CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Center for Economic Research and Graduate Education

Dissertation Thesis

2025 **Baisalova Aisha**

CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Center for Economic Research and Graduate Education

Aisha Baisalova

Spatial Heterogeneity in Consumer Responses to Excise Taxes: Evidence from Cross-State Cigarette Purchases in the United States

Dissertation Thesis

Prague 2025

Author: **Aisha Baisalova**
Supervisor: **prof. Nikolas Mittag, Ph.D.**.
Year of the defense: 2026

Dissertation Committee

Nikolas Mittag, Ph.D. (CERGE-EI, Chair)

Alena Bicakova, Ph.D. (CERGE-EI)

Stepan Jurajda, Ph.D. (CERGE-EI)

Referees

Rusty Tchernis, Ph.D. (Andrew Young School of Policy Studies - Georgia State University)

Michael F. Pesko, Ph.D. (University of Missouri)

Declaration

1. I hereby declare that I have compiled this thesis using the listed literature and resources only.
2. I hereby declare that my thesis has not been used to gain any other academic title.
3. I fully agree to my work being used for study and scientific purposes.

In Prague on September 30, 2025 Aisha Baisalova

Table of Contents

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Nikolas Mittag, for his constructive feedback, ongoing support, invaluable advice, and patience throughout this long process. I am also grateful to the members of my Dissertation Committee, Alena Bicakova and Stepan Jurajda, for their comments and insightful feedback.

I wish to thank Jan Hanousek for his helpful suggestions and detailed comments on the first chapter of my dissertation. My sincere appreciation goes to my referees, Rusty Tchernis and Michael F. Pesko, for their time and thoughtful evaluation.

I am thankful for the support provided by the CERGE-EI Academic Skills Center, particularly Deborah Novakova, Gray Krueger, and Andrea Downing, for their careful editing and valuable comments. I would also like to acknowledge the assistance of the Study Affairs Office with various academic and administrative matters.

Lastly, I gratefully recognize the financial support provided by the Grant Agency of Charles University, Project No. 440120.

In Prague on September 30, 2025 Aisha Baisalova

Abstract

This dissertation examines the impact of geographic proximity to lower-tax jurisdictions referred to as the “border effect” on consumer responses to cigarette excise taxes in the United States. Using NielsenIQ Consumer Panel data, the three chapters collectively analyze how this proximity influences the tax sensitivity of cigarette consumption, consumer behavioral responses, and the pass-through of excise taxes to retail cigarette prices.

The first chapter examines the bias created by border effects and shows that tax sensitivity of cigarette consumption is systematically attenuated for consumers residing near the borders of states with lower cigarette taxes. This bias is present across all demographic groups and declines with the distance from the lower-tax state border, reflecting greater opportunities for tax avoidance among those residing near lower-tax state borders.

The second chapter further investigates how border effects influence the tax sensitivity of cigarette consumption. We suggest a novel analytical framework that uses a threshold regression model with location, and time fixed effects to explicitly model the “border effect” as a linearly decreasing function of distance to the nearest lower-tax state, with a maximal influence at the border and a vanishing effect beyond a specified cutoff distance. Our analysis of demographic heterogeneity shows that tax sensitivity declines with income: low-income consumers are the most responsive to excise taxes, while high-income consumers are the least responsive. Although all groups are affected by border effects, for high-income consumers, the effect is limited to short distances, and disappears beyond 62 kilometers from the lower-tax border.

The third chapter examines how border effects influence the pass-through of taxes to cigarette prices. We analytically formulate tax pass-through as the “true” pass-through rate, adjusted by the “border effect”. We model the border effect as an exponential function that decreases with distance from the lower-tax state, reaches its highest value at the border itself and diminishes to zero at greater distances. To estimate the parameters of the “border effect” function, we employ an exponential regression model with location, time, and UPC fixed effects. Our analysis of demographic heterogeneity shows that high-income consumers exhibit the highest pass-through rates. In contrast, low-income consumers and those who are not engaged in paid employment are more likely to exploit cross-border purchasing and other tax avoidance opportunities, resulting in a reduced tax pass-through for these demographic groups.

Overall, the findings highlight that geographic proximity to lower-tax state borders significantly undermines the intended impact of excise tax policies. We show that cigarette excise taxes, while effective on average, can operate unevenly across geographies and demographic groups. Policymakers should consider both geographic and demographic factors in the evaluation and design of excise tax policy, particularly in a tax system with heterogeneous tax regimes such as the US.

Introduction

In the United States, cigarette excise taxes represent a significant source of government revenue and serve as a policy instrument with direct public health implications. Research shows that higher excise taxes lead to reductions in smoking rates, thereby improving public health outcomes. Nevertheless, the effectiveness of these tax policies is not uniform. Consumer responses to tax increases are influenced by various factors, including socioeconomic status, geographic location, and access to alternative purchasing options. Among these factors, the opportunity for cross-state-border shopping is particularly important.

In the context of the United States, where we can track the variability of excise taxes across states, geographic proximity to lower-tax states creates opportunities for tax arbitrage. Consumers residing near state borders can mitigate the impact of local tax hikes by purchasing cigarettes in neighboring states where the tax burden is lower. Moreover, driven by profit motives, shops close to borders may adjust prices to smooth the unfavorable tax difference to a certain extent. As a result, these "border effects" lead to spatial variation in both the tax sensitivity of cigarette consumption and the pass-through of taxes to retail prices, which tends to be systematically lower in areas close to lower-tax states. This effect diminishes as the distance from the lower-tax state border increases. While the research literature acknowledges the existence of cross-border purchasing, fewer studies have provided an analytical framework of how border effects shape the relationships between excise taxation, cigarette consumption, and tax pass-through rates. Employing comprehensive NielsenIQ Consumer Panel data, this dissertation presents a new theoretical framework that quantifies the extent of border effects in relation to both the tax sensitivity of consumption and the pass-through rate of cigarette taxes to prices.

The first chapter examines the bias arising from border effects and investigates how sensitivity to cigarette excise taxes and the size of bias vary for different demographic groups. We find that border effects create a bias in the estimate of consumption sensitivity to an increase in the excise tax rate, which is present for all demographic groups. Tax sensitivity increases with the average distance to the lower tax state border, implying that residence near such a border decreases the impact of excise tax policy interventions on consumer choice.

The second chapter examines in greater detail how border effects shape the responsiveness of cigarette consumption to taxation. We analytically formulate the "border effect" as a linear function that decreases with distance from the closest lower-tax state. We reasonably assume that the "border effect" reaches a maximum at the border with the lower-tax state and then linearly decreases with distance from the border, eventually reaching zero after a certain cutoff distance. We estimate the parameters of the "border effect" function, employing a threshold regression model with location and time fixed effects. As a robustness check, we also run a segmented regression using separate tax sensitivity estimates for a range of distance intervals. We verify that the estimates from segmented regression align with the linear pattern derived from the threshold model. Further, we enhance the "border effect" function by incorporating a difference between the home state tax and the closest lower-tax state tax as an additional factor, and then compare the estimation results for the two specifications. Beyond geographic variation, we also examine how the tax sensitivity of cigarette consumption differs across demographic groups. Our analysis shows that tax sensitivity varies by income level: high-income consumers are the least responsive to excise tax increases, while low-income consumers are the most responsive, with middle-income consumers falling in between. All income groups are influenced by the "border effect", but for high-income consumers this effect is only present at distances up to 62 kilometers from the lower-tax state. Our analysis based on employment status indicates that both employed and non-employed consumers display a similar shape in the "border effect" function; however, non-employed consumers show significantly higher tax sensitivity than employed consumers.

The third chapter analyzes how proximity to a lower-tax state affects the pass-through of tax to cigarette prices. We analytically formulate tax pass-through to prices as the "true" pass-through rate attenuated by the "border effect". In this chapter the "border effect" refers to the impact of cross-state purchasing behavior on the extent to which excise taxes are passed on to cigarette prices. We model the border effect as an exponential function that decreases with distance from the lower-tax state, reaching the highest value at the border itself and diminishing to zero at large distances. We estimate the parameters of the "border effect" function by employing an exponential regression model with location, time, and UPC fixed effects. The results of the robustness check, where we estimate a segmented regression using separate tax pass-through estimates for a range of distance intervals, support the linear pattern observed in the exponential model. We also extend the model to account for the tax differential between the home state and the nearest lower-tax state and perform a comparative analysis of the two model specifications. In addition to geographic variation, we also analyze how the pass-through of cigarette taxes varies across different demographic groups. High-income households face the highest tax pass-through and are largely unaffected by border proximity, while middle-income households are affected by the border effects only when the distance from the lower-tax state does not exceed 90 kilometers. Low-income households remain sensitive to border effects at greater distances, though their responsiveness declines beyond 200 kilometers. Moreover, consumers who are not engaged in paid employment exhibit significantly lower pass-through, suggesting greater scope for tax avoidance through flexible shopping behavior.

Taken together, the three chapters advance understanding of spatial variability in consumer responses to excise taxes within a tax system with heterogeneous tax regimes, such as the US. The dissertation makes several contributions. First, it examines how border effects can bias estimates of consumer sensitivity to changes in excise taxes. Second, the dissertation develops a theoretical framework that models how the impact of border effects decreases with distance from the nearest lower-tax state. This framework analyzes border effects from two complementary angles: the consumption responses to taxation and the pass-through of excise taxes to cigarette prices. Finally, beyond geographic heterogeneity, we also investigate how the tax sensitivity of cigarette consumption, the tax pass-through rate, and border effects differ among various demographic groups.

1 Exploring Border Effects: Sensitivity of Cigarette Consumption to Excise Tax

This study was supported by Charles University, GAUK project No 440120. Published as CERGE-EI Working Paper Series No 726.

Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1.1 Introduction

The majority of existing studies maintain that an increase in tax rates reduces the consumption of taxable products. Excise taxation is particularly important in the case of alcohol and cigarettes, as "sin" goods exhibit pronounced negative externalities that result in direct implications for public health.

Demographic composition is one of the various observable and unobservable factors that determines the consumer response to increases in excise taxes. According to statistics provided by the Centers for Disease Control and Prevention (hereinafter CDC) in the US, people with low levels of income and education have higher smoking rates than the average population. Heckley et al. (2017) show that people with a higher level of income and a college degree in total consume more alcohol compared to non-educated low-income adults, who are, however, are more exposed to binge drinking. All these facts indicate that heterogeneity in population significantly influences the social and welfare outcomes of measures taken by policymakers. Identifying what groups of population benefit or suffer from a specific tax policy can help policymakers to achieve their social goals. In general, a policy can be targeted not only to the aggregate population but also to a particular demographic group; for example, government policy to reduce alcohol consumption among youth. If the government's primary objective is the equal distribution of social benefits among the population, a significant increase in excise tax on cigarettes can lead to the opposite effect since poor consumers have higher propensity to smoke, resulting in a higher tax burden on this less fortunate stratum of the population. Regardless of what goals policymakers want to achieve, it is impossible to construct an appropriate public policy measure without knowing how welfare and public health implications will differ among various demographic groups.

