coefficients in a substantial and significant manner to approximately 18%. The results mirror the findings seen in the representative cross-section and sibling samples. Estimates from models with twin by year fixed effects in the third column show a significant reduction in earnings of 19%. Like estimates from the sibling models controlling for schooling does not influence the magnitude of the earnings reduction as it remains consistent around 18%. Lastly, in both sibling and twin samples, being a former-smoker has a statistically indistinguishable effect on earnings.

On the whole, estimates from the family fixed-effects models are consistent with those from Lång and Nystedt (2018) who use a twin by year fixed-effects research design with data from the Swedish Twin Registry to examine the effect of smoking on earnings. They estimated separate regressions by zygosity and found the effect of smoking on earnings to be negative, but the confidence intervals for the effect of smoking tended to be large. Thus, while they found the effect of smoking to be negative, zygosity did not appear to play a big role as estimates from separate regressions were statistically similar to each other.

The main takeaway from Tables 5 and 6 is that the earnings reduction for smokers is large and statistically significant. It theoretically consists of a multitude of factors (Table 1) with healthcare costs and addiction-related productivity declines being the most important. Another contributing explanation for the large earning disparity is that respondents in MIDUS are middle aged and therefore have reached the flat portion of the earnings profile. The secondary takeaway pertains to the econometric analysis. The family fixed-effects models appear robust to the controls for schooling and might be successful in controlling for other unobserved variables such as character skills, tastes, and preferences. Consistent with the research design, the point estimates on the effect of smoking on earnings from the family models are smaller vis-à-vis the full sample. Yet, the confidence intervals overlap between the point estimates from the family sample and the full sample, which suggests that genetic factors do not have a large influence on the earnings reduction of smokers. Furthermore, the earning penalty appears to dissipate for former-smokers both in the full and family samples.

5.2 **The role of ESHI on the earnings of smokers**

Separating the effect of higher health expenses from productivity changes poses a challenge beyond addressing the selection problem without imposing additional assumptions. To differentiate between the addiction-related productivity declines versus higher health insurance explanations, I separate the sample into individuals with ESHI versus individuals without

ESHI. Estimates from these models separate the effect of ESHI on earnings because both smokers with ESHI and without ESHI should be afflicted with addiction-related productivity declines under an assumption that firms do not discriminate in hiring smokers. Support for plausibility of the anti-discrimination assumption is based on the fact that insurance is offered at the firm level and not at the individual worker level (similar to Cowan and Schwab, 2011). Furthermore, conditional comparisons between former-smokers and never-smokers provide auxiliary evidence that current health costs and productivity, rather than past health costs or addiction, influence the differences in earnings. Since twins and siblings abate selection concerns, the estimates from Table 7, which compares across ESHI status, crudely function as a test of whether worker productivity or health costs cause the reduction in earnings of smokers.

For individuals without ESHI, I present findings in Panel A of Table 7. In all specifications for the full sample and within siblings and twins, I find the effect of smoking on earnings to be largely negative but imprecisely estimated. It is still possible that smoking affects productivity through smoking breaks and time off in terms of illness and sick days but that these negative effects of smoking are small and are undetected at the conventional levels of significance. Two notable difficulties arise in this analysis for heterogeneity. First, this regression analysis by ESHI suffers from the problem of limited sample size. Second, the conceptualization of earnings with bins introduces measurement error for the dependent variable and thereby inflates

Table 7 The earnings impact of smoking by ESHI

	ESHI					
	1	2	3	4	5	6
	Full sample		Siblings		Twins	
Panel A: no ESHI						
Smoker	−0.214***	−0.159***	−0.218	−0.202	−0.148	−0.141
	(0.037)	(0.039)	(0.307)	(0.311)	(0.326)	(0.327)
n	4,029	4,029	1,851	1,851	1,253	1,253
Panel B: ESHI						
Smoker	−0.211***	−0.154***	−0.282***	−0.269***	−0.209**	−0.202**
	(0.021)	(0.023)	(0.104)	(0.102)	(0.098)	(0.098)
n	4,946	4,946	2,229	2,229	1,521	1,521
Panel C: DiD						
Smoker*ESHI			−0.094	−0.086	−0.092	−0.084
			(0.163)	(0.147)	(0.157)	(0.0158)
n			4,368	4,368	3,062	3,062
Covariates						
Education	No	Yes	No	Yes	No	Yes
Family Fixed Effects (FE)	No	No	Yes	Yes	Yes	Yes

Source: National Survey of Midlife Development (1996, 2006, and 2014).

