Household Influences on Smoking Behavior: Evidence from Portugal*

Glaucia Possas da Motta†

September 1, 2025

Abstract

This paper examines how household smoking influences individual smoking behavior in Portugal. Using nationally representative health survey microdata, we estimate models that separate the decision to smoke from the intensity of cigarette consumption. Our results show that co-residing with a smoking spouse is strongly associated with both a higher likelihood of smoking and greater cigarette consumption, consistent with assortative matching and mutual reinforcement. In contrast, living with a smoking parent does not significantly affect smoking participation but increases consumption among smokers, suggesting an influence on intensity rather than initiation. The findings highlight the role of intra-household spillovers in sustaining smoking behavior and suggest that tobacco control policies could be more effective if they target households as a unit—for example through family-based cessation programs or initiatives promoting smoke-free homes.

JEL classification codes: I12

Keywords: Cigarette consumption; household composition; social interactions; double-hurdle model; Portugal

*I thank my advisor, Pedro Pita Barros, for helpful suggestions and discussions. I also gratefully acknowledge funding from Fundação para a Ciência e a Tecnologia (2021.08329.BD). All errors are my own.

†Nova School of Business and Economics. Email: glaucia.possas@novasbe.pt

1 Introduction

It is well known that an individual's behavior is strongly influenced by the habits of other members of the same household. However, most studies on the determinants of smoking have overlooked how household composition and the smoking habits of cohabitants affect one's smoking decisions. For example, a spouse who is a heavy smoker may influence their partner's probability of smoking and the number of cigarettes the partner consumes, or smokers may simply select each other as partners. Similarly, a young adult living with a parent who smokes might be influenced to start smoking or to smoke more (conditional on smoking at all). These household factors could play a significant role in smoking behavior, beyond the individual determinants traditionally studied.

Few studies directly focus on household composition and smoking habits. Notably, Clark and Etilé (2006) and Li and Gilleskie (2021) examine interactions between an individual's smoking and a partner or other household contacts. However, these studies have limitations: (i) they do not control for several individual and household characteristics that have been shown elsewhere to be important determinants of smoking (e.g., health status and education, as in Yen and Jones, 1996; Jones, 1989a; Aristei and Pieroni, 2008), risking omitted variable bias; and (ii) they focus only on the binary decision to smoke, not on the intensity of smoking. Ignoring the consumption level can lead to misspecification, as Jones (1989b) explains, because factors might affect smoking participation and intensity differently.

In this paper, we use data from two waves of the Portuguese National Health Survey (1995 and 1998) to analyze the effects of household composition and the smoking habits of household members on individual smoking decisions. To capture household influence, we construct two key variables: *Smoking Spouse* (whether the individual lives with a spouse/partner who smokes) and *Smoking Parent* (whether the individual lives with a parent who smokes). These allow us to distinguish between a partner's effect — which could reflect both selection into relationships and mutual influence — and a parent's effect, which we expect to reflect mainly social influence (since one does not "choose" one's parents). We then jointly model the decision to smoke and the level of cigarette consumption using a double-hurdle model, including our household composition indicators and a rich set of control variables such as demographics, socioeconomic status, and health indicators.

We find evidence of intra-household spillovers in smoking behavior. An individual with a smoking spouse has a substantially higher probability of being a smoker and also smokes more cigarettes per day on average than if their spouse does not smoke. This points to both a selection effect — smokers tending to partner with smokers — and an influence effect — one spouse's smoking behavior encouraging the other's. Meanwhile, living with parents who

smoke does not significantly increase the likelihood of being a smoker, but it does correspond to a higher cigarette consumption among those who do smoke — indicating an influence effect on smoking intensity, though not on initiation.

At a broad level, these results suggest that anti-smoking policies and programs targeting entire households or social groups could complement individual-focused measures. For example, smoking bans in shared environments (workplaces, public spaces, and by extension encouragement of smoke-free homes) can induce mutual reinforcement: if one household member quits or cuts down due to external restrictions, the others may follow, amplifying the policy's impact. Conversely, interventions aimed solely at individuals might be less effective if that individual remains in a pro-smoking household environment. In our context, a smoker trying to quit may struggle if their spouse continues to smoke at home.

In addition to our main household variables, we uncover effects that are largely consistent with established findings in the literature. Prior studies have similarly documented that individuals with private health insurance are less likely to smoke, yet conditional on smoking tend to consume more cigarettes—a pattern consistent with a moral hazard interpretation. Evidence is also in line with earlier work showing that people with chronic health conditions (e.g., bronchitis, asthma, hypertension, obesity) are less likely to smoke, but those who do smoke despite such conditions consume more cigarettes on average, suggesting a more nicotine-dependent subgroup. Finally, our finding that earlier smoking initiation is associated with higher current consumption echoes the well-documented persistence of addiction and long-term impact of early initiation.

Our main contribution is the empirical evidence showing how the household composition (whether it is spouses or parents who smoke) affects smoking habits, using a double-hurdle model. Previous research, like Jones (1989a), found that living with smokers influences smoking behavior but didn't differentiate the effects of partners from parents. This distinction matters because a partner's influence involves selection and peer effects, whereas a parent's impact (particularly on an adult) tends to be one-directional. Additionally, we make two further contributions. This study is the first to evaluate the factors influencing smoking behavior in Portugal using microdata, thereby contributing to global smoking research. Secondly, by employing retrospective data from the 1995 NHS, we estimate Portugal's historical smoking prevalence from 1970, providing insights into long-term smoking initiation and cessation trends.

The paper is organized as follows. Section 2 reviews the relevant literature on economic models of smoking and social influence in smoking behavior. Section 3 describes the data and provides background on anti-smoking policies and smoking prevalence in Portugal during the sample period. Section 4 lays out the double-hurdle methodology. Section 5 presents the

estimation results and discussion. Finally, Section 6 concludes with policy implications and suggestions for future research.

2 Literature Review

Early economic research on smoking behavior grappled with the question of whether smokers act myopically or rationally. Pollak (1970, 1976) portrayed addictive consumption as largely myopic, with individuals not fully considering future consequences. In contrast, Stigler and Becker (1977) and later Becker and Murphy (1988) developed a rational addiction framework based on stable preferences. In the rational addiction model, current and past consumption are complements: an increase in past consumption raises the marginal utility of current consumption (often termed "adjacent complementarity"). An individual is considered addicted if past consumption significantly increases current consumption. Empirically, this model implies that forward-looking smokers will respond to expected future prices or health costs of smoking. Indeed, many studies have found evidence consistent with rational addiction in smoking. For example, Chaloupka (1991) and Becker et al. (1994) show that smokers take future consequences into account, and that cigarette demand is sensitive to price changes. Similarly, Labeaga (1993), for Spain, finds that higher cigarette prices reduce consumption, indicating that smokers are not entirely myopic.

