

# Education and smoking behavior in Brazil: decision to smoke and daily cigarette consumption intensity

*Educação e comportamento tabagista no Brasil: decisão de fumar e intensidade de consumo diário de cigarros*

*Marcelo Justus* <sup>(1)</sup>

*Elder G. Sant'Anna* <sup>(2)</sup>

*Eloá S. Davanzo* <sup>(3)</sup>

*Gustavo C. Moreira* <sup>(4)</sup>

<sup>(1)</sup> University of Campinas

<sup>(2)</sup> University of São Paulo

<sup>(3)</sup> University of Campinas

<sup>(4)</sup> Federal University of São João

Del-Rei

## Abstract

This study analyses the relationship between education and smoking behavior. Our empirical models were estimated using data for Brazilian citizens from the 2008 Special Survey on Tobacco Addiction. We estimated probit, sample-selection, and zero-inflated negative binomial models. Empirical analyses for women and men were conducted separately. We found that higher education levels are associated with lower smoking probability and lower daily consumption intensity.

## Keywords

smoking, tobacco, addiction, health, human capital.

**JEL Codes** I12, D12, C50.

## Resumo

*Este estudo analisa a relação entre educação e o comportamento tabagista. Os modelos empíricos foram estimados a partir de dados de cidadãos brasileiros participantes da Pesquisa Especial de Tabagismo de 2008. Os modelos estimados foram o probit, seleção amostral e binomial negativo de zeros inflados. Análises empíricas para mulheres e homens foram conduzidas separadamente. Os resultados mostram que níveis de educação mais elevados estão associados com menor probabilidade de fumar e menor intensidade de consumo diário.*

## Palavras-chave

*tabagismo, tabaco, vício, saúde, capital humano.*

**Códigos JEL** I12, D12, C50.

## 1 Introduction

According to the 2013 National Health Survey, there are about 21.9 million smokers in Brazil. The negative externalities of tobacco consumption are evident. There is consensus that smoking increases spending on healthcare and decreases productivity due to increased morbidity and premature death, which in turn substantially reduce the human capital (World Bank, 1999). Also considering indirect costs, smoking accounts for losses estimated at US\$ 1 trillion worldwide every year (World Bank, 2015). The World Health Organization considers smoking as the main cause of preventable death in the world, but an absurdly high number of deaths recorded every year are still related to tobacco use. This means that smoking is a public health problem! However, the harms of tobacco use outweigh its negative effects on health. In Brazil, for example, smoking reduces earnings in the labor market (Almeida and Araújo Júnior, 2017; Justus *et al.*, 2019).

Investing in human capital leads to higher real income in the future for an individual (Becker, 1962). This investment includes schooling, professional training, acquisition of information about the economic system, and also healthcare. All investments of this kind in human beings could improve their physical and mental skills, making it possible to predict with greater certainty that their real income will increase in the future. Workers can invest in any aspect that can enhance and improve their skills (schooling and/or training), thus improving their human capital and, consequently, raising their marginal productivity and earnings.

Much of what is referred to as consumption is actually investment in human capital: spending on education, healthcare, internal migrations in search of better job opportunities, among other expenditures. By investing in themselves, individuals can expand the set of choices available to them, increasing their well-being (Schultz, 1961). Healthcare is very important as an element of human capital: longevity, provision of direct utility, and time that may be devoted to working (Galama and van Kippersluis, 2015).

Are health-related behaviors, such as smoking, affected by education in Brazil? What we know is that tobacco addiction is prevalent among less educated Brazilians and lower-income groups in Brazil (IBGE, 2014; Pinto and Ugá, 2010). In this context, our main objective is to investigate empirically whether there is a relationship between education and smoking behavior (the decision to smoke or not to smoke) and daily consumption intensity.

Our main contribution is that of investigating the relationship between education and smoking using a large random sample, which is rarely available for empirical studies in other countries. Moreover, the unprecedented aspect of our study lies in that we investigate whether the connection between smoking and education can be driven by better use of information on the risks of smoking behavior and of exposure to cigarette ads.

The next section provides details about the empirical modeling. Section 3 presents our main results and a brief discussion. Our concluding remarks are presented in Section 4.

## 2 Empirical modeling

### 2.1 Data

Our empirical models are based on data for Brazilian citizens from the 2008 Special Survey on Tobacco Addiction (PETab, in the Brazilian acronym) – which is a subsample of the 2008 National Household Sample Survey (PNAD, in the Brazilian acronym)<sup>1</sup>. The survey was conducted through a partnership between the Brazilian Institute for Geography and Statistics (IBGE, in the Brazilian acronym), the Ministry of Health, the National Cancer Institute (INCA, in the Brazilian acronym), the Health Surveillance Secretariat (SVS, in the Brazilian acronym), and the National Health Surveillance Agency (ANVISA, in the Brazilian acronym).

The PETab survey is carried out in Brazil as part of an initiative launched by the World Health Organization and by the Centers for Disease Control and Prevention. This partnership was established with the aim of carrying out part of a survey conducted in 14 countries, including Brazil, entitled *Global Adult Tobacco Survey*.

The sample used in the PETab survey is made up of individuals aged 15 and above living in about 51,000 Brazilian households. The individuals included in that subsample answered questions related to tobacco use, their attempts to quit smoking, their exposure to smoke, and their access to awareness-raising campaigns and information on smoking risks, among other aspects. For other individuals interviewed in the 2008 PNAD, infor-

1 Available at: <https://www.ibge.gov.br/estatisticas/downloads-estatisticas.html>.

mation is only available for smoking behavior, type of tobacco product used, and consumption intensity.