Notes: Huber–White clustered standard errors are given in parentheses. All individuals are between 25 and 66 years, and all regressions include controls for race, gender, and age. Panel A contains the effect of smoking for individuals without ESHI. Panel B contains the effect of smoking for individuals with ESHI. Panel C contains the DiD estimates of individuals with ESHI who also smoke. Statistical significance is denoted by the following: $^{**}P < 0.05$ and $^{***}P < 0.01$.

Abbreviations: ESHI, employer-supplied health insurance; FE, fixed effects; DiD, difference-in-differences.

standard errors. Both factors contribute to the large confidence intervals for the estimates on the effect of smoking on earnings.

On the other hand, estimates for smoking on individuals who do have ESHI are large in magnitude and statistically distinguishable from zero. In column 1 of Panel B, without controls for schooling, ESHI smokers earn nearly 21% less than nonsmokers. Controlling for schooling in the second column reduces this gap to approximately 15% of earnings. In columns 3 and 4 for siblings, the effects are large and negative at approximately 27%. For twins in columns 5 and 6, the effects are 20% respectively. The reduction in earnings for smokers with ESHI are large and statistically different from zero, despite the modest sample size.

In Panel C of Table 7, I set up a DiD research design to rigorously estimate the health insurance/expense effects of smoking on earnings. An advantage of the formal DiD research design is that it does increase sample size because now twin and sibling sets where at least one member has ESHI are also included. In both the twin and the sibling samples across columns 3 to 6, the coefficient on the DiD estimator is negative but statistically significant. While the negative point estimates are modest in size ranging between 7% and 10% due to the large confidence intervals, inference is limited.

To provide a pattern of suggestive evidence in support of the hypothesis that healthcare contributes to the earning penalty for smokers, I examine for heterogeneity with correlates of disproportionate healthcare usage. I separate the sample by age and gender to examine for heterogeneity on the dimension of healthcare costs. Although a priori it is unknown whether the effect of nicotine addiction varies with age and gender, it is a stylized fact that health insurance costs do vary considerably on the dimensions of age and gender. I divide the sample into two groups based on the midpoint age of 45 years, and I present these results in Table 8.

Table 8 Does the earnings impact vary by age?

	Age					
	1	**2**	**3**	**4**	**5**	**6**
	Full sample		**Siblings**		**Twins**	
Panel A: old						
Smoker	−0.273***	−0.218***	−0.283***	−0.275***	−0.253***	−0.320***
	(0.031)	(0.033)	(0.095)	(0.095)	(0.097)	(0.095)
n	5,255	5,255	2,431	2,431	1,582	1,582
Panel B: young						
Smoker	−0.192***	−0.119***	−0.051	−0.049	−0.096	−0.090
	(0.030)	(0.031)	(0.117)	(0.117)	(0.109)	(0.110)
n	3,729	3,729	1649	1,649	1,192	1,192
Covariates						
Education	No	Yes	No	Yes	No	Yes
Family Fixed Effects (FE)	No	No	Yes	Yes	Yes	Yes

Source: National Survey of Midlife Development (1996, 2006, and 2014).

Notes: Huber–White clustered standard errors are given in parentheses. All individuals are between 25 and 66 years, and all regressions include controls for race, gender, and age. Panel A contains individuals who are between 25 and 45 years, and Panel B contains individuals who are between 46 and 66 years. Statistical significance is denoted by the following: ***$P < 0.01$.

Evaluating the smokers' wage penalty by differences in age shows that in the full sample, older smokers are likely to have higher penalties than younger smokers. In fact, as column 2 of Panels A and B shows, the earning penalty for older smokers is approximately twice as large as the penalty for younger smokers. For the sibling sample, the earning penalty is approximately 28% in the older sample and nonexistent in the younger sample. The most perplexing findings are for older twins as a similar pattern of results is seen with the twin sample. Unlike young twin estimates, which are smaller than estimates from the full sample, the estimates for the earning penalty are substantially larger in the twin sibling subsample. It is possible that twin-specific idiosyncratic factors contribute to these estimates or that the subsample lacks variation in the data.

Since healthcare costs also vary by gender, I reestimate the models by gender in Table 9. Panel A shows the outcomes for men. The full sample with controls for schooling in column 2 produces effect sizes that are similar to the impacts of smoking estimated from sibling by year fixed-effects models both with and without controls for schooling at approximately 20%. Twin by year fixed-effects estimates are also similar in magnitude at 17%, respectively, but they are imprecisely estimated. Panel B presents the outcomes for women and the pattern of estimates follows that of men. All columns of estimates for the effect of smoking on earnings are negative for women smokers in the full sample and the family sample. Estimates for the earning penalty range from 14% to 20%, and again, the within-family models produce larger estimates than the full sample. Although the earning penalty appears smaller for female smokers than male smokers, the effects are statistically similar due to large standard errors. A similar pattern follows in the within-family and the cross-sectional analyses.

