



Smoking Behavior of Older Adults: A Panel Data Analysis Using HRS

Kajal Lahiri¹ · Xian Li¹

Published online: 22 January 2020
© The Indian Econometric Society 2020

Abstract

Using longitudinal data from Health and Retirement Surveys over 1992–2010, this paper analyzes decisions by older American to continue smoking and the number of cigarettes to consume using two-part hurdle models with correlated effects. We build on the existing literature by incorporating a myriad of factors including cigarette prices, health shocks and smoke-free laws in one econometric framework. Our estimates indicate that higher cigarette prices play an important role in both reducing participation and the intensity of consumption even for this adult population. In addition, health shocks, as measured by newly diagnosed diseases, raise the probability of quitting, highlighting the ‘curative’ aspects of cessation. However, we find very little effect of health on smoking intensity if an older adult does not quit after a health shock. Per capita cigarette consumption in the US declined by over 64% during the period. We show that increased cigarette prices and health shocks together contribute almost equally to explain nearly 86% of the decline, with little that can be attributed to smoking bans and anti-smoking sentiment.

Keywords Smoking cessation · Price elasticity · Health shocks · Hurdle model · Correlated random effects · Self-selection · Anti-smoking laws

JEL Classification C23 · C25 · I13

Introduction

Smoking continues to be one of the leading causes of preventable mortality and a major contributor of various types of cancer, chronic pulmonary and cardiovascular diseases. This epidemic is one of the biggest public health threats the world is facing, killing over 8 million people a year, see Jha and Peto (2014). Moreover, approximately 80% of the 1.1 billion smokers live in low- and middle-income countries,

✉ Kajal Lahiri
klahiri@albany.edu

¹ Department of Economics, University at Albany, SUNY, Albany, NY 12222, USA

where the burden of tobacco-related illness and death is the heaviest. About half the world's male smokers live in 3 Asian countries: China, India, and Indonesia. Mishra et al. (2012) reported that, with a 42.4% male prevalence rate of tobacco use in India, around 1 million deaths a year can be attributable to smoking by the mid-2010s. Till very recently, India was the second largest consumer of tobacco in the world, second only to China. With 20% of the world's population, China produces and consumes about 30% of the world's cigarettes, and suffers over a million deaths a year from tobacco use, see Jha (2019) and Yang et al. (2019).

Ever since a causal relationship between cigarette smoking and coronary heart disease was reported at Mayo Clinic, the smoking prevalence rate has steadily declined in the US from 40% in 1965 to 13.7% in 2018 due to a variety of aggressive anti-smoking measures spearheaded by the courts and government at different levels. Even then, morbidity and ill health due to cigarette smoking and exposure to second hand smoke account for over 480,000 deaths per year.¹ Furthermore, the direct health care expenditure and lost productivity due to smoking add up to approximately \$290 billion a year.² Not surprisingly, academic researchers and public health officials are paying increasing attention to find ways to reduce smoking and make smokers quit all together.

Most policy makers consider cigarette taxes as an effective way to reduce tobacco consumption, and many studies have estimated its elasticity and corroborated its significance. Additionally, health risks as perceived by smokers have a substantial effect on smoking. Research has shown that people who possess correct health information have a tendency to change their addictive habits—the so-called 'preventive' motive for quitting to avoid being sick in the future. In addition, many smokers quit after diagnosed with smoking-related diseases—the 'curative' motive for quitting. To the extent that smoke-free laws in public places deter smoking and influence public perception, these laws can also have a direct effect on smoking through the preventative channel. However, most of the previous studies have focused primarily on one aspect of the mechanism, not considering the effect of cigarette prices, health shocks and smoke-free laws simultaneously. In this paper, we try to disentangle the separate effects of these three factors on smoking reduction under one framework using a longitudinal data from the US. Health and Retirement Study (HRS).³ In our preferred estimation, we propose a two-stage hurdle model to estimate the determinants of smoking participation and conditional demand simultaneously. Our

¹ Centers for Disease Control and Prevention. Smoking and tobacco use. http://www.cdc.gov/tobacco/data_statistics/fact_sheets/fast_facts. Accessed August 18, 2016.

² Centers for Disease Control and Prevention. Economic facts about US tobacco production and use. http://www.cdc.gov/tobacco/data_statistics/fact_sheets/economics/econ_facts. Accessed August 18, 2016.

³ The biennial Indian HRS known as LASI (Longitudinal Ageing Study in India) is currently in its 4th round, which started in 2010. In China, HRS sister survey CHARLS was implemented for a national representative sample of persons 45 years of age or older since 2011. These surveys are harmonized with HRS in the U.S. Thus, the methodology adopted in this paper can be implemented in India and China using LASI and CHARLS, respectively.

hope is that the lessons learnt from the US experience can be implemented to other countries.

Section 2 reviews the rich existing literature on the impact of health shocks, cigarette prices and smoke-free laws on smoking and on modeling selectivity in nonrandom samples. Section 3 discusses our empirical strategies, followed by descriptions of data in Sect. 4. Section 5 is devoted to the interpretation of results in the context of models of smoking participation and consumption. The paper concludes in Sect. 6.

The Existing Literature on Elderly Smoking

The 1990 US Surgeon General Report is the first comprehensive statement on the health benefits of smoking cessation and concluded that smoking cessation improves immediate and long-term health and increases longevity, even for those who already suffer from smoking-related illness. For instance, smoking cessation by persons with diagnosed heart problems markedly reduces the risk of heart attack and cardiovascular death. In many studies, this reduction in risk has estimated to be more than 50%. As for mortality, Taylor et al. (2002) found that individuals who quit enjoy prolonged lives, relative to those who continue to smoke. Although the gain in longevity is largest when quit at younger ages, it remains substantial at older ages too. Ostbye and Taylor (2004) further found that smoking cessation leads not only to increases in years of life but also in years of healthy life (YHL) by reducing smoking-related illness.

Based on the classic health demand model of Grossman (1972) and the rational addiction model of Becker and Murphy (1988), there is now a rich empirical literature on the determinants of initiation and cessation of smoking. Given the obvious public policy implications, most of these studies have focused on the price elasticity of participation and conditional demand for cigarettes by the youth. However, studying the smoking behavior of the older adults is also important because cigarette taxes interact with health problems that are associated with continued smoking and ageing in general. The deterioration of health in the elderly due to smoking has serious implications for the ever-burgeoning public health insurance programs like Medicaid and Medicare in the US.

In one of the earliest studies on elderly smoking behavior, Lewit and Coate (1982) found little evidence that taxes reduce the smoking participation or consumption of adults, their estimated price elasticities for participation and conditional demand were not significant at the conventional significance levels. Other studies, using varied data sources, found no systematic evidence that higher prices reduce smoking prevalence among older adults. For example, Evans and Farrelly (1998), and Farrelly et al. (2001) defined an older adult as an individual at least 40 years old and found little evidence that they would change their behavior to respond to cigarette prices. Adda and Cornaglia (2013) found no significant tax elasticity for the age group 45 and older using data from National Health and Nutrition Examination Survey (NHANES) during 1988–2006. In another recent study Maclean et al. (2016) estimated the tax elasticity in the elderly population,

using the Health and Retirement Study 1992–2008, and found that participation is not responsive to a tax increase. However, using cross-sectional data from Behavioral Risk Factor Surveillance System (BRFSS) over 2000–2005, DeCicca and McLeod (2008) estimated the price elasticity of participation to be around -0.3 for the age group 45–59, and -0.2 for the group 45–64, which were significantly different from zero.

Unlike elasticity of participation, relatively more consistent results have been obtained regarding the responsiveness of conditional consumption to cigarette prices for different age groups. For example, Evans and Farrelly (1998) found a significant -0.498 price elasticity of conditional demand at the extensive margin. Maclean et al. (2016) also found that smokers respond modestly to a tax increase, with elasticity in the range -0.03 to -0.04 . Nesson (2017) obtains elasticities around -0.15 , using NHANES data.

A number of studies have estimated the price effect for adult smoking, not focusing on the older people specifically. Tauras (2006), using a sample of individuals with 46 years as average age, found a small price participation elasticity of -0.126 and price demand elasticity of -0.07 . DeCicca et al. (2008)'s estimates indicate that higher taxes have no overall effect on young adults' cessation and participation decisions when anti-smoking sentiment is controlled for. Liu (2010) found that a significant participation elasticity of -0.159 for the age group 45–64, but the effect was not statistically significant for the age group 65 and more. Pesko et al. (2016) considered the influence of intra-state price variation on the estimates of price elasticity on the 18 and older age group, and showed that using local price variation could increase the price elasticity of consumption substantially, from -0.06 to -0.25 .

Health problems, especially smoking-related diseases, are important for older smokers in their decisions to continue smoking and how many. Indeed, studies have found that smokers adjust smoking behavior based on revealed health risk and information. Although Jones (1994) found that doctor advice cannot effectively increase the probability of quitting, Lahiri and Song (2000) and Khwaja et al. (2006) provided evidence that individuals quit smoking in response to major health shocks. Furthermore, Arcidiacono et al. (2007) showed that models of forward-looking behavior explain the pattern of heavy smoking better than a myopic model, which sheds some light on preventive quitting. Sundmacher (2012), using German data, investigated the effect of health shocks on smoking and obesity, and confirmed that a contemporary health shock has a significant positive impact on the probability of quitting.