Tax avoidance opportunities can serve as another important determinant of a consumer's purchase decision in response to an excise tax increase. Indeed, cross-state purchasing in the nearest lower-tax state decreases the impact of excise tax policy measures. Moreover, because of profit motives, shops close to borders may adjust prices to smooth the unfavorable tax difference to a certain extent. Ignoring these "border effects" leads to a biased estimate of the tax elasticity of consumption. The bias is particularly large for border residents, since the cost of traveling to the nearest lower-tax state to purchase taxable goods at the lower price increases with the distance to the state border. In this study, we estimate the bias arising from border effects and investigate how sensitivity to cigarette excise tax and the size of bias vary for different demographic groups. We specifically concentrate on excise taxation of cigarettes in the US, where we can track the variability of state excise taxes across states.

We use NielsenIQ Consumer Panel data for the years from 2004 to 2019 to explore whether ignoring border effects will bias the estimate of tax elasticity. We estimate the bias arising from cross-state tax avoidance opportunities by constructing a regression of cigarette consumption on the excise tax rate and other explanatory variables, and comparing the estimation results to the same regression specification with additional variables related to cross-border purchasing. Furthermore, we analyze how the tax sensitivity of cigarette consumption and the size of bias vary among households with different demographic compositions.

Our results show that the consumer response to a cigarette tax increase varies substantially between households with different demographic characteristics. We observe higher tax elasticity for the low income group. Higher tax sensitivity estimated for unemployed consumers and consumers without college degree can be potentially explained by the fact that, on average, these demographic groups have lower income. Furthermore, we identify that estimated tax sensitivity increases with smoking intensity, in contrast to Lee (2008) and Cotti et al. (2018), who show that heavy smokers do not respond to excise tax policy measures. Finally, we find that border effects create a bias in the estimate of consumption sensitivity to an increase in the excise tax rate, which is present for all demographic groups. Tax sensitivity increases with the average distance to the lower tax state border, implying that border residence decreases the impact of excise tax policy interventions on the purchase decision of consumers.

1.2 Previous Literature

The negative effect of excise tax increases on tobacco consumption has been discussed in numerous studies. Sensitivity of cigarette consumption to a tax increase is an important question for policy makers from the perspective of public health implications and tax revenue effects.

Using data from telephone survey conducted from April to July 2004 in 23 major cities and counties in Taiwan, Lee (2008) evaluates the effect on cigarette consumption of a large increase in cigarette tax of NT$22 per pack, which is equivalent to a 44% price increase. The study analyzes how price elasticity varies among different socio-demographic groups and finds that price sensitivity decreases with income and is higher for female smokers, moderate smokers, and smokers who purchase mid- and low-price cigarettes.

Cotti et al. (2018) use NielsenIQ Consumer Panel data for the years 2011 through 2015 to investigate how tobacco control polices, such as excise taxes and smoke-free laws, affected purchases of cigarettes, electronic cigarettes and smoking cessation products. The authors analyze the impact of these policy measures on the probability that a household purchases tobacco products and on the quantity of cigarettes purchased. The results indicate that excise taxes decrease both these parameters, and smoke-free air laws decrease the quantity of tobacco products consumed. Cotti et al. (2018) investigate the heterogeneity of these effects for various demographic groups in order to understand what subgroups respond more strongly to tobacco control measures. According to the results, older households are more responsive to excise tax increases in cigarette consumption and, conversely, younger households respond more strongly in e-cigarette consumption. Furthermore, analysis of heterogeneity depending on the household's cigarette purchase level shows that light smokers decrease cigarette consumption in response to an excise tax increase, while for heavy smokers the effect of this tax policy measure is insignificant. Moreover, low income smokers are more sensitive to an excise tax increase compared to high income consumers, which is consistent with economic theory assumptions.

Pesko et al. (2020) find evidence that higher traditional cigarette tax rates reduce adult traditional cigarette use and increase adult e-cigarette use.The estimates are based on the data from the Behavioral Risk Factor Surveillance System and National Health Interview Survey over the period from 2011 to 2018. The effects were examined across demographic sub-groups. The study shows that younger consumers have higher own- and cross-tax responsiveness as younger adults are much more likely to use e-cigarettes than other groups of adults.

One limitation of these studies is that they do not take into consideration the fact that tax sensitivity can be affected by possible tax avoidance actions of consumers, such as stockpiling if the future increase of taxes is known in advance or cross-border purchasing in the nearest lower-tax state. Since the consumer decision is determined by the final purchase price, imperfect tax pass-through to prices may bias the estimate of tax sensitivity and decrease the applicability of the obtained results.

In their study, Harding et al. (2012) show that in the US cigarette taxes are less than fully passed through to prices mainly due to cross border purchasing. Using information on consumer location provided in NielsenIQ scanner data for the years 2006–2007, the authors show that tax avoidance opportunities create significant differences in the pass-through rate of taxes to prices. Kim & Lee (2020), employing a similar estimation strategy to that used by Harding et al. (2012), find that cigarette taxes are shifted significantly less to consumer prices in cities with large minority (black and Hispanic) populations. The estimates are obtained using NielsenIQ scanner data on cigarette sales for the years 2009–2011 from 1,687 stores across the US. Xu et al. (2014) investigate how tax pass-through rate differs between premium and generic brands of cigarettes and conclude that for premium brands consumers bear a full tax burden with an additional premium, i.e. pass-through rate is higher than 100%, whereas consumers of generic brands pay only 30-83 cents for every 1$ tax increase.

Imperfect tax pass-through stemming from potential tax avoidance opportunities creates a bias in the estimate of consumption sensitivity to an increase in the excise tax rate. A number of studies analyze the impact of cross-state purchasing on smoking behavior.

For example, Lovenheim (2008) examines the impact of border effects on price elasticity using data from Current Population Survey Tobacco Supplements spanning from September 1992 to February 2002. The study finds that demand elasticities with respect to the home state price are indistinguishable from zero on average and vary significantly with the distance individuals live to a lower–price border. However, when tax avoidance opportunities are eradicated, the price elasticity is negative but still inelastic.

Using CPS Tobacco Use Supplement (TUS) data for February, June, and November 2003, Chiou & Muehlegger (2008) introduce a discrete choice model to examine tax avoidance and state border crossing in the market for cigarettes. The authors estimate a consumer's tradeoff between distance and price when choosing a location to maximize utility, which allows them to simulate tax avoidance under alternative cigarette excise tax levels.

1.3 Data

We obtain historical data on state cigarette excise taxes from the Centers for Disease Control and Prevention.

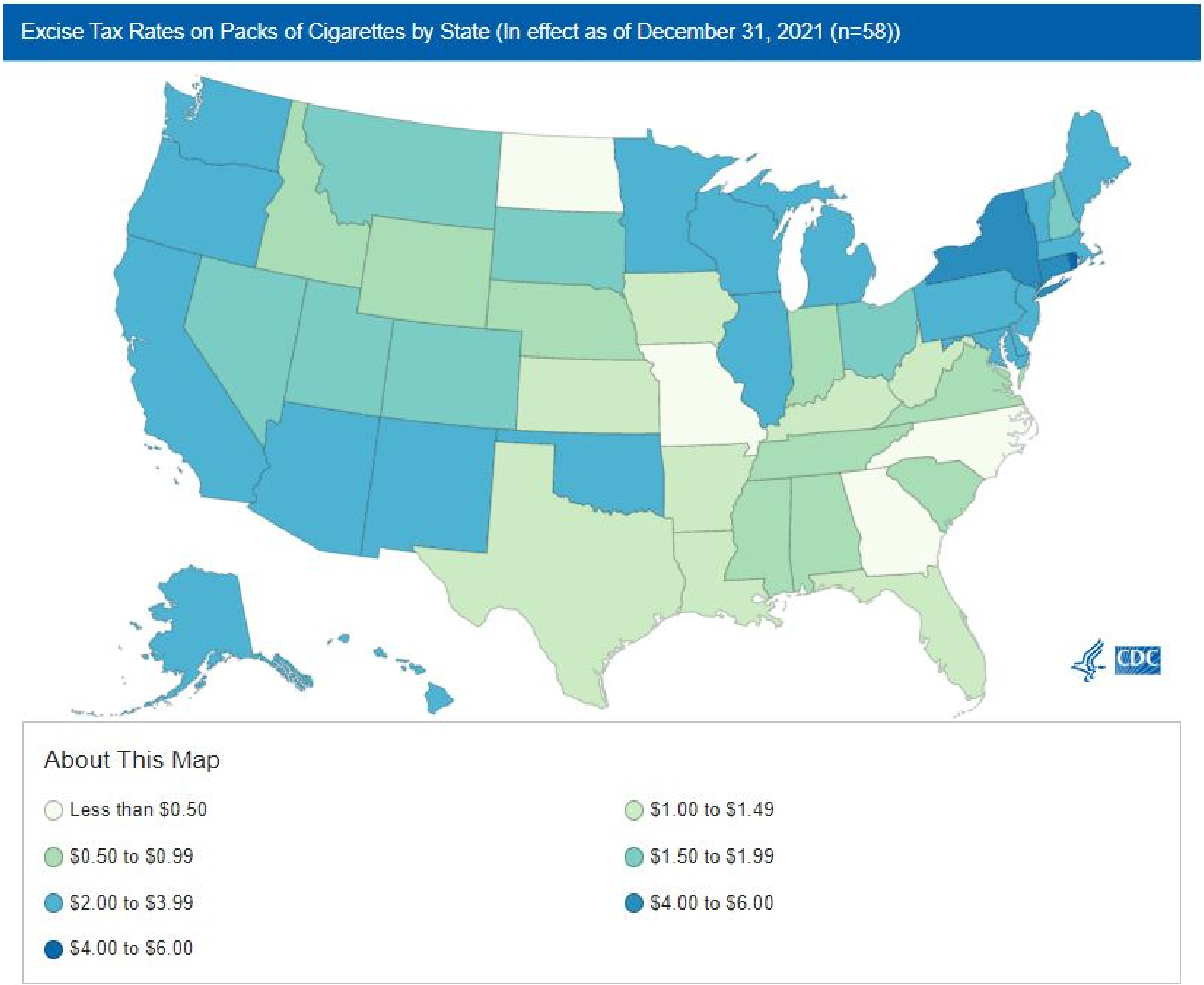


Figure 1: Excise Tax Rates on Packs of Cigarettes by State.