Another important insight from the literature is that smoking involves a two-stage decision: whether to smoke at all, and if so, how much to smoke. Several studies have argued and shown that the determinants of smoking participation can differ from those of consumption intensity (conditional on smoking). Jones (1989a) introduced and applied a double-hurdle or two-part model to smoking in the UK, finding, for instance, that income, education, and health concerns affected the decision to smoke differently than they affected cigarette demand among smokers. Yen and Jones (1996) and Aristei and Pieroni (2008) likewise used two-part models (with Box-Cox corrections for non-normality in the latter case) for smoking and found it superior to single-equation models like Tobit. These studies typically find that higher education and health awareness reduce the probability of smoking uptake (e.g., Jones, 1989a), while among smokers, factors like income and addiction metrics play a larger role in how much is consumed. For example, Jones (1989a) included a measure of social interaction ("number of other smokers in the household") and found that the presence of multiple smokers in a household increased individual consumption and made quitting harder. This suggests there are social reinforcement effects within households. Our work builds on this idea by explicitly including spouse and parent smoking status as covariates in both the participation and intensity equations.

Beyond individual rationality and two-part decisions, a growing literature examines *social*

and peer influences on smoking. Christakis and Fowler (2008) demonstrated in a social network context that if someone's friend or spouse quits smoking, it significantly increases the probability that the individual will quit, highlighting the diffusion of behavior through social ties. Focusing on households, Clark and Etilé (2006) ("Don't Give Up on Me Baby") analyzed spousal correlations in smoking using British panel data. They found that a spouse's smoking status is strongly correlated with one's own, and argued that much of this correlation is due to assortative mating (i.e., smokers tending to marry smokers) and shared environment, though there was also some evidence of learning or influence between partners over time (for instance, one spouse's successful quitting was associated with a higher chance of the other quitting). In a more recent study, Li and Gilleskie (2021) modeled the smoking outcomes of pairs of individuals (spouses, friends, siblings, etc.) as a simultaneous game. Their analysis suggests that one person's smoking can causally influence the other's, although identification is challenging. They also emphasize, as we do, that unobserved factors lead smokers to partner with smokers, which is why separating "influence" from "selection" requires careful thought or additional data.

Intergenerational influence is another relevant aspect: children (even in adulthood) may be influenced by parental smoking behavior observed while growing up. Public health studies have consistently found that having a parent who smokes is associated with a higher risk of smoking initiation for the child (e.g., higher youth smoking uptake in households with smoking parents). This effect can carry into adulthood in the form of ingrained attitudes or nicotine exposure. That said, if an adult is still living with a parent, the dynamics might differ; it could be that those who have not left the parental home by adulthood are systematically different (selection) or that parental smoking mainly normalizes higher consumption once the individual is a smoker (influence on intensity rather than initiation). These nuances have not been fully explored in prior economic literature, which motivates our distinction between spouse and parent influences.

In summary, previous studies indicate that *(a)* smokers show some forward-thinking tendencies and their behavior aligns with rational addiction models, *(b)* smoking involvement and intensity should be analyzed separately using double-hurdle models to minimize bias, and *(c)* the social environment, particularly within households, significantly affects smoking behavior through group selection and peer pressure. Our research integrates these views by employing a double-hurdle model (addressing point b and incorporating addiction metrics per point a) and explicitly factoring in household smoking variables to assess spousal and parental impact (addressing point c). To our knowledge, no prior research has concurrently explored the effects of spousal and parental smoking on both smoking decisions and intensity, specifically in a Portuguese context. Our results will broaden the understanding of the interplay between

social context and individual addiction in shaping smoking behaviors.

3 Data and Context

We use microdata from the Portuguese National Health Survey (Inquérito Nacional de Saúde – INS) for the years 1995, 1998, 2005, 2014, and 2019. These nationally representative cross-sectional surveys contain detailed information on individuals' health behaviors (including smoking), health status, and socio-demographic characteristics.

Our outcome variables are (i) an indicator for whether the individual is a current smoker (defined as smoking daily or occasionally at the time of the survey), and (ii) the number of cigarettes smoked per day (zero for non-smokers). The key explanatory variables of interest—*Smoking Spouse* and *Smoking Parent*—can only be constructed in the 1995 and 1998 waves, which collected detailed information on household members and their smoking status. Later waves lack relationship identifiers, so household composition effects cannot be directly measured. However, they remain valuable for estimating smoking participation and intensity, testing robustness, and documenting long-run prevalence trends.

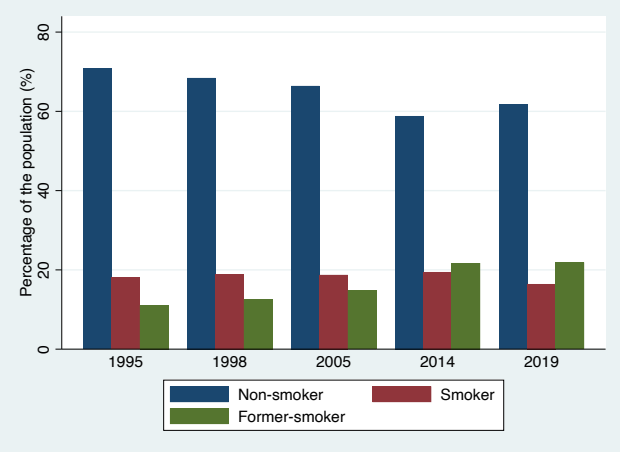
3.1 Smoking prevalence in Portugal

Before turning to the household composition analysis, we briefly document national trends in smoking prevalence and intensity. Figure 1a shows that overall smoking prevalence declined only modestly between 1995 and 2019, falling from about 19% to 16.5%. In comparative terms, this is a relatively high level: Portugal's prevalence remains above that of many Western European countries, where smoking rates are closer to 12–14% in recent years. By contrast, cigarette intensity among smokers declined sharply (Figure 1b), reflecting reductions in daily consumption among continuing smokers.

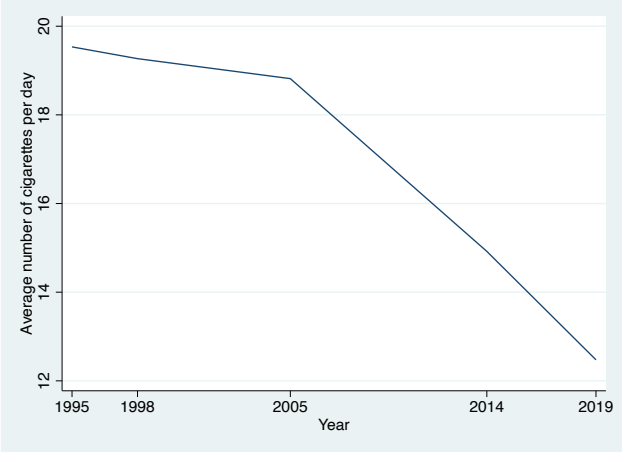
Retrospective data from the 1995 wave allow us to reconstruct historical trends since the mid-1970s (Figure 2). These series indicate that prevalence declined only slowly over several decades, while cessation increased markedly. The gender dimension is particularly striking: women accounted for less than 15% of smokers in the 1970s, but nearly 40% by 2014 (Figure 3). This convergence reflects rising female participation and slower declines in male prevalence.

Taken together, these descriptive facts highlight two key features of the Portuguese smoking epidemic: (i) prevalence has remained persistently high, despite some decline in intensity; and (ii) gender convergence has amplified the challenge for public health. Against this backdrop, examining the household context helps explain why prevalence is stubbornly high: intra-household influences may reinforce smoking and slow down quitting dynamics.

Figure 1: Percentage of the population of Smokers, former smokers, and non-smokers and average cigarette consumption time series in Portugal



(a) Percentage of the population of Smokers, former smokers, and non-smokers



(b) Smokers' average cigarette consumption

Figure 2: Smoking prevalence before 1995 - a synthetic approach

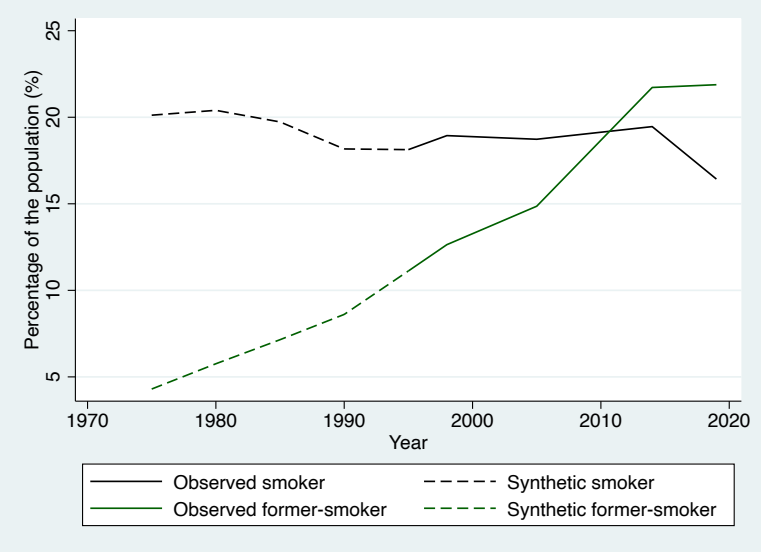
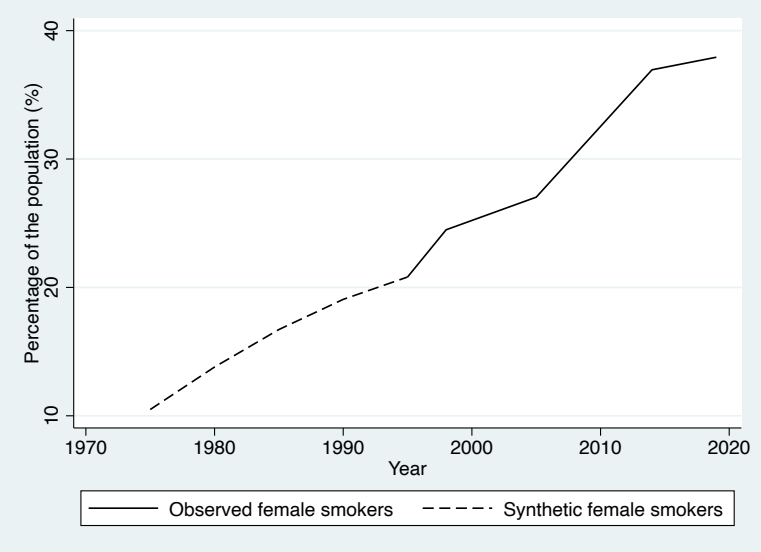


Figure 3: Smoking Prevalence per Gender (Base year – 1995)



3.2 Descriptive statistics

Across survey years, around 17–20% of Portuguese adults report being current smokers, with average daily consumption among smokers declining from more than 15 cigarettes in the mid-1990s to under 10 by 2019. Women represent just over half of the sample, and over time smoking has become increasingly concentrated among women, the less educated, and those without private health insurance. Structural shifts are also visible in the sample: the population ages considerably, the prevalence of obesity rises from single to double digits, and the share with private health insurance increases from below 5% in 1995 to over 20% by 2019. Household smoking variables are relatively rare: in 1995 and 1998, about 4–5% of respondents lived with a smoking spouse and 2–3% with a smoking parent.

Detailed descriptive statistics, including breakdowns by demographics, education, and health indicators, are provided in Appendix A. These confirm the broad gradients noted above and illustrate how household smoking measures are only available in the first two survey waves, which is why our intra-household analysis focuses on 1995 and 1998.

4 Methodology

We model individuals' smoking behavior using a double-hurdle model (Cragg, 1971; Jones, 1989b) to account for the two-tier decision process: (1) the decision to smoke at all (participation), and (2) the decision on how much to smoke (intensity), given participation. This approach allows the explanatory variables to have different effects on each decision and ac-

commodates the pile-up of zeros in consumption data (non-smokers).

Formally, let the latent propensity to smoke for individual i be $y^*_{1,i}$, and the latent desired consumption (cigarettes per day) be $y^*_{2,i}$. The double-hurdle model is specified as:

$$y^*_{1,i} = x_{1,i}\beta_1 + u_{1,i}, \tag{1}$$

$$y^*_{2,i} = x_{2,i}\beta_2 + u_{2,i}, \tag{2}$$

with the observed outcome (cigarettes smoked per day) defined by:

$$y_i = \begin{cases} y^*_{2,i} & \text{if } y^*_{1,i} > 0 \text{ (the individual is a smoker);} \\ 0 & \text{if } y^*_{1,i} \leq 0 \text{ (the individual does not smoke).} \end{cases} \tag{3}$$

In other words, the individual must clear two hurdles to have a positive y_i: first, $y^*_{1,i} > 0$ corresponds to the binary decision $D_i = 1$ (smoker vs. non-smoker); second, conditional on being a smoker, the observed consumption y_i is given by the continuous latent $y^*_{2,i}$ (assumed to be positive if defined). If $D_i = 0$, then $y_i = 0$. The vectors $x_{1,i}$ and $x_{2,i}$ may overlap but need not be identical – one can include certain covariates only in one equation if theory suggests they affect only participation or intensity. In our specification, we include most controls in both equations, but, for example, variables like "age started smoking" are only defined for smokers and thus enter only the intensity equation.

We list the main variables included:

- $y_{1,i}$: binary indicator =1 if individual i is a current smoker, 0 otherwise.
- $y_{2,i}$: number of cigarettes smoked per day by individual i.
- $x_{1,i}$: covariates in the participation equation – includes age, age squared, gender, marital status, education dummies, region dummies, health indicators (insurance, chronic conditions), *Smoking Spouse* and *Smoking Parent*.
- $x_{2,i}$: covariates in the intensity equation – includes a similar set as $x_{1,i}$, but instead of age squared we often just include age (if non-linear effects are insignificant), and includes additional variables defined only for smokers (such as years of smoking or age of initiation, to capture addiction length).

We assume the error terms $(u_{1,i}, u_{2,i})$ are jointly normally distributed:

$$(u_{1,i}, u_{2,i}) \sim N(0, \Sigma),$$

with $\Sigma = \begin{pmatrix} 1 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$. We normalize the variance of $u_{1,i}$ to 1 (since $y^*_{1,i}$ is not observed on

a natural scale) and allow $\sigma_{12} = \sigma_{21}$ to be non-zero, i.e., the two equations' errors can be correlated with correlation $\rho = \sigma_{12}/\sigma_2$. If $\rho \neq 0$, there is correlation between the unobservable factors affecting participation and those affecting consumption intensity, which justifies the use of a full likelihood approach rather than estimating two separate models. If $\rho = 0$, the model effectively reduces to a two-part model (independent equations).

The log-likelihood for the double-hurdle (following Engel and Moffatt, 2014 and others) can be written as the sum of contributions from observations with $y_i = 0$ (non-smokers or ex-smokers) and those with $y_i > 0$ (current smokers):

$$\log L = \sum_{y_i=0} \ln \Big[1 - \Psi\big(x_{1,i}\beta_1,\ \frac{x_{2,i}\beta_2}{\sigma_2},\ \rho\big)\Big] \ + \sum_{y_i>0} \ln \Big\{ \Psi\Big(\frac{x_{1,i}\beta_1 + \frac{\rho}{\sigma_2}(y_i - x_{2,i}\beta_2)}{\sqrt{1-\rho^2}}\Big) - \ln(\sigma_2) + \ln\Big[\phi\Big(\frac{y_i - x_{2,i}\beta_2}{\sigma_2}\Big)\Big]\Big\},$$

where $\Psi(\cdot,\cdot,\rho)$ is the bivariate standard normal CDF with correlation ρ, and $\phi(\cdot)$ is the standard normal PDF. The first term corresponds to the probability of observing a zero outcome (the individual did not smoke), which occurs when the latent utility to smoke is below zero *and* the latent consumption is below the threshold (the minimum of the two hurdles is ≤ 0). The second term is for smokers ($y_i > 0$), which includes the joint probability of clearing both hurdles and the density of the observed consumption.

We estimate the model via maximum likelihood, using starting values from separate probit and truncated regression estimates. Standard errors are clustered by survey year (since observations within a year share the same policy environment). We also estimate models separately for each year to see how marginal effects change over time.

Our focus will be on the marginal effects of key variables, particularly the presence of a smoking spouse or parent. For ease of interpretation, we report marginal effects rather than raw coefficients for the probit part. For continuous outcomes, we present effects evaluated at the mean of covariates. All results tables present marginal effects: for the participation equation, the change in the probability of smoking (in percentage points), and for the consumption equation, the change in the expected number of cigarettes (conditional on being a smoker). Statistical significance is indicated by asterisks.

5 Results

Table 1 and Table 2 display the estimated marginal effects for the smoking participation and intensity equations, respectively, across the five survey waves. We first discuss our primary variables of interest (household composition), then turn to other covariates and how their

effects align with expectations.

Table 1: Marginal effects: Participation equation.

Variables	**1995**	**1998**	**2005**	**2014**	**2019**
Sex (Women)	0.887	0.781	0.075***	0.098***	0.097***
	(18.510)	(14.700)	(0.015)	(0.017)	(0.020)
Age	-0.017***	-0.014***	-0.051***	-0.061***	-0.051***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)
Age Square	0.000***	0.000***			
	(0.000)	(0.000)			
Single	0.093***	0.138***	0.121***	0.103***	0.101***
	(0.027)	(0.035)	(0.015)	(0.018)	(0.021)
Separated/Divorced	0.101**	0.088**	0.097***	0.122***	0.116***
	(0.046)	(0.041)	(0.021)	(0.018)	(0.022)
Widow	-0.015	0.008	0.032	0.102***	0.069**
	(0.027)	(0.024)	(0.022)	(0.025)	(0.032)
Level of Schooling 1	-0.012	-0.008	-0.003	0.032	
	(0.023)	(0.022)	(0.015)	(0.025)	
Level of Schooling 2	-0.028	-0.019	0.001	0.002	
	(0.025)	(0.025)	(0.021)	(0.030)	
Level of Schooling 3	-0.051	-0.018	-0.062***	-0.065**	
	(0.035)	(0.037)	(0.023)	(0.031)	
Centro	0.002	0.028	-0.032*	0.021	0.031
	(0.021)	(0.021)	(0.018)	(0.023)	(0.031)
Lisboa e Vale do Tejo	-0.037**	-0.008	0.003	0.014	-0.037
	(0.016)	(0.015)	(0.017)	(0.023)	(0.028)
Alentejo	0.037*	0.020	0.018	0.043*	0.011
	(0.022)	(0.020)	(0.017)	(0.025)	(0.034)
Algarve	-0.013	-0.045**	-0.007	0.031	0.017
	(0.020)	(0.018)	(0.017)	(0.023)	(0.028)
Açores			0.059***	0.059**	0.070**
			(0.017)	(0.025)	(0.029)
Madeira			0.104***	-0.006	0.009
			(0.019)	(0.025)	(0.030)
Health Insurance	-0.031	-0.010	-0.061***	-0.029*	-0.034*
	(0.024)	(0.028)	(0.016)	(0.017)	(0.020)
Bronchitis	-0.020	-0.037	-0.007	0.008	0.008
	(0.022)	(0.023)	(0.023)	(0.028)	(0.033)
Asthma	-0.081***	-0.005	-0.053***	0.001	0.032
	(0.031)	(0.024)	(0.021)	(0.033)	(0.038)
Hypertension	-0.034*	-0.038**	-0.061***	-0.047***	-0.037**

Continued on next page

Variables	1995	1998	2005	2014	2019
	(0.018)	(0.015)	(0.011)	(0.015)	(0.018)
Underweight	-0.018	-0.009	0.083*	0.177**	0.024
	(0.045)	(0.046)	(0.044)	(0.071)	(0.072)
Pre-obesity	-0.097***	-0.078***	-0.116***	-0.114***	-0.096***
	(0.015)	(0.014)	(0.010)	(0.014)	(0.017)
Obesity class I	-0.155***	-0.147***	-0.181***	-0.169***	-0.172***
	(0.019)	(0.018)	(0.014)	(0.019)	(0.023)
Obesity class II	-0.162***	-0.101**	-0.190***	-0.231***	-0.279***
	(0.039)	(0.045)	(0.031)	(0.037)	(0.050)
Obesity class III	-0.138	-0.112	0.001	-0.233***	-0.329***
	(0.124)	(0.098)	(0.060)	(0.082)	(0.077)
Mother within 1 year	-0.017	0.053			
	(0.044)	(0.071)			
Mother within 3 years	0.005	0.025			
	(0.038)	(0.050)			
Mother within 5 years	0.000	-0.019			
	(0.030)	(0.028)			
Smoking spouse	0.318***	0.306***			
	(0.065)	(0.067)			
Smoking parent	0.571	0.416			
	(96.480)	(85.660)			
Observations	11248	12207	11329	5684	4046

Notes: Marginal effects; standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Marginal effects: quantity equation.

Variables	1995	1998	2005	2014	2019
Sex (Women)	-6.364***	-6.641***	-10.110***	-6.450***	-10.280***
	(0.442)	(0.414)	(0.472)	(0.514)	(0.850)
Age	-0.175**	-0.172**	1.638***	1.067***	2.219***
	(0.079)	(0.078)	(0.163)	(0.194)	(0.324)
Age Square	-0.013***	-0.013***			
	(0.001)	(0.001)			
Single	5.113***	5.531***	-2.694***	-1.042*	-4.038***
	(0.572)	(0.572)	(0.490)	(0.556)	(0.877)
Separated/Divorced	6.062***	6.695***	0.474	0.261	-3.169***
	(0.940)	(0.832)	(0.755)	(0.704)	(1.041)
Widow	9.936***	10.590***	-2.372*	0.580	-1.230
	(1.183)	(1.188)	(1.211)	(1.053)	(1.248)
Level of Schooling 1	1.525*	1.183	0.518	-0.283	

Continued on next page

Variables	1995	1998	2005	2014	2019
	(0.846)	(0.987)	(0.774)	(1.037)	
Level of Schooling 2	0.926	0.361	-1.952**	-2.106*	
	(0.886)	(1.015)	(0.891)	(1.118)	
Level of Schooling 3	1.361	-0.066	-1.860*	-1.565	
	(1.025)	(1.113)	(0.979)	(1.177)	
Centro	-2.353***	-2.478***	0.989	2.711***	-0.033
	(0.558)	(0.519)	(0.725)	(0.733)	(1.000)
Lisboa e Vale do Tejo	0.030	-0.356	3.243***	1.189	1.421
	(0.463)	(0.427)	(0.648)	(0.731)	(0.965)
Alentejo	-0.605	-0.565	1.978***	1.408*	-0.861
	(0.580)	(0.553)	(0.654)	(0.794)	(1.108)
Algarve	1.283**	-0.050	3.807***	0.855	1.034
	(0.599)	(0.547)	(0.634)	(0.735)	(0.920)
Açores			4.356***	2.596***	2.945***
			(0.634)	(0.760)	(0.977)
Madeira			4.059***	2.172***	0.668
			(0.703)	(0.794)	(0.991)
Health Insurance	-0.244	1.165*	0.366	-0.741	1.115*
	(0.731)	(0.622)	(0.662)	(0.508)	(0.638)
Bronchitis	-0.523	2.214**	1.193	1.916**	0.949
	(0.869)	(1.102)	(1.071)	(0.975)	(1.138)
Asthma	0.848	-0.456	0.057	-1.324	-4.874***
	(1.259)	(0.816)	(0.859)	(1.082)	(1.169)
Hypertension	-3.176***	-1.009	1.207*	0.831	1.528**
	(0.700)	(0.639)	(0.617)	(0.629)	(0.686)
Underweight	1.270	3.208***	0.827	-0.385	-0.653
	(1.019)	(0.962)	(0.955)	(1.084)	(1.604)
Pre-obesity	0.019	-0.098	2.772***	1.079*	3.626***
	(0.402)	(0.384)	(0.485)	(0.560)	(0.806)
Obesity class I	0.692	1.323*	6.094***	3.060***	7.047***
	(0.839)	(0.786)	(0.834)	(0.878)	(1.378)
Obesity class II	2.490	-3.750**	7.042***	4.766***	10.270***
	(1.932)	(1.760)	(1.745)	(1.769)	(2.735)
Obesity class III	-5.417***	-6.065***	0.251	1.949	15.150***
	(1.385)	(1.431)	(2.342)	(3.547)	(3.563)
Mother within 1 year	-0.885	-2.267***			
	(0.937)	(0.793)			
Mother within 3 years	-1.674**	-0.781			
	(0.772)	(0.751)			
Mother within 5 years	-0.250	-0.483			
	(0.719)	(0.675)			