6 Conclusion

Using data from multiple survey years of the National Survey of Midlife Development, this paper investigates the earning differences between smokers and nonsmokers. The main esti-

Table 9 Does the earnings impact vary by gender?

	Gender					
	1	2	3	4	5	6
	Full sample		Siblings		Twins	
Panel A: men						
Smoker	−0.257***	−0.203***	−0.222**	−0.198*	−0.176	−0.162
	(0.026)	(0.031)	(0.113)	(0.114)	(0.124)	(0.123)
n	4,417	4,417	1,894	1,894	1,282	1,282
Panel B: women						
Smoker	−0.219***	−0.140***	−0.200	−0.204	−0.201	−0.204
	(0.036)	(0.036)	(0.131)	(0.130)	(0.141)	(0.142)
n	4,558	4,558	2,186	2,186	1,492	1,492
Covariates						
Education	No	Yes	No	Yes	No	Yes
Family Fixed Effects (FE)	No	No	Yes	Yes	Yes	Yes

Source: National Survey of Midlife Development (1996, 2006, and 2014).

Notes: Huber–White clustered standard errors are given in parentheses. All individuals are between 25 and 66 years, and all regressions include controls for race and age. Statistical significance is denoted by following: *$P < 0.10$, **$P < 0.05$, and ***$P < 0.01$.

mates from within-family models show that smokers earn approximately 15% to 16% less than nonsmokers. The estimates from the family models are statistically similar to traditional models using cross-section data that show a reduction in earnings of smokers by 16% to 18%. Importantly, the earning difference is statistically nonexistent in jobs without ESHI, but instead the difference seems to be driven by smokers who are in jobs with ESHI. Analysis by demographic subgroups reveals that the earning differences for smokers vary with age and gender in a manner consistent with variation in health costs across age and gender.[8] The earning differences dissipate for former-smokers.

The ESHI-driven reductions in earnings for smokers have important policy implications for both models of workers' compensation and health insurance. If transactions costs are low, then firms can adjust the price of health insurance to reflect the differential costs of providing health benefits to smokers. In practice, transactions costs are high, and the practical provision of healthcare benefits preclude individualizing health insurance based on each person's healthcare costs. Despite the inability of firms to charge distinct prices to insure smokers, they can and do differentiate how they compensate employees. Since addiction-related productivity declines do not appear to significantly influence earnings of smokers and non-monetary benefits besides health insurance are unlikely to be influenced by smoking, a plausible case exists that the provision of health benefits causes a reduction in earnings for smokers. Such a finding implies that firms adjust compensation on overall worker quality; thereby, firms adjust for frictions from the insurance market on the labor market.

In contrast, standard economic models for insurance (Rothschild and Stiglitz, 1976) have assumed a distinct market for health insurance without spillovers into the labor market. In such markets for health insurance, firms are inflexible because of asymmetric information or high transaction costs that prevent adjustable premiums for unhealthy workers such as smokers. Under such a framework, insurance markets might function inefficiently and converge to a pooling equilibrium in which healthier workers such as nonsmokers end up paying too much and under insure while unhealthier workers such as smokers over insure and end up paying too little. Another possibility is that the entire market might cease to exist if the market results in adverse selection and healthier workers exit the market. The findings in this paper show that smoking-related earning differences, a type of compensating differentials for health investments, mitigate inefficiencies that arise from asymmetric information in healthcare markets.

Lastly, the findings of this paper indicate that the incidence of smoking does fall on smokers to an extent, as smokers appear to pay for their behavior by earning less than nonsmokers. The pass-through of the incidence remains an empirical question. A possibility remains that the shifting of healthcare costs onto smokers through lower earnings is not entirely complete, and the possibility of negative spillovers on nonsmokers might still exist.

Declarations

Availability of data and material

Individuals interested in data with earnings can apply for it from the National Survey of Midlife Development by visiting the following site: http://midus.wisc.edu/.

8 Table A3 contains disaggregated results by gender and schooling. The pattern of descriptive results does indicate that more educated individuals face a higher wage penalty as would be predicted by a model of compensating differentials (healthcare costs that are shifted onto employers).

Competing interests
Moiz Bhai has no conflicts of interest to disclose.

Funding
Not applicable.

Authors' contributions
Not applicable.