Source: Centers for Disease Control and Prevention (CDC), 2021

Excise tax rate data is available on quarterly frequency. The main advantage of using US data is that in the US excise taxes are not uniform and exhibit significant variability across states. This allows us to to take into account not only changes in excise taxes over time, but also state-level heterogeneity. Figure 1 displays the variation in cigarette excise taxes across US states as of December 2021.

We use NielsenIQ Consumer Panel Data containing information about the purchase history of 40,000-60,000 households (varies by year) who continually provide information to NielsenIQ about their demographic characteristics, products they buy, as well as timing and location where they make purchases in a longitudinal study. Consumer panelists use in-home scanners to record all of their purchases intended for personal, in-home use. Panelists are geographically dispersed and demographically balanced (James M. Kilts Center for Marketing, NielsenIQ datasets, n.d.).

The scanner data covers 3,158,152 cigarette purchase transactions made by 52,726 households spanning from 2004 until 2019. Further, the transactional data set was transformed to panel data by aggregating the data to the household-quarter level. The frequency of the panel data set coincides with the frequency of historical cigarette tax data obtained from the CDC database.The resulting panel data set comprises 378,101 observations of quarterly cigarette purchases. The data set covers the demographic characteristics of the households, including income range, size, gender composition, presence and age of children, marital status, type of residence, race, and Hispanic origin. Additionally, it includes geographic characteristics, such as the panelist's zip code and product characteristics, which contain UPC code, description, brand, multi-pack, and size. The geographies of the data cover the entire United States (James M. Kilts Center for Marketing, NielsenIQ datasets, n.d.).

The major advantage of the NielsenIQ database is that it monitors the residence address of panelists. This allows us to incorporate geographic controls in our estimation strategy. Figures 2 and 3 show the geographic distribution of household locations based on proximity to lower-tax state borders. Specifically, they depict panelist ZIP codes located within 20 kilometers and more than 100 kilometers of such borders, respectively.

We measure the distance to the nearest lower-tax state using Census TIGER/Line shape files provided by United States Census Bureau. We estimate the distance between consumers and lower-tax borders as the distance from the household's census tract of residence provided in the data to the border of the closest lower tax state. The household's census tract of residence is defined as the centroid of the panelist's ZIP code. The lower tax state does not need to be a border state. We identify the coordinates of boundaries for each US state and calculate the distance from each panelist ZIP code to the state boundaries of every US state.

We estimate the distance to the lower tax state for each time period and consumer zip code as the closest distance to the border of the state with the lower state cigarette tax. Further, we match the tax rate with the corresponding lower tax state.

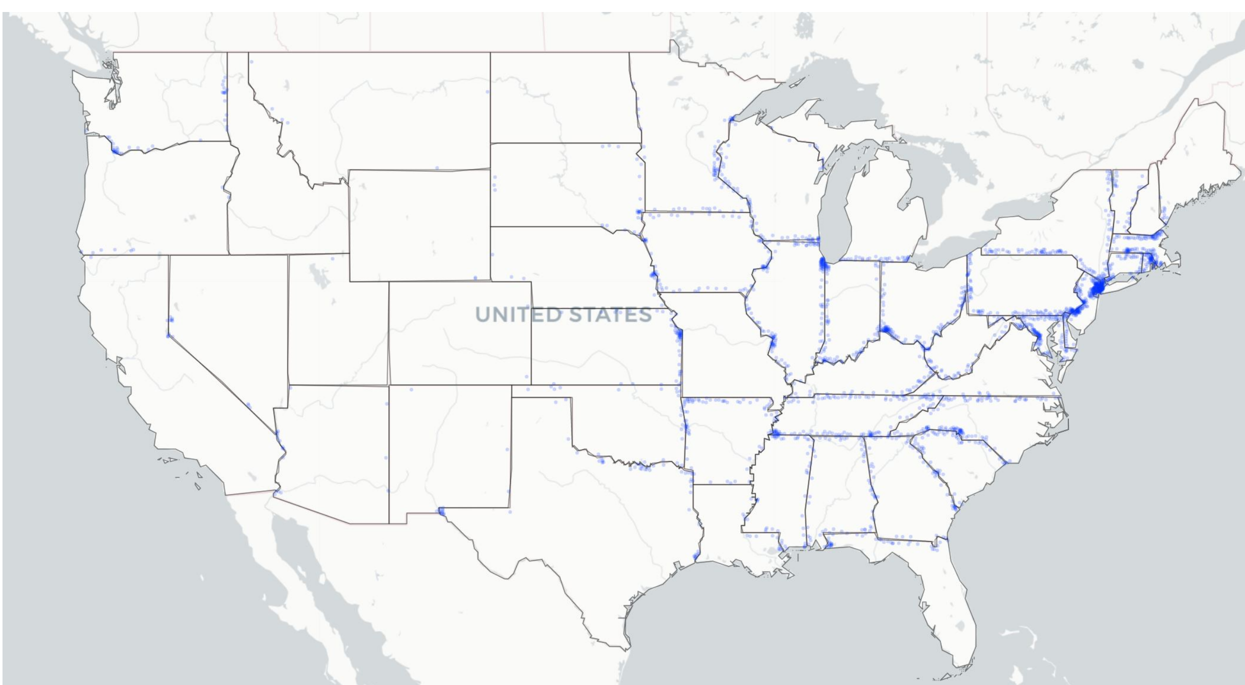


Figure 2: Distribution of Panelists Residing Near the Border of a Lower Tax State.

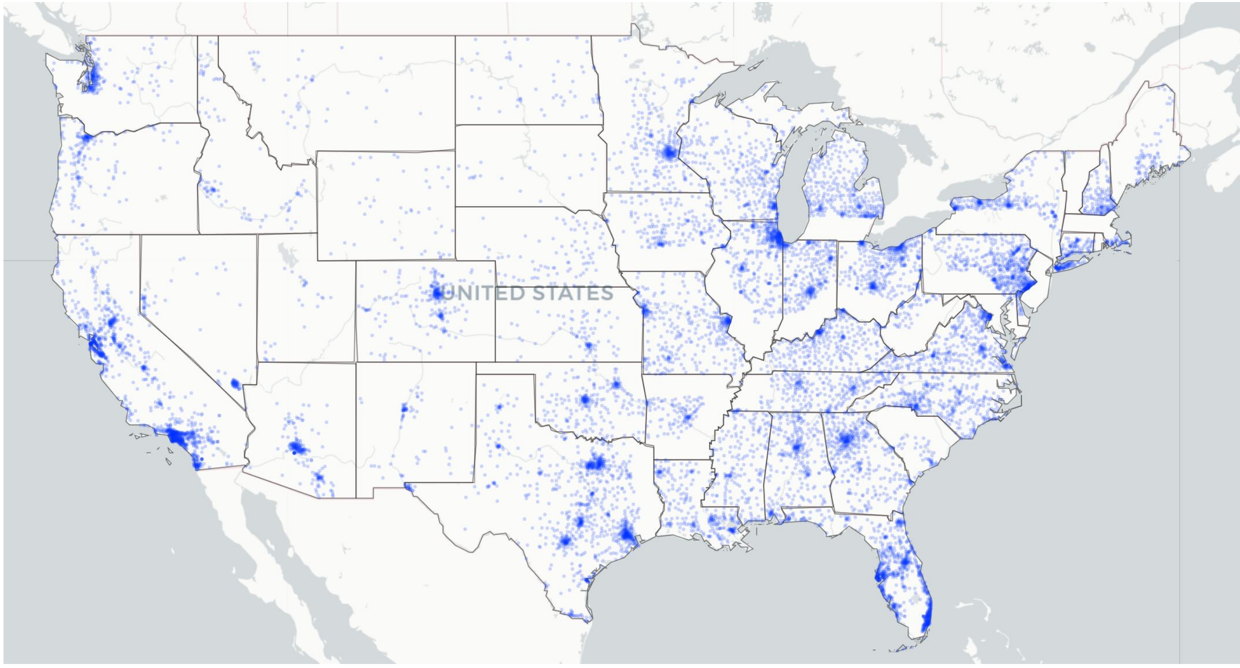


Figure 3: Distribution of Panelists Residing Far from the Border.

Since we measure the distance to the lower tax state for each time period, we are able to properly capture the state and time level heterogeneity in cigarette taxes and the cost of cross-border purchasing. One restriction imposed on the data used for model estimation is that we have not taken into account distances to lower-tax neighboring countries, such as Mexico. Although these cross-border effects may also be substantial, they are not considered in the present analysis.

Table 1 summarizes the descriptive statistics of analysis variables in the created panel data set.

Table 1: Descriptive Statistics of Analysis Variables

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Total packs purchased	46.144	63.484	0	5	64	2,234
Price per pack	4.575	3.598	0	3.2	5.5	600
Lower tax state value	0.704	0.608	0.025	0.300	0.995	4.250
Lower tax state distance	189.487	207.521	0.000	52.342	249.092	1,218.755
Tax value	1.210	0.852	0.025	0.550	1.600	4.350
Tax difference	0.506	0.469	0.000	0.130	0.810	2.810
Tax distance interaction	76.484	100.273	0.000	13.214	102.538	1,050.534
Smoking rate [1]	46.144	52.101	0	9	65.4	669

We create the following rules for the construction of categorical analysis variables. The majority of these variables are demographic characteristics. The distribution of variables by categories is presented in Table 2.

[1]Smoking rate is calculated as the average number of cigarette packs consumed during a quarter

Table 2: Rules for Construction of Categorical Variables

Category	N
Household Income	
High: Annual income≥70,000$	$83,242$
Middle: Annual income 30.000$ - 69.999$	$174,766$
Low: Annual income < 30.000$	$120,093$
Household size	
1: 1 member	$89,722$
2: 2 members	$159,819$
3: 3 members	$61,672$
4: 4 members	$39,480$
5: 5 members	$16,640$
6 plus: $\geq$ 6 members	$10,768$
Head Employment [2]	
≤35 hours	$38,886$
35+ hours	$198,079$
Not employed	$141,136$
Head Education [2]	
BA plus	$94,314$
Some college	$130,103$
High school graduate or lower	$153,684$
Head Age [2]	
< 35 years	$18,481$
35-49	$104,087$
$\geq$ 50	$255,533$
Presence of children	
0: No children	$298,333$
1: Children present	$79,768$
Gender composition	
Female and male head	$231,856$
Female head only	$104,700$
Male head only	$41,545$
Border residence	
Residence $\leq$ 25 km. from the lower tax state border	$53,824$
Residence $\succ$ 25 km. from the lower tax state border	$324,277$
Smoking rate	
Heavy smoker: $\geq$ 80th percentile	$156,655$
Average smoker: 30th percentile - 80th percentile	$186,861$
Light smoker: $\leq$ 30th percentile	$34,585$

[2]Note: The sample includes only those households in the NielsenIQ Homescan data sample that make at least

1.4 Estimation Strategy

This section considers the econometric model that measures the household's tax sensitivity with regard to cigarette consumption. Tax avoidance opportunities, such as cross-border purchasing in the closest lower-tax state, may have a considerable impact on the estimate of tax elasticity. Households buy taxable goods in the shops belonging to the nearest lower-tax state if the transportation costs are lower than the benefits from buying taxable goods at a lower price. Moreover, because of profit motives, shops close to the borders may adjust prices to smooth the unfavorable tax difference to a certain extent. The combination of these factors constitutes "border effects". Omitting the border effects from estimation creates a bias in the estimate of consumption sensitivity to an increase in the excise tax rate. The bias arising from the border effects is estimated by constructing a regression of cigarette consumption on the excise tax rate and household demographic characteristics and comparing estimation results to the same regression specification with additional variables related to border effects.