Selectivity can potentially jeopardize elasticity estimates if unobserved factors in the sequential decisions of participation and consumption are correlated. Lahiri and Song (2000) formulated such a sequential self-selection model in a cross-sectional context to study the effect of smoking on cancer incidence. To model selectivity in count data models, Ophem (2000) suggested an estimable model by transforming the underlying processes to the bivariate normal distribution, in which the two regimes are characterized by potentially two different data generating processes. Min and Agresti (2005) proposed a two-part hurdle model with correlated random effects, which comprehensively handles the zero observations and the positive counts, and takes into account the correlation between measurements for the same individual at

different time periods. In our context, this two-regime framework allows for relapses after quitting, and the self-selection underlying the sequential decisions.

Empirical Approaches

We employ two approaches to estimate participation and consumption decisions. Section 3.1 uses a two-part model including Probit model with random effects to analyze the smoking participation decision, and a zero-truncated count model with random effect for the conditional demand. Section 3.2 proposes a two-part hurdle model with correlated random effects to analyze these two decisions simultaneously, and addresses the selectivity problem affecting both. This is a prime innovation in our paper. Most studies in the literature use state dummies to control for unobserved effects without recognizing the possibility of unobserved individual effects. Additionally, apart for losing many degrees of freedom in a short panel like ours, the state dummies absorb all inter-state variation in taxes and anti-smoking legislations, making it difficult to capture their independent role in smoking cessation.⁴ However, the assumption that the individual effect is independent of observed regressors is restrictive and unrealistic. We follow an intermediate solution by specifying a control function for the unobserved individual effects in terms of the individual observed means; the random error in this specification is assumed uncorrelated with observed regressors.⁵ Thus, in models described below, we add the averages of statistically significant individual level regressors, like income and exercise habit, to both participation and demand equations. As explained in Chatterji et al. (2014), with a nested error structure where each individual nests within a specific state, controlling for individual effects will control for unobserved state effects, but not the other way round.

Two-Part Model of Probit and Zero-Truncated Negative Binomial Regression

We model smoking participation and conditional consumption decisions separately as outcomes of a smoker's utility maximization using a random utility model. Heckman (1979) first postulated this two-part model based on the fundamental premise that the baseline and induced risk factors of cigarette smoking are not always same for all individuals. After perceiving the risk from health shocks, the retail price and the local anti-smoking environment, smokers determine their behavior based on updated subjective beliefs concerning possible health deterioration and monetary cost.

In each period in our observation window, the individual faces a decision to make between two alternatives: (1) continue to smoke or (2) quit. Let y_{it} be the binary variable which indicates the response to the question “Do you smoke cigarettes now?”

⁴ Many studies have demonstrated the bias in fixed effects estimation of non-linear models with short panels, see Fernandez-Val (2009), Hahn and Kuersteiner (2011) and others.

⁵ See Chamberlain (1984), Wooldridge (1995), Mundlak (1978) and Labeaga (1999).

for the individual i at time t . If the answer is yes, then $S_{it} = 1$, otherwise $S_{it} = 0$. With the latent variable S_{it}^* , we can construct the following Probit model as the participation decision:

$$\begin{aligned} S_{it} &= 1 && \text{if } S_{it}^* > 0 \\ &= 0 && \text{if } S_{it}^* \leq 0 \end{aligned} \tag{1}$$

where

$$S_{it}^* = x_{it}^T \beta_X + P_{it} \beta_p + H_{it}^T \beta_H + \mu_i + u_{it} \tag{2}$$

and x_{it} is a $K_1 \times 1$ vector, including individual characteristics; P_{it} is the retail price of cigarettes per pack for the individual i at time t ; H_{it} is a $K_2 \times 1$ vector which includes a set of measures of health (including health shocks) for individual i at time t . β_X , β_p and β_H are $K_1 \times 1$, 1×1 and $K_2 \times 1$ parameter vectors to be estimated.

The random effect μ_i is used to address the correlation across time with the following standard assumption

$$\mu_i \sim \text{IIN}\left(0, \sigma_\mu^2\right)$$

which is independent of u_{it} and (x_{it}, H_{it}) .

We estimate this panel data model via maximizing log-likelihood with random-effects. The conditional probability of smoking, $y_{it} = 1$, is

$$\begin{aligned} \Pr(S_{it} = 1 | x_{it}, P_{it}, H_{it}, \mu_i) &= \Pr(\varepsilon_{it} > -(x_{it}^T \beta_X + P_{it} \beta_p + H_{it}^T \beta_H + \mu_i) | x_{it}, P_{it}, H_{it}, \mu_i) \\ &= \Phi(x_{it}^T \beta_X + P_{it} \beta_p + H_{it}^T \beta_H + \mu_i) \end{aligned}$$

Then the panel-level marginal likelihood L_i is given by

$$L_i(B, \sigma_{12}^s, \sigma_\mu) = \int_{-\infty}^{\infty} \prod_{t=1}^T [\Pr(y_{it} = 1 | x_{it}, P_{it}, H_{it}, \mu_i)^{y_{it}} \Pr(y_{it} = 0 | x_{it}, P_{it}, H_{it}, \mu_i)^{(1-y_{it})}] f(\mu_i) d\mu_i$$

where $f(\cdot)$ is density function of μ_i defined above and $B = (\beta_X^T, \beta_p, \beta_H^T)^T$.

The integral is approximated by Adaptive Gaussian quadrature, with the number of quadrature points 20. Then the likelihood function is written as

$$L(B, \sigma_{12}^s, \sigma_\mu) = \prod_{i=1}^N L_i(B, \sigma_{12}^s, \sigma_\mu)$$

A follow up question on “how many cigarettes do you smoke per day?” is queried from the daily smokers with responses as positive integers. In specifying the conditional demand function, we employ a (truncated) count model due to the discrete

nature of the number of cigarettes.⁶ Poisson distribution and negative binomial distribution are commonly used to model count data, but the latter distribution is preferred because it allows for richer heterogeneity and over-dispersion.

Specifically, y_{it} here is the number of cigarettes smoked per day by individuals who self-report as smokers. Then the function is

$$P(y_{it}|V_{it}, \mu_i, Smoking = Yes) = \frac{g(y_{it}; \lambda_{it}(\mu_i))}{[1 - g(0; \lambda_{it}(\mu_i))]}$$

where μ_i , the usual random effect, is included to allow for correlation across time; V_{it} is the covariates vector that includes $\{x_{it}, H_{it}, P_{it}\}$. The conditional mean of dependent variable $\lambda_{it}(\mu_i)$ is parameterized as

$$\lambda_{it}(\mu_i) = \exp(x_{it}^T \zeta_x + P_{it} \zeta_p + H_{it}^T \zeta_H + \mu_i)$$

with the negative binomial probability density function as

$$g(y_{it}; \lambda_{it}(\mu_i)) = \frac{\Gamma(\alpha^{-1} + y_{it})}{\Gamma(\alpha^{-1}) \Gamma(y_{it} + 1)} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \lambda_{it}(\mu_i)} \right)^{\alpha^{-1}} \left(\frac{\lambda_{it}(\mu_i)}{\alpha^{-1} + \lambda_{it}(\mu_i)} \right)^{y_{it}} .$$

MLE is used in the estimation with likelihood function of

$$L(\zeta_x, \zeta_p, \zeta_H, \sigma^2, \alpha) = \int \prod_{i=1}^N \prod_{t=1}^T \frac{g(y_{it}; \lambda_{it}(\mu_i))}{[1 - g(0; \lambda_{it}(\mu_i))]} f(\mu_i) d\mu_i .$$

As with the Probit part, the integral is approximated by Adaptive Gaussian quadrature, and the number quadrature points was 20.

Hurdle Model with Correlated Random Effects

Our preferred specification consists of a hurdle model with correlated Probit and negative binomial regressions, cf. Min and Agresti (2005). The demand for cigarettes can now be analyzed jointly with the participation decision by considering smoker’s previous smoking trajectory and the degree of addiction. Nesson (2017) has emphasized the need to control for previous smoking history in determining cessation. However, due to extreme state dependence, it was problematic to include lagged smoking intensity in our participation equation, see also Maclean et al. (2016). In our model, the propensity to quit utilizes the information on smoking intensity through the correlated (unobserved) random effects. To the best of our knowledge, no study has analyzed participation in smoking and cigarettes consumption jointly while being cognizant of self-selection in the two switching regimes.

⁶ Although previous studies have typically used linear regressions (cf. Maclean et al. 2016), our comparison shows that the count model outperforms linear regression in our application in terms of model fit. Detailed results are available upon request. See also Wang and Heitjan (2008).

The Hurdle model allows for two correlated data generating processes for the binary and the positive count variables in one framework. The Probit model is used to model the data generating process (DGP) for dichotomous variable indicating whether the individual quits, and a zero-truncated negative binomial distribution to model the probability for each possible count for smokers. Selectivity prevails because the observations are distributed across the two regimes by an endogenous self-selection process. Ophem (2000) showed that the endogenous selectivity could be modeled in the two regimes of a switching-count model by allowing error terms from the two components to follow a bivariate normal distribution. This simplifies our framework to a hurdle model with correlated random effects in the two regimes. Thus, the first regime is formulated using Probit with random effect μ_{1i} and a zero-truncated negative binomial distribution with random effect μ_{2i} is used to model the data generating process of positive values.