Continued on next page

Variables	1995	1998	2005	2014	2019
Smoking spouse	6.923***	6.464***			
	(0.491)	(0.524)			
Smoking parent	2.950**	1.586			
	(1.293)	(1.494)			
Years being a smoker	1.408***	1.443***			
	(0.028)	(0.027)			
Age Start			-0.311***	-0.210***	-0.129***
			(0.035)	(0.036)	(0.040)
mill ratio		-2.702*	-13.350***	-5.692***	-18.090***
		(1.509)	(1.658)	(1.744)	(3.322)
Observations	11248	12207	11329	5684	4046

Notes: Marginal effects; standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Household Composition Effects – Spouse and Parent. Our main results confirm that living with other smokers significantly influences individual smoking behavior. Having a smoking spouse is associated with a higher likelihood of being a smoker and a higher cigarette consumption. In 1995, for example, living with a smoking spouse increased an individual's probability of smoking by about 31.8 percentage points ($p<0.01$), and in 1998 by 30.6 points (both effects are large, reflecting the strong assortative matching and influence between partners). In later years, due to data limitations, we could not estimate this effect, but it is reasonable to assume a smoking spouse would continue to have a substantial impact. The dual significance of spouse smoking on both participation and quantity suggests a *selection effect*: smokers likely tend to partner with smokers (which raises the probability equation), and once partnered, they may reinforce each other's consumption habits (raising the intensity). By contrast, living with a smoking parent does not significantly affect whether an individual (who is presumably a young adult in our sample) smokes – the marginal effect on participation is near zero and insignificant. However, conditional on smoking, if one's parent smokes, the individual consumes roughly 6–7 more cigarettes per day (significant in 1995 and 1998). This indicates an *influence effect on intensity*: a household culture of smoking (parents smoking at home) might encourage heavier use among those who do smoke, even if it doesn't draw non-smokers into smoking in the first place. These results broadly match our hypotheses: a spouse's smoking has both matching and peer influence components, whereas a parent's smoking mainly provides an additional social cue or tolerance for smoking that affects how much one smokes.

It is important to note that these are associations – as discussed earlier, we cannot fully disentangle causality. The spouse effect in particular likely captures an element of selection (people often marry those with similar habits). What we observe as a spouse "influence" on

consumption could partly be that smokers who have not quit are married to fellow smokers, and they might jointly decide to smoke at home, etc. Unfortunately, without longitudinal data on smoking initiation and partner selection, we treat these effects as reduced-form indications of intra-household correlation. The large magnitude and statistical strength, however, underscore that ignoring household smoking status would greatly underestimate one's own smoking propensity and intensity.

Gender and the Rise of Female Smoking. In 1995 and 1998, we find no significant difference between men and women in the probability of being a smoker, holding other factors constant (the marginal effect of female is small and not statistically significant). By 2005, a noteworthy change occurs: being female is associated with a *higher* probability of smoking (+7.4 percentage points in 2005, +9.8 points in 2014, and +9.7 points in 2019, all $p<0.01$). This result mirrors the descriptive trends – historically, Portuguese women had much lower smoking rates than men, but in recent decades female smoking increased. Our findings imply that among later cohorts, women are actually more likely to smoke than men once we control for age, education, etc. This could be due to social changes (greater workforce participation, changing norms around women smoking). It corroborates the evidence reported by Leite et al. (2019) that the female share of smokers rose substantially. Interestingly, despite higher uptake, women smokers tend to consume fewer cigarettes per day than men. Across all years, the marginal effects in the consumption equation for the female indicator are negative (and statistically significant), indicating that, conditional on smoking, women smoke fewer cigarettes than men (on average, a female smoker consumes about 1–2 fewer cigarettes per day than a male smoker with similar characteristics). This aligns with common findings that even when women's smoking prevalence catches up, their intensity often remains slightly lower, possibly due to biological or behavioral differences, or different preferences (e.g., women might more often be light or social smokers).

Age Effects. Age and age-squared (when included) show the typical inverted-U relationship in the participation equation: the probability of smoking increases with age in early adulthood and then decreases for older individuals. In 1995/1998, the age effect is significantly negative while age-squared is positive, implying that smoking prevalence peaks somewhere in middle age. In 2005 onward, we simplified the specification (the quadratic term was dropped as it was not improving fit once cohort effects were absorbed by year dummies). Essentially, younger individuals (teens and 20s) are less likely to smoke than those in their 30s, but then smoking rates decline again for those in their 50s and above (likely due to quitting or cohort differences). In the consumption equation, we found it more meaningful to include metrics of smoking history (like years smoking or age of start) rather than a raw age effect – see "Addiction and history" below for discussion.

Marital Status. Marital status can proxy lifestyle differences and stress factors. We used "Married" as the reference category. The results indicate that single individuals are significantly *more likely* to smoke than married individuals, by about 10–14 percentage points (across all years, Single is positive and significant). Separated/divorced individuals also have higher smoking probabilities than married ones (by roughly 9–12 points). Widowed individuals show no clear pattern or a slight increase in later years. These findings suggest that being married might have a protective effect (possibly due to spousal influence to stay healthy, or because married people, especially with children, have more incentives to quit). Conversely, single and divorced people may have social or stress factors that lead to higher smoking rates. These patterns are consistent with many epidemiological findings and also align with Aristei and Pieroni (2008) who found Italian single-adult households had different smoking behavior. In our intensity equation, marital status effects are generally not significant (once someone is a smoker, how much they smoke doesn't depend strongly on marital status, except perhaps widowed smokers in 2014 smoked slightly more). This implies marital status mainly affects the decision to smoke, not how heavy of a smoker one is, again hinting at social influence on initiation/cessation.

Education. Education is a strong predictor of smoking behavior. We see that higher education levels consistently reduce the probability of being a smoker. For example, in 1995 a person with secondary or higher education had a lower probability of smoking (the omitted category is no schooling; the negative signs on the schooling dummies confirm the gradient). By 2014, the marginal effect for having a college education is about -6.5 percentage points on smoking participation (significant). These results mirror well-known patterns that more educated individuals are less likely to start or continue smoking, likely due to better health knowledge and different social norms. Jones (1989a) and Yen and Jones (1996) found similar negative education effects. In our consumption equation, we expected that among smokers, higher education might correlate with smoking slightly more (since higher-income, more-educated smokers can afford more cigarettes, or one might hypothesize they are "social smokers" who smoke less – it's ambiguous). In our data, the effect of education on consumption was not very large or consistent. We found a slight positive effect only in one case: in 1995, having primary education (as opposed to none) increased daily cigarettes among smokers by a small amount. But generally, the education of smokers did not significantly affect how much they smoked – perhaps because heavy smoking cuts across education levels among those who are addicted.

Health Insurance. Having private health insurance is associated with a significantly lower probability of being a smoker in the 2005 and 2014 waves (e.g., about -6.1 percentage points in 2005). In earlier years it was negative but not significant (and note that very

few people had private insurance in 1995, only 4%). By 2014, about 17% had some private health plan, and those individuals were 2.9 points less likely to smoke ($p<0.10$). This likely reflects a selection effect: higher-income or more health-conscious people both buy insurance and avoid smoking. It could also be that insurers indirectly encourage healthier behavior. However, conditional on smoking, those with insurance actually consumed more cigarettes (the coefficient in the intensity equation is positive in multiple years, significantly so in 1998 and 2014). This paradoxical pattern is what we referred to as a "moral hazard" effect: once insured, individuals might feel more shielded against medical costs of smoking, or it might simply be that wealthier smokers can afford to smoke more. By 2014, smokers with insurance consumed about 1.1 more cigarettes per day than comparable smokers without insurance (significant at 5%). This combination – insurance deterring some from smoking but leading others to smoke more – is intriguing, though the net effect is still beneficial (fewer people smoke). We mention in passing that health insurance typically does not cover certain smoking-related illnesses fully (e.g., life insurance premiums are higher for smokers, and some cancers might not be fully covered), so the moral hazard interpretation has its limits. Nonetheless, the pattern is consistent with the idea that risk mitigation (insurance) can lead to riskier behavior among those already engaging in it.

Health Conditions. We included dummy variables for chronic respiratory conditions (bronchitis or asthma) and for hypertension. These conditions can be caused or exacerbated by smoking, so there is endogeneity concern. Our results likely reflect a mix of two effects: those who have these health issues may quit smoking (so we see fewer smokers among them – a negative association in participation), but the ones who continue to smoke despite illness might be heavier smokers (positive association in intensity). Indeed, having bronchitis or asthma significantly reduces the probability of smoking in 1995 and 2005 (by about 2–4 points), although in other years the effect was not significant. Hypertension shows a stronger and consistent negative effect on smoking participation: in all survey years, individuals with hypertension were about 3–10 percentage points less likely to be smokers, controlling for other factors. This suggests that people diagnosed with high blood pressure often quit or never start smoking, which aligns with medical advice and possibly reverse causality (smoking contributes to hypertension, so many smokers may develop it; however, the ones we observe with hypertension who still don't smoke could be those who quit after diagnosis). Regarding intensity, bronchitis/asthma had no significant effect on number of cigarettes consumed among the remaining smokers (all coefficients were statistically zero). Hypertension had an ambiguous effect: in 1995 it was associated with slightly lower consumption (perhaps those hypertensive smokers tried to cut down), but in 2005 it was associated with higher consumption. None of these were very large or consistent, and we suspect sample differences or unobserved hetero-

geneity (the "hard-core" smokers with hypertension might be particularly nicotine-dependent in some years). Interestingly, Yen and Jones (1996) found similarly mixed results for health conditions in the UK – health issues strongly deterred participation but had unclear effects on intensity, as smokers who persist despite illness might be atypical. In summary, our findings confirm that chronic health problems are a strong deterrent to smoking uptake (or a motivator to quit), which is reassuring from a public health perspective.

Body Mass Index (BMI). Obesity and smoking have a complex relationship. We included categories for BMI (using WHO classifications: underweight, normal, pre-obese, class I/II/III obese). We find that higher BMI categories are generally associated with a lower probability of smoking. For instance, individuals classified as obese (especially Class I obesity, BMI 30–34.9) were significantly less likely to smoke in all years. The marginal effect for Class I obesity is around -2 to -4 percentage points on smoking participation ($p<0.01$). This might reflect lifestyle or preference differences – some people substitute food for cigarettes or vice versa, and non-smokers tend to have higher BMI on average. Yen and Jones (1996) similarly noted an inverse relationship between smoking and obesity prevalence. On the consumption side, our results for BMI were inconclusive; none of the obesity categories had a robust effect on number of cigarettes smoked. It appears that BMI influences whether one smokes (perhaps via lifestyle and health consciousness), but among smokers, body weight does not predict how many cigarettes they consume.

Motherhood and Young Children. A novel aspect of our study is examining whether having young children in the household affects smoking behavior, especially for women. We created indicators for mothers who had given birth within the last 1 year, 3 years, or 5 years. These variables are, of course, only applicable for women; we include them mainly in the intensity equation (since being a recent mother might affect how much a mother smokes if she hasn't quit). The results show no significant effect of recent motherhood on the probability of being a smoker – likely because many women who become pregnant or have young babies choose to quit entirely (and thus exit the smoker category, which is already captured in the participation decision). However, conditional on continuing to smoke, there are signs that mothers of young children smoke less. In 1998, being a mother of an infant (within 1 year) is associated with a reduction in daily cigarettes (the coefficient corresponds to about 1.2 fewer cigarettes per day, significant at 5%). Similarly, in 1995, being a mother of a child under 3 years old is associated with lower consumption. Although these specific estimates vary by year and are not always significant, the pattern aligns with intuition and prior evidence: when there are young children at home, parents (especially mothers) often reduce or regulate their smoking, possibly by smoking outdoors, less frequently, or quitting. This is consistent with findings in Aristei and Pieroni (2008) for Italy, where the presence of young children in the

household was linked to lower tobacco expenditure. Our results lend support to the idea that family formation is a critical moment for smoking cessation efforts – many people successfully quit around the time of having children, which is also a rationale for policies targeting smoking in families (e.g., pediatrician interventions, smoke-free home campaigns).

Addiction and Smoking History. For smokers, we included measures of addiction intensity. In 1995 and 1998, the surveys recorded how long the person had been smoking (years of smoking). We find that each additional year of smoking history increased daily consumption by about 0.14 cigarettes (in 1995 and 1998, significant at 1%). This may seem small per year, but over, say, 10 years it's 1.4 more cigarettes per day – indicating that longer exposure leads to heavier smoking, consistent with addiction deepening over time. In 2014 and 2019, unfortunately, the surveys did not collect age of initiation or years smoked, so instead we included the age the person started smoking (available for 2005 and some earlier data). The coefficient on "Age started smoking" is significantly negative: the older someone was when they began smoking, the fewer cigarettes they smoke now. In other words, those who start very young (e.g., in their teens) tend to end up as heavier smokers than those who start in their twenties. This aligns with the medical understanding that early initiation leads to stronger nicotine dependence. Our findings here echo the concept of "adjacent complementarity" from rational addiction theory (Becker and Murphy, 1988): past consumption (starting earlier, smoking for more years) raises current consumption. It also matches the empirical evidence in Yen and Jones (1996) (who included a "peak consumption" variable to capture addiction stock and found it significant for both participation and conditional consumption). In practical terms, this underscores the importance of preventing youth smoking initiation – not only does it reduce the number of smokers, but it also averts the emergence of heavy smokers later on.

To summarize the regression results: All the above factors collectively explain smoking behavior reasonably well. The household smoking variables remain impactful even after accounting for everything else, which is a central result of this paper. Moreover, our model's structure seems validated by the data: the significant correlation we found between the two error terms (not reported in tables for brevity) indicates the double-hurdle model is appropriate – for instance, unobservables like an individual's general risk tolerance likely affect both the decision to smoke and how much to smoke, and our likelihood-ratio tests rejected the independence assumption ($\rho = 0$). This justifies modeling the two hurdles jointly rather than separately.

6 Conclusion

This study provides new evidence on how household composition shapes smoking behavior. Using nationally representative survey microdata for Portugal, we find that co-residing with a smoking spouse is strongly associated with both the probability of smoking and the number of cigarettes consumed, while co-residing with a smoking parent increases intensity among smokers but does not affect participation. These patterns are consistent with assortative matching between smokers as well as intra-household influence.

The results underscore that tobacco control policies can be more effective when they explicitly target social units rather than individuals alone. Family-based cessation programs, joint interventions for couples, and campaigns promoting smoke-free homes may leverage intra-household spillovers to achieve larger reductions in smoking. Conventional measures aimed only at individual smokers risk being undermined if the household environment continues to reinforce smoking behavior.

More broadly, the persistence of smoking in Portugal—despite declining intensity among smokers—suggests that addressing household influences is central to reducing prevalence. Policies that acknowledge the family context, and not just individual choice, may thus play a key role in sustaining progress in tobacco control.

References

Aristei, D. and Pieroni, L. (2008). A double-hurdle approach to modelling tobacco consumption in italy. *Applied Economics*, 40(19):2463–2476.

Becker, G. S., Grossman, M., and Murphy, K. M. (1994). An empirical analysis of cigarette addiction. *The American Economic Review*, 84(3):396–418.

Becker, G. S. and Murphy, K. M. (1988). A theory of rational addiction. *Journal of Political Economy*, 96(4):675–700.

Chaloupka, F. (1991). Rational addictive behavior and cigarette smoking. *Journal of Political Economy*, 99(4):722–742.

Christakis, N. A. and Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *New England Journal of Medicine*, 358(21):2249–2258.

Clark, A. E. and Etilé, F. (2006). Don't give up on me baby: Spousal correlation in smoking behaviour. *Journal of Health Economics*, 25(5):958–978.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, 39(5):829–844.

Engel, C. and Moffatt, P. G. (2014). Dhreg, xtdhreg, and bootdhreg: Commands to implement double-hurdle regression. *The Stata Journal*, 14(4):778–797.

Jones, A. (1989a). The uk demand for cigarettes 1954–1986, a double-hurdle approach. *Journal of Health Economics*, 8(1):133–141.

Jones, A. M. (1989b). A double-hurdle model of cigarette consumption. *Journal of Applied Econometrics*, 4(1):23–39.

Labeaga, J. M. (1993). Individual behaviour and tobacco consumption: a panel data approach. *Health Economics*, 2(2):103–112.

Leite, A., Machado, A., Pinto, S., Nunes, B., and Dias, C. M. (2019). Daily tobacco consumption and associated socioeconomic factors in the portuguese population: National health interview survey data 1987-2014. *Revista Portuguesa de Cardiologia*, 38(8):583–593.

Li, C. and Gilleskie, D. B. (2021). The influence of endogenous behaviors among social pairs: Social interaction effects of smoking. *Journal of Health Economics*, 80:102547.

Pollak, R. A. (1970). Habit formation and dynamic demand functions. *Journal of Political Economy*, 78(4, Part 1):745–763.

Pollak, R. A. (1976). Habit formation and long-run utility functions. *Journal of Economic Theory*, 13(2):272–297.

Stigler, G. J. and Becker, G. S. (1977). De gustibus non est disputandum. *The American Economic Review*, 67(2):76–90.

Yen, S. T. and Jones, A. M. (1996). Individual cigarette consumption and addiction: a flexible limited dependent variable approach. *Health Economics*, 5(2):105–117.

A Descriptive Statistics

Table A.1: Sample Descriptive Statistics by Survey Year.

	1995	1998	2005	2014	2019
Smoking outcomes:					
Current Smoker (%)	19.0	19.8	19.5	17.0	16.5
Cigarettes per day (all)	3.1	3.3	3.2	2.5	1.8
Demographics:					
Female (%)	52.0	52.0	52.0	56.0	56.0
Mean age (years)	41.1	41.8	42.0	52.0	53.1
Married (%)	54.0	55.0	56.0	52.0	50.0
Single (%)	36.0	35.0	33.0	25.0	26.0
Education:					
No schooling (%)	13.0	11.0	10.0	8.0	7.5
Primary (1–4 years)	54.0	51.0	64.0	57.0	55.0
Secondary (5–11 years)	26.0	30.0	12.0	16.0	18.0
Higher (12+ years)	7.0	8.0	14.0	19.0	19.5
Health and insurance:					
Private insurance (%)	4.0	9.5	8.0	17.0	21.0
Hypertension (%)	16.0	17.0	23.0	30.0	35.0
Obesity (BMI $\geq$30) (%)	9.0	11.0	13.0	15.0	16.0
Household smoking:					
Smoking spouse (%)	4.0	5.0	—	—	—
Smoking parent (%)	3.0	2.0	—	—	—

Notes: Sample means by survey year. Empty cells indicate data not available in that wave. Smoking outcomes: "Current smoker" includes daily and occasional smokers; "Cigarettes per day" includes zeros for non-smokers. Education categories reflect highest level completed.