In the first regression specification (1), we regress cigarette consumption on excise tax rate, distance to the nearest border of lower-tax state, difference in the tax rate between state of residence and lower-tax state, interaction between distance and tax difference, and household's demographic characteristics. In addition, the econometric model should consider the fact that different states are heterogeneous by their nature and vary by economic factors, such as GDP per capita, poverty rate, cultural factors, smoke-free laws and many others, as well as the fact that individuals can have different search costs, attitudes towards stockpiling behavior, etc. This fact should be incorporated in our model through state level and household fixed effects. Household fixed effects control for unobservable individual-level heterogeneity and, therefore, reduce heterogeneity bias. State-level fixed effects represent geographic controls. The analytical formulation of the panel data regression is the following:

$$cig_{ijt} = \alpha_0 + \alpha_1 \tau_{jt}^h + \alpha_2 \left(\tau_{jt}^h - \tau_{jt}^b\right) + \alpha_3 D_{ijt} + \alpha_4 D_{ijt} \times \left(\tau_{jt}^h - \tau_{jt}^b\right) + \beta X_i + \sigma_i + \omega_j + \epsilon_{ijt}, \quad (1)$$

where cig_{ijt} is the number of cigarette packs consumed by a household i in state j and time t;
τ_{jt}^h is the home state tax;
τ_{jt}^b is the closest lower-tax state's tax;
D_{ijt} is the distance to the closest lower-tax state;
X_i is a vector of household demographic characteristics;
σ_i and ω_j are individual and state level fixed-effects.

The coefficient α_1 represents the tax sensitivity of cigarette consumption, measuring the reduction in quarterly household cigarette purchases associated with a one-dollar increase in the state cigarette tax. The coefficient α_2, associated with the tax difference between the home state tax and the closest lower tax state, reflects the attenuation of this tax effect when a neighboring state has a low excise tax rate. The coefficients α_3 and α_4 represent the costs associated with cross-border purchasing and show the marginal increase in tax sensitivity of cigarette consumption with each one-kilometer increase in distance from the nearest lower-tax state.

one cigarette purchase. "Head Age" and "Head Education" refer to male household head if a male household head is present. In the cases in which no male household head is present, these variables refer to the female household head. This is in line with the study by the National Institute of Drug Abuse (April 2021) that finds men tend to use tobacco products at higher rates than women, and therefore men are more likely to be the primary buyers of cigarettes in grocery stores in a two-headed household.

In the second regression specification (2), we use the same model but without variables related to border effects, which are distance to the nearest border of lower-tax state, difference in the tax rate between state of residence and lower-tax state, interaction between distance and tax difference.

$$cig_{ijt} = \alpha_0 + \alpha_1 \tau^h_{jt} + \beta X_i + \sigma_i + \omega_j + \epsilon_{ijt}, \tag{2}$$

where cig_{ijt} is the number of cigarette packs consumed by a household i in state j and time t;
τ^h_{jt} is the home state tax;
X_i is a vector of household demographic characteristics;
σ_i and ω_j are individual and state level fixed-effects.

Estimated sensitivity of cigarette consumption to an increase in the tax rate is compared in these two model specifications. The difference between two estimates of the coefficient on the excise tax rate constitutes a bias arising from omitting variables related to border effects. The regression specification was separately estimated for each demographic group in order to test for the presence of bias related to cross-state purchasing. This allows us to analyze how the size of bias and estimated tax sensitivity vary among households with different demographic compositions. The set of robustness checks for the proposed model specification is summarized in Section 1.6.

It is worth noting that we employed a "within" fixed effects model that measures the within-individuals and within-states variability in the variables, since we are using household level and state level fixed effects as their own controls. Therefore, coefficients on demographic characteristics represent the marginal change in cigarette consumption associated with changes in socioeconomic characteristics for a particular household, for example, an increase in the household's size, change in head employment status, household has a child, etc. Coefficient on home state tax measures the consumer response to a change in the home state tax as state-level fixed effects absorb state level heterogeneity. A similar interpretation applies to coefficient on difference in the tax rate between state of residence and lower-tax state and distance to the closest lower-tax state. Distance to the lower-tax state border can change due to a change in the panelist's residence or tax change in the neighboring states. Note that distance to the lower tax state was estimated for each time period and panelist's zip code, which allows us to properly capture time level heterogeneity in the cost of cross-state purchasing. Table 3 shows the distribution of households that experienced a change in the distance to the closest lower-tax state.

Table 3: Distance to the Closest Lower-tax State

Distance change	Number of households
Distance to the closest lower-tax state not changed	40,054
Distance to the closest lower-tax state changed	12,672

NielsenIQ tracks households that change their residential address. Consequently, state fixed effects are not absorbed by individual fixed effects. Table 4 presents the distribution of households that relocated, i.e., their distance to the closest lower-tax state changed due to relocation. As a supplementary robustness check, households that changed their residential address were excluded from the sample, and the regression was estimated using only households that did not relocate. We observe that the regression results have not changed significantly. The estimation results for non-relocating households, excluding state fixed effects, are reported in the Appendix A.

Table 4: Distribution of Households by ZIP Code Change

ZIP code change	Number of households
ZIP code not changed	$49,376$
ZIP code changed	$3,350$

The choice of the "within" fixed effects model was determined by the large heterogeneity bias in the regression specification without household-level fixed effects, which resulted in the coefficient estimates on demographic characteristics being inconsistent with economic theory assumptions. Nevertheless, estimated tax sensitivity and variables related to border effects are within a similar range. Estimation results of an alternative regression specification without household-level fixed effects are presented in the Appendix B.

1.5 Estimation Results

Table 5 summarizes the fixed effects regression model results of quarterly cigarette consumption for the following two specifications. In column (1), we estimate the model specification with variables related to border effects, which are distance to the nearest border of lower-tax state, difference in the tax rate between states of residence and lower-tax state, interaction between distance and tax difference. In column (2), we removed variables related to border effects in order to assess the presence of omitted variable bias. In addition to demographic characteristics, we added household and state-level fixed effects for both model specifications (1) and (2) to control for individual-level and geographic heterogeneity.

We observe that tax sensitivity in the model specification with variables related to border effects is larger than in the similar specification excluding these variables. Moreover, variables related to border effects are statistically significant in the model specification (1). Therefore, estimate of tax elasticity is biased when variables related to "border effects" are omitted from the model. As a result, sensitivity of cigarette consumption to a change in cigarette tax in model specification (2) is underestimated.

We observe that a one-dollar tax increase decreases households' quarterly cigarette consumption by approximately 14 packs. The positive coefficient on the tax difference between the home state tax and the closest lower tax state reflects the attenuation of this tax effect when a neighboring state has a low excise tax rate. Negative coefficients on distance variables related to "border effects" show that households residing at greater distances from the nearest lower-tax state exhibit higher sensitivity of cigarette consumption to tax changes compared to border residents.

The estimation results show how demographic groups vary by intensity of cigarette consumption. For example, smoking intensity is lower for older consumers and those with higher income, implying that young low-income consumers contain the largest share of heavy smokers. Households with children have lower cigarette consumption, which is consistent with the existing studies. Lin H. (2020) evaluates the existence of the upward inter-generational effect of the presence of children on parents' smoking behavior in China. The estimation results show that the number of children is significantly inversely associated with smoking behavior. Households with a single female head have a lower amount of quarterly cigarette purchases than households with a single male head. This is in line with the study by the National Institute of Drug Abuse (April 2021) that men tend to use tobacco products at higher rates than women.

Further, we estimate the same regression specification for different demographic groups. This allows us to analyze how tax elasticity and the size of bias vary among households with different demographic compositions. The results with estimated tax sensitivity among heterogeneous consumer groups for two regression specifications with and without variables related to border effects are presented in Table 6.

We find that border effects create a bias in the estimate of consumption sensitivity to an increase in excise tax rate, which is present for all demographic groups. Border effects affect all demographic groups, which is confirmed by the presence of bias when omitting variables related to border effects. The bias is particularly large for border residents, since the cost of traveling to the nearest lower-tax state to purchase taxable goods at a lower price increases with the distance to the state border. Therefore, border residence may decrease the impact of excise tax policy interventions on the consumer's purchase decision.

Table 6 demonstrates that estimated elasticities are larger for the low income group. Higher tax sensitivity estimated for unemployed consumers and consumers without college degree can be potentially explained by the fact that, on average, these demographic groups have lower income. Moreover, from the estimation results of the panel regression presented in Table 5, we observe a decreasing pattern of cigarette consumption with age. Lower sensitivity of young consumers to a cigarette tax increase can be partially attributed to a lower reaction of this demographic group to policy measures and smoking bans as opposed to adult consumers. This result is in line with Lee (2008), who shows that adolescent smokers under 18 years of age have lower cigarette price elasticity. Nevertheless, a possible reason for the "irregular" coefficient on the tax rate can stem from the small population sample of young consumers, which comprises only $18,481$ observations. Furthermore, we identify that estimated tax elasticity increases with smoking intensity in contrast to Lee (2008) and Cotti et al. (2018), who show that heavy smokers do not respond to excise tax policy measures.

Table 5: Estimation of Baseline Model on the Whole Population Sample.