The (conditional) hurdle model is then formulated as

$$P(y_{it} | V_{it}, \mu_{1i}, \mu_{2i}) = \begin{cases} (1 - P_{it}(\mu_{1i})) & \text{if } y_{it} = 0 \\ P_{it}(\mu_{1i}) \frac{g(y_{it}; \lambda_{it}(\mu_{2i}))}{[1 - g(0; \lambda_{it}(\mu_{2i}))]} & \text{if } y_{it} > 0 \end{cases}$$

where

$$V_{it} = \{x_{it}, H_{it}, P_{it}\}$$

$$\text{probit } P_{it}(\mu_{1i}) = x_{it}^T \beta_x + P_{it} \beta_p + H_{it}^T \beta_H + \mu_{1i} \tag{3}$$

$$g(y_{it}; \lambda_{it}(\mu_{2i})) = \frac{\Gamma(\alpha^{-1} + y_{it})}{\Gamma(\alpha^{-1}) \Gamma(y_{it} + 1)} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \lambda_{it}(\mu_{2i})} \right)^{\alpha^{-1}} \left(\frac{\lambda_{it}(\mu_{2i})}{\alpha^{-1} + \lambda_{it}(\mu_{2i})} \right)^{y_{it}} \tag{4}$$

x_{it} , P_{it} , and H_{it} are defined in Sect. 3.1, with their corresponding coefficient vectors ξ_x , ξ_p , and ξ_H . The term μ_{1i} is the random effect in the first regime.

To parameterize the mean and/or variance in the second regime, we assume

$$\lambda_{it}(\mu_{2i}) = \exp(x_{it}^T \zeta_x + P_{it} \zeta_p + H_{it}^T \zeta_H + \mu_{2i}) \tag{5}$$

where μ_{2i} is the random effect in the second regime.

Those two regimes of the model are tied together by assuming that the random effects are jointly normal and correlated as,

$$(\mu_{1i}, \mu_{2i}) \sim \text{Normal} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \right) \tag{6}$$

Let B be the vector of all parameters to estimate: $(\beta_x^T, \beta_p, \beta_H^T, \zeta_x^T, \zeta_p, \zeta_H^T, \sigma_1^2, \sigma_2^2, \sigma_{12}, \alpha)$. Then the marginal likelihood function for this hurdle model can be written as:

$$L(\mathbf{B}) = \prod_{i=1}^N L_i(\mathbf{B})$$

where

$$L_i(\mathbf{B}) = \int \prod_{t=1}^T [1 - P_{it}(\mu_{1i})]^{I_{y_{it}=0}} \prod_{t=1}^T \left\{ P_{it}(\mu_{1i}) \frac{g(y_{it}; \lambda_{it}(\mu_{2i}))}{[1 - g(0; \lambda_{it}(\mu_{2i}))]} \right\}^{I_{y_{it}>0}} \phi(\mu_{1i}, \mu_{2i}) d(\mu_{1i}, \mu_{2i})$$

$$= \prod_{t=1}^T [1 - P_{it}]^{I_{y_{it}=0}} \prod_{t=1}^T \left\{ P_{it} \frac{g(y_{it}; \lambda_{it})}{[1 - g(0; \lambda_{it})]} \right\}^{I_{y_{it}>0}}$$

and $\phi(\mu_{1i}, \mu_{2i})$ is the joint density function of (μ_{1i}, μ_{2i}) defined in (6).

The procedure NLMIXED in SAS enables us to do the estimation as a nonlinear mixed model by maximizing an approximation to the likelihood integrated over the random effects. The popular software Stata will also have the corresponding code to estimate this model jointly. Different integral approximations and optimization techniques are available for this procedure, and after some experimentation, we use adaptive Gaussian quadrature for the integration, and Quasi-Newton algorithm for the likelihood maximization.

Data and Variables

The Health and Retirement Study (HRS) is a longitudinal household survey data set for the study of retirement and health among the elderly in the United States. The study interviews approximately 22,000 Americans ages over 50 and their spouses every 2 years on domains like health care, housing, assets, pensions, employment and disability. Furthermore, the RAND Center for the Study of Aging, with funding and support from the National Institute on Aging (NIA) and Social Security Administration (SSA), created the RAND (L version) HRS data files. The RAND HRS we use is a concise subset of the HRS, over ten waves (1992, 1993/1994, 1995/1996, 1998, 2000, 2002, 2004, 2006, 2008, and 2010). This study uses the main body of RAND version matched with sensitive administrative information on state and race/ethnicity from different waves of HRS. We use basically the same data as Maclean et al. (2016), but with subtle differences in model specifications.

Many of the health measures like newly diagnosed heart disease are obtained by differencing two consecutive responses to the interview question, ‘Have you ever been told by a doctor that you have heart disease?’ Therefore, the effective number of waves of our panel is nine. Due to this reason, missing data and natural attritions, our final sample contains 98,941 observations on 18,177 respondents. Table 1 has the description of all variables, including demographics, health, and smoking-relevant variables at both individual and state levels.

Table 1 Description of variables

Variables	Description
<i>Time-invariant variables (demographics) panel A</i>	
Cohort: AHEAD	Born before 1924
Cohort: CODA	Born during depression, 1924–1930
Cohort: HRS	Born during 1931–1941
Cohort: WB	Born during war, 1942–1947
Cohort: EBB	Born during Early baby boomers, 1948–1953
Catholic	Religious preference, Catholic
Protestant	Religious preference, Protestant
Black	Binary indicator for Africa-American
White	Binary indicator for Caucasian
Hispanic	Binary indicator for Hispanic
Male	Binary indicator for male
Foreign born	Not born in the US
South born	Born in southern states
Western born	Born in western states
North born	Born in northeast states
BMI	Body Mass Index, normalized by sample mean
BMI ²	Squared BMI
Exercise	Vigorous physical activity at least three times per week
Veteran	Ever been in military service
Service	The industry associated with the longest-held job is service (business/repair service/personal services and entertainment)
Sales	The industry associated with the longest-held job is sales (wholesale/retail)
Technology	The industry associated with the longest-held job is technology (Transportation/professional/related service)
Finance	The industry associated with the longest-held job is Finance (finance/insurance/real estate)
Blue collar	The industry associated with the longest-held job is blue collar (mining and construction/manufacture)
Income	Household income divided by the size of family
Plan_Hor_Short	Financial planning horizon is next few month or shorter
School year	Years of schooling normalized by sample mean
School year ²	Squared schooling year
Father Educated	Binary indicator for schooling year of father ≥ 12
Mother Educated	Binary indicator for schooling year of mother ≥ 12
<i>Time-variant variables (Demographics) Panel B</i>	
Age	Age when interviewed
Age ²	Squared age

Table 1 (continued)

Variables	Description
Married	Binary indicator if respondent is married when interviewed
Spouse smoking	Binary indicator if married and spouse is smoking when interviewed
Retired	Binary indicator if respondent is retired when interviewed
IRA	Binary indicator if value of IRA/Keogh account is positive
Wave dummies	Binary indicator if interview year = 1994, 1996, 1998, 2000, 2002, 2004, 2006, 2008, 2010
<i>Time-variant variables (health measure dummies) Panel C</i>	
HC: heart disease	Binary indicator: respondent was ever told by a doctor that she/he had a heart problem ^a
HC: cancer	Binary indicator: respondent was ever told by a doctor that she/he has cancer ^b or a malignant tumor
HC: hypertension	Binary indicator: respondent was ever told by a doctor that she/he has hypertension/high blood pressure
HC: SAH	Self-assessed health from level 1 (excellent) to level 5 (poor)
HS: hospitalization	Binary indicator: at least 1-day stay in hospital during last 2 years or since the last interview
HS: doctor visits	Binary indicator: visited doctors at least 5 times during last 2 years or since the last interview
HS: SAH	Binary indicator: self-assessed health is decreased by at least 2 levels during last 2 years or since the last interview
HS: heart disease	Binary indicator for a newly diagnosed heart disease after last interview and <i>HC: heart disease</i> is 0 before this wave
HS: cancer	Binary indicator for a newly diagnosed cancer after last interview and <i>HC: cancer</i> is 0 before this wave
HS: hypertension	Binary indicator for a newly diagnosed hypertension after last interview and <i>HC: hypertension</i> is 0 before this wave
<i>Time-variant variables (smoking-related) Panel D</i>	
Smoke ever	Binary indicator if ever initiate smoking, dependent variable in initiation stage
Smoke now	Binary indicator if smoking now, dependent variable in participation model
Smoke intensity	Number of cigarettes smoked per day if smoking, dependent variable in conditional demand model
Cigarette price	Weighted average price per package (20 cigarettes) including federal and state taxes

Table 1 (continued)

Variables	Description
Sentiment	Anti-smoking sentiment, measured by the percentage of respondent who thinks smoking should not be allowed anywhere in bars in state level, from TUS-CPS
Smoking bans	Binary indicator if there is effective state-level smoking ban on one of the locations including restaurants, private worksites and bars

^aIncluding a heart attack, coronary heart disease, angina, congestive heart failure, or others

^bExcluding minor skin cancer

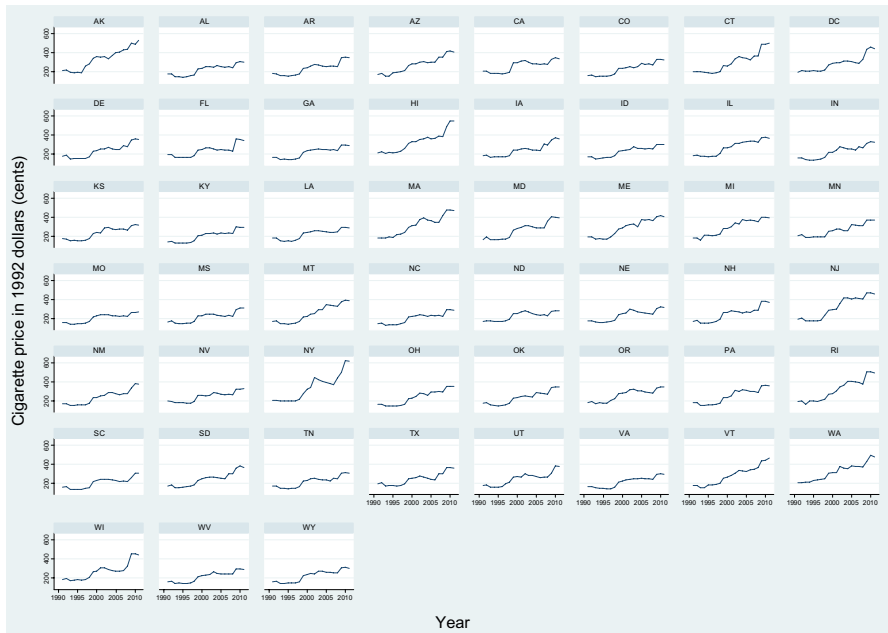
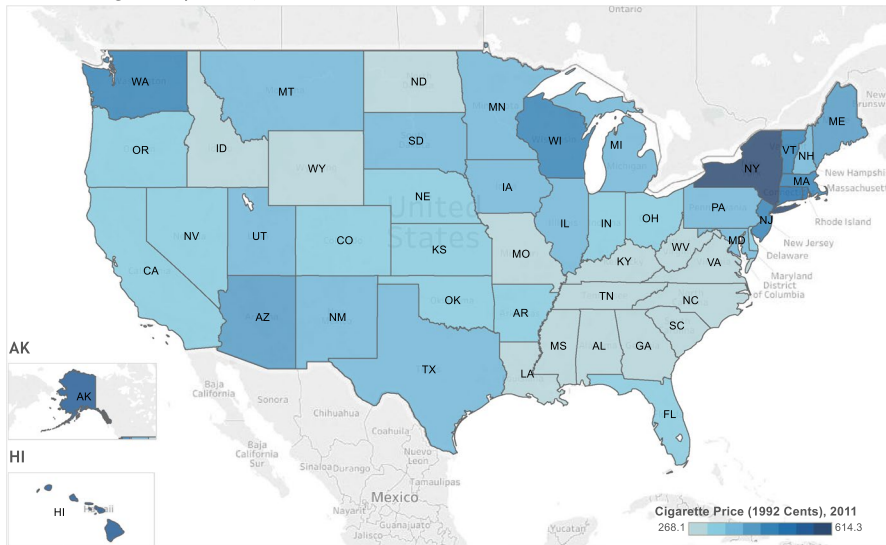


Fig. 1 Cigarette price by state from 1991 to 2011 Source: The Tax Burden on Tobacco (2012), 2012

Cigarette Price and Smoke-Free Laws

We merge information on cigarette prices from the different editions of *Tax Burden on Tobacco* (Orzechowski and Walker 2012), which report taxes and retail prices by state and year, with administrative data on state identifiers of respondents. We found our results to be very similar whether use price or tax. Here we choose the price (including federal and state taxes) because this is what smokers face in their decision-making process. In the spirit of Pesko et al. (2016) to maximize the price variation across states and time, we deflate the variable using the regional consumer price indexes, even though the use of the conventional national CPI data did not

US Price of Cigarettes per Pack, 2011



Note: This map is produced using shape file provided by Tableau Desktop v10.0, and the *Tax Burden on Tobacco* (2012)

Fig. 2 Spatial variation of cigarette price by states, 2011

make much difference in our elasticity estimates. Since HRS collects information over the whole year, while the *Tax Burden on Tobacco* reports the retail price of cigarettes as of November every year, we use last year's retail price in our models. Figure 1 displays the cigarette price per pack between 1991 and 2011 by state, adjusted for regional inflation. It is interesting to note that of the two dimensions of variation in the price data, we found the cross-sectional variation between states is 29.28% of the total variation in our panel data, thus most of the variation in prices comes from within-state variation over time, see Fig. 2.

Both models also include controls for the state-level smoke-free air laws and sentiment. Following the literature and after extensive experimentation, we used venue-specific smoking ban data from the *ImpacTeen* project (1991–2008) and Centers for Disease Control and Prevention (2009–2011), and then match the data to the HRS data set by state and year. Considering the consistency of measures over time, we defined a dichotomous variable for smoking ban, which is equal to one if there is an effective state-level smoking ban in one of the locations including restaurants, private worksites and bars. The sentiment measure comes from the TUS-CPS (Tobacco Use Supplement to the Current Population Survey)⁷, from which we use the percentage of respondents who think smoking should not be allowed anywhere

⁷ The Tobacco Use Supplement to the Current Population Survey (TUS-CPS) is a national survey of tobacco use as part of the US Census Bureau's Current Population Survey in 1992–1993, 1995–1996, 1998–1999, 2000, 2001–2002, 2003, 2006–2007, and 2010–2011.

in bars as anti-smoking sentiment proxy at the state-level. Since smoke-free laws in public spaces may change perception and attitudes towards smoking, the sentiment measure may pick up additional effect of smoke-free legislations on quitting and the intensity of smoking, see Carton et al. (2016).

Health Shocks and Other Health Measures

We use two types of health information. One type captures health shocks (HS) that an individual experienced recently (within last 2 years). This is similar to health shocks defined in Smith et al. (2001) and Sundmacher (2012). Specifically, if the individual stayed in a hospital overnight, visited doctors an unexpected number of times, or lowered one's self-rated health substantially, then a health shock (HS) is considered to have occurred to the individual. By inspecting the relevant histograms, we defined doctor visits more than five times a year, and lowering of the self-rated health by two or more categories as a health shock. This way we hope distinguish between sudden unexpected changes in health status, net of expected trends due to aging and associated morbidities. In addition, if there is a newly diagnosed heart disease, cancer, or high blood pressure during last 2 years and has no history of that disease before, then a health shock is also identified to have occurred. For example, if the respondent reported a heart disease in wave 5 that occurred after the survey of wave 4, and reported no history of heart disease before wave 4, then this is recorded as a health shock. In summary, the first type of health shock (HS) is defined by new and substantive changes health status.

The other type of health information is health conditions (HC) that individual ever had, which includes previously diagnosed heart diseases, cancers, and hypertension. We control for both HS and HC since it is reasonable to assume that people might have a different reaction to newly diagnosed diseases and old health morbidities. Most studies attempting to estimate price elasticity do not use such extensive controls for health shocks and morbidity.

Summary Statistics

The smoking prevalence in the US has been steadily declining, especially after the release of *The Surgeon General Report (1964)*.

Figure 3 shows this declining smoking rate for Whites, Blacks and Hispanics, over 1992–2010. While the prevalence of the white population dropped from 24% in 1992 to 11% in 2010, the same for blacks dropped from 32% to 17% during the same time. The average cigarette consumption per capita per year decreased approximately from 100 packs in 1991 to 50 packs in 2011. If we look at the smoking intensity for smokers in our sample, we find that the distribution of the number of cigarettes smoked per day shifted to the left steadily, which indicates that not only the prevalence but also the intensity fell (see Fig. 4). Concurrently, the (deflated) total tax collection including federal and states tax started to increase after 1991 that accelerated after 2008.



Fig. 3 Smoking prevalence in the US by race from 1992 to 2010. Source: The Tax Burden on Tobacco (2012) and authors' calculation using HRS

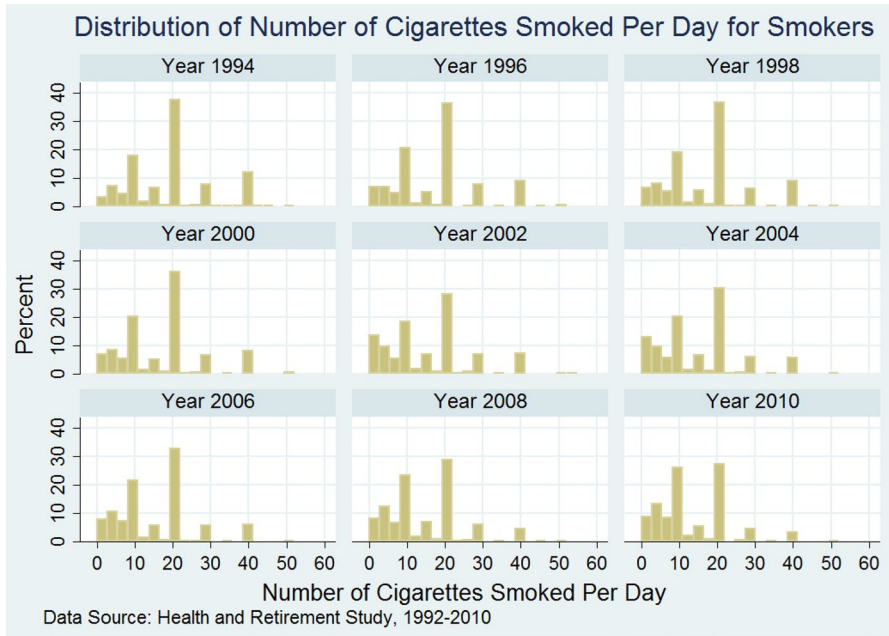


Fig. 4 Distribution of number of cigarettes smoked per day for smokers only

Figure 5 shows the age distribution of smokers and number of years since quitting at the beginning of our sample. Consistent with other studies, we find that former smokers, on the average, quit smoking in their 40s. Table 2 compares

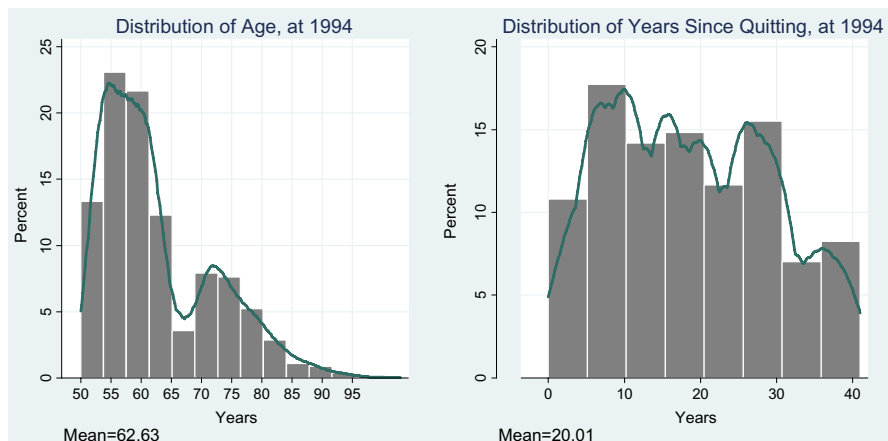


Fig. 5 Distribution of age and years since quitting for quitters. Source: Authors' calculation using HRS survey in 1994

Table 2 Summary statistics—by observation

By observations	Current-smokers	Quitter and non-smokers	Full sample
Obs.	16,375	82,395	98,770
Age	60.92***	65.00***	64.33
School year	11.85***	12.71***	12.56
Household size	2.33***	2.27***	2.28
SAH	2.96***	2.68***	2.73
Male	45.38%*	44.36%*	44.53%
Married	58.49%***	70.19%***	68.25%
Divorced/widowed	32.50%***	23.90%***	25.33%
Retired	45.29%***	55.80%***	54.06%
Catholic	26.64%	26.49%	26.52%
Protestant	62.97%***	65.03%***	64.69%
HS: hospital	20.24%***	23.04%***	22.57%
HS: doctor	48.85%***	58.72%***	57.08%
HS: SAH	5.40%***	4.74%***	4.85%
HS: heart disease	2.56%***	3.26%***	3.14%
HS: cancer	1.61%***	2.20%***	2.11%
HS: hypertension	4.89%	4.55%	4.61%
HC: heart disease	20.81%***	24.56%***	23.94%
HC: cancer	10.72%***	13.40%***	12.96%
HC: hypertension	47.95%***	56.38%***	54.99%

P value for the test $H_0 : P_1 = P_2$ or $\mu_1 = \mu_2$ versus $H_1 : P_1 \neq P_2$ or $\mu_1 \neq \mu_2$

* P value < 0.05, ** P value < 0.01, *** P value < 0.001

Table 3 Summary statistics—by individual

	By individual	Ex-smokers and current smokers	Non-smokers	Full sample
Ind.		11,328	6784	18,112
Age		53.09	52.14	52.73
School year		12.35***	12.89***	12.55
Male		53.10%***	33.21%***	45.65%
Catholic		27.24%***	25.27%***	26.50%
Protestant		62.10%***	66.56%***	63.77%
HS: hospital		60.13%***	56.50%***	58.77%
HS: doctor		83.83%	84.71%	84.16%
HS: SAH		23.90%***	21.33%***	22.94%
HS: heart disease		18.04%***	15.63%***	17.14%
HS: cancer		12.19%***	10.30%***	11.48%
HS: hypertension		24.76%	25.74%	25.13%

P value for the test $H_0 : P_1 = P_2$ or $\mu_1 = \mu_2$ versus $H_1 : P_1 \neq P_2$ or $\mu_1 \neq \mu_2$

* P value < 0.05 , ** P value < 0.01 , *** P value < 0.001

the demographic characteristics of current smokers in the full sample, whereas Table 3 compares the sample of ever-smoked individuals with never-smokers.

Table 2 shows that current smokers are significantly younger, have better self-assessed health and less number of schooling years, which may suggest self-selection. At the same time, current smokers are significantly more likely to be unmarried, working, divorced and widowed. When we compare the sample of ever-smoked individuals with non-smokers, these two groups do not have significantly different age, but non-smokers have higher years of education, are substantially less likely to be a male, and more likely to be a protestant. As for the prevalence of health shocks (HS), we observe fewer health shocks for current smokers. This phenomenon is possibly caused by a (curative) behavioral change in former smokers following a health shock, and could be explained by Table 3, in which former smokers are seen to have a significantly higher prevalence of health shocks including hospitalizations, newly diagnosed heart diseases and cancers compared to smokers, and non-smokers.

Empirical Results

Main Estimates: Price Elasticity and Effect of Health Risks

We summarize our baseline estimates for the two models in Tables 4, 5, 6 and 7. Left panel reports estimates of the participation equation, where the dependent variable equals 1 if the individual is a daily smoker, and the right panel reports estimates of the cigarette consumption equation. In both left and right panels, we present the estimates of price-responsiveness of smoking in terms of price elasticity of demand.

Table 4 Estimated effect of cigarette price on smoking

Full sample	Probability of daily smoking		Average number of cig. per day	
	(1) Probit with RE	(2) Hurdle model- Regime 1	(3) ZTNB with RE	(4) Hurdle model-Regime 2
Real price	– 0.017 (0.051)	– 0.147** (0.060)	– 0.051*** (0.017)	– 0.061*** (0.018)
Elasticity	– 0.066	– 0.569**	– 0.132***	– 0.159***
Dep. Var. mean	0.166	0.166	16.821	16.821
Mean price	2.579	2.579	2.579	2.579
<i>N</i>	98,770	98,770	16,221	16,221

Standard error in parentheses

For Probit model and zero-truncated negative binomial model (ZTNB), the random effects are estimated using empirical Bayes

The marginal effects in regime 1 of hurdle models are calculated as $\hat{\beta} \cdot \phi(z)$, where $z = \Phi^{-1}(p)$ and p is the sample mean of the indicator variable for daily smoking, $\hat{\beta}$ is the estimated coefficient, $\phi(\cdot)$ is the standard normal probability density function, and $\Phi^{-1}(\cdot)$ is the inverse of the standard normal cumulative density function

The marginal effect in regime 2 of hurdle model is $\hat{\beta} \cdot \bar{Y}$, and the elasticity is $\hat{\beta} \cdot \bar{X}$, where \bar{Y} and \bar{X} are sample mean of dependent variable and cigarette price, respectively

Standard errors of marginal effects and elasticities are calculated using Delta-Method

P* value < 0.1, *P* value < 0.05, ****P* value < 0.01

The results reported in these four columns are, respectively, Probit model with random effects, Probit part from hurdle model with correlated random effects, zero-truncated negative binomial regression with random effect, and count part from hurdle model. With a similar structure, Tables 5 and 6 show the estimated effects of a number of other health conditions and shocks on the two smoking decisions.

As showed in Table 4, we find evidence that higher cigarette price reduces daily smoking prevalence in our second specification. In particular, our estimate implies that a \$1 increase in cigarette price will reduce the probability of smoking to about 13.0% on average. Relative to a base of 16.6%, this translates to 21.7% reduction in probability and participation elasticity of – 0.569, which is large relative to previous studies (reviewed in Sect. 2). Interestingly, as in Maclean et al. (2016), this elasticity is insignificant in the uncorrelated effects model. The correlation between the two random effects of the two equations is 0.71 and highly significant. This high correlation between the two errors suggests that more addicted smokers tend not to quit, and former heavy smokers are more likely to relapse into smoking again. Moreover, given the formula of the marginal effect of price (see Table 4, footnote 4), and the value of z evaluated at the sample means (– 0.97), the strong positive correlation implies that more addicted smokers respond more to quit in response to increases in cigarette prices, supporting the economic model of rational addiction, cf. Maclean et al. (2016). Given the substantial heaping of number of cigarettes smoked at 20, heavy smokers in our sample possibly mean those smoking 20 or more cigarettes a day (see Fig. 4). Under appropriate incentives provided by increasing taxes coupled

Table 5 Estimated effect of health shocks on smoking

Full sample	Probability of daily smoking		Average number of Cig. per day	
	(1)	(2)	(3)	(4)
	Probit with RE	Hurdle model-Regime 1	ZTNB with RE	Hurdle model-Regime 2
HS: Hospitalization	− 0.398*** (0.035) [− 0.080]	− 0.449*** (0.038) [− 0.088]	− 0.0732*** (0.012) [− 1.187]	− 0.0800*** (0.013) [− 1.292]
HS: Doctor Visits	− 0.255*** (0.031) [− 0.056]	− 0.214*** (0.035) [− 0.048]	0.0142 (0.011) [0.37]	− 0.0140 (0.011) [− 0.234]
HS: SAH	− 0.108** (0.055) [− 0.025]	− 0.140** (0.060) [− 0.033]	− 0.00120 (0.019) [− 0.020]	− 0.00826 (0.020) [− 0.138]
Dep. Mean	0.166	0.166	16.821	16.821
N	98,770	98,770	16,221	16,221

Standard error in parentheses

Marginal effect is in brackets

For Probit model and zero-truncated negative binomial model (ZTNB), the random effects are estimated using empirical Bayes, and marginal effects are calculated at the individual level

The marginal effect in regime 1 of hurdle model is calculated as $\Phi(z + \hat{\beta}) - p$, where $z = \Phi^{-1}(p)$ and p is the sample mean of the indicator variable for daily smoking, $\hat{\beta}$ is the estimated coefficient

The marginal effect in regime 2 of hurdle model is $(\exp(\hat{\beta}) - 1) \cdot \bar{Y}$, where \bar{Y} is the sample mean of the dependent variable

Standard errors of marginal effects are calculated using Delta-Method

* P value < 0.1, ** P value < 0.05, *** P value < 0.01

with looming health shocks for the elderly, these highly addicted smokers control themselves to quit. Furthermore, in cigarette (conditional) demand part, we find consistent and significant evidence that higher cigarette price reduces the amount smoked. Specifically, a \$1 increase in cigarette price reduces the average number of cigarettes consumed per day by nearly 1.03 in the hurdle model and 0.85 in ZTNB model, which are about 6.1% and 5.1% of the average intensive margin, and implies a price elasticity of -15.9% and -13.2% , respectively. Our estimates are very similar to those in Nesson (2017) and Maclean et al. (2016). Since with inelastic demand, consumers bear the majority of tax incidence, our estimates suggest that most of the tax burden will be borne by the elderly smokers when tobacco taxes rise.⁸

⁸ Pesko et al. (2016) use intra-state variations in cigarette prices due to local taxes to show, though for a much younger sample, that the resultant price elasticity is almost tripled compared to state-level prices. Since most of these local taxes are overwhelmingly concentrated in Alabama, Missouri and Virginia, we included a dummy variable representing these three states and its interaction with the state price variable in our cessation and conditional consumption equations. These additional controls were statistically insignificant in our estimations.

Table 6 Estimated effect of health shocks on smoking

Full sample	Probability of daily smoking		Average number of Cig. per day	
	(1) Probit with RE	(2) Hurdle model-Regime 1	(3) ZTNB with RE	(4) Hurdle model-Regime 2
HS:	– 0.623***	– 0.754***	– 0.068	– 0.032
Heart disease	[– 0.110] {– 66.54%}	[– 0.124] {– 74.52%}	[– 1.107] {– 6.58%}	[– 0.538] {– 3.20%}
Previously diagnosed heart disease	– 0.397*** [– 0.080] {– 48.33%}	– 0.487*** [– 0.093] {– 56.30%}	– 0.037** [– 0.613] {– 3.64%}	– 0.035* [– 0.579] {– 3.44%}
HS:	– 0.435***	– 0.560***	0.027	0.022
Cancers	[– 0.086] {– 51.8%}	[– 0.103] {– 62.1%}	[0.453] {– 2.69%}	[0.359] {– 2.13%}
Previously diagnosed cancer	– 0.372*** [– 0.076] {– 45.9%}	– 0.470*** [– 0.091] {– 54.9%}	– 0.033 [– 0.544] {– 3.24%}	– 0.032 [– 0.524] {3.12%}
HS:	– 0.306***	– 0.344***	0.019*	– 0.002
Hypertension	[– 0.065] {– 39.2%}	[– 0.071] {– 43.1%}	[0.328] {1.95%}	[– 0.187] {– 1.11%}
Previously diagnosed hypertension	– 0.446*** [– 0.088] {– 52.8%}	– 0.360*** [– 0.074] {– 44.7%}	– 0.015 [– 0.252] {– 1.50%}	– 0.022 [– 0.364] {– 2.17%}
Dep. Mean	0.166	0.166	16.821	16.821
<i>N</i>	98,770	98,770	16,221	16,221

Marginal effect is in brackets

The marginal effect in percentage in braces

Diseases with prefix HS are newly diagnosed

Marginal effects and elasticity are calculated as Table 5

Standard errors of marginal effects are calculated using Delta-Method

P* value < 0.1, *P* value < 0.05, ****P* value < 0.01

Estimates for the effect of health shocks are uniformly significant in participation equations across models (Tables 5, 6). Health shocks defined as hospitalizations and high frequency of doctor visits, on average, could reduce the probability to smoke by 0.08 and 0.06, respectively, according to hurdle model's estimates. A downgrade of self-assessed health, which is a proxy of subjective health shocks, also shows a significant effect on smoking participation, in both models. However, unlike in participation, only hospital stay shows an effective reduction in cigarette consumption, and specifically, one hospitalization would reduce average conditional demand by about 7.7% or 1.3 cigarettes.

Table 7 Estimates of other parameters

Full sample	Probability of daily smoking		Average number of Cig. per day	
	(1) Probit with RE	(2) Hurdle model-Regime 1	(3) ZTNB with RE	(4) Hurdle model-Regime 2
Exercise	– 0.0469 (0.0333)	– 0.007 (0.037)	0.0145 (0.0116)	0.012 (0.012)
Age	0.0256 (0.0328)	0.072* (0.038)	0.0211* (0.0112)	0.025** (0.012)
Age ²	– 0.00101*** (0.000255)	– 0.001*** (0.000)	– 0.000214** (0.0000870)	– 0.00024*** (0.000)
Married	– 1.034*** (0.0491)	– 0.761*** (0.060)	– 0.111*** (0.0156)	– 0.113*** (0.017)
Spouse smoking	1.502*** (0.0589)	1.097*** (0.065)	0.116*** (0.0165)	0.133*** (0.018)
Retired	– 0.0724** (0.0352)	– 0.150*** (0.040)	0.00458 (0.0115)	0.012 (0.012)
Plan_Hor_Short	0.423*** (0.0748)	0.356*** (0.106)	– 0.0465** (0.0195)	– 0.001 (0.022)
Sales industry	0.447*** (0.0762)	0.373*** (0.110)	0.0289 (0.0206)	0.062*** (0.024)
Technology Industry	0.225*** (0.0676)	0.003 (0.100)	0.0377** (0.0187)	0.042** (0.021)
School year	– 0.526*** (0.0407)	– 0.298*** (0.063)	– 0.0410*** (0.0117)	– 0.050*** (0.013)
School year ²	– 0.177*** (0.0228)	– 0.017 (0.036)	– 0.0230*** (0.00681)	– 0.036*** (0.008)
Veteran	0.514*** (0.0855)	0.621*** (0.125)	– 0.0108 (0.0231)	0.037 (0.026)
Male	0.423*** (0.0756)	0.321*** (0.112)	0.147*** (0.0211)	0.214*** (0.024)
Africa-American	– 0.199** (0.0815)	0.039 (0.111)	– 0.457*** (0.0225)	– 0.502*** (0.026)
Hispanic	– 1.129*** (0.117)	– 1.380*** (0.262)	– 0.522*** (0.0349)	– 0.533*** (0.040)
Sentiment	– 0.128 (0.282)	– 0.300 (0.338)	– 0.295*** (0.0922)	– 0.171* (0.100)
Smoking bans	0.0509 (0.0502)	– 0.012 (0.057)	0.0189 (0.0173)	0.018 (0.018)
$\overline{\text{SAH}}$	0.718*** (0.0387)	0.270*** (0.054)	0.0285*** (0.0103)	0.058*** (0.012)
$\overline{\text{Exercise}}$	– 1.006*** (0.117)	– 0.891*** (0.181)	– 0.178*** (0.0353)	– 0.173*** (0.040)

Table 7 continued

Full sample	Probability of daily smoking		Average number of Cig. per day	
	(1) Probit with RE	(2) Hurdle model-Regime 1	(3) ZTNB with RE	(4) Hurdle model-Regime 2
<i>Income</i>	– 0.114 (0.0795)	– 0.100 (0.167)	0.00779 (0.0115)	– 0.003 (0.015)
<i>N</i>	98,770	98,770	16,221	16,221

Standard error in parentheses

All models also contain the wave dummies

P* value < 0.1, *P* value < 0.05, ****P* value < 0.01

When it comes to specifically diagnosed diseases (Table 6), results have a similar pattern as general health shocks. A newly diagnosed heart disease, cancer and hypertension effectively reduces the participation probability by, respectively – 0.124, – 0.103 and – 0.071, which implies – 74.52%, – 62.1% and – 43.1% declines relative to the average probability of smoking. The effects are similar, but slightly smaller, with heart disease and cancers diagnosed at least 2 years before, suggesting decaying effect over time. Previously diagnosed hypertension shows a significant impact on participation. If we compare the immediate effects of these three types of health shocks, heart disease and cancer show a stronger effect on quitting, whereas high blood pressure shows a smaller but more persistent impact. These findings provide a sound evidence on curative motive for quitting for the elderly, and underscores the need to include these health shock variables in specifying equations for cigarette consumption both at the intensive and extensive margins. The difference in effects of the newly diagnosed diseases compared with previously diagnosed diseases reflects the additional incentive to quit based on new health information or, in other words, decaying effect of old illness.

The effect of health shocks on conditional consumption is very different. We did not find much evidence on reduced consumption after health shocks. Only recently-diagnosed heart disease and hypertension have significant coefficients on conditional demand of cigarettes, implying that once the smoker decides to continue smoking even after a serious health shock including cancer, the number of cigarettes smoked does not tend to decrease. Interestingly, hospital stays induce less smoking, which is possibly caused by smoking bans in hospitals and the resultant incapacitation from smoking. Qualitatively, this result is remarkably similar to that in Wang and Heitjan (2008)'s count regression.

Consistent with recent studies on the general population of adults, we find little evidence that state-level smoke-free laws and anti-smoking sentiment reduces smoking prevalence for the elderly (Bitler et al. 2010; Jones et al. 2015; Carton et al. 2016; Maclean et al. 2016). However, we find some evidence that the state-level anti-smoking sentiment is associated with lower average smoking intensity (Table 7). As evidenced in Nesson (2017, 2017), allowing for endogeneity in anti-smoking sentiment in our model will possibly wipe out this effect too. Note that in their carefully designed study, Carton et al. (2016) did not find any effect of state-level smoke-free policies on smoking behavior of the elderly.