Acknowledgments
I thank my dissertation committee members Tim Classen, Ben Feigenberg, Darren Lubotsky, Deirdre McCloskey, and advisor Frank Chaloupka for their feedback. I also thank Jesse Cunha, Pavel Dramski, Young Jo, Sean Mulholland, and conference participants at the Midwestern Economics Association, Western Economic Association, Illinois Economic Association, and the editor and three anonymous referees for helpful comments. Dylan Wright provided excellent research assistance. All remaining errors are my own.

Disclaimer
The research presented in this paper uses data from The National Survey of Midlife Development in the United States. The MIDUS 1 study (Midlife in the US) was supported by the John D. and Catherine T. MacArthur Foundation Research Network on Successful Midlife Development. The waves of MIDUS 2 and 3 research were supported by a grant from the National Institute on Aging (P01-AG020166) to conduct a longitudinal follow-up of the MIDUS I investigation. The opinions reflected in this article are of the author only.

Author's information
Moiz Bhai is an Assistant Professor of Economics in the Department of Economics and Finance at the University of Arkansas at Little Rock. Email: mxbhai@ualr.edu

References
Anger, S.; M. Kvasnicka (2010): Does Smoking Really Harm Your Earnings so Much? Biases in Current Estimates of the Smoking Wage Penalty. Applied Economics Letters 17(6), 561-564.

Auld, M.C. (2005): Smoking, Drinking, and Income. Journal of Human Resources 40(2), 505-518.

Berman, M.; R. Crane; E. Seiber; M. Munur (2014): Estimating the cost of a smoking employee. Tobacco Control 23(5), 428-433.

Brim, O.G.; P.B. Baltes; L.L. Bumpass; P.D. Cleary; D.L. Featherman; W.R. Hazzard; R.C. Kessler; M.E. Lachman; H.R.Markus; M.G. Marmot; A.S. Rossi; C.D. Ryff; R.A. Shweder (2011): National Survey of Midlife Development in the United States (MIDUS), 1995-1996.

Card, D. (2001): Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. Econometrica 69(5), 1127-1160.

Centers for Disease Control and Prevention. (2015): Current Cigarette Smoking Among Adults—United States, 2005–2014. Morbidity and Mortality Weekly Report 64(44), 1233-1240.

Center for Substance Abuse Prevention (US). (1997): Reducing Tobacco Use Among Youth: Community Based Approaches (Reference Guide). Rockville (MD): Substance Abuse and Mental Health Services Administration (US). (Prevention Enhancement Protocols System (PEPS), No. 1.) 1, The Problem of Tobacco Use Among Youth. Available from: http://www.ncbi.nlm.nih.gov/books/NBK52574/.

Chaloupka, F.J.; K.E. Warner (2000): The Economics of Smoking. Handbook of Health Economics 1, 1539-1627.

Cowan, B.; B. Schwab (2011): The Incidence of the Healthcare Costs of Smoking. Journal of Health Economics 30(5), 1094-1102.

Currie, J.; B.C. Madrian (1999): Health, Health Insurance and the Labor Market. Handbook of Labor Economics 3, 3309-3416.

Grafova, I.B.; F.P. Stafford (2009): The Wage Effects of Personal Smoking History. Industrial & Labor Relations Review 62(3), 381-393.

Gruber, J. (2000): Health Insurance and the Labor Market. Handbook of Health Economics 1, 645-706.

Halpern, M.T.; R. Dirani; J.K. Schmier (2007): Impacts of A Smoking Cessation Benefit Among Employed Populations. Journal of Occupational and Environmental Medicine 49(1), 11-21.

Kennedy, P.E. (1981): Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations. The American Economic Review 71(4), 801.

Lång, E.; P. Nystedt (2018): Blowing up Money? The Earnings Penalty of Smoking in the 1970s and the 21st Century. Journal of Health Economics 60, 39-52.

Levine, P.B.; T.A. Gustafson; A.D. Velenchik (1997): More Bad News for Smokers? The Effects of Cigarette Smoking on Wages. Industrial & Labor Relations Review 50(3), 493-509.

Luce, B.R.; S.O. Schweitzer (1978): Smoking and Alcohol Abuse: A Comparison of their Economic Consequences. New England Journal of Medicine 298(10), 569-571.

Lundborg, P. (2007): Does Smoking Increase Sick Leave? Evidence Using Register Data on Swedish Workers. Tobacco Control 16(2), 114-118.

Manning, W.G.; E.B. Keeler; J.P. Newhouse; E.M. Sloss; J. Wasserman (1989): The Taxes of Sin: Do Smokers and Drinkers Pay their Way? JAMA 261(11), 1604-1609.

Nikaj, S. (2017): Peer Effects and Youth Smoking in the European Global Youth Tobacco Survey. Review of Economic Perspectives 17(3), 219-238.

Oster, G., G.A. Colditz; N.L. Kelly (1984): The Economic Costs of Smoking and Benefits of Quitting for Individual Smokers. Preventive Medicine 13(4), 377-389.

Rothschild, M.; J. Stiglitz (1976): Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. The Quarterly Journal of Economics 90(4), 629-649.

Ryff, C.; D.M. Almeida; J. Ayanian, et al. (2012): Midlife in the United States (MIDUS 2), 2004-2006. Inter-university Consortium for Political and Social Research [distributor].

Ryff, Carol; David Almeida; John Ayanian; Neil Binkley; Deborah Carr; Christoper Coe; Richard Davidson; Joseph Grzywacz; Arun Karlamangla; Robert Krueger; Margie Lachman; Gayle Love; Marsha Mailick; Daniel Mroczek; Barry Radler; Teresa Seeman; Richard Sloan; Duncan Thomas; Maxine Weinstein; David Williams (2013-2014): National Survey of Midlife Development in the United States (MIDUS 3).

Surgeon General. (1964): Report of the Advisory Committee to the Surgeon General of the Public Health Service. US Department of Health, Education and Welfare, Washington, DC.

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(5), 863-886.

Appendix

Table A1 Does smoking or ever-smoking influence labor market participation?

	Labor supply					
	1	2	3	4	5	6
	Full sample		Siblings		Twins	
Panel A: full time						
Smoker	−0.001 (0.011)	0.018 (0.011)	0.060 (0.045)	0.059 (0.045)	0.054 (0.056)	0.052 (0.056)
Former-smoker	−0.011 (0.010)	−0.004 (0.010)	−0.033 (0.073)	−0.032 (0.074)	−0.035 (0.087)	−0.034 (0.088)
Panel B: part time						
Smoker	−0.021** (0.007)	−0.017* (0.007)	−0.008 (0.031)	−0.008 (0.031)	0.006 (0.038)	0.007 (0.038)
Former-smoker	−0.004 (0.007)	−0.001 (0.007)	0.054 (0.056)	0.055 (0.056)	0.040 (0.056)	0.041 (0.056)
Covariates						
Education	No	Yes	No	Yes	No	Yes
Family Fixed Effects (FE)	No	No	Yes	Yes	Yes	Yes

Source: National Survey of Midlife Development (1996, 2006, and 2014).

Notes: Huber–White clustered standard errors are given in parentheses. All individuals are between 25 and 66 years, and all regressions include controls for race, gender, and age. Statistical significance is denoted by the following: *$P < 0.10$ and **$P < 0.05$.

Abbreviations: FE, fixed effects.

Table A2 The effect of smoking on earnings by zygosity

	Twin sample			
	1	2	3	4
Panel A: monozygotic				
Smoker	−0.322*** (0.066)	−0.280*** (0.074)	−0.213* (0.111)	−0.213* (0.111)
Panel B: dizygotic – same				
Smoker	−0.262*** (0.063)	−0.178** (0.069)	−0.198* (0.117)	−0.199* (0.117)
Panel B: dizygotic – different				
Smoker	−0.232*** (0.070)	−0.138** (0.079)	−0.223 (0.150)	−0.193 (0.152)
Education	No	Yes	No	Yes
Twin Fixed Effects (FE)	No	No	Yes	Yes

Source: National Survey of Midlife Development (1996, 2006, and 2014).

Notes: Huber–White clustered standard errors are given in parentheses. All individuals are between 25 and 66 years, and all regressions include controls for race, gender, and age. Statistical significance is denoted by the following: *$P < 0.10$, **$P < 0.05$, and ***$P < 0.01$.

Table A3 The effect of smoking on earnings by gender and schooling

	College		No college	
	Male	**Female**	**Male**	**Female**
Full sample	−0.220*** (0.041)	−0.134*** (0.045)	−0.190*** (0.048)	−0.144 (0.059)
Siblings	−0.271*** (0.049)	−0.239*** (0.047)	−0.234*** (0.050)	−0.249*** (0.048)
Twins	−0.313***	−0.252***	−0.239***	−0.115
	0.083	0.070	0.082	0.103

Source: National Survey of Midlife Development (1996, 2006, and 2014).

Notes: Huber–White clustered standard errors are given in parentheses. All individuals are between 25 and 66 years, and all regressions include controls for race and age. Statistical significance is denoted by the following: ***$P < 0.01$.