## 2.2 Filtered sample

The percentage of young individuals (aged 15-29) in the group of non-smokers is higher than those in other age groups, and many of those young people are still students. Considering this fact and that there may be significant differences between ex-smokers and those who made an attempt to quit smoking with regard to unobserved characteristics, we chose to keep only individuals aged 30 and above in the sample. Our final sample was thus reduced to 6,895 men and 9,865 women. Analyses for women and men were conducted separately.

## 2.3 Variables

Table 1 shows the variables used for modeling the decision to smoke and daily cigarette consumption intensity. The specifications of our empirical models were based on a theoretical framework and, mainly, on previous studies, which will be discussed in Section 3

We also use education levels instead of years of schooling to check for any nonlinearity in the relationship between education and smoking behavior (see Zhu *et al.*, 1996).

The individual characteristics that we used included controls for race/color, current work status, age bracket, and per capita household income. As for geographic characteristics, we included dummy variables for urban area and metropolitan region. We also included dummies to capture the heterogeneity of Brazil's regions.

We controlled for whether the individual had access to information of some kind about the potential risks of smoking and for whether he or she had been exposed to any cigarette ads. The former, labeled as warning, is a dummy variable that takes value 1 if the individual is aware of at least one of the following risks: smoke from cigarettes causes serious illnesses, smoking causes serious illnesses, smoking causes heart diseases, smoking causes lung cancer, and using smokeless tobacco causes serious diseases,

and 0 otherwise. The latter, labeled as marketing, is a dummy variable that takes value 1 if the individual was exposed to cigarette ads through any type of media, and 0 otherwise.

The percentage of smokers in the households (the individual is not considered) is another relevant control variable, labeled as smokers%.

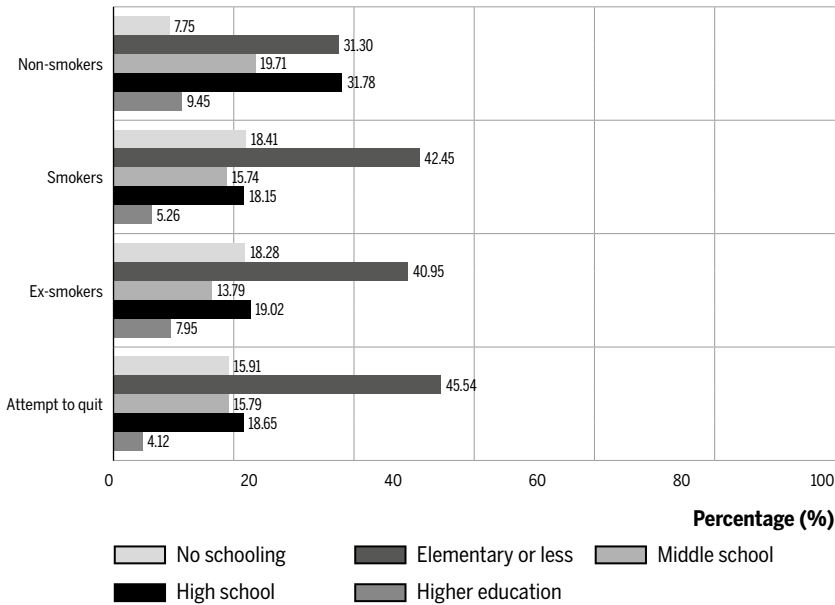
Table 1 Definition and summary statistics of the variables

Variable	Definition	Mean	Std. Dev.
no schooling	1 if the individual has no education and 0 otherwise	0.11	0.32
elementary or less	1 if the individual did not complete primary education and 0 otherwise	0.35	0.48
primary education	1 if the individual completed elementary school and 0 otherwise	0.18	0.38
high school	1 if the individual completed high school and 0 otherwise	0.27	0.45
higher education	1 if the individual is a college graduate and 0 otherwise	0.08	0.28
income(ln)	Logarithm for per capita household income	6.02	1.01
works	1 if the individual works and 0 otherwise	0.67	0.47
young	1 if the individual is a young person (aged 15-29) and 0 otherwise	0.31	0.46
adult	1 if the individual is an adult (aged 30-59) and 0 otherwise	0.53	0.50
elderly	1 if the individual is elderly (aged 60 and above) and 0 otherwise	0.16	0.36
white	1 if the individual is Caucasian or Oriental and 0 otherwise	0.47	0.50
householder	1 if the individual is the family head and 0 otherwise	0.52	0.50
smokers%	Percentage of smokers in the household	0.06	0.14
warning	1 if the individual is aware of the risks of tobacco addiction and 0 otherwise	0.97	0.17
marketing	1 if the individual saw cigarette ads and 0 otherwise	0.40	0.49
urban	1 if the individual lives in an urban area and 0 otherwise	0.85	0.36
metropolis	1 if the individual lives in a large city and 0 otherwise	0.38	0.48
cigarettes smoked	Number of cigarettes smoked in a day (if smoker)	1.96	5.19
smoking decision	1 if the individual smokes manufactured cigarettes and 0 otherwise	0.84	0.37

Note: men + women = 35,601 observations.

Figure 1 shows the level of education according to the group of non-smokers, smokers, ex-smokers, and those who made an attempt to quit smoking. Education levels are clearly higher in the population of non-smokers. We also observed that there is no clear difference between smokers and ex-smokers in terms of their level of education. This fact reinforces our hypothesis that education is an important factor for deciding to smoke or not to smoke. Nevertheless, the relationship between education and the decision to stop smoking or to make an attempt to quit smoking is not clear.

**Figure 1 Education levels according to the group of non-smokers, smokers, ex-smokers and those who made an attempt to quit smoking**



Note: Note: men + women = 35,601 observations.

Source: prepared with data from the 2008 PETA Survey (Brazil).

A descriptive analysis for other relevant characteristics of the groups is presented in Table 2. Mean difference tests were carried out for those characteristics and for each pair of groups (see Table A.1).

Smokers live in households with a lower per capita income than people who do not smoke (non-smokers and ex-smokers). Therefore, economically vulnerable families are more likely to suffer smoking-related harms.

Table 2 Average characteristics of non-smokers, smokers, ex-smokers and those who made an attempt to quit smoking

Characteristics	Non-smokers		Smokers		Ex-smokers		Att. To quit smoking	
	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.
income (ln)	6.04	1.01	5.88	1.02	6.08	1.01	5.79	1.01
works	0.66	0.47	0.73	0.44	0.62	0.49	0.75	0.43
young	0.39	0.49	0.21	0.41	0.12	0.33	0.25	0.43
adult	0.49	0.50	0.67	0.48	0.58	0.49	0.65	0.48
elderly	0.12	0.33	0.14	0.35	0.29	0.46	0.11	0.31
man	0.40	0.49	0.60	0.49	0.54	0.50	0.55	0.45
white	0.49	0.50	0.43	0.49	0.47	0.45	0.41	0.49
householder	0.44	0.45	0.63	0.48	0.67	0.47	0.62	0.48
smokers%	0.05	0.13	0.12	0.18	0.06	0.14	0.10	0.17
warning	0.97	0.16	0.94	0.24	0.97	0.17	0.98	0.14
marketing	0.39	0.49	0.43	0.49	0.36	0.48	0.47	0.5
urban	0.86	0.35	0.80	0.40	0.84	0.37	0.82	0.38
metropolis	0.39	0.49	0.37	0.48	0.36	0.48	0.33	0.47
cig. smoked			12.15	6.81			11.29	6.72
n		22,967		3,444		6,713		2,477

Note: men + women = 35,601 observations.

The percentage of smokers is also higher in the group of workers (73%) and heads of family (63%), suggesting that working may be a proxy for a stressful life, as pointed out in Kouvonen *et al.* (2005). Regarding household characteristics, we can see that, on average, a smokers' household tends to have a higher number of smokers. This reveals an externality of smoking habits that affects other members of the household. A study by Health and Social Care Information Centre (2018) showed that young people living with smoking adults in the same household are more likely to start smoking.

Most smokers are adults (65%) and men (60%). Young and elderly groups have the lowest willingness to quit smoking (24% and 11%, respectively) and most ex-smokers are adults and elderly people (58% and 29%, respectively). As expected, the group of smokers is the one that is most exposed to cigarette ads and, on average, smokers are less conscious of the risks of smoking than non-smokers. It is noteworthy that the aver-

age consumption of cigarettes among smokers is approximately 12 cigarettes a day, a higher number than that consumed by the group of those willing to quit smoking (approximately 11).

Lastly, Table A.1 also indicates that the difference between smokers and those who attempted to quit smoking is not so clear. But we can at least infer that there is a difference in non-observable characteristics, i.e. those who made an attempt to quit smoking may have more discipline and will-power (Almeida and Araujo Júnior, 2017). The same argument is plausible for explaining differences observed between smokers and ex-smokers.

## 2.4 Estimation procedures

We applied three estimation procedures (see Cameron and Trivedi, 2009). First, the decision to smoke has two possibilities: to smoke or not to smoke. Therefore, the response variable of the empirical model we used to investigate the relationship between education and the decision to smoke is dichotomous (smoking decision if the individual smokes manufactured cigarettes and = 0 otherwise). Here, we fit a probit regression.

Second, once the process involved in the individual choice of smoking was investigated, we began to investigate the addiction intensity aspect, i.e. the number of cigarettes smoked per day. However, this number is conditional on the decision to smoke, which may lead to the well-known sample-selection bias. To avoid this possibility, Heckman's maximum likelihood method (see Heckman, 1979; Cameron and Trivedi, 2009) was applied to correct the sample selectivity bias in the main equation (consumption equation), so the two processes were investigated concomitantly. The hypothesis that the first and second equations are independent was rejected at the level of 1 percent.

Third, we used count-data models to estimate the relationship between education and the number of cigarettes consumed daily (tobacco consumption intensity). Since the sample contained a large number of zeros due to the presence of non-smokers, we ran Vuong's LR test to discriminate between the negative binomial model and the zero-inflated negative binomial (ZINB) model. The results showed that the ZINB model is better. We used the same regressors in the two parts of the ZINB model, except the dummy variables for regions, where they were omitted in the inflated component.



In short, the usual probit model and the first step of the sample-selection models (which is also a probit model) were used for modeling the decision to smoke; and the second stage of the sample selection and the count part of the ZINB models were used for the daily consumption intensity. These models were estimated using the sample expansion factors associated to each observation.

### 3 Results and discussion

Tables A.2 and A.3 show the coefficients estimated from the probit, sample-selection, and ZINB models for women and men, respectively. Table 3 shows the marginal effects (at mean) calculated from these coefficients. We omitted the results for the inflated part in the ZINB models because they are useless for our purposes here. For this reason, these estimates were omitted. Standard errors for the estimated coefficients were also omitted, but their significance levels were presented.

For variables related to education levels, using no schooling as a benchmark we found evidence that the higher the level of education, the lower the probability of smoking and the lower the daily cigarette consumption. Thus, a negative relationship between education and smoking behavior was observed.

The results suggest that the probability of women smoking is about 0.1 percentage points (p.p.) lower for more educated individuals (variable: higher education) as compared to unschooled individuals. Moreover, among women smokers, highly educated women consume about 1 cigarette less a day than those unschooled. For more educated men, we found a reduction of about 0.2 p.p. in the probability of smoking and a decrease of three cigarettes in the daily average consumption as compared to the benchmark.

The evidence that there is a negative relationship between education and the decision to smoke corroborates the findings by Cutler and Lleras-Muney (2006), Kendler *et al.* (1999), Madden (2008), and Grimard and Parent (2007). However, it is not in tune with the findings by Koning *et al.* (2015), who after controlling for endogeneity found no relationship between education and the decision to smoke. Regarding daily cigarette consumption, our results corroborate the findings by Cutler and Lleras-Muney (2006) and De Walque (2007).

Table 3 Marginal effects calculated by gender and model

Variables	Women (n = 9,865)				Men (n = 6,895)				
	Probit model		Selection-sample model		Probit model		Selection-sample model		
	selection eq.	consumption eq.	selection eq.	consumption eq.	selection eq.	consumption eq.	selection eq.	consumption eq.	
elementary or less	-0.047*** (0.016)	-0.549** (0.218)	-0.041*** (0.014)	-0.392** (0.194)	-0.073*** (0.024)	-0.073*** (0.024)	-0.073*** (0.024)	-0.779* (0.455)	-0.705** (0.352)
middle school	-0.050*** (0.018)	-0.448 (0.276)	-0.042** (0.018)	-0.390* (0.229)	-0.125*** (0.028)	-0.125*** (0.028)	-0.126*** (0.028)	-1.443*** (0.546)	-1.301*** (0.417)
high school	-0.084*** (0.017)	-1.049*** (0.237)	-0.076*** (0.016)	-0.823*** (0.215)	-0.206*** (0.026)	-0.206*** (0.026)	-0.206*** (0.026)	-2.975*** (0.491)	-2.538*** (0.376)
higher education	-0.077*** (0.019)	-0.983*** (0.268)	-0.073*** (0.018)	-0.864*** (0.234)	-0.222*** (0.030)	-0.222*** (0.030)	-0.222*** (0.030)	-3.472*** (0.538)	-2.855*** (0.418)
warning	-0.079** (0.031)	-0.762* (0.394)	-0.052** (0.025)	-0.609** (0.284)	-0.118*** (0.045)	-0.118*** (0.045)	-0.118*** (0.045)	-1.612 (1.087)	-1.053 (0.655)
elderly	-0.062*** (0.008)	-0.888*** (0.119)	-0.063*** (0.009)	-0.779*** (0.109)	-0.005 (0.025)	-0.005 (0.025)	-0.005 (0.025)	-0.309 (0.490)	-0.200 (0.363)
income(ln)	-0.006 (0.005)	-0.056 (0.067)	-0.005 (0.005)	0.002 (0.059)	-0.015* (0.008)	-0.015* (0.008)	-0.014* (0.008)	0.209 (0.164)	0.036 (0.114)
works	0.002 (0.008)	-0.040 (0.116)	0.003 (0.008)	-0.042 (0.104)	0.019 (0.019)	0.019 (0.019)	0.019 (0.019)	0.155 (0.371)	0.266 (0.264)
white	-0.018** (0.008)	-0.172 (0.111)	-0.015** (0.008)	-0.032 (0.096)	-0.016 (0.014)	-0.016 (0.014)	-0.016 (0.014)	-0.401 (0.258)	-0.242 (0.192)

<b>householder</b>	0.053*** (0.008)	0.055*** (0.008)	0.840*** (0.128)	0.667*** (0.108)	0.019 (0.016)	0.019 (0.016)	0.047 (0.306)	0.131 (0.231)
<b>smokers</b>	0.286*** (0.022)	0.269*** (0.021)	3.875*** (0.308)	3.484*** (0.283)	0.654*** (0.050)	0.655*** (0.050)	11.306*** (0.927)	8.941*** (0.693)
<b>marketing</b>	0.032*** (0.008)	0.028*** (0.008)	0.426*** (0.119)	0.351*** (0.105)	0.045*** (0.013)	0.045*** (0.013)	1.046*** (0.266)	0.804*** (0.200)
<b>urban</b>	-0.001 (0.011)	0.005 (0.009)	0.097 (0.135)	0.195 (0.120)	-0.027 (0.018)	-0.027 (0.018)	-0.409 (0.348)	-0.182 (0.250)
<b>metropolis</b>	0.012 (0.008)	0.016** (0.008)	0.178 (0.116)	0.261** (0.103)	0.001 (0.014)	0.001 (0.014)	0.294 (0.286)	0.244 (0.218)
<b>north</b>	-0.045*** (0.014)	-0.050*** (0.013)	-0.765*** (0.193)	-0.335*** (0.090)	-0.012 (0.026)	-0.012 (0.026)	-0.911* (0.503)	-0.643*** (0.208)
<b>northeast</b>	-0.040*** (0.012)	-0.039*** (0.011)	-0.594*** (0.173)	-0.198*** (0.071)	-0.003 (0.020)	-0.003 (0.020)	-0.582 (0.405)	-0.402*** (0.144)
<b>southeast</b>	-0.025** (0.011)	-0.023** (0.011)	-0.327** (0.165)	-0.065 (0.059)	-0.004 (0.018)	-0.004 (0.018)	-0.520 (0.371)	-0.226* (0.134)
<b>midwest</b>	-0.044*** (0.014)	-0.037*** (0.014)	-0.586*** (0.198)	0.057 (0.088)	-0.023 (0.023)	-0.023 (0.023)	-0.370 (0.495)	-0.007 (0.177)

Notes: standard errors in parentheses; \* significance at 10%; \*\* significance at 5%; and \*\*\* Significance at 1%.

Table 4 Marginal effects of warning variable calculated by level of education, gender, and by age groups: adult and elderly

Variables	Adult						Elderly					
	Probit model		Selection-sample model		ZINB model		Probit model		Selection-sample model		ZINB model	
	selection	eq.	consumption	eq.	consumption	eq.	selection	eq.	consumption	eq.	consumption	eq.
<b>Women</b>												
no schooling	-0.054 (0.075)	-0.035 (0.058)	-0.570 (1.023)	-0.084 (0.789)	-0.203*** (0.063)	-0.148*** (0.046)	-2.583*** (0.878)	-1.754** (0.759)				
elementary or less	-0.197*** (0.066)	-0.182*** (0.061)	-3.500** (1.385)	-2.024** (1.009)	-0.054 (0.043)	-0.044 (0.035)	-0.608 (0.491)	-0.307 (0.288)				
middle school	-0.139 (0.102)	-0.116 (0.086)	-2.174 (1.766)	-1.551 (1.489)	-0.046 (0.065)	-0.025 (0.052)	-0.164 (0.805)	-0.027 (0.473)				
high school	0.069*** (0.018)	0.074*** (0.016)	1.002*** (0.209)	0.976*** (0.127)	-0.109 (0.151)	-0.069 (0.101)	-0.691 (1.274)	0.386 (0.286)				
higher education	-0.293** (0.143)	-0.177** (0.079)	-2.846* (1.568)	-0.932 (0.714)	-0.514* (0.295)	-0.282** (0.142)	-5.985* (3.478)	-1.677 (1.239)				

<b>Men</b>									
<b>no schooling</b>	-0.051 (0.075)	-0.050 (0.075)	0.039 (1.147)	-0.459 (1.075)	-0.132 (0.097)	-0.133 (0.098)	-4.108 (3.010)	-2.225 (1.529)	
<b>elementary or less</b>	-0.038 (0.065)	0.076 (0.065)	0.039 (1.197)	0.013 (0.964)	-0.259** (0.121)	-0.257** (0.121)	-4.727* (2.445)	-4.115** (2.000)	
<b>middle school</b>	0.080 (0.093)	0.080 (0.093)	2.476** (1.017)	1.710* (1.023)	0.044 (0.156)	0.044 (0.156)	0.388 (3.978)	0.055 (3.033)	
<b>high school</b>	-0.264* (0.136)	-0.263* (0.135)	-6.739 (5.593)	-3.435 (2.681)	-0.307 (0.207)	-0.305 (0.206)	-4.873 (3.610)	-5.297 (3.595)	
<b>higher education</b>	-0.135 (0.154)	-0.135 (0.154)	0.219 (1.200)	0.092 (0.958)	-0.504* (0.299)	-0.504* (0.299)	-7.701* (4.291)	-8.227* (4.527)	

Notes: standard errors in parentheses; \* significance at 10%; \*\* Significance at 5%; and \*\*\* Significance at 1%.

It is quite plausible that there is a negative association between education and smoking behavior. This relationship may be due to several factors, such as: a higher level of education influences the way individuals think and make decisions (Cutler and Lleras-Muney, 2006); education prevents people from adopting harmful health behaviors (Koning *et al.*, 2015); and more education implies a better understanding of other costs associated with smoking besides its well-known negative impact on health.

Education levels are the regressors of interest here. Nevertheless, controlling for others relevant variables we found other interesting results that deserve to be highlighted.

We found a positive sign for the variable that controlled for exposure to any tobacco advertising strategies (variable: marketing) for both the decision to smoke and daily cigarette consumption. This evidence, which was found for both genders, suggests that being exposed to tobacco ads is positively correlated to smoking. Fortunately, tobacco advertising strategies are regulated by law in Brazil.

The results for both genders reveal a negative relationship between knowledge of the risks of smoking (warning) and smoking behavior. This relationship may be due to many factors, two of which are very plausible. The first factor is education: more educated individuals can better understand the health risks posed by tobacco use. Access to information is the second one: individuals with access to information or who seek information are more likely to use it in connection with their decision to smoke.

In order to investigate whether the link between smoking and education was determined by making better use of information on the risks of smoking behavior, we interacted the level of education with the variable that captures knowledge about the risks of smoking behavior. Table 4 shows estimates for two age variables: adult and elderly. The result has not shown a clear pattern to corroborate the hypothesis regarding the above-mentioned factors.

The theoretical literature describes a persistent link between education and health in several ways. There is evidence that education increases access to important information for people to take into account health-related aspects when making decisions. More educated people respond to new information faster than less educated or uneducated ones. Education also affects health through changes in behaviors and income opportunities (World Bank, 1999; Cutler and Lleras-Muney, 2006; de Walque, 2010;

Feinstein, 2002; Grossman, 1972).

The empirical literature suggests that the effects of education on health may be due to three factors: economic factors, such as income and employment; factors associated with health-related behaviors; and psychosocial factors. The strong relationship between education and health, even after controlling for income, was found to be robust in empirical studies (see Fuchs, 1982; Deaton and Paxson, 2004; Grossman, 2004; Lleras-Muney, 2005; Cutler and Lleras-Muney, 2006; Grossman, 2006). Previous evidence suggests that there is a strong relationship between lower schooling and greater likelihood of smoking (Wetter *et al.*, 2005) and that more educated people are less likely to smoke and, even when they do smoke, they are more likely to quit smoking (De Walque, 2007). Grimard and Parent (2007) also observed the effect of education on starting to smoke, but they found no evidence that more educated individuals are more likely to quit smoking. It should be highlighted that only few studies did not detect an association between education and tobacco use (see Tenn *et al.*, 2010) and between education and the decision to start smoking (see Koning *et al.*, 2015) or found a nonlinear relationship between smoking and years of schooling (see Zhu *et al.*, 1996). But there is a consensus on the positive externalities of education in health conditions. It is thus expected, for example, that more educated individuals will not smoke or, if they do smoke, that their intensity of smoking will be lower than if they were less educated.

Education can affect the decision to quit smoking and more education reduces smoking initiation and addiction to nicotine. Indirectly, if education makes people more patient, it reduces their propensity to indulge in short-term pleasures with long-term costs (De Walque, 2010; Kendler *et al.*, 1999; Koning *et al.*, 2015; Sander, 1995).

#### **4 Concluding remarks and future studies**

First of all, we would like to emphasize that our findings reflect the association between smoking behavior and the variables that were controlled for in the regressions. However, although the robustness of the estimates was checked in the light of different specifications and estimation procedures, any inferences of causality are only suggestive and tentative. Notwithstanding, our results are consistent with those of previous studies and

suggest that more educated individuals are less likely to smoke and, even if they do smoke, they tend to consume fewer cigarettes daily.

Our findings suggest that the relationship between education and smoking behavior is more pronounced for men compared to women. We also observed a negative association between exposure to cigarette ads and the decision to smoke, and also with the number of cigarettes consumed daily. Our results suggest that the law enacted in Brazil to control cigarette advertising has been working.

Even though our results provide good inputs for policymakers, there is still room for further investigation. For example, we have not yet investigated the effect of schooling on cessation of addiction, on attempts of cessation of addiction, and on the duration of addiction. It is very important to conduct advanced studies to estimate the causal effect of education on smoking behavior.

## References

- ALMEIDA, A. T. C. d.; I. T. d. ARAÚJO JÚNIOR. Tabagismo e penalização salarial no mercado de trabalho brasileiro. *Economia Aplicada* 21(2), 249, 2017.
- BECKER, G. S. Investment in human capital: A theoretical analysis. *Journal of Political Economy* 70(5, Part 2), 9–49, 1962.
- CAMERON, A. C.; TRIVEDI, P. K. *Microeconometrics using Stata* (2 ed.). College station: Stata Press, 2009.
- CUTLER, D. M.; LLERAS-MUNEY, A. *Education and health: evaluating theories and evidence*. Technical report, National Bureau of Economic Research. 2006.
- DE WALQUE, D. Does education affect smoking behaviors? Evidence using the Vietnam draft as an instrument for college education. *Journal of Health Economics* 26(5), 877–895, 2007.
- DE WALQUE, D. Education, information, and smoking decisions evidence from smoking histories in the united states, 1940–2000. *Journal of Human Resources* 45(3), 682–717, 2010.
- DEATON, A. S.; PAXSON, C. Mortality, income, and income inequality over time in Britain and the United States. In: *Perspectives on the Economics of Aging*, pp. 247–286. University of Chicago Press, 2004.
- FEINSTEIN, L. *Quantitative Estimates of the Social Benefits of Learning, 2: Health (Depression and Obesity)*. Wider Benefits of Learning Research Report, ERIC, 2002.
- FUCHS, V. R. Introduction to economic aspects of health. In: *Economic Aspects of Health*, pp. 1–12. University of Chicago Press, 1982.
- GALAMA, T. J.; KIPPERSLUIS, H. VAN. *A theory of education and health*. Working Paper No.



- 2015-007 (2015-007), 1–78, 2015.
- GRIMARD, F.; PARENT, D. Education and smoking: Were Vietnam war draft avoiders also more likely to avoid smoking? *Journal of Health Economics* 26(5), 896–926, 2007.
- GROSSMAN, M. On the concept of health capital and the demand for health. *Journal of Political economy* 80(2), 223–255, 1972.
- GROSSMAN, M. Education and Personal Behavior. In: *Handbook of the Economics of Education*. Forthcoming Amsterdam: Elsevier Science, 2004.
- GROSSMAN, M. Education and nonmarket outcomes. In: *Handbook of the Economics of Education*, Volume 1, Chapter, Elsevier, pp. 577–633, 2006.
- HEALTH AND SOCIAL CARE INFORMATION CENTRE. *Statistics on smoking – smoking patterns among young people*. Technical report, National Health Service, 2018.
- HECKMAN, J. J. Sample selection bias as a specification error. *Econometrica* 47(1), 153–61, 1979.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Pesquisa Nacional de Saúde 2013: percepção do estado de saúde, estilos de vida e doenças crônicas. Brasília, 2014.
- JUSTUS, M.; SANT’ANNA, E. G.; DAVANZO, E. S.; MOREIRA, G. C. The wage penalty of smoking in Brazil: Evidence from the special survey on tobacco addiction. *Economia Aplicada*, v. 23, n. 2, 2019, pp. 127-144, 2019.
- KENDLER, K. S.; NEALE, M.; SULLIVAN, P.; COREY, L.; GARDNER, C.; PRESCOTT, C. A population-based twin study in women of smoking initiation and nicotine dependence. *Psychological Medicine* 29(02), 299–308, 1999.
- KONING, P.; WEBBINK, D.; MARTIN, N. G. The effect of education on smoking behavior: new evidence from smoking durations of a sample of twins. *Empirical Economics* 48(4), 1479–1497, 2015, 2015.
- KOUVONEN, A.; KIVIMÄKI, M.; VIRTANEN, M.; PENTTI, J.; VAHTERA, J. Work stress, smoking status, and smoking intensity: an observational study of 46 190 employees. *Journal of Epidemiology & Community Health*, BMJ Publishing Group Ltd, 59, 63-69, 2005.
- LLERAS-MUNEY, A. The relationship between education and adult mortality in the United States. *The Review of Economic Studies* 72(1), 189–221, 2005.
- MADDEN, D. Sample selection versus two-part models revisited: The case of female smoking and drinking. *Journal of Health Economics* 27(2), 300–307, 2008.
- PINTO, M.; UGÁ, M. A. D. Os custos de doenças tabaco-relacionadas para o sistema único de saúde. *Cad. Saúde Pública* 26(6), 1234–1245, 2010.
- SANDER, W. Schooling and quitting smoking. *The Review of Economics and Statistics*, 191–199, 1995.
- SCHULTZ, T. W. Investment in human capital. *The American Economic Review*, 1–17, 1961.
- TENN, S.; HERMAN, D. A.; WENDLING, B. The role of education in the production of health: An empirical analysis of smoking behavior. *Journal of Health Economics* 29(3), 404–417, 2010.
- WETTER, D. W.; COFTA-GUNN, L.; FOULADI, R. T.; IRVIN, J. E.; DAZA, P.; MAZAS, C.; WRIGHT, K.; CINCIRIPINI, P. M.; GRITZ, E. R. Understanding the associations among

education, employment characteristics, and smoking. *Addictive Behaviors* 30(5), 905–914, 2005.

WORLD BANK. *A epidemia do tabagismo: Os governos e os aspectos econômicos do controle do tabaco*. The World Bank, 1999.

WORLD BANK. *Taxes on tobacco can save a life every six seconds*. Technical report, The World Bank, 2015.

ZHU, B.-P.; GIOVINO, G. A.; MOWERY, P. D.; ERIKSEN, M. P. *The relationship between cigarette smoking and education revisited: implications for categorizing persons' educational status*. *American Journal of Public Health* 86(11), 1582–1589, 1996.

### About the authors

*Marcelo Justus – mjustus@unicamp.br*

Institute of Economics, University of Campinas, Campinas, São Paulo, Brazil.

ORCID: <https://orcid.org/0000-0001-8660-1779>.

*Elder G. Sant'Anna – eldergenerozo@gmail.com*

University of São Paulo, Piracicaba, São Paulo, Brazil.

ORCID: <https://orcid.org/0000-0001-6988-4383>.

*Eloá S. Davanzo – eloadavanzo@gmail.com*

Institute of Economics, University of Campinas, Campinas, São Paulo, Brazil.

ORCID: <https://orcid.org/0000-0003-2394-4031>.

*Gustavo C. Moreira – gustavocmoreira@hotmail.com.br*

Federal University of São João Del-Rei, São João Del-Rei, Minas Gerais, Brazil.

ORCID: <https://orcid.org/0000-0001-9526-0267>.

The first author thanks the National Council for Technological and Scientific Development (CNPq) for its financial support to conducting this research (process number 442483/20147) and for his Productivity in Research Grant. Preliminary versions of this paper were presented at the 14<sup>th</sup> International Conference on Urban Health (ISUH International Society for Urban Health, 2017) and at the 45<sup>th</sup> Brazilian Meeting of Economics (ANPEC, 2017).

### About the article

Submission received on August 06, 2018. Approved for publication on September 13, 2019.

## APPENDIX

Table A1 *p*-value of the mean difference tests (smokers, ex-smokers, and who made an attempt to quit smoking)

Charac- teristics	Non- smokers			Smokers		Ex- smokers
	Smokers	Ex- smokers	Att. To quit smok	Ex- Smokers	Att. To quit smok	Att. To quit smok
income(ln)	0.00	0.03	0.00	0.00	0.00	0.00
works	0.00	0.00	0.00	0.00	0.13	0.00
adult	0.00	0.00	0.00	0.00	0.91	0.00
elderly	0.00	0.00	0.04	0.00	0.00	0.00
man	0.00	0.00	0.00	0.00	0.00	0.00
white	0.00	0.04	0.00	0.00	0.20	0.00
householder	0.00	0.00	0.00	0.00	0.30	0.00
smokers%	0.00	0.00	0.00	0.00	0.00	0.00
warning	0.00	0.50	0.08	0.00	0.00	0.00
marketing	0.00	0.00	0.00	0.00	0.00	0.00
urban	0.00	0.00	0.00	0.00	0.05	0.00
metropolis	0.14	0.00	0.00	0.36	0.00	0.36
cig. smoked					0.00	

Note: men + women = 35,601 observations.

Table A2 Coefficients estimated from probit, sample-selection, and ZINB models for women ( $n = 9,865$ )

Variables	Probit model	Selection-sample model		ZINB model
		selection eq.	consump- tion eq.	consump- tion eq.
elementary or less	0.318	0.361	8.648	0.203
middle school	0.153	0.162	4.396	0.216
high school	-1.453***	-1.423***	-30.157***	-0.876***
higher education	0.434	0.165	3.724	-0.610*
warning	-0.203	-0.139	-2.781	0.244
elementary or less × warning	-0.508	-0.523*	-11.642*	-0.102
middle school × warning	-0.326	-0.317	-7.451	-0.195
high school × warning	0.998**	1.013**	21.688**	0.923***

(continues on the next page)

Table A2 (continuation)

Variables	Probit model	Selection-sample model		ZINB model
		selection eq.	consump- tion eq.	consump- tion eq.
higher education × warning	-0.952**	-0.670*	-13.522*	0.588*
elderly	0.246	0.137	3.220	-0.035
elementary or less × elderly	-1.127**	-1.047***	-23.816***	-0.435
middle school × elderly	-1.082*	-1.005*	-21.463*	-0.471
high school × elderly	0.930	0.900	15.955	-3.301***
higher education × elderly	0.321	0.169	5.422	-0.186
warning × elderly	-0.601*	-0.505*	-11.992**	-0.172
elementary or less × warning × elderly	0.892*	0.803**	18.048**	0.267
middle school × warning × elderly	0.695	0.723	21.570	0.900
high school × warning × elderly	-0.895	-0.867	-15.901	3.406***
higher education × warning × elderly	-0.030	0.168	-0.609	0.253
income(ln)	-0.035	-0.030	-0.499	0.039
works	0.014	0.017	-0.493	-0.078*
white	-0.108**	-0.089**	-1.527	0.078*
householder	0.305***	0.310***	7.158***	0.050
smokers	1.741***	1.626***	35.943***	0.143
marketing	0.186***	0.162***	3.780***	-0.024
urban	-0.009	0.032	0.955	0.215***
metropolis	0.072	0.093**	1.555	0.092**
north	-0.267***	-0.300***	-7.301***	-0.340***
northeast	-0.231***	-0.228***	-5.233***	-0.186***
southeast	-0.135**	-0.126**	-2.577**	-0.057
midwest	-0.255***	-0.211***	-5.184***	0.048
constant	-0.669**	-0.824***	-17.971***	1.797***

Notes: \* significance at 10%; \*\* significance at 5%; and \*\*\* significance at 1%.

Table A3 Coefficients estimated from probit, sample-selection, and ZINB models for men ( $n = 6,895$ )

Variables	Probit model	Selection-sample model		ZINB model
		selection eq.	consumption eq.	consumption eq.
elementary or less	-0.188	-0.184	1.064	-0.037
middle school	-0.786*	-0.785*	-4.296	-0.265
high school	0.022	0.020	8.614	0.027
higher education	-0.351	-0.350	-6.701	-0.726
warning	-0.142	-0.140	2.206	0.039
elementary or less × warning	0.026	0.021	0.153	0.093
middle school × warning	0.453	0.451	5.253*	0.315
high school × warning	-0.758*	-0.756*	-8.408	-0.058
higher education × warning	-0.379	-0.380	6.386	0.707
elderly	0.318	0.321	2.893	-0.030
elementary or less × elderly	0.115	0.104	-1.570	0.178
middle school × elderly	-0.225	-0.227	13.850**	0.786***
high school × elderly	-0.062	-0.067	-8.969	0.278
higher education × elderly	0.578	0.574	4.342	0.950
warning × elderly	-0.210	-0.214	-6.996	-0.183
elementary or less × warning × elderly	-0.429	-0.416	2.500	-0.110
middle school × warning × elderly	0.069	0.070	-3.901	-0.419
high school × warning × elderly	0.189	0.196	14.282	0.053
higher education × warning × elderly	-0.929	-0.926	-5.236	-1.036
income(ln)	-0.049*	-0.049*	2.042***	0.082***
works	0.064	0.065	-0.726	0.011
white	-0.056	-0.056	-0.605	-0.018
householder	0.064	0.065	-1.243	-0.040
smokers	2.223***	2.226***	2.881	0.163*
marketing	0.151***	0.151***	1.409**	0.076**
urban	-0.091	-0.091	0.163	0.063
metropolis	0.002	0.002	1.306	0.077**
north	-0.041	-0.041	-3.302**	-0.235***
northeast	-0.011	-0.011	-2.392**	-0.140***
southeast	-0.013	-0.013	-2.062**	-0.076*
midwest	-0.080	-0.080	0.224	-0.002

(continues on the next page)

Table A3 (continuation)

Variables	Probit model	Selection-sample model		ZINB model
		selection eq.	consumption eq.	consumption eq.
constant	-0.158	-0.161	4.487	2.013***

Notes: \* significance at 10%; \*\* significance at 5%; and \*\*\* significance at 1%.