From the perspective of model selection, estimates for the conditional demand of these two models are similar in both significance and size of coefficients. The main difference between the uncorrelated two-part model with the correlated hurdle model lies in the participation equation, as the probability to continue smoking depends on the extent of addiction and smoking intensity, as emphasized by Nesson (2017). As a result, our correlated hurdle model yields a large and statistically significant price elasticity of participation of 0.05%.

Decomposition of Various Factors on Cigarette Consumption

DeCicca and Kenkel (2015) have cautioned against the possibility of over-estimation of price elasticity of total consumption in the literature. Based on a meta-analysis of 86 studies they reported a mean price elasticity of -0.48 . According to them, this is inconsistent with the fact that during 1995–2010 the actual average number of cigarettes smoked decreased by around 25%, but the cigarette prices increased by more than 100%. This underscores the importance of specifying the model correctly generating the price elasticities.

With this in mind, we follow a novel approach to decompose the separate contributions of our three broad groups of factors, viz., cigarette prices, health problems, and state-level smoking bans (including time dummies) in explaining the observed drop in smoking participation and cigarette consumption over our sample period.

To make our results comparable to other studies on smoking behavior, we predicted the decline in smoking prevalence, conditional consumption and (unconditional) consumption counterfactually due to the three groups of factors separately based on our estimated elasticities. The average number of cigarettes per capita (or total consumption) is the product of the probability of smoking and conditional consumption of cigarettes per day. Let $\Pr(\cdot)$, $\lambda(\cdot)$ and $h(\cdot)$ be the probability functions of smoking, conditional consumption and total consumption, respectively. Then the estimated change in these three outcome variables caused by variable j alone between time $t-1$ and t can be calculated as

$$\Delta \hat{\Pr}(\cdot)_{jt} = \Pr(G_{t-1} + \Delta x_{jt} \hat{\beta}_j) - \Pr(G_{t-1})$$

$$\Delta \hat{\lambda}(\cdot)_{jt} = \text{Exp}(D_{t-1} + \Delta x_{jt} \hat{\beta}_j) - \text{Exp}(D_{t-1})$$

$$\Delta \hat{h}(\cdot)_{jt} = \Pr(G_{t-1} + \Delta x_{jt} \hat{\beta}_j) \text{Exp}(D_{t-1} + \Delta x_{jt} \hat{\beta}_j) - \Pr(G_{t-1}) \text{Exp}(D_{t-1})$$

for $j = 1 \dots, J$ and $t = 1 \dots, T$

where $G_{t-1} = \Phi^{-1}(P_{t-1})$, $D_{t-1} = \ln(\lambda_{t-1})$, P_{t-1} is the sample mean of the probability of smoking at time $t-1$, and λ_{t-1} is the sample mean of conditional consumption at time $t-1$. Hence, the relative shares of predicted decline in the probability of smoking, conditional consumption and total consumption due to each of the exogenous variables j are calculated as

$$s_j^p = \frac{\sum_{t=1}^T \Delta \widehat{Pr}(\cdot)_{jt}}{\sum_{j=1}^J \sum_{t=1}^T \Delta \widehat{Pr}(\cdot)_{jt}}$$

$$s_j^\lambda = \frac{\sum_{t=1}^T \Delta \widehat{\lambda}(\cdot)_{jt}}{\sum_{j=1}^J \sum_{t=1}^T \Delta \widehat{\lambda}(\cdot)_{jt}}$$

$$s_j^h = \frac{\sum_{t=1}^T \Delta \widehat{h}(\cdot)_{jt}}{\sum_{j=1}^J \sum_{t=1}^T \Delta \widehat{h}(\cdot)_{jt}}$$

During 1994–2010, the overall smoking prevalence dropped from 23.60 to 12.03% in our sample, i.e., by 11.57 percentage points. The average cigarette price per pack increased from 1.71 to 3.62 (in 1992 dollars) during the same period, which implies a 6.8% reduction, using the methodology described above. In other words, about 59% of the observed reduction in the prevalence rate can be attributed to increased prices. The state-level smoking bans and anti-smoking sentiment variables show hardly any effect on the individual level participation behavior. The time dummies, which can possibly capture some of the unobserved effects of anti-smoking legislations, were statistically insignificant in this regression. On the other hand, coefficients for the health variables are substantial both statistically and economically. Thus, we attribute the other nearly 40% of predicted decline in smoking prevalence to the occurrence of health shocks and health morbidities in our sample population.

Conditional on smoking, the number of cigarettes smoked per day decreased from 19.77 to 13.85, over 1994–2010. As the price of cigarettes per pack increased from 1.71 to 3.62 (in 1992 dollars), a reduction of cigarettes consumption by 1.99 is predicted by the price elasticity of demand of 0.159 using the methodology described above. This is about 34% of the total decrease in smoking intensity. Health problems, unlike in participation decision, were barely found to have any effect on conditional demand. On the other hand, we found a decline of 3.68 cigarettes consumption per capita is due to anti-smoking sentiment and aggregate trends, i.e. 62% of the total decline. We attribute at least part of the unobserved changes in anti-smoking environment in the society, not captured by our two anti-smoking variables, to the time dummies included in the model.

In our sample, unconditional consumption decreased from 4.67 to 1.67 cigarettes per capita during the year 1994–2010. Cumulatively, increased cigarette prices explained 47% of the total drop, higher prevalence of health shocks and morbidities explained 49%, and only 4% of the drop is explained by the variables representing restricted smoking bans and anti-smoking sentiment in our model (see Fig. 6).

To summarize the decomposition results on the two sequential smoking decisions, we find that higher prices helped to reduce both participation and the number of cigarettes smoked quite considerably in this elderly population. Health

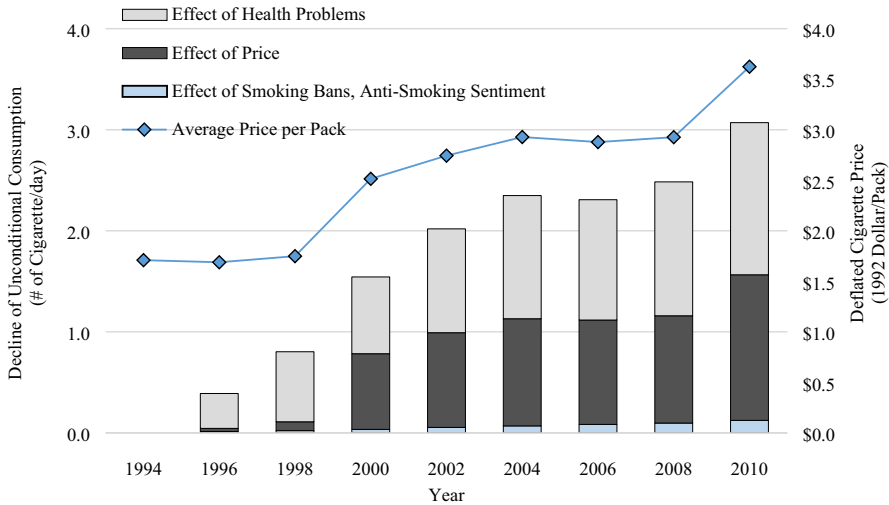


Fig. 6 Decomposition results for unconditional demand for cigarettes. Source: The Tax Burden on Tobacco (2012) and authors' calculation

problems, including health shocks and old health problems induce smokers to quit very significantly, but not the intensity. More restricted smoke-free air law and anti-smoking sentiment including aggregate trends seem to affect very significantly the conditional consumption demand, but not smoking participation. The sum of the individual effects reconciles well with the observed dynamics in smoking prevalence and the amount of cigarette consumption.

Other Covariates

We have included a large number of controls in our regressions. We find individuals with spouses are less likely to smoke, and smoke fewer cigarettes if they do. However, we also found that a smoking spouse increases not only the probability of smoking but also the intensity of cigarettes consumption by the partner. Table 7 shows that compared with females, males carry a higher probability of smoking, and consume more cigarettes conditional on smoking. Using Whites as the reference, Africa-Americans do not have significantly different probability to smoke, and, if they smoke, they smoke less. Hispanics shows a lower propensity and less demand for cigarettes conditional on smoking. Consistent with the finding of Khwaja et al. (2006), income has no statistically significant effect either on the probability of smoking or on the level of consumption, after controlling for individual effects.

With hurdle model, our estimates also suggest that both probability and intensity of smoking are non-linear functions of age. Given the age range in our sample, the probability of participation decreases and smokers tend to smoke less as people age. One interesting point to note is that the predicted probability of participation by age is slightly concave that peaks at age 36, which confirms the findings in previous studies that smokers usually quit around 1940s. The predicted conditional demand as a function of age is also concave, but with a peak at age 50.⁹ In our hurdle model specification, years of schooling and its quadratic term are included. Similar to the effect of aging, higher education reduces the probability of smoking and the conditional consumption.

Conclusions

This study analyzes the effect of cigarette prices/taxes, health shocks and other factors including smoking bans on the smoking participation and consumption behavior among older adults. Specifically, participation and cigarettes demand are estimated both separately and jointly by using two-part and hurdle models with correlated random effects.

We found evidence that higher cigarette prices reduce both smoking prevalence and intensity even for the older adults. The effects are significant and substantial. Specifically, our estimated participation elasticity is large relative to the general consensus for this population, which has ranged from -21 to -56% . Although smaller than the elasticity associated with smoking participation, our results present consistently significant and accurate price elasticity of demand at the extensive margin that ranged from -0.13 to -0.16 . After controlling for health shocks and smoking bans, we find that higher cigarette price explained 59% drop in smoking participation and 34% drop in smoking intensity over our sample period.

When we compare the effect of prices to that of health shocks, we find that the latter factor has a much stronger effect on cessation than on demand conditional on smoking. Individual health shocks have marginal effects on the probability of quitting anywhere from 20 to 53%, whereas heart diseases, cancers and hypertension reduce the probability of participation ranging from 41 to 75%, depending on which disease is diagnosed. On the contrary, health shocks show little impact on the smoking intensity for the smokers who choose to continue smoking even after being diagnosed with smoking-related health problems; only hospitalizations seem to reduce their smoking intensity.

Unconditional consumption, which is a product of the probability of continuing to smoke and the number of cigarettes smoked per day conditional on smoking,

⁹ We should point out that since HRS does include few individuals in their 1930s and 1940s (except due to the inclusion of younger spouses), the peaks of age profiles at 36 or 50 should be taken with due caution.

decreased from 4.67 to 1.67 per capita during years 1994–2010. Our decomposition analysis shows that cumulatively over our sample period, increased cigarette prices and health shocks explain 47% and 44% this decline respectively, and only 4% to more restricted smoking bans and anti-smoking sentiment at the state level. The empirical characterization of smoking behavior of the elderly at both intensive and extensive margins, and their strong response to higher prices established in this paper underscore the continued role of cigarette excise taxes to curb smoking, particularly in low- and middle-income countries.

Acknowledgements This research was supported by the National Institute on Minority Health and Health Disparities, National Institutes of Health (Grant number 1 P20 MD003373). The content is solely the responsibility of the authors and does not represent the official views of the National Institute on Minority Health and Health Disparities or the National Institutes of Health. We thank an anonymous referee for making a number of useful comments that has helped to revise our paper.

Compliance with Ethical Standards

Conflict of interest Both authors Kajal Lahiri and Xian Li declare that they have no conflict of interest with the results of this study.

Ethical standards This article does not contain any studies with human participations or animals performed by any of the authors (University at Albany Institutional Review Board Protocol number 13269-01).

References

- Adda J. and Cornaglia F., (2013). “Taxes, Cigarette Consumption, and Smoking Intensity: Reply,” *American Economic Review*, *American Economic Association*, vol. 103(7), pages 3102–14, December.
- Arcidiacono, P., H. Sieg, and F. Sloan. 2007. Living rationally under the volcano? An empirical analysis of heavy drinking and smoking. *International Economic Review* 48: 37–65.
- Becker, G.S., and K.M. Murphy. 1988. A Theory of Rational Addiction. *Journal of Political Economy* 96 (4): 674–700.
- Bitler, M.P., C.S. Carpenter, and M. Zavodny. 2010. Effects of venue-specific state clean indoor air laws on smoking-related outcomes. *Health Economics* 19: 1425–1440.
- Carton, T.W., M. Darden, J. Leventis, S.H. Lee, and I. Rickett. 2016. Comprehensive Indoor Smoking Bans and Smoking Prevalence. *American Journal of Health Economics* 2 (4): 535–556.
- Chamberlain, G. (1984). Panel data, chapter 22 in Z. Griliches and M. Intriligator, eds., *Handbook of Econometrics*, (North Holland, Amsterdam), 1247–1318.
- Chatterji, P., D. Kim, and K. Lahiri. 2014. Birth weight and academic achievement in childhood. *Health Economics* 23: 1013–1035.
- DeCicca, Philip, Don Kenkel, and Alan Mathios. 2008. Cigarette taxes and the transition from youth to adult smoking: Smoking initiation, cessation, and participation. *Journal of Health Economics* 27 (2008): 904–917.
- DeCicca, Philip, and Logan McLeod. 2008. Cigarette taxes and older adult smoking: Evidence from recent large tax increases. *Journal of Health Economics* 27 (2008): 918–929.
- DeCicca, Philip, and Don Kenkel. 2015. Synthesizing Econometric Evidence: The Case of Demand Elasticity Estimates. *Risk Analysis* 35 (6): 1073–1085.
- Evans, N. William, and C. Matthew Farrelly. 1998. The compensating behavior of smokers: taxes, tar and nicotine. *RAND Journal of Economics* 29 (3): 578–595.
- Farrelly, C.M., W.J. Bray, T. Pechacek, and T. Woollery. 2001. Responses by adults to increases in cigarette prices by sociodemographic characteristics. *Southern Economic Journal* 68 (1): 156–165.
- Fernandez-Val, I. 2009. Fixed effects estimation of structural parameters and marginal effects in panel probit models. *Journal of Econometrics* 150: 71–85.

- Smoking: Duration Analysis of British Data. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, Vol. 164, No. 3(2001), pp. 517–547.
- Grossman, Michael. 1972. On the concept of health capital and the demand for health. *Journal of Political Economics* 80: 223–255.
- Hahn, J., and G. Kuersteiner. 2011. Bias reduction for dynamic nonlinear panel models with fixed effects. *Econometric Theory* 27: 1152–1191.
- Heckman, James J. (1979): Sample Selection Bias as a Specification Error. *Econometrica*, Vol. 47, No. 1 (Jan., 1979), pp. 153–161.
- ImpacTeen.org. (2009). Tobacco Control Policy and Prevalence Data: 1991–2008. In: ImpacTeen.org.
- Jha, P., and R. Peto. 2014. Global Effects of Smoking, of Quitting, and of Taxing Tobacco. *New England of Medicine* 370: 60–68. <https://doi.org/10.1056/nejmra1308383>.
- Jha, P. 2019. Smoking cessation and e-cigarettes in China and India. *BMJ* 367: l6016. <https://doi.org/10.1136/bmj.l6016>.
- Jones, Andrew M. 1994. Health, addiction, social interaction and the decision to quit smoking. *Journal of Health Economics* 13: 93–110.
- Jones, A.M., A. Laporte, N. Rice, and E. Zucchelli. 2015. Do public smoking bans have an impact on active smoking? Evidence from the UK. *Health Economics* 24: 175–192.
- Khawaja, A., F. Sloan, and S. Chung. 2006. Learning about individual risk and the decision to smoke. *International Journal of Industrial Organization* 24 (2006): 683–699.
- Labeaga, J.M. 1999. A double-hurdle rational addiction model with heterogeneity: Estimating demand for Tobacco. *Journal of Econometrics* 93: 49–72.
- Lahiri, Kajal, and Song, G.J: Song. 2000. The effect of smoking on Health using a sequential self-selection model. *Health Economics* 9: 491–511.
- Lewit, M.Eugene, and Douglas Coate. 1982. The potential for using excise taxes to reduce smoking. *Journal of Health Economics* 1: 121–145.
- Liu, Feng. 2010. Cutting through the smoke: separating the effect of price on smoking initiation, relapse and cessation. *Applied Economics* 42: 2921–2939.
- Maclean, C.Johanna, S.Asia Kessler, and S.Donald Kenkel. 2016. Cigarette Taxes and Older Adult Smoking: Evidence from the Health and Retirement Study. *Health Econ.* 25: 424–438.
- Min, Y., and A. Agresti. 2005. Random effect models for repeated measures of zero-inflated count data. *Statistical Modelling* 5: 119.
- Mishra, G.A., S.A. Pimple, and S.S. Shastri. 2012. An overview of the tobacco problem in India. *Indian Journal of Medical and Paediatric and Oncology* 33 (3): 139–145.
- Mundlak, Y. 1978. On the pooling of time series and cross section data. *Econometrica* 46 (1): 69–85.
- Erik, Nesson. 2017. Heterogeneity in Smokers' Responses to Tobacco Control Policies. *Health Economics* 26 (Dec): 206–225.
- Ophem, V.Hans. 2000. Modeling Selectivity in Count-Data Models. *Journal of Business & Economic Statistics* 18 (4): 503–511.
- Orzechowski, W., and R.C. Walker. 2012. *The Tax Burden on Tobacco, Historical Compliance*. Arlington: Orzechowski and Walker.
- Ostbye, T., and D.H. Taylor. 2004. The effect of smoking on 'Years of Healthy Life' lost among middle-aged and older Americans. *Health Services Research* 39 (3): 531–551.
- Pesko, F. Michael, Taurus, A. John, Huang, Jidong, Chaloupka, J. Frank (2016): The Influence of Geography and Measurement in Estimating Cigarette Price Responsiveness. NBER Working Paper 22296.
- Smith, V.Kerry, Donald H. Taylor Jr., Frank A. Sloan, F.Reed Johnson, and William H. Desvousges. 2001. Do smokers respond to health shocks? *Review of Economics and Statistics* 83 (4): 675–687.
- Sundmacher, L. 2012. The effect of health shocks on smoking and obesity. *Europe. J. Health Econ.* 12: 451–460.
- Taurus, J.A. 2006. Smoke-free Air Laws, Cigarette Prices, and Adult Cigarette Demand. *Economic Inquiry* 44: 333–342.
- Taylor, D.H., V. Hasselblad, J. Henley, M.J. Thurn, and F.A. Sloan. 2002. Benefits of smoking cessation for longevity. *American Journal of Public Health* 92: 990–996.
- US Department of Health, Education, and Welfare. Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service. Washington, DC: US Department of Health, Education, and Welfare, Public Health Service; 1964. (Public Health Service Publication No. 1103).

- Wang, H., and D. Heitjan. 2008. Modeling heaping in self-reported cigarette counts. *Statistics in Medicine* 27: 3789–3804.
- Wooldridge, J.M. 1995. Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics* 68 (1): 115–132.
- Yang, J.J., D. Yu, W. Wen, et al. 2019. Tobacco Smoking and Mortality in Asia. *JAMA Network Open* 2 (3): e191474.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.