	Dependent variable:	
	Total packs purchased	
	(1)	(2)
Tax difference	4.858^{***}	
	(0.484)	
Lower tax state distance	-0.004^{***}	
	(0.001)	
Tax distance interaction	-0.024^{***}	
	(0.002)	
Tax value	-13.902^{***}	-12.294^{***}
	(0.300)	(0.221)
Factor: Low income	1.560^{***}	1.705^{***}
	(0.472)	(0.472)
Factor: Middle Income	1.401^{***}	1.487^{***}
	(0.355)	(0.355)
Factor: Household size 2	2.399^{***}	2.521^{***}
	(0.453)	(0.453)
Factor: Household size 3	5.336^{***}	5.512^{***}
	(0.547)	(0.547)
Factor: Household size 4	4.294^{***}	4.532^{***}
	(0.657)	(0.657)
Factor: Household size 5	8.387^{***}	8.670^{***}
	(0.831)	(0.832)
Factor: Household size 6 plus	5.164^{***}	5.410^{***}
	(0.985)	(0.985)
Factor: Head employment 35+ hours	3.246^{***}	3.317^{***}
	(0.387)	(0.387)
Factor: Head employment Not employed	-1.123^{***}	-1.284^{***}
	(0.397)	(0.397)
Factor: Head education HS graduate or lower	-2.402^{***}	-2.390^{***}
	(0.528)	(0.528)
Factor: Head education Some college	-1.350^{***}	-1.295^{***}
	(0.428)	(0.428)
Factor: Head age $\geq$50	-9.561^{***}	-10.073^{***}
	(0.896)	(0.896)
Factor: Head age 35-49	-2.767^{***}	-2.957^{***}
	(0.844)	(0.844)
Factor: Presence of children = yes	-2.032^{***}	-1.979^{***}
	(0.440)	(0.440)
Factor: Gender composition Female head only	-8.270^{***}	-8.319^{***}
	(0.551)	(0.551)
Factor: Gender composition Male head only	-4.720^{***}	-4.725^{***}
	(0.859)	(0.859)
Consumer fixed effects:	*yes*	*yes*
State fixed effects:	*yes*	*yes*
Observations	378,101	378,101
R^2	0.018	0.017
F Statistic	87.288^{***} (df = 67; 325308)	86.855^{***} (df = 64; 325311)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 6: Estimated Tax Sensitivity among Heterogeneous Consumer Groups

Demographic Group	Coefficient Estimate on τ^h (1)	(2)	Omitted Variable Bias	Omitted Variable Bias as % of $\tau^h_{(2)}$	D	$\tau^h - \tau^b$	$D \times (\tau^h - \tau^b)$
Border resident	-19.929***	-9.639***	10.290	107%	-0.394**	15.247***	1.920
Not border resident	-14.225***	-12.798***	1.427	11%	-0.011***	2.725***	-0.380***
Heavy smoker	-24.079***	-20.863***	3.216	15%	-0.007***	9.203***	-1.107***
Average smoker	-6.451***	-5.402***	1.049	19%	-0.001**	3.022***	-0.272***
Light smoker	-0.271***	-0.176***	0.095	54%	-0.0001	0.151**	0.005
High income	-13.294***	-11.832***	1.462	12%	-0.004**	5.014***	-0.667***
Middle income	-13.600***	-11.689***	1.911	16%	-0.004***	5.563***	-0.026***
Low income	-15.243***	-13.606***	1.637	12%	-0.004***	4.108***	-0.472***
Head employment: 35+ hours	-12.106***	-10.509***	1.597	15%	-0.003**	5.158***	-0.680***
Head employment: ≤35 hours	-11.741***	-9.158***	2.583	28%	0.0005	7.790***	-0.537***
Head employment: Not employed	-16.019***	-14.069***	1.950	14%	-0.009***	4.150***	-0.458***
Head education: HS graduate or lower	-15.975***	-13.493***	2.482	18%	-0.009***	4.978***	-0.488***
Head education: Some college	-12.556***	-11.541***	1.015	9%	0.002	5.370***	-0.730***
Head education: BA +	-11.426***	-10.100***	1.325	13%	-0.003**	4.895***	-0.703***
Head age: ≥50	-14.310***	-13.011***	1.299	10%	-0.003***	4.852***	-0.705***
Head age: 35-49	-13.756***	-11.183***	2.573	23%	-0.004**	5.787***	-0.445***
Head age: < 35 years	-6.843***	-5.418***	1.425	26%	0.003	6.794***	-1.475***
Presence of children: yes	-12.079***	-9.181***	2.898	32%	-0.005***	6.052***	-0.418***
Presence of children: no	-14.298***	-12.913***	1.385	11%	-0.004***	4.380***	-0.600***
Gender composition: Female head only	-12.354***	-10.523***	1.831	17%	-0.001	5.541***	-0.485***
Gender composition: Female and male head	-14.764***	-13.266***	1.498	11%	-0.004***	4.761***	-0.705***
Gender composition: Male head only	-12.536***	-10.615***	1.920	18%	-0.006***	4.627***	-0.532***

Note: *p<0.1; **p<0.05; ***p<0.01; $\tau^h_{(2)}$ refers to τ^h estimate in spec (2); Omitted Variable Bias is estimated as $\tau^h_{(2)} - \tau^h_{(1)}$; D, $\tau^h - \tau^b$, $D \times (\tau^h - \tau^b)$ refer to distance to the nearest border of lower-tax state, difference in the tax rate between states of residence and lower-tax state, interaction between distance and tax difference respectively

1.6 Robustness analysis

As a robustness check, we want to ensure that tax sensitivity τ^h in model specification (2) on average exhibits a decreasing pattern when we subsequently remove households residing near a lower-tax state border from the estimation. We start with the whole population sample and estimate the tax elasticity of cigarette demand for each demographic group. Further, we subsequently exclude border residents residing less than 5, 10, 15, ..., 50 kilometers away from the border and re-estimate the tax sensitivity for each population group. We performed the same exercise for the aggregate sample. The decreasing pattern of the negative coefficient on the home state tax τ^h implies that the cost of cross-border purchasing increases with the distance to the lower-tax state border. Therefore, the tax sensitivity estimate gradually converges to the unbiased estimate when border effects are eliminated.

Figure 4 demonstrates the expected decreasing pattern of tax sensitivity for the aggregate sample when we increase the cost of cross-border purchasing.

We performed the same robustness check for each demographic group. We observe a similar pattern in the evolution of tax sensitivity for unemployed and low-income consumers. Tax elasticity reaches its minimum at the 15^{th} and 25^{th} kilometer from the lower-tax state border and then demonstrates an increasing pattern. One potential explanation can be that the impact from the gradual decrease of the population sample size outperforms the influence of border effects on the tax sensitivity estimate for these demographic groups. For the remaining demographic groups by head education, head age, gender composition, we observe a decreasing pattern when average distance to the lower tax state border is gradually increasing. It is worth noting that tax sensitivity for households with a female head reaches its minimum at the 15^{th} kilometer, compared to the male head households at the 25^{th} kilometer, which may indicate that the border effect for female consumers reaches its maximum at the lower distance. Tax sensitivity by smoking intensity demonstrates an expected decreasing pattern with a more pronounced effect for heavy and average smokers. Tax elasticity of light smokers still preserves a bias related to border effects; nevertheless, it is comparatively smaller as a measure of predicted decrease in quarterly cigarette consumption in response to a 1$ tax increase.

The performed robustness check confirmed the validity of border effects and the direction of bias. Tax sensitivity, on average, gradually increases in absolute value when we subsequently remove border residents and increase the average distance to the lower tax state border.

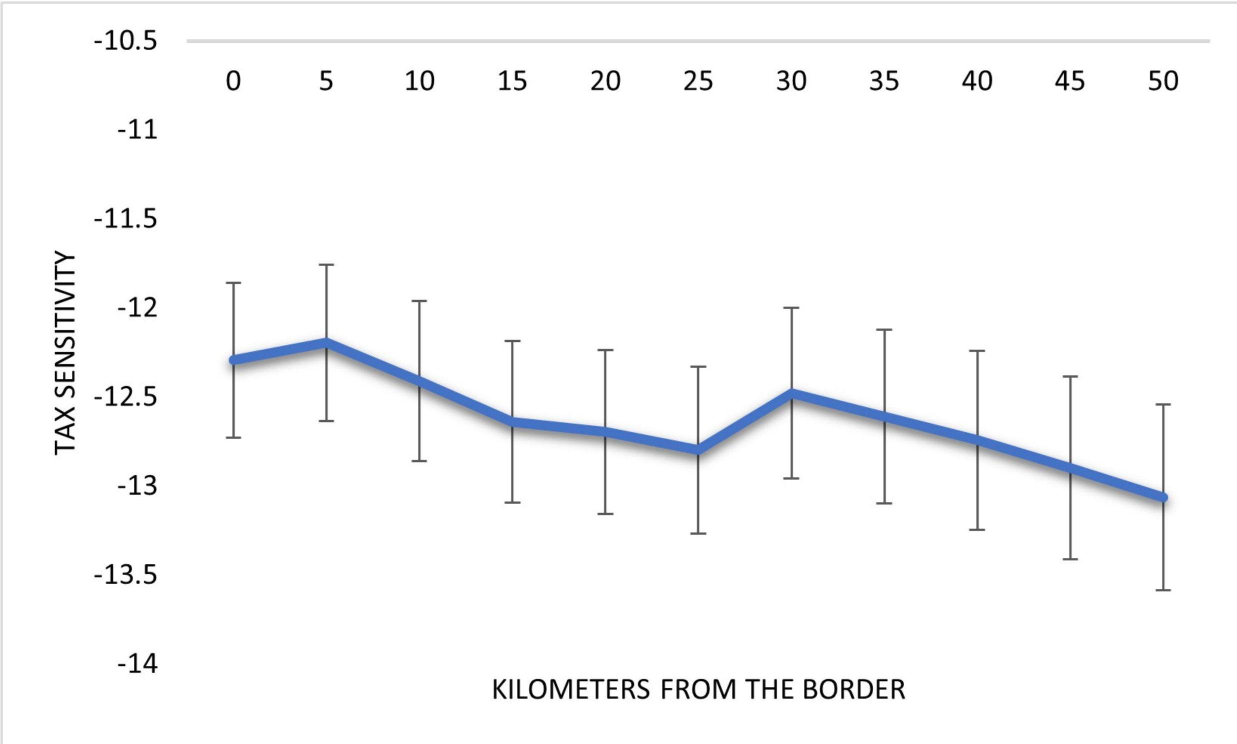


Figure 4: Excluding households residing near the border: Aggregate Sample

[*] The error bands show the bounds of the 95 percent confidence interval.

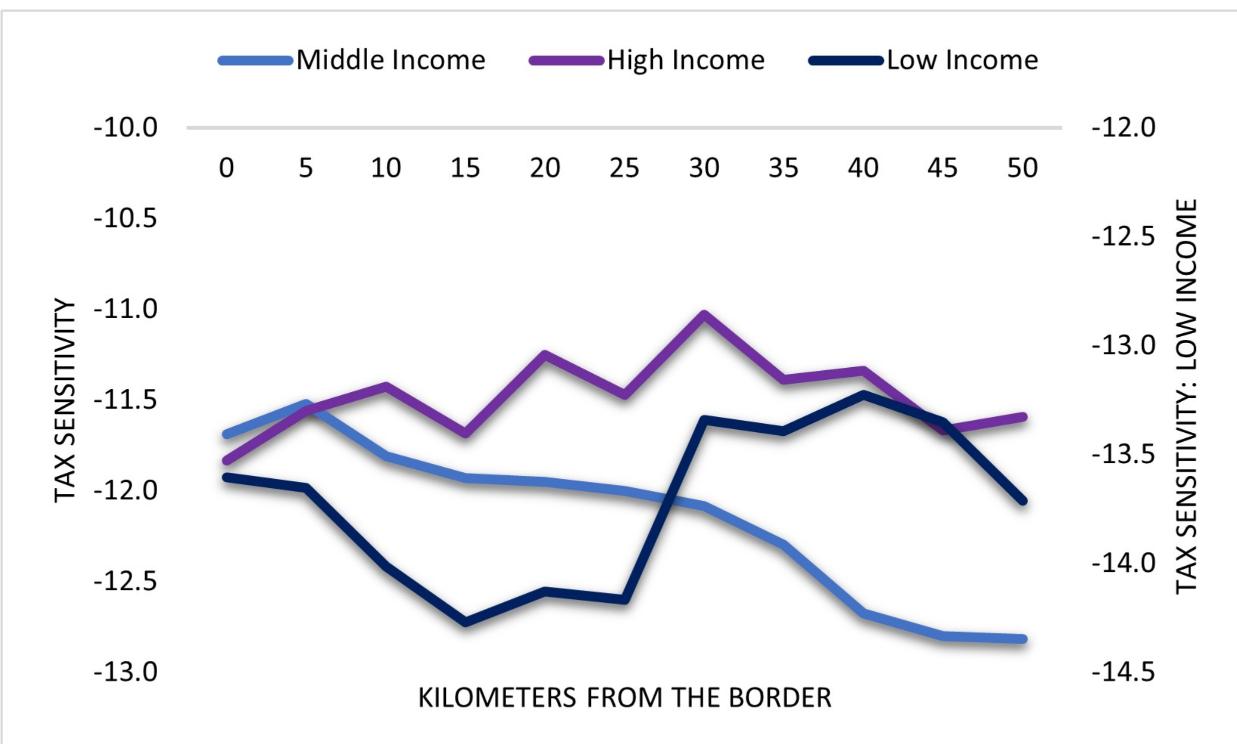


Figure 5: Excluding households residing near the border: Household Income

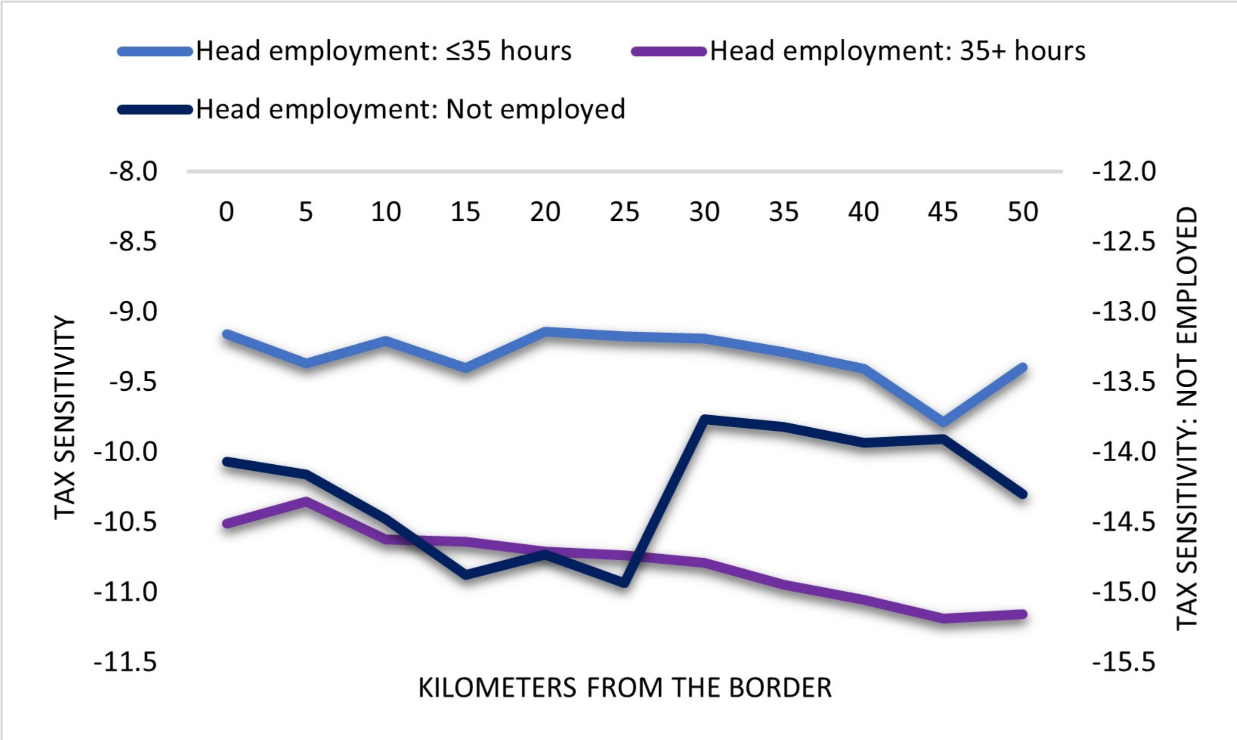


Figure 6: Excluding households residing near the border: Head Employment

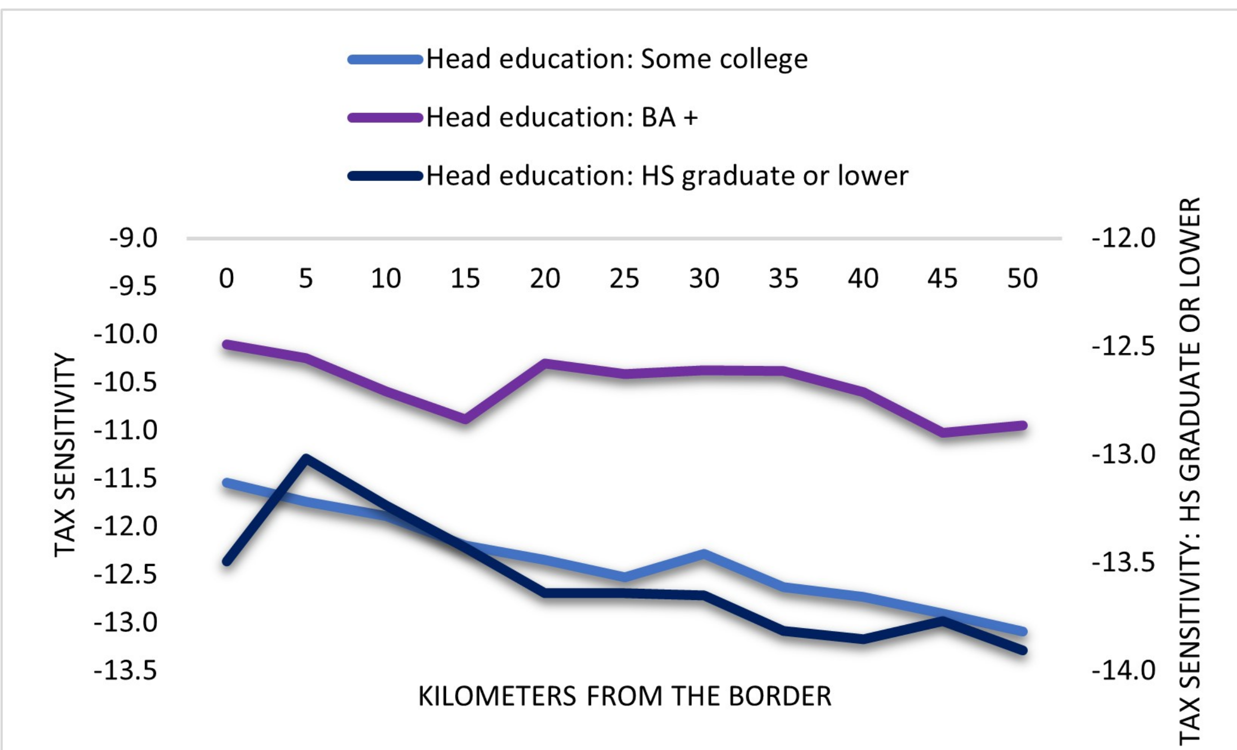


Figure 7: Excluding households residing near the border: Head Education

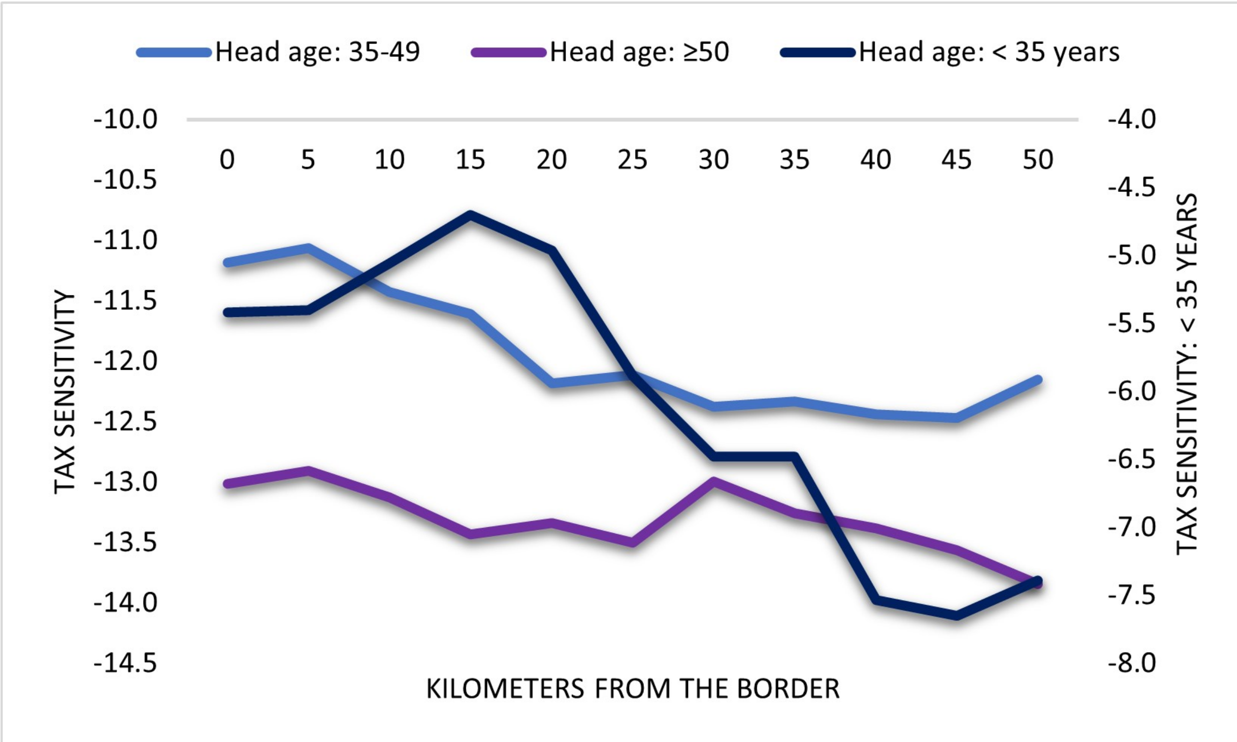


Figure 8: Excluding households residing near the border: Head Age

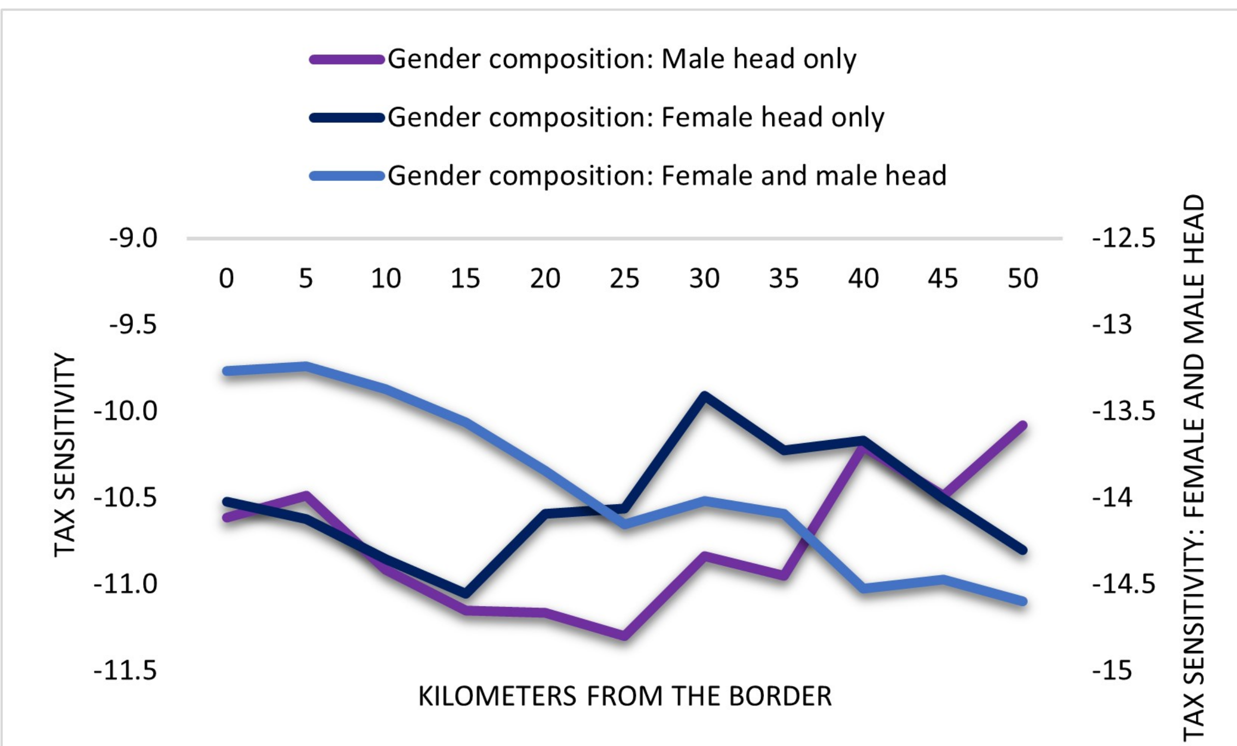


Figure 9: Excluding households residing near the border: Gender Composition

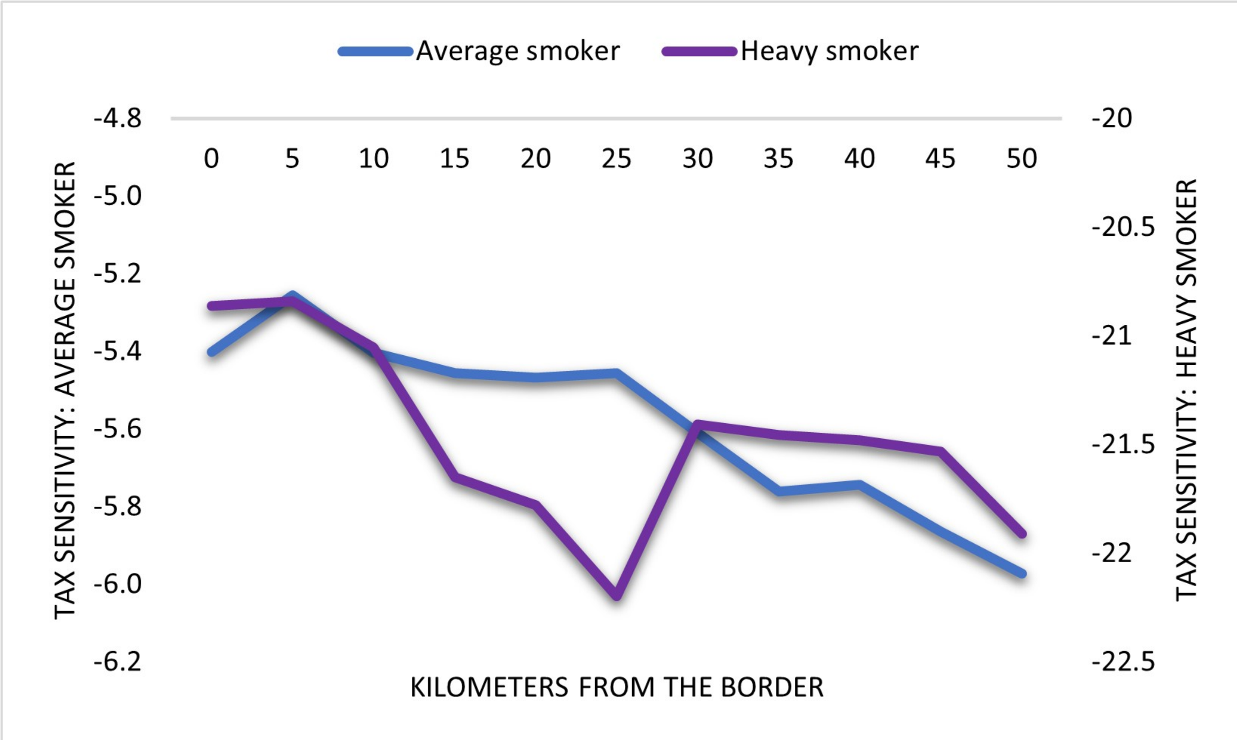


Figure 10: Excluding households residing near the border: Smoking Intensity by Heavy and Average Smokers

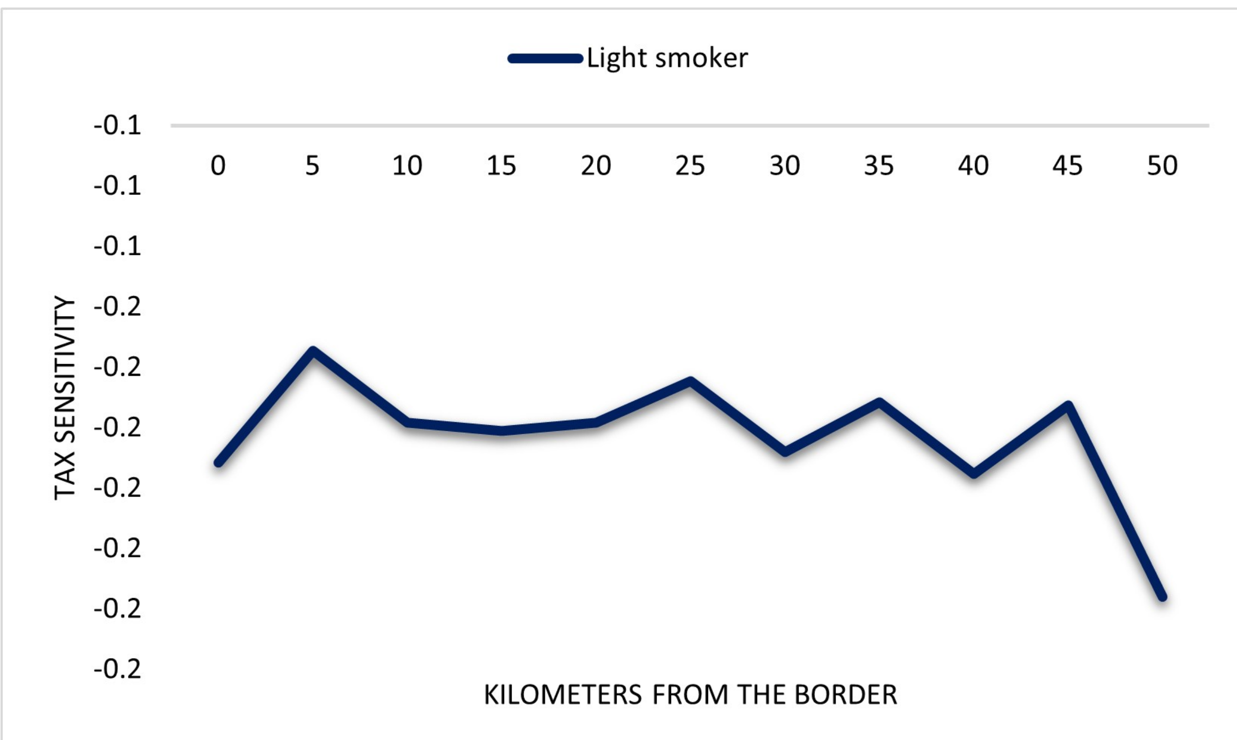


Figure 11: Excluding households residing near the border: Smoking Intensity by Light Smokers

1.7 Conclusion

Using NielsenIQ Consumer Panel data for the years from 2004 to 2019, this study explores whether ignoring tax avoidance opportunities will bias the estimate of tax elasticity. We find that border effects create a bias in the estimate of tax elasticity, which is present for all demographic groups. The bias is particularly large for border residents, since the cost of traveling to the nearest lower-tax state to purchase taxable goods at a lower price increases with the distance to the state border. This implies that ignoring possible tax avoidance actions of consumers, in particular cross border purchasing, results in a biased estimate of tax sensitivity and decreases the applicability of the obtained results. The fact that residing near a lower tax state border decreases the impact of excise tax policy interventions should be considered by policy makers.

Moreover, we analyze how the consumer response to a cigarette tax increase varies between households with different demographic compositions. We observe higher tax elasticity for the low income group. Higher tax sensitivity estimated for unemployed consumers and consumers without college degree can be potentially explained by the fact that, on average, these demographic groups have lower income. Furthermore, we identify that estimated tax sensitivity is statistically significant for heavy smokers and increases with smoking intensity, which can be beneficial from the perspective of potential public health implications, unlike Lee (2008) and Cotti et al. (2018), who show that heavy smokers do not respond to excise tax policy measures.

2 Spatial Heterogeneity in Tax Sensitivity: Evidence from Cross-State Cigarette Purchases

This study was published as CERGE-EI Working Paper Series No 803.

Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

2.1 Introduction

In the United States, cigarette excise taxes represent a significant source of government revenue and serve as a policy instrument with direct public health implications. Previous studies show that cigarette tax hikes reduce tobacco consumption, thereby contributing to the improvement of public health outcomes. However, the effectiveness of these tax measures can be considerably undermined by various forms of tax avoidance, including cross-border purchasing in the nearest lower-tax states or Indian reservations, smuggling, and Internet purchasing.

In the context of the United States, where we can track the variability of excise taxes across states, geographic proximity to the lower-tax states creates opportunities for tax arbitrage. Consumers residing near state borders may mitigate the impact of local tax hikes by purchasing cigarettes in neighboring states where the tax burden is lower. Moreover, because of profit motives, shops close to borders may adjust prices to smooth the unfavorable tax difference to a certain extent. Ignoring these "border effects" leads to a biased estimate of the tax elasticity of consumption.

Although several studies investigate cross-border purchasing from various perspectives, limited attention has been paid to analytically quantifying and modeling the spatial heterogeneity in tax responsiveness across consumers. Most existing studies focus on estimating cigarette demand elasticities or analyzing specific forms of tax avoidance behavior, such as smuggling or purchases in an Indian reservation. While informative, these approaches often fail to capture how the intensity of tax avoidance varies continuously with geographic location and tax differentials.

This paper suggests a novel analytical framework that explicitly models the "border effect"—the attenuation in tax sensitivity induced by spatial proximity to a lower-tax state. We formalize this effect as a linearly decreasing function of distance to the nearest lower-tax border, with a maximal influence at the boundary and a vanishing effect beyond a specified cutoff distance. Our approach leverages comprehensive NielsenIQ Consumer Panel Data covering the period from 2004 to 2019, allowing us to track household-level cigarette purchases over time and the panelist's geographic location.

To quantify the magnitude of the border effect, we estimate a threshold regression model with location and time fixed effects, identifying the critical distance beyond which the border effect is no longer present. As a robustness check, we also run a segmented regression using separate tax sensitivity estimates for a range of distance intervals. We verify that the segmented estimates align with the linear pattern derived from the threshold model. Further, we enhance the "border effect" function by adding a difference between the home-state tax and the closest lower-tax state tax as an additional factor, and compare the estimation results for the two specifications.

In addition to geographic variation, we also analyze how the tax sensitivity of cigarette consumption differs across demographic groups. By linking household-level purchase data with demographic characteristics from the NielsenIQ Consumer Panel, we assess whether the level and functional form of tax sensitivity differ among population subgroups. Our findings indicate that tax sensitivity decreases as income increases: high-income consumers are the least responsive to increases in excise taxes, low-income consumers are the most responsive, and middle-income consumers fall somewhere in between. This trend aligns with economic expectations. All income groups are influenced by the "border effect", but for high-income consumers this effect is only present at distances up to 62 kilometers from the lower-tax state. Furthermore, our analysis based on employment status reveals that both employed and non-employed consumers exhibit a similar pattern in the "border effect". However, non-employed consumers demonstrate significantly higher sensitivity to increases in excise taxes compared to employed consumers.

Our findings contribute to the literature in several ways. First, we provide empirical evidence showing that there is a spatial heterogeneity in the tax sensitivity of cigarette consumption within the United States. Second, we quantify how this sensitivity changes with distance and differences in state tax rates. Third, we examine how the tax sensitivity of cigarette consumption varies across different demographic groups. Together, these findings highlight the need to consider border effects when evaluating and designing excise tax policies in a tax system with heterogeneous tax regimes, such as the US.

2.2 Previous Literature

Previous literature suggests that excise tax hikes reduce tobacco consumption. The sensitivity of cigarette consumption to tax increases is an important question for policymakers for two main reasons: taxes can aid in achieving public health goals by reducing cigarette consumption and generating tax revenue. However, the effectiveness of these tax measures can be considerably undermined by tax avoidance strategies such as cross-border purchases made in nearby lower-tax states or on Indian reservations, smuggling, and Internet purchases.

A number of studies demonstrate that these tax avoidance actions result in imperfect tax pass-through to cigarette prices and may reduce the effectiveness of excise tax policy interventions. For example, Harding et al. (2012) find that excise taxes in the US are less than fully passed on to cigarette prices, primarily due to cross-border purchases. Using NielsenIQ scanner data from 2006–2007, which includes information on consumer locations, they show that opportunities for cross-border purchases lead to significant variations in the tax pass-through rate. Kim & Lee (2020), employing a similar estimation strategy to that used by Harding et al. (2012), find that cigarette taxes are shifted significantly less to consumer prices in cities with large minority (black and Hispanic) populations. They obtain their estimates using NielsenIQ scanner data on cigarette sales for the years 2009–2011 from 1,687 stores across the US. Xu et al. (2014) further show that the tax pass-through rate differs significantly by price minimizing strategy. Consumers who buy premium brands outside Indian reservations face a full tax burden with an additional premium, i.e., the pass-through rate is higher than 100%. In contrast, carton buyers likely to make purchases on Indian reservations pay only 30-83 cents for every 1$ tax increase. These insights point to considerable variation in consumer responses to tax changes depending on purchase behavior. Hanson & Sullivan (2009) analyze Wisconsin's $1 cigarette tax hike using micro-level data on cigarette prices from retail locations in Wisconsin and states that share its border. They conclude that consumers pay the entire amount of the tax as well as a premium of between 8–17 cents per pack of cigarettes. In addition, geo-coded data for locations near the borders of states with different tobacco taxation show that the premium amount is lower for stores located near a lower-tax state border.

A related strand of literature examines how tax avoidance influences cigarette demand elasticities and the design of optimal excise tax policy. For example, Lovenheim (2008) develops and estimates a cigarette demand model that accounts for cross-border purchases using data from the Current Population Survey Tobacco Supplements (TUS-CPS) spanning from September 1992 to February 2002. He finds that demand elasticities concerning the home state price are indistinguishable from zero on average and vary significantly with the distance from which individuals live to a lower-price state border. When opportunities for tax avoidance are removed, the price elasticity becomes negative, although it remains inelastic. Using the same data source from TUS-CPS for February, June, and November 2003, Chiou & Muehlegger (2008) introduces a discrete choice model to examine tax avoidance and state border crossing in the cigarette market. The authors estimate a consumer's tradeoff between distance and price when choosing a location to maximize utility, which allows them to simulate tax avoidance under alternative cigarette excise tax amounts. Expanding on the welfare implications of tax avoidance, DeCicca et al. (2013a) develop an extension of the standard formula for the optimal Pigouvian corrective tax to incorporate the possibility of cross-border purchases. To provide a key parameter to this formula, they estimate a structural endogenous switching regression model of border-crossing and cigarette prices using data from the 2003 and 2006–2007 cycles of the same data source, TUS-CPS. They conclude that, after considering tax avoidance in many states, the optimal tax is smaller than the standard Pigouvian tax. These three studies use the Current Population Survey Tobacco Supplements (TUS-CPS) dataset. The main advantage of this dataset is that consumers directly report the location of their most recent cigarette purchase. However, the authors acknowledge that the estimates could be potentially affected by several sources of reporting bias. First, an individual might not report cross-border purchasing if she perceives cross-border purchases as being quasi-illegal. Second, an individual may report their home state for internet cross-border purchases or the state in which an Indian reservation is located, rather than reporting that they purchased cigarettes on an Indian reservation. Thirdly, even though the last purchase can be considered a random draw from the distribution of each smoker's purchases, consumers might not respond to this question accurately but instead base their responses on their typical purchase location. The authors of these three studies performed a set of robustness checks to ensure the validity of the results obtained.

Alternative empirical strategies have also been employed to detect and quantify tax avoidance. Merriman (2010) used littered cigarette packs in Chicago, and treated cigarette packs without a local tax stamp as direct evidence of tax evasion. He concluded that large tax differentials with neighboring jurisdictions decrease the probability of a local stamp by almost 60 percent, and a one-mile increase in distance to the lower-tax state border increases the likelihood of a pack with a local stamp by about one percent. Using data on county-level sales tax remittances from cigarette retailers in Kansas (2001–2005), Nicholson et al. (2014) estimate the extent of smuggling activity and the revenue effects resulting from increases in cigarette excise tax rates. The authors find substantial sales tax revenue leakage near low-tax borders in response to rising cigarette excise tax rates, particularly affecting tobacco shops. This leakage diminishes as the distance from the border increases. The results indicate significantly negative taxable-sales elasticities—reflecting the responsiveness of a state's taxable cigarette sales to changes in the cigarette excise tax rate—at the border, highlighting meaningful cross-border substitution behavior. Gruber et al. (2003) estimate price elasticity of cigarette demand in Canada controlling for cigarette smuggling. They present two approaches to correcting the bias to estimate elasticities from smuggling. The first is to use legal sales data and exclude regions and years in which the smuggling problem was the worst.

The second is to use micro-data on consumer cigarette expenditures and compare the elasticity estimates with the estimates derived from sales data after applying corrections for smuggling bias. However, Stehr (2005) later discusses in his study that this approach requires researchers to know the years and provinces in which smuggling occurred to verify that the difference in elasticities is due to smuggling and not to other differences between the datasets. Further, they also show that the sensitivity of smoking to price is much larger among lower-income demographic groups. Stehr (2005) investigates how increases in US state cigarette taxes lead to reduced consumption and increased tax avoidance through smuggling, cross-border purchases, and Internet purchases. The author compares cigarette sales data from the publication Tax Burden on Tobacco to cigarette consumption data from the Behavioral Risk Factor Surveillance System (BRFSS). Stehr (2005) reasonably assumes that if tax avoidance exists, then the elasticity of sales concerning price should be larger than that of consumption concerning price. He shows that after subtracting percent changes in consumption, residual percent changