tobacconomics

Economic Research Informing
Tobacco Control Policy

The Impact of *Cigarette Price Increases on the Prevalence of Daily Smoking and Initiation in Brazil*

INSTITUTE FOR
HEALTH RESEARCH
AND POLICY
UIC

Citation

Franco-Churruarin F and Gonzalez-Rozada M. The impact of cigarette price increases on the prevalence of daily smoking and initiation in Brazil. A Tobacconomics Research Report. Chicago, IL: Tobacconomics, Health Policy Center, Institute for Health Research and Policy, University of Illinois Chicago, 2022. *www.tobacconomics.org*

Authors

This Research Report was written by Martin Gonzalez-Rozada PhD, Professor of Econometrics, Department of Economics, Universidad Torcuato Di Tella, Buenos Aires, Argentina, and Fiona Franco-Churruarin, Researcher, Department of Economics, Universidad Torcuato Di Tella, Buenos Aires, Argentina. It was peer-reviewed by Corné van Walbeek, PhD, Professor of Economics, University of Cape Town, Cape Town, South Africa; and and Guillermo Paraje, PhD, Professor, Business School, Universidad Adolfo Ibáñez, Santiago, Chile.

This Research Report is funded by Bloomberg Philanthropies.

About Tobacconomics

Tobacconomics is a collaboration of leading researchers who have been studying the economics of tobacco control policy for nearly 30 years. The team is dedicated to helping researchers, advocates, and policymakers access the latest and best research about what's working—or not working—to curb tobacco consumption and its economic impacts. As a program of the University of Illinois Chicago, Tobacconomics is not affiliated with any tobacco manufacturer. Visit *www.tobacconomics.org* or follow us on Twitter *www.twitter.com/tobacconomics*.

Key Findings

Increasing cigarette excise taxes that result in increases in cigarette prices reduces the initiation of daily smoking in Brazil.

Increasing cigarette prices would reduce daily smoking prevalence across the population of Brazil.

Price increases delay the age of daily smoking initiation. An increase of 10 percent in cigarette prices delays smoking initiation by almost two and a half years. Delaying the age at which individuals start smoking makes initiation itself less likely, as fewer people take up smoking as they get older.

Daily smoking prevalence decreases with wealth and is higher for men than women.

Executive Summary

There is abundant evidence documenting the negative consequences of smoking. Over the last 30 years Brazil has implemented effective policies to reduce tobacco consumption, and smoking prevalence has decreased in the past 14 years. In Brazil, smoking usually starts in adolescence. On average, people start smoking daily at 17 years old, with many starting as young as the age of ten. Moreover, early smoking initiation predicts long-term nicotine dependence, affecting smoking behavior for the duration of a person's life. In practice, while it is possible to predict how certain groups may be more likely to start smoking, it is not possible to predict which individuals will start smoking and which ones will not. Since avoiding the habit of sustained smoking can have substantial health benefits, there is a compelling need to address the issue of onset and prevalence of smoking in young individuals by the means of population-level policies that work across groups of people.

To the authors' knowledge, there is no study quantifying how changes in cigarette prices affect smoking onset nor prevalence in Brazil. This research report contributes to the understanding of the determinants of smoking onset and prevalence in Brazil by estimating daily prevalence price and smoking onset elasticities.

This study uses data from the National Health Survey 2013 (Pesquisa Nacional de Saúde, or PNS) to analyze the determinants of daily smoking prevalence and smoking onset. According to this sample, 10.42 percent of Brazilians smoke every day. Prevalence is lower for women than for men, and it is also negatively associated with wealth. This study finds that the price elasticity of daily smoking prevalence is -0.264. This implies that a 10-percent increase in cigarette prices would induce a reduction of 2.6 percent in smoking prevalence. The price elasticity of smoking prevalence does not differ significantly between genders[1], age groups, nor wealth quartiles.

While this study finds daily smoking prevalence is affected by an increase in prices, there is no evidence this affects prevalence among both daily and occasional smokers. This suggests that the effect of an increase in prices is more likely to discourage smokers from smoking every day. The data show that Brazilians start smoking daily mostly around the ages of 16 and 17, with men

[1]This research study is based on the PNS 2013 data, which uses gender and sex terms interchangeably. For this reason, this research study follows the same convention. But at the same time, the researchers recognize the important difference between gender and sex terms, as defined in more recent surveys in Brazil. For example, the PNS 2019 clearly defines sex as a biological term (for example, male).

starting, on average, earlier than women. However, the youngest smokers start between the ages of 8 and 10. This study's estimation of the price elasticity of smoking onset is around 3.5. In other words, an increase of 10 percent in prices delays the age of smoking initiation by 35 percent—at mean smoking starting age, which implies that smoking would be delayed by almost two and a half years. This result is robust across different specifications.

This research report finds that real cigarette prices, gender, age, and wealth are important determinants of daily smoking prevalence in Brazil. Individuals who belong to the highest quartile of wealth have a lower smoking prevalence than those located in lower-wealth quartiles. On average, daily smoking prevalence is lower for women than men and higher for younger individuals. Poor men (in terms of wealth) are less responsive, in terms of prevalence, to a change in prices. Also, increments in prices delay the starting age of smoking, suggesting that a policy of increasing excise could be very effective to reduce smoking initiation.

The evidence presented in this research report suggests that an increase in cigarette excise taxes that results in higher retail prices would reduce daily smoking prevalence and induce a delay in smoking initiation.

1. Introduction

Nicotine addiction is the fundamental reason that individuals persist in using tobacco products, and this persistent tobacco use contributes to many diseases (USDHHS, 2010). Global evidence shows that nicotine dependence symptoms can manifest soon after onset in some adolescents, often well before they start smoking daily or even regularly (DiFranza et al., 2000; DiFranza et al., 2007; Gervais et al., 2006; O'Loughlin et al., 2003; O'Loughlin et al., 2009) and that early onset predicts long-term adult smoking (Chassin et al., 1990). Since it is not possible to identify those individuals who, after first use of tobacco, will adopt the habit of sustained smoking, the need to prevent that first use is compelling (Klein, 2006; Gervais et al., 2006).

Moreover, public policies that increase excise taxes, leading to a rise in cigarette prices are effective for long-term smokers, but evidence suggests that they are even more effective for short-term smokers (Gonzalez-Rozada & Montamat, 2019). Naturally, younger people tend to be short-term smokers. This evidence highlights the importance of addressing he tobacco epidemic through control policies targeted at early ages, since delaying the age at which individuals start smoking by even a few years can have substantial health benefits.

Over the last decade, smoking bans, taxation, and public health tobacco control campaigns induced a decline in smoking prevalence in several Latin American countries. In Brazil, there has been progress in reducing smoking prevalence through the years. The proportion of smoking adults[2] has decreased from more than 16 percent in 2006 to currently about 10 percent of the population (Divino et al., 2019). This evidence suggests that tax increases that lead to higher prices and other tobacco control measures have been very effective in reducing prevalence. Nevertheless, there is no quantitative measure of the magnitude of the impact of increasing cigarette prices on smoking prevalence and smoking initiation. This research fills that gap.

This research report analyzes the determinants of daily smoking initiation—

[2]This is measured using data from VIGITEL (Risk Factor Surveillance and Protection for Chronic Diseases by Telephone Survey), that measures prevalence of smoking as "five or more days a week" (Bernal et al, 2017).

in particular, the impact of increasing the price of cigarettes via increased cigarette excise taxes—on smoking prevalence and the onset of cigarette use. In Brazil, cigarettes are subject to excise taxes with ad valorem and specific components. To examine the effectiveness of increasing cigarette prices through taxes on those who are most likely to become addicted, this study focuses on the determinants of smoking initiation and smoking onset among daily smokers. There is substantial evidence that, among those individuals who have ever tried smoking, about one-third become daily smokers (USDHHS, 1994). And among those smokers who try to quit, less than five percent are successful at any time (CDC, 2002, 2004). Consequently, any efforts to reduce tobacco initiation must consider the addictive potential of cigarettes.

This research report is organized as follows. Section 2 describes the data used in estimations. Section 3 discusses the methodology and presents the split population model. Results are presented in Section 4. Finally, Sections 5 and 6 discuss these results and present the conclusions. Appendices provide further details on the estimation procedures and other analyses.

2. Data

2.1. Survey data

To estimate price elasticities for smoking prevalence and onset, this study uses data from the 2013 edition of the National Health Survey (Pesquisa Nacional de Saúde, or PNS) carried out by the Brazilian Institute of Statistics and Geography (IBGE).[3] This survey has the objective of producing data on the health situation and lifestyle of the Brazilian population. Module P of the survey includes questions related to current and past smoking behavior. In particular, it asks about whether the individual smokes or not, frequency of smoking, quantity of cigarettes (and other tobacco products) smoked per week, starting age, price and quantities bought in the last purchase, and quitting age. Key questions included in the Global Adult Tobacco Survey (GATS) are included as questions on the PNS 2013.

The PNS has 205,546 individual records but Module P, featuring only lifestyle questions, was designed only to be answered by those individuals over the age of 18, yielding a total of 60,202 observations that answered the questions related to smoking (there are 162,183 individuals of all ages that did not answer any question of this module). The reference month of the survey is July 2013. Table 1 describes the data.

The definition of a daily smoker adopted in this study includes individuals who self-report as smokers and who report they smoke at least one cigarette every day. According to this definition, daily smoking prevalence is 10.42 percent in Brazil. While it is true that prevalence is higher for men than for women (13.13 percent versus 8.34 percent, respectively), the difference is less marked than in other middle-income countries such as Vietnam, South Africa, and Mexico (see Guindon, 2014; Vellios & van Walbeek, 2016; Gonzalez-Rozada & Franco Churruarin, 2020).

A broader definition of a smoker includes anyone who reported they smoked at the time of the survey, regardless of the frequency. Using this measure, in Brazil the proportion of the population that smokes is more than 15 percent—almost 20 percent for men and just below 12 percent for women.

The average age at which people start smoking daily is 17 years old. For men, the

[3]There is a 2019 edition of this survey, but the module with smoking data (Module P) was not yet released at the time of this study.

Table 1

Descriptive statistics of the PNS survey

Variables Sample size	Total 60,202	Men 25,920	Women 34,282
Daily smoker	10.42% (0.0021)	13.13% (0.0036)	8.34% (0.0024)
Daily and less than daily smoker	15.13% (0.0024)	19.59% (0.0042)	11.73% (0.003)
Age of daily smoking initiation	17.03 (0.08)	16.68 (0.1)	17.46 (0.12)
Price per pack (20 cigarettes) (R$)	4.38 (0.04)	4.41 (0.05)	4.35 (0.06)
Highest level of education attained			
No formal education	40.93% (0.0049)	42.01% (0.0066)	40.12% (0.0053)
Primary	14.91% (0.0026)	15.64% (0.0037)	14.36% (0.0033)
Secondary	31.09% (0.0037)	30.52% (0.0053)	31.53% (0.0044)
Tertiary and university	13.06% (0.0038)	11.83% (0.0046)	14.00% (0.0043)
Employment status			
Employed	60.68% (0.0037)	74.52% (0.0049)	50.12% (0.0046)
Unemployed	3.08% (0.0012)	2.64% (0.0016)	3.42% (0.0016)
Out of labor force	36.25% (0.0037)	22.84% (0.0047)	46.46% (0.0047)
Marital status			
Married	42.18% (0.004)	45.84% (0.0058)	39.38% (0.0047)
Separated	2.93% (0.0013)	2.69% (0.0018)	3.11% (0.0017)
Divorced	5.43% (0.0015)	4.66% (0.0021)	6.02% (0.0021)
Widowed	9.08% (0.002)	3.91% (0.0019)	13.02% (0.0031)
Single	40.38% (0.004)	42.90% (0.0058)	38.46% (0.0047)

Table 1

Descriptive statistics of the PNS survey (cont.)

Variables Sample size	Total 60,202	Men 25,920	Women 34,282
Ethnicity			
White (Branca)	47.93% (0.006)	47.41% (0.0071)	48.33% (0.0066)
Black (Preta)	9.25% (0.0025)	9.19% (0.0031)	9.29% (0.0031)
Yellow (Amarela)	0.92% (0.0007)	0.84% (0.0009)	0.99% (0.0009)
Brown (Parda)	41.47% (0.0053)	42.18% (0.0064)	40.93% (0.006)
Indigenous (Indígena)	0.42% (0.0004)	0.38% (0.0005)	0.46% (0.0005)
Wealth index (x100)			
1st quartile (poorest)	26.55 (0.17)	25.84 (0.21)	27.18 (0.17)
2nd quartile	44.93 (0.07)	44.95 (0.11)	44.91 (0.09)
3rd quartile	64.34 (0.08)	64.42 (0.12)	64.29 (0.11)
4th quartile (richest)	83.58 (0.12)	83.64 (0.17)	83.54 (0.13)
Age at survey	44.44 (0.13)	43.75 (0.18)	44.98 (0.17)

Note: Standard errors in parentheses
Source: Authors' calculations based on PNS 2013

average age of initiation is 16 years and 8 months, and for women it is 17 years and 5 months. At the time of the survey in 2013 smokers paid, on average, around R$ 4.38 in their last purchase, which is around US$ 1.94 at the average exchange rate of the reference month. There is little variability in the price paid by gender: US$ 1.96 for men and US$ 1.93 for women.

Regarding education, 40.93 percent of the surveyed population have not completed any level of formal education (14.35 percent of the survey participants have no level of instruction whatsoever, whereas 26.58 percent started but did not finish primary school). Meanwhile, 14.91 percent completed primary school (fundamental school or *ensino fundamental*, which spans 9 years), 31.09 percent finished secondary school (12 years of education total), and the remaining 13.06 percent of the population achieved higher studies. Disaggregation by gender shows that men tend to have lower levels of education than women, as more woman have completed secondary and tertiary studies.

Considering employment status, 60.68 percent of the surveyed individuals are employed, 3.08 percent are unemployed, and 36.25 percent are outside of the labor force. As is common in many low- and middle-income countries, the proportion of employed individuals is higher for men than for women, whereas the proportion of individuals outside of the labor force is higher for women than for men.

The survey also includes information about marital status and ethnicity. Regarding marital status: 42.18 percent of respondents are married, and 40.38 percent are single. The rest are divided as follows: 9.08 percent are widowed, 5.43 percent are divorced, and 2.93 percent are separated. The main difference between genders is that there are more widowed women than men.

Regarding ethnicity/race, almost half of respondents identify as "white." The category with the second-most answers is "brown," selected by 41.47 percent of the survey population. The next survey category with the highest proportion is "black," with 9.25 percent of the population. Lastly, 0.92 percent answered as "yellow" and 0.42 percent as "indigenous."

Due to the high amount of missing data in reported monthly income, the authors constructed a wealth index using principal components analysis (PCA). Weights for this index are defined with the first principal component. The variables included in the PCA are binary and reflect socioeconomic characteristics of the person surveyed, such as education beyond secondary school and household possessions. The index ranges from 0 to 1 and is higher for individuals with more characteristics. In these data, the proportion of women is slightly higher in the first three wealth categories, and the proportion of men in the lowest wealth category is 3.5 percentage points higher than for women.

The average age at survey time in the sample is 44 years old. Men in the sample have a lower average age than women. The minimum age is 18, and the maximum observed age is 101 (although the 99th percentile of the distribution of age is 85). The 25th percentile of the age distribution is 30. This is relevant for the estimation methodology of onset price elasticity, which uses the expanded cross-sectional data from PNS to create a pseudo-panel. Since it requires the use of a time-varying measure of cigarette prices, the availability of this price series is the binding constraint on the age of the individuals that can be kept in this pseudo-panel. The procedure is further explained in Section 3.

Table 2 shows prevalence of daily smoking by age. In the group of people aged 18–24 years old in 2013, 7.43 percent smoke. Prevalence increases in the age group of 25–44 to 9.59 percent, and in the age group of 45–64 it increases to 14.86 percent. Smoking prevalence decreases sharply in the group of people aged 65 and older to 5.54 percent. Disaggregation by gender shows that men have a higher daily smoking prevalence than women in the same age group. The largest difference between groups is in the group of ages 18–24, in which prevalence is more than twice as large for men than for women, and the difference between genders is less marked in the group of ages 45–64.

Table 3 shows daily prevalence of smoking by wealth. Individuals in the bottom half of the wealth index distribution have the highest smoking prevalence—on average, 11.7 percent. Smoking decreases as wealth increases: in the first three quartiles, prevalence is between 10 and 12 percent, whereas in the fourth quartile prevalence is 8 percent. Disaggregation by gender shows that smoking prevalence is relatively more stable across wealth quartiles for women than for men. The high daily smoking prevalence in the bottom half of the wealth distribution is largely driven by men's daily smoking prevalence of about 15.5 percent. Differences among men and women in

Table 2

Daily smoking prevalence by age groups

Age group	Aggregate	Men	Women
18 to 24 years old	7.43% (0.0051)	10.71% (0.0094)	4.64% (0.005)
25 to 44 years old	9.59% (0.003)	12.52% (0.005)	7.33% (0.0034)
45 to 64 years old	14.86% (0.0045)	17.05% (0.0071)	13.16% (0.0055)
65 years and older	5.54% (0.0039)	7.91% (0.008)	4.01% (0.0041)
Total	10.42% (0.0021)	13.13% (0.0036)	8.34% (0.0024)

Note: Standard errors in parentheses
Source: Authors' calculations based on PNS 2013

Table 3

Daily smoking prevalence by wealth quartiles

Wealth quartile	Prevalence	Men	Women
1st quartile (poorest)	11.61% (0.0044)	15.27% (0.0067)	8.41% (0.0053)
2nd quartile	11.83% (0.0048)	15.78% (0.0082)	9.09% (0.0049)
3rd quartile	10.17% (0.0042)	12.00% (0.0072)	8.81% (0.005)
4th quartile (richest)	8.03% (0.0038)	9.36% (0.0068)	7.04% (0.0046)
Total	10.42% (0.0021)	13.13% (0.0036)	8.34% (0.0024)

Note: Standard errors in parentheses
Source: Authors' calculations based on PNS 2013

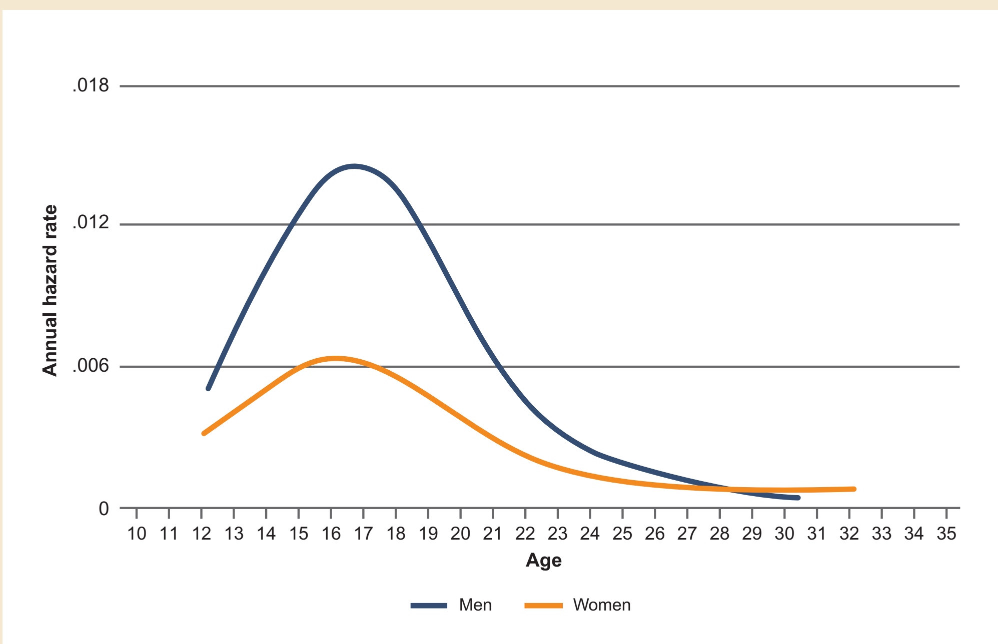
every quartile are all significant at the usual levels considering that variances are unequal between them. Between wealth groups, prevalence does not vary significantly between the first two quartiles in every case (aggregate, men, and women). However, there is evidence that prevalence decreases in the third and fourth quartiles for men but only in the fourth quartile for women.

Delaying the age at which individuals start daily smoking is associated with substantial health benefits (Institute of Medicine, 2015). Therefore, this report studies the impact of increasing cigarette prices on the onset of daily smoking. An "increase in the onset of smoking" means that the age at which individuals start smoking daily is delayed. For a first look at this issue, Figure 1 shows the risk of initiating the habit of daily smoking. As shown in the figure, people have a positive risk of initiating smoking daily from around age 12 or 13. Accordingly, in the modelling below, an individual is considered to be at risk of starting to smoke daily at the age of 10. Young men around the age of 17 have the highest risk of starting to smoke daily, while for women the highest risk is around 16 years old.

Figure 2 shows the cumulative hazard function of initiating daily smoking by gender. The cumulative hazard function describes the total amount of risk of initiating smoking (from this point on "initiating smoking" means initiating daily smoking) that has been accumulated up to each age in the x-axis. The cumulative hazard of starting smoking begins to increase around 13 years old, while in the case of women it seems to start later, around the age of 14.

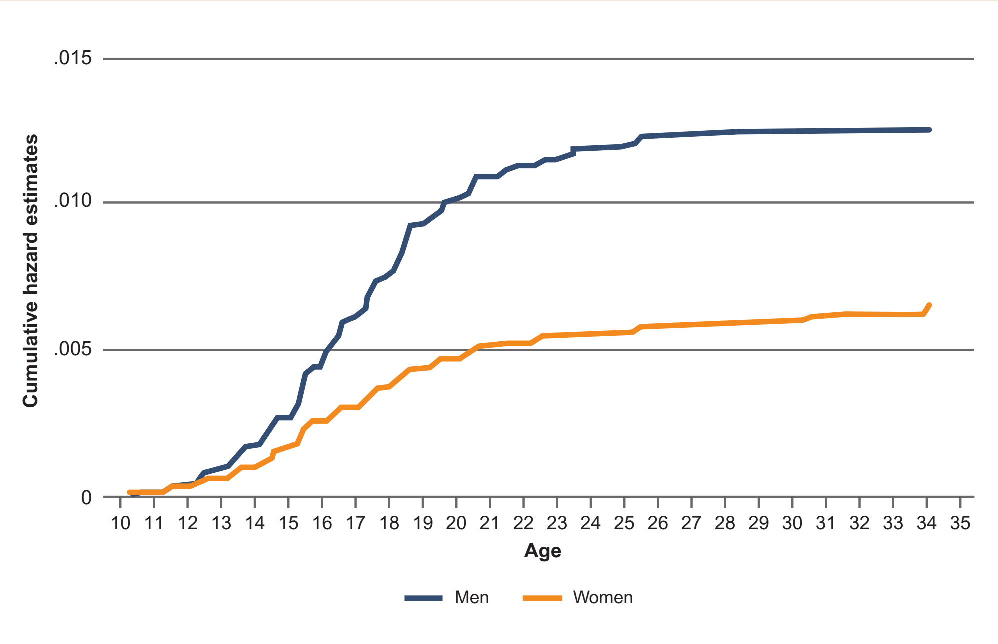
Around 21 years of age, the figure shows that the cumulative hazard of starting smoking among men is more than twice as high as that of women, and this relationship holds at older ages. Moreover, the slope of both curves is different, suggesting that between 13 and 20 years old the risk of initiating smoking for men increases at a faster rate (steeper slope) than for women. For women, the acceleration in the risk of starting to smoke is slower than for men, but after the age of 21 the slope of both cumulative hazard curves stabilizes. From the age of 25 men's risk of taking up smoking is consistently more than two times higher than the risk for women.

Figure 1
Smoothed hazard function



Note: Authors' elaboration based on PNS 2013

Figure 2
Nelson-Aalen cumulative hazard estimates



Source: Authors' elaboration based on PNS 2013

2.2. The price variable

The PNS survey contains self-reported cigarette prices. A common concern when estimating smoking prevalence is the potential endogeneity of this variable. To address this potential problem two different price variables are constructed. The first price variable assigns to smokers the self-reported price paid for the last purchase and uses a random regression imputation (sometimes called stochastic regression imputation) to assign a price for those non-smokers in the sample. The second price variable assigns to smokers and non-smokers the average of the self-reported price by primary sampling unit (PSU).

Following the recommendations in the *Economics of Tobacco Toolkit: Economic analysis of demand using data from the Global Adult Tobacco Survey (GATS)* John et al. (2019) note that this report checks for endogeneity of the self-reported price using the Rivers-Vuong (1988) test statistic. The Rivers-Vuong procedure is similar to the Hausman (1978) test for endogeneity in the linear model but applied to the probit prevalence estimation. It consists of two steps. The first step consists of estimating by least squares the reduced form of the potential endogenous self-reported price variable on the instrument and all exogenous variables of the model and generating the residuals of this estimation. The second step consists of estimating the prevalence equation using a probit model with the residuals of the first stage as an explanatory variable. In this second step, the Rivers-Vuong test for endogeneity consists of testing if the coefficient accompanying the residuals is equal to zero. Rejection of this hypothesis would suggest that the self-reported price is endogenous.

Before applying the test, since the survey does not contain cigarette prices for non-smokers, the authors have to impute a price for them as if they had been smokers. This is done by using a stochastic random regression imputation. This price variable is constructed as follows. First, the authors estimate the regression equation for the smokers in the sample using as a dependent variable the self-reported price paid for the last purchase and as explanatory variables: gender (women=1), age, labor and education categories, wealth quartiles, binary variables for students and marital status, and survey strata fixed effects. Then, the authors input prices for non-smokers using the predicted price from this regression plus a random draw from a normal distribution with mean and standard deviation equal to the mean and standard deviation of the residuals (Table A1 in the Appendix shows the complete procedure). The average of this "random imputation price" per pack of 20 cigarettes in the sample is US$ 1.88 (R$ 4.24).

The second variable (instrumental variable used in the Rivers-Vuong test procedure) is constructed by assigning to both smokers and non-smokers the average price by primary sampling unit (PSU). This instrumental price variable is constructed by regressing the self-reported price per pack on binary variables for PSUs. The predicted price from this regression assigns the average price by PSU to each individual in the sample, whether they are smokers or nonsmokers. The average of this "average price per PSU" is US$ 1.75 (R$ 3.94).

Table 4 summarizes the distribution of these measures of prices by showing the average price by deciles of each variable. The first two columns refer to the random imputation price, whereas the last two columns refer to the average price by PSU. In both cases, the first column shows the price variable expressed in logs, and the second column shows the actual price (in this case, in Brazilian reais (BRL) of 2013). From the comparison of both price variables, it seems that the stochastic random regression imputation price estimates smaller prices at the lower deciles but larger prices at the higher deciles of prices. This evidence suggests

Table 4
Average price by deciles

Deciles	Average random imputation price		Average price by PSU	
	Logged price	Actual price	Logged price	Actual price
1	0.5	1.69	0.59	1.84
2	0.88	2.42	1.06	2.88
3	1.09	2.99	1.34	3.84
4	1.26	3.54	1.38	3.99
5	1.41	4.1	1.39	4
6	1.57	4.8	1.4	4.01
7	1.67	5.31	1.55	4.76
8	1.8	6.08	1.61	5.01
9	2	7.39	1.72	5.6
10	2.38	11.13	1.97	7.37

Source: Authors' elaboration based on PNS 2013. All prices are in logs and measured in BRL.

the average price by PSU has a lower variability than the random regression imputation price.

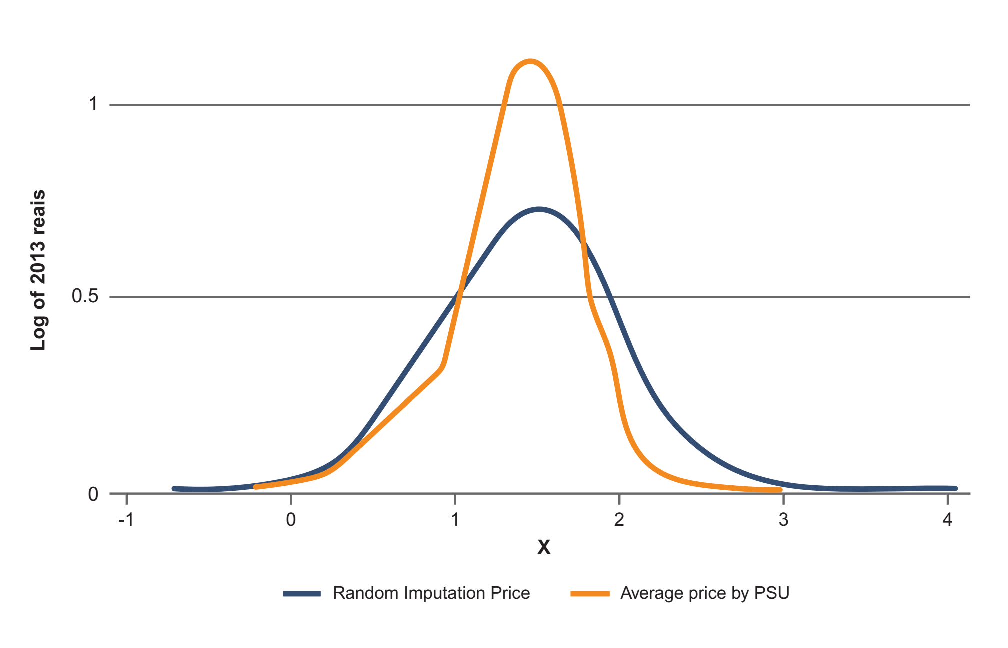
Figure 3 plots the kernel density estimates of the two price variables (measured in logs of Brazilian Reais). It is clear from the figure what Table 4 suggests: the variability of the random regression imputation price is larger than the average price by PSU.

Table 5 shows the results of the Rivers-Vuong second-step probit estimation. The first column of the table shows the exogenous explanatory variables, the natural logarithm of the stochastic random regression imputation price, *log(p)*, and the residuals from the first-stage estimation. The coefficient accompanying the residuals is statistically different from zero at the usual significance levels. This result indicates that the random imputation price is an endogenous variable, suggesting that the smoking prevalence should be estimated using an instrumental variable (IV) probit model. Results of the first step of the Rivers-Vuong test procedure can be found in the Appendix, in Table A2. These results also hold when considering daily and less than daily smokers.

For the estimation of the impact of cigarette prices on smoking onset, the data need to be transformed into a pseudo-panel in order to assign to each smoker the cigarette price at their smoking initiation date. For this exercise the authors use a monthly index for the real price of cigarettes constructed with Consumer Price Index (CPI) data and the cigarette disaggregation from IBGE, which is available from June 1989. Figure 3 shows the evolution of this index.

Figure 3

Kernel density estimates of prices



Source: Authors' elaboration based on PNS 2013

Table 5

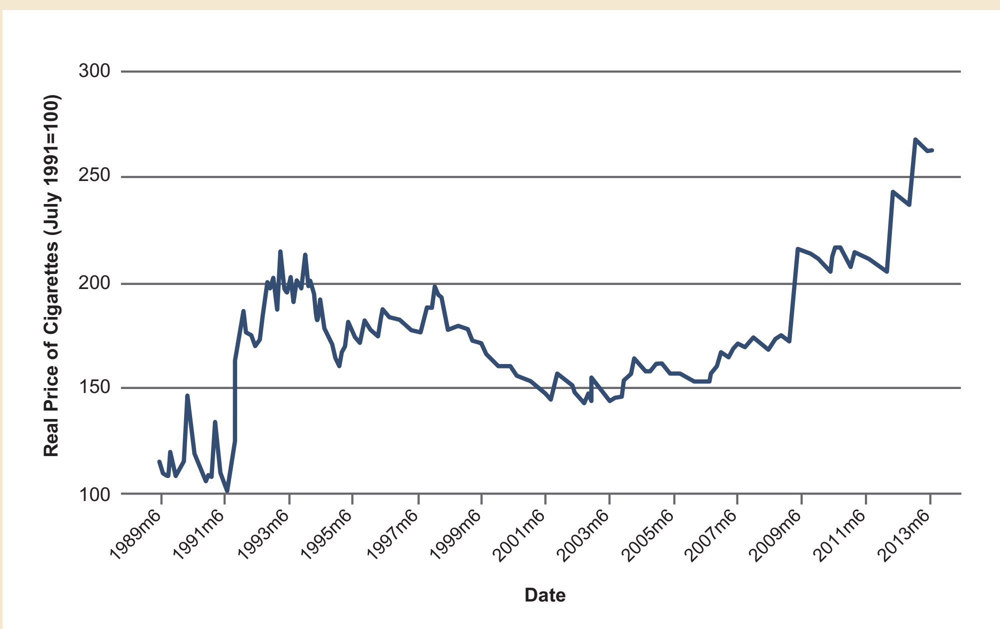
Rivers-Vuong endogeneity test

Dependent variable: Daily smokers	Estimated coefficient	Linearized standard error	t-statistic	p-value
Price of cigarettes (in logs)	0.0493	0.0291	1.69	0.09
Gender (women=1)	-0.2676	0.024	-11.13	0
Wealth index	-1.0906	0.069	-15.8	0
Residuals	**-0.1165**	**0.0247**	**-4.71**	**0**
Age categories				
25-44 years old	0.187	0.0422	4.43	0
45-64 years old	0.4523	0.043	10.51	0
65 and older	-0.1813	0.0554	-3.27	0.001
Labor categories				
Unemployed	0.0324	0.0618	0.53	0.599
Out of labor force	-0.0793	0.0299	-2.65	0.008
Constant	-1.2087	0.1056	-11.44	0
Province fixed effects	Yes			

Source: Authors' estimations based on PNS 2013

Figure 4

Evolution of the real price of cigarettes



Source: Authors' estimation based on IBGE

3. Methodology

This study uses survival analysis estimation focusing not only on the probability of smoking but also on the onset of cigarette use. For smoking prevalence the authors estimate a probit model, and for smoking onset a split-population model is used (Schmidt & Witte, 1989).

Since the survey used in this study has a single record per individual for the age they started smoking, a pseudo panel is constructed. Based on the reported age of initiation, a duration spell is created for each individual. Duration refers to the time that elapses between the risk age of smoking onset and the age of starting. Therefore, a spell begins at the risk age (assumed to be 10 years old) and either ends at the year the individual reported to have started smoking or at the survey date if they never started.

The main idea behind the use of a split-population model is to account for the fact that not all individuals who have an incomplete spell will eventually start smoking, as opposed to the traditional assumption of standard duration models that they all will. The duration process applies then only to those individuals who are predicted to eventually "fail." The likelihood of each observation is weighted with the probability that the individual will ever start smoking. Formally expressed, the log-likelihood function to be maximized is:

$$ln\,(L) = \sum w_i\,\{c_i\, ln\,[\phi(\alpha' z_i)\, f(t|s_i = 1, x_i(t)] +$$
$$(1 - c_i)\;\; ln\,[1 - \phi(\alpha' z_i) + \phi(\alpha' z_i)\, S(t|s_i = 1, x_i(t)]\} \qquad (1)$$

where c_i is a dummy variable equal to 1 if individual i ever smoked and 0 otherwise, s_i is another dummy equal to 1 if the individual will eventually start smoking and 0 if they never do. Φ is the standard normal cumulative function, and z^i is a vector of time-invariant covariates. f refers to the chosen conditional density function to model duration, S is the respective survival function, and w_i is a survey weight. $xi(t)$ are time-varying covariates, including the price of cigarettes.

The contribution to the log likelihood (1) for individual i observed smoker in the sample (c_i = 1, uncensored observations) is simply the natural logarithm of the probability of daily smoking, $\Phi(\alpha' zi)$, multiplied by the probability density function of starting smoking at the observed starting age $f(t/s_i = 1, x_i(t))$. For those i observed not starting smoking (c_i = 0, censored observations) the contribution is the natural logarithm of the probability of no daily smoking, $1 - \Phi(\alpha' zi)$, plus the probability of starting after the age observed in the survey, $\Phi(\alpha' zi) S(t/s{=}1, x_i(t))$ (Forster & Jones, 2001).

Notice that in the traditional split-population model the probability to start smoking is constant for all individuals, $\Phi(\alpha' zi) = k$, while here with a more general setup not all individuals have the same probability of starting to smoke. Smoking prevalence depends on the socioeconomic characteristics of the individuals. That is,

$$Pr(yi = 1 \mid zi) = \Phi(\alpha' zi) \qquad (2)$$

where $y_i{=}1$ indicates that individual i smokes and z_i is a vector of explanatory variables including the log of the imputed self-reported cigarette price; the wealth index; a dummy for women, rural residence, and being a student; and labor and age categories.

Using (2) as part of the log likelihood (1) means that, instead of estimating a single coefficient k for smoking prevalence, as in the traditional split-population model, it is necessary to estimate the coefficients of a nonlinear function. This makes the log likelihood (1) to be maximized highly nonlinear and difficult to fit because the convergence to a maximum is more likely to fail (Jenkins, 2001). To avoid this problem, the strategy adopted here is first using a probit model to estimate equation (2), $\widehat{\Phi}(\hat{\alpha}' z_i)$ and then introducing this estimation into equation (1) to estimate the duration coefficients. This procedure has the advantage of allowing the authors to compute the prevalence elasticity directly from equation (2), using

$$\epsilon_i = \frac{\partial \Phi(\alpha' z_i)}{\partial \ln(cp_i)} \times \frac{1}{\Phi(\alpha' z_i)} \tag{3}$$

where ln (cp_i) is the log of the imputed self-reported cigarette price. Equation (3) is a function that gives a different elasticity for each *i*. Therefore, when reporting the estimated elasticity, the average prevalence price elasticity is presented over the relevant group of people.

This study follows Forster and Jones (2001), who also use a split-population model to study the effect of tobacco taxes on smoking initiation, choosing the distribution of duration time to be log-logistic. This means that the density function in (1) is:

$$f\left(t|s=1, x_i(t)\right) = \frac{1}{\gamma} \frac{\psi^{1/\gamma} t^{1/\gamma - 1}}{\left[1 + (\psi t)^{1/\gamma}\right]^2} \tag{4}$$

where $\psi = \exp\left(-\beta' x_i(t)\right)$. The authors refer to γ as the "shape parameter" because it governs the shape of the density and the hazard. The hazard function of the log-logistic model is:

$$\lambda\left(t|s=1, x_i(t)\right) = \frac{1}{\gamma} \frac{\psi^{1/\gamma} t^{1/\gamma - 1}}{\left[1 + (\psi t)^{1/\gamma}\right]} \tag{5}$$

The log-logistic model belongs to the continuous time accelerated failure time (AFT) class of models. Since this study uses monthly data, and the event of interest happens years after starting to be at risk, the assumption of continuous time is a reasonable one. The AFT class of models leads to an intuitive interpretation of coefficients because they are interpreted as the proportional change in survival time for a unit change in the regressor (Jenkins, 2005). In the case of regressors measured in logarithms, the coefficient accompanying it is an elasticity. The authors seek to estimate the price elasticity of daily smoking onset η_p, which is:

$$\eta_p = \frac{\partial \ln (T)}{\partial \ln (p)} = \beta_1 \tag{6}$$

so this study's results can be interpreted as indicating that a one-percent increase in prices (in real terms) leads to a β_1% increase in daily smoking onset. As mentioned previously, an increment in smoking onset suggests a delay in the age at which individuals start smoking. The delay is calculated in "months after the risk age of 10," which is the (dependent) time variable in the model. Thus, the delay in months at a given age *a* and risk age *r* (both in years) after a given price change of Δ_p is:

$$D(\beta_1, \Delta_p, a, r = \beta_1 \cdot \Delta_p \cdot 12(a - r) \tag{7}$$

where $\Delta_p = (p_1 - p_o) / p_o$. After calculating this, it is easy to recover the delay in years. It is important to acknowledge that the delay cannot be compared to the results of studies in which the individuals are assumed to be at risk at other starting ages (Guindon, 2014).

In order to estimate the split-population model with time-varying covariates, the authors expand the survey data from the risk age of smoking onset up to the date of the survey. For those individuals that started smoking daily, the cigarette price is linked to the calendar month-year in which they started to smoke daily. That is, if the person is 25 years old at the date of the survey and began smoking daily at 15 years old, the authors assign the cigarette price the person faced when he/she was 15 years old.

Assignment of months is randomized due to the fact that people report the age in years at which they began smoking daily. Since the survey asks only about the age at which individuals started daily smoking, the authors input the price of a month of that year at random. This cannot be done for those individuals who had not started smoking at the time of the survey. The solution to this problem adopted here is to attribute these individuals the cigarette price at the time of the survey. This procedure is constrained by the availability of cigarette price data. Prices are available from June 1989. For those observations whose age of starting smoking corresponds to a calendar month year before June 1989 the authors do not have any price to assign. If they were to be included, these observations would not be seen since the beginning of the time at risk.

The other explanatory variables in the duration part of the model are time invariant. The authors assign the value of the covariate at the date of the survey for each individual *i* in the new database. Thus, covariates vary between individuals but are fixed in time. This study shows estimates of more than one specification of the model. The first specification to be estimated uses baseline covariates, the second specifications include additional covariates, and the third specification includes state (in Brazil, *federative units*) fixed effects. An assumption the authors implicitly make is that there are no movements between states over time.

As mentioned before, the starting risk age assumed here is 10 years old, and since the series of real cigarette prices started in June 1989, the data set keeps all individuals who, at the time of the survey, were between 18 and 34 years old. There are 1,483 individuals in the survey in between those ages, which accounts for 36.12% of the sample size. The consequence of this is that the results apply mostly to young individuals, but this is not a limitation given that one of the main problems of smoking onset is that it occurs at young ages. The data set is expanded replicating each observation to create a pseudo-panel of monthly frequency with only time-invariant covariates and combine it with time-varying covariates such as the price of cigarettes.

4. Results

4.1. Daily smoking prevalence

In this section the research report presents the results of the daily prevalence price elasticity estimation using the full cross-section PNS sample. Table 6 shows four alternative estimations. The first two columns of the table show estimations for daily smokers, whereas the last two columns use daily and less than daily smokers. Column (1) presents a probit estimation using the average price by PSU, while column (2) shows the estimation of an IV probit model, allowing for potential endogeneity of the random imputation self-reported price described in section 2.2. This column uses the average price by PSU as the instrument. Columns (3) and (4) show the estimation using the same models as in column (1) and (2), respectively, but for daily and less than daily smokers. Since the empirical evidence presented in Section 2.2 suggests that the random regression imputation price is an endogenous variable, the authors' preferred specification is shown in column (2) of the table.

In all cases, the authors control for the following covariates: gender (women=1), wealth index, age categories (18–24 as the base group, 25–44, 45–64, and 65 and older), and employment status (employed as the base group, unemployed, and out of the labor force).

As shown in the table, the prevalence price elasticity is negative and statistically significant in both specifications for daily smokers, suggesting an increment in cigarette prices would reduce daily smoking prevalence. In the authors' preferred specification, the estimated daily prevalence price elasticity is -0.264, implying that a 10-percent increase in cigarette prices would be associated with a reduction of 2.6 percent in daily smoking prevalence.

Table 6

Prevalence price elasticity estimation

	Daily smoking		Daily and less than daily smoking	
	Average price by PSU	**Instrumental variables probit**	**Average price by PSU**	**Instrumental variables probit**
Price of cigarettes (in logs)	-0.0516 (0.029)* [-0.0089]	-0.1374 (0.065)** [-0.0232]	0.0054 (0.024) [0.0012]	0.0985 (0.058)* [-0.0202]
Gender (women=1)	-0.2397 (0.024)*** [-0.0415]	-0.2549 (0.015)*** [-0.0410]	-0.3142 (0.022)*** [-0.0686]	-0.3347 (0.014)*** [-0.0704]
Wealth index	-0.328 (0.053)*** [-0.0568]	-0.2746 (0.043)*** [-0.0438]	-0.9495 (0.049)*** [-0.2073]	-1.0361 (0.037)*** [-0.2064]
Age categories				
25-44 years old	0.1611 (0.041)*** [0.0237]	0.2039 (0.026)*** [0.0276]	0.2227 (0.037)*** [0.0417]	0.2456 (0.023)*** [0.0421]
45-64 years old	0.4444 (0.042)*** [0.0792]	0.437 (0.028)*** [0.0697]	0.5079 (0.038)*** [0.1112]	0.5129 (0.025)*** [0.1006]
65 and over	-0.0957 (0.053)* [-0.0117]	-0.0084 (0.036) [-.0011]	-0.0127 (0.046) [-.0021]	0.1073 (0.031)*** [0.0127]
Labor categories				
Unemployed	0.07 (0.061) [0.0131]	0.0663 (0.041) [0.0114]	0.1172 (0.06)* [0.0273]	0.0737 (0.037)** [0.0125]
Out of labor force	-0.0814 (0.029)*** [-0.0138]	-0.0955 (0.018)*** [-0.0151]	-0.0333 (0.026) [-0.0072]	-0.0578 (0.016)*** [-0.0142]
Intercept	-1.0868 (0.061)***	-1.0632 (0.092)***	-0.6551 (0.051)***	-0.7975 (0.08)***
Prevalence price	-0.0921 {.051}*	-0.2642 {0.1182}**	0.0087 {0.038}	-0.1566 {0.0952}

*Note: Standard errors in parentheses. Statistical significance * 10%, ** 5%, and *** 1%. Marginal effects in brackets. Delta method standard errors in braces.*
Source: Authors' estimations based on PNS 2013

However, when considering not only daily but also occasional smokers in this study's measure of smoking prevalence, prevalence price elasticity is not statistically significant—meaning that there is no evidence of an effect of a relationship between prices and smoking prevalence. This suggests that an increase in prices might induce frequent smokers to smoke less and become occasional smokers rather than to quit smoking altogether. This evidence suggests that, irrespective of the construction of the price variable and the definition of prevalence, an increase in prices is associated with a decline in smoking probability. However, the magnitude of the effect is different depending on the prevalence measure adopted. For the daily prevalence of smoking the price variable coefficients are statistically significant, while for the daily and less than daily prevalence they are not.

To give a measure of this impact on smoking prevalence, the last row in the table shows the implied average prevalence price elasticity. This elasticity is negative and significant but depends on the price variable used. When using the average price by PSU the prevalence elasticity is -0.0921, while when considering the instrumental variable, it is -0.2642, suggesting that an increment of 10 percent in cigarette prices is expected to reduce daily smoking prevalence between 1.0 and 2.6 percent. If one considers the effect on the level of prevalence, if the price were to increase 10 percent, smoking prevalence would decrease from 10 percent to 9.74 percent. Considering that prevalence in Brazil is 10.42 percent, it would be expected to decrease to 10.14 percent. The average price prevalence elasticity for daily and less than daily smokers is not statistically significant.

The coefficient of gender is negative and statistically significant indicating that, on average, smoking prevalence is lower for women than for men. Table 6 also suggests that prevalence is negatively associated with wealth. This means that smoking prevalence diminishes with wealth. Age is also an important determinant of smoking prevalence. Results in the table suggest that prevalence is higher for those individuals 45–64 years of age.

Table 7 shows the prevalence price elasticity by categories for the IV probit estimation for daily smokers. Estimates of prevalence price elasticity are higher, in absolute value, for women than for men. Elasticity increases in absolute value with

Table 7
Daily prevalence price elasticity by categories

Categories	Prevalence price elasticity	Delta method standard error	p-value
Gender			
Men	-0.2444	0.1094	0.0255
Women	-0.2791	0.1249	0.0254
Age			
18-24 years old	-0.2911	0.1303	0.0255
25-44 years old	-0.2657	0.1189	0.0254
45-64 years old	-0.2369	0.1060	0.0255
over 65 years old	-0.2952	0.1321	0.0255
Wealth quartiles			
Q1 (poorest)	-0.2545	0.1140	0.0255
Q2	-0.2630	0.1177	0.0254
Q3	-0.2687	0.1201	0.0253
Q4 (richest)	-0.2736	0.1225	0.0255

Source: Authors' estimations based on PNS 2013

wealth, and it is slightly higher for younger than older people. However, using delta method standard error, accounting for the fact that an IV probit model is used, the null hypothesis that prevalence price elasticity does not vary significantly between groups cannot be rejected.

4.2 Smoking onset

Table 8 presents the duration estimation results that show the estimates for the time that elapses between the risk age of daily smoking onset and the age of starting. The results for the split-population model (Equation 1) use the IV probit daily prevalence equation presented in column (2) of Table 6. The duration component of the model is presented in accelerated failure time format, and, therefore, the estimated coefficients can be interpreted as regression coefficients for the logarithm of time until failure. For an explanatory variable expressed in natural logarithm its coefficient can be interpreted as an elasticity (see Forster & Jones, 2001). A positive coefficient indicates that higher values of the explanatory variable delay the initiation in smoking.

Column (1) is the baseline model using as explanatory variables gender (women=1), wealth quartiles, age, education, and labor categories. Column (2) is the baseline

Table 8

Split population estimates using real cigarette price index

	(1)	(2)	(3)
Real price of cigarettes (in logs)	3.5465 [0.85]***	3.5503 [0.901]***	3.4846 [0.729]***
Gender (women=1)	0.1481 [0.099]	0.1453 [0.102]	0.1657 [0.096]*
Wealth quartiles (base: Q1)			
Q2	0.1196 [0.126]	0.1218 [0.13]	0.0817 [0.13]
Q3	0.1440 [0.129]	0.1349 [0.133]	0.0859 [0.133]
Q4	0.0123 [0.168]	0.0055 [0.179]	-0.0419 [0.164]
Age categories	YES	YES	YES
Labor categories	YES	YES	YES
Education categories	YES	YES	YES
Ethnicity/race	NO	YES	NO
Fixed effects by federative units	NO	NO	YES
Intercept	-13.9583 [4.433]***	-13.9810 [4.697]***	-13.3043 [3.763]***
Shape	0.3370 [0.02]***	0.3369 [0.019]***	0.3314 [0.018]***

*Note: Bootstrapped standard errors in brackets. Statistical significance * 10%, ** 5%, and *** 1%.*
Source: Authors' estimations based on PNS 2013

specification adding ethnicity categories. Since there are federal and state taxes affecting cigarette prices, retail prices differ among the 27 Brazilian states; column (3) adds to the baseline model federative unit fixed effects. In all specifications the price elasticity of delay of smoking initiation is around 3.5. This elasticity is statistically significant at the usual levels, and it suggests that an increase in cigarette prices of 10 percent would delay the age of smoking initiation by 35 percent. This result implies that, at the mean starting age of 17 years, an increment of 10 percent in prices would delay daily smoking initiation by around 2 years and 5 months. This is calculated as 35 percent of the 12 x 7 months after the person turned 10, at which age it is assumed that they are at risk of starting to smoke. The coefficient on the gender variable is not statistically significant, and thus indicates that women in this sample initiate smoking approximately at the same age as men. The estimation of the shape parameter of the hazard rate is positive and statistically less than one, implying that the smoking hazard rate first rises with time and then falls monotonically, as suggested in Figure 1 above.

Table 9 inquires if changes in cigarette prices have a different effect on the

Table 9

Split-population estimates using price-gender interactions

	(1)	(2)	(3)
Price-men interaction (in logs)	3.5341 [0.813]***	3.5380 [0.876]***	3.4708 [0.721]***
Price-women interaction (in logs)	3.5637 [0.822]***	3.5671 [0.887]***	3.5038 [0.729]***
Wealth quartiles (base: Q1)			
Q2	0.1184 [0.125]	0.1205 [0.13]	0.0804 [0.131]
Q3	0.1428 [0.126]	0.1335 [0.132]	0.0845 [0.128]
Q4	0.0117 [0.165]	0.0047 [0.178]	-0.0429 [0.164]
Age categories	YES	YES	YES
Labor categories	YES	YES	YES
Education categories	YES	YES	YES
Marital status categories	NO	YES	NO
Ethnicity/race	NO	YES	NO
Fixed effects by federative units	NO	NO	YES
Intercept	-13.8940 [4.231]***	-13.9179 [4.567]***	-13.2334 [3.718]***
Shape	0.0196 [0.019]***	0.3367 [0.019]***	0.3313 [0.018]***

*Note: Bootstrapped standard errors in brackets. Statistical significance * 10%, ** 5%, and *** 1%.*
Source: Authors' estimations based on PNS 2013

smoking initiation of women versus men. The structure of the table is similar to that of Table 8; the only difference is that instead of using the logarithm of the real price it shows the interaction of this price variable with gender. As demonstrated in the table, there are no meaningful differences between men and women. An increase of 10 percent in cigarette prices delays smoking initiation by 35 percent for both men and women.

Table 10 shows the results of the split-population model, where the logarithm of the cigarette price is interacted with a dummy variable adopting the value of one for those individuals in the lowest quartile of wealth (price-poor interaction). As in the case of gender, the table shows no significant difference between an increase in prices over the smoking initiation of wealth-poor people and those individuals located in the non-poor quartiles of wealth. This means that an increase in prices affects poor and non-poor people in the same way.

Table 10
Split-population estimates using price-wealth interactions

Dependent variable: Months after risk age	Specifications (1)	(2)	(3)
Price-poor interaction (in logs)	3.5314 [0.885]***	3.7412 [1.732]**	3.4666 [0.72]***
Price-not-poor interaction (in logs)	3.5520 [0.822]***	3.4108 [0.559]***	3.4895 [0.713]***
Gender (women=1)	0.1568 [0.100]	0.1431 [0.094]	0.1719 [0.094]*
Age categories	YES	YES	YES
Labor categories	YES	YES	YES
Education categories	YES	YES	YES
Marital status categories	NO	YES	NO
Ethnicity/race	NO	YES	NO
Fixed effects by federative units	NO	NO	YES
Intercept	0.3375 [0.020]***	-0.0587 [0.300]	0.3315 [0.018]***
Shape	-13.8621 [4.554]***	0.0312 [0.086]	-13.2493 [3.673]***

*Note: Bootstrapped standard errors in brackets. Statistical significance * 10%, ** 5%, and *** 1%.*
Source: Author' estimations based on PNS 2013

5. Discussion

The results in the previous section indicate that increments in cigarette prices would reduce daily smoking prevalence and delay smoking initiation. Therefore, a policy of increasing excise taxes which would lead to an increase in prices is an effective strategy to reduce the prevalence of people who smoke and would increase the average age of smoking initiation. In particular, a 10-percent increase in cigarette prices would induce a reduction of 2.6 percent in smoking prevalence and would delay smoking initiation by almost two and a half years. Since the data set includes all individuals who, at the time of the survey were between 18 and 34 years old, this finding on smoking initiation applies mostly to young individuals.

In Brazil, there are currently four tobacco taxes charged at the federal level and one excise tax charged at the state level. The four federal taxes are: industrialized products tax (IPI), tax for Social Integration program financing (PIS), tax for Social Security financing (COFINS), and an import duty (II). The only subnational tax is the Merchandise and Service Circulation Tax (ICMS), which varies depending on the state. The general rule for the industrialized product tax is an ad-valorem excise tax (for a more detailed description of the cigarette tax structure see Ribeiro & Pinto, 2019). The industrialized product tax is the fiscal policy instrument that most likely would induce an increment in cigarette prices across regions in Brazil. Using this policy instrument to increase retail prices would reduce daily smoking by delaying or dissuading smoking initiation.

This study does not find differences in the impact of increasing cigarette prices on the age of starting smoking between men and women nor a difference between wealth-poor and non-poor individuals. In all cases an increment in cigarette prices would delay the age of smoking initiation, independent of gender and wealth status. An increment of 10 percent in retail prices would delay the smoking initiation age by almost two and a half years. This evidence indicates that using excise taxes to induce increments in prices would be effective to deter smoking initiation of young people in Brazil.

On the other hand, increments in retail cigarette prices would reduce daily smoking prevalence regardless of the income group, the age group, or the gender of the population considered. An increase of 10 percent in prices would induce a reduction in smoking prevalence of around 2.6 percent for individuals between 18 and 24 years of age.

Some limitations of this study are: (i) a potential recall error since individuals have to remember when they started smoking daily; and (ii) the authors do not account for price variation across brands, because the PNS database does not include questions about the cigarette brand smoked. Therefore, they do not capture any potential substitution between cheaper or illegal brands and more expensive brands when there is an increment in ad valorem excise taxes leading to a rise in cigarette prices.

6. Conclusion

In this research report the authors estimate the impact of increasing cigarette prices on daily smoking prevalence and on the age of starting smoking in Brazil. A split-population model is used to specify daily smoking participation and smoking onset equations. The authors estimate these equations to obtain prevalence and onset price elasticities and to evaluate the importance of other determinants of smoking probability and factors affecting the starting age of daily smoking.

The empirical evidence presented suggests that prices, gender, age, and wealth are important determinants of daily smoking prevalence in Brazil. On average, daily smoking prevalence is lower for women than men and is negatively associated with wealth; it is also higher for those individuals 45 to 64 years of age.

The addictive nature of tobacco products is at the center of many health problems, and adolescence is a key phase in which addiction may develop. The evidence presented in this report suggests that increases in cigarette prices are, on average, linked to lower prevalence of daily smoking and a delay in the development of the habit of daily smoking. Delaying or reducing smoking at young ages is expected to improve health outcomes over the life course. Hence, a policy of increasing excise taxes with the objective of increasing cigarette prices could be very effective to reduce or delay smoking initiation.

References

Bernal, Regina Tomie Ivata et al. (2017). Sistema de Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (Vigitel): mudança metodologia de ponderação. Epidemiologia e Serviços de Saúde. v. 26, n. 4, pp. 701-712. ISSN 2237-9622. https://doi.org/10.5123/S1679-49742017000400003.

Chassin, L., Presson, C. C., Sherman, S. J., & Edwards, D. A. (1990). The natural history of cigarette smoking: predicting young-adult smoking outcomes from adolescent smoking patterns. Health Psychology, 9(6), 701–716. http://dx.doi.org/10.1037/0278-6133.9.6.701

DiFranza, J. R., Rigotti, N. A., McNeill, A. D., Ockene, J. K., Savageu, J. A., St Cyr, D., Coleman, M. (2000). Initial symptoms of nicotine dependence in adolescents. Tobacco Control, 9(3), 313–319. http://dx.doi.org/10.1136/tc.9.3.313

DiFranza, J.R., Savageau, J.A., Fletcher, K.E., O'Loughlin, J., Pbert, L., Ockene, J. K., McNeill, A. D., Hazelton, J., Friedman, K., Dussault, G., Wood, C., Wellman, R. J. (2007). Symptoms of tobacco dependence after brief intermittent use: the Development and Assessment of Nicotine Dependence in Youth study II. Archives of Pediatrics & Adolescent Medicine, 161(7), 704–710. http://dx.doi.org/10.1001/archpedi.161.7.704

Divino J. A., Philipp Ehrl, Osvaldo Candido and Marcos Valadão. (2020). An extended cost-benefit analysis of tobacco taxation in Brazil. UCB Policy Report.

Forster, M. and Jones, A. (2001). The role of tobacco taxes in starting and quitting smoking: Duration analysis of British data. Journal of the Royal Statistical Society Series A, 164, issue 3, p. 517-547.

Gervais, A., O'Loughlin, J., Meshefedjian, G., Bancej, C., & Tremblay, M. (2006). Milestones in the natural course of onset of cigarette use among adolescents. Canadian Medical Association Journal, 175(3), 255–261. http://dx.doi.org/10.1503/cmaj.051235

Gonzalez-Rozada, M and Franco Churruarin, F. (2020). Increasing Cigarette Excise Tax Would Delay Smoking Initiation in Mexico. A Tobacconomics Policy Brief, University of Illinois at Chicago.

Gonzalez-Rozada, M., & Montamat, G. (2019). How increasing tobacco prices affects the decision to start and quit smoking: evidence from Argentina. International Journal of Environmental Research and Public Health, 16(19), 3622. https://doi.org/10.3390/ijerph16193622

Guindon GE. (2014). The impact of tobacco prices on smoking onset in Vietnam: duration analysis of retrospective data. Euro J Health Econ. 2014;15:19–39.

Hausman, J. (1978). Specification Tests in Econometrics. *Econometrica*, 46(6), 1251-1271. doi:10.2307/1913827

Institute of Medicine (2015). Public Health Implications of Raising the Minimum Age of Legal Access to Tobacco Products. Washington, DC: The National Academies Press.https://doi.org/10.17226/18997.

Jenkins, S. (2005). Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK.

John R., Chelwa G., Vulovic V., Chaloupka F., (2019) Using Household Expenditure Surveys for Research in the Economics of Tobacco Control. A Tobacconomics Toolkit. Chicago, IL: Tobacconomics, Health Policy Center, Institute for Health Research and Policy, University of Illinois at Chicago, 2019. www.tobacconomics.org

Klein, J. (2006). Adolescents and smoking: the first puff may be the worst. Canadian Medical Association Journal, 175(3), 262–263.

O'Loughlin, J., DiFranza, J. R., Tyndale, R. F., Meshefedjian, G., McMillan-Davey, E., Clarke, P. B. S., Hanley, J, Paradis, G. (2003). Nicotine-dependence symptoms are associated with smoking frequency in adolescents. American Journal of Preventive Medicine, 25(3), 219–225. http://dx.doi.org/10.1016/S0749-3797(03)00198-3

O'Loughlin, J., Gervais, A., Dugas, E., & Meshefedjian, G. (2009). Milestones in the process of cessation among novice adolescent smokers. American Journal of Public Health, 99(3), 499–504. http://dx.doi.org/10.2105/AJPH.2007.128629

Ribeiro, L. and Pinto, V. (2019). Accelerating Effective Tobacco Taxes in Brazil: Trends and Perspectives. Red Sur Country Studies Series "Tobacco taxes in Latin America". Country Study N° 3/2019.

Rivers D and Quang H. V. (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics*, vol. 39, issue 3, 347-366

Schmidt, P. and Witte, A.D. (1989). Predicting criminal recidivism using 'split population' survival time models. Journal of Econometrics, 40(1): 141-159.

Vellios N, van Walbeek C. (2016). Determinants of regular smoking onset in South Africa using duration analysis. BMJ Open 2016;6: e011076.

doi:10.1136/bmjopen-2016-011076

U.S. Department of Health and Human Services. (2010). How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

World Health Organization. (2010). Economics of tobacco toolkit: economic analysis of demand using data from the Global Adult Tobacco Survey (GATS). ISBN 978 92 4 150016 6

Appendix A

Table A1

First step of random regression price imputation

	Price of 20-cigarette pack (in logs)
Gender (women=1)	0.0173 (0.017)
Age categories	
25-44 years old	-0.0538 (0.033)
45-64 years old	-0.1034 (0.036)***
65 and older	-0.0670 (0.050)
Education categories	
Fundamental	0.1277 (0.025)***
High school	0.2256 (0.023)***
University	0.3321 (0.025)***
Labor categories	
Unemployed	-0.1259 (0.046)***
Out of labor force	-0.0405 (0.024)*
Student	-0.1078 (0.032)***
Monthly income	0.0000 (0.000)***
Federative units	
Acre	-0.3427 (0.076)***
Amazonas	0.3385 (0.076)***
Roraima	0.3413 (0.069)***
Pará	0.1725 (0.087)**
Amapá	0.3454 (0.072)***
Tocantins	0.1065 (0.074)

Table A1

First step of random regression price imputation (cont.)

	Price of 20-cigarette pack (in logs)
Maranhão	-0.0771 (0.087)
Piauí	-0.0488 (0.086)
Ceará	-0.0400 (0.073)
Rio Grande do Norte	-0.0417 (0.082)
Paraíba	0.0085 (0.095)
Pernambuco	0.1151 (0.07)
Alagoas	0.0756 (0.077)
Sergipe	0.2242 (0.086)***
Bahia	0.1137 (0.087)
Minas Gerais	-0.0013 (0.073)
Espírito Santo	0.2344 (0.07)***
Rio de Janeiro	0.2327 (0.066)***
São Paulo	0.0044 (0.067)
Paraná	-0.2987 (0.079)***
Santa Catarina	0.0053 (0.074)
Rio Grande do Sul	0.0692 (0.067)
Mato Grosso do Sul	-0.5560 (0.076)***
Mato Grosso	-0.1445 (0.09)
Goiás	-0.0180 (0.073)
Distrito Federal	0.1856 (0.069)***

Table A2

First stage of the Rivers-Vuong test

	Price of 20-cigarette pack (in logs)
Instrument for price	1.8337 (0.938)*
Gender (women=1)	0.0159 (0.004)***
Age categories	
25-44 years old	-0.0554 (0.007)***
45-64 years old	-0.1092 (0.007)***
65 and older	-0.0832 (0.009)***
Education categories	
Fundamental	0.1230 (0.006)***
High school	0.2155 (0.005)***
University	0.3186 (0.007)***
Labor categories	
Unemployed	-0.1015 (0.011)***
Out of labor force	-0.0336 (0.005)***
Student	-0.1028 (0.007)***
Monthly income	0.0000 (0.000)***
Fixed effects by PSU	YES

www.tobacconomics.org

INSTITUTE FOR
HEALTH RESEARCH
AND POLICY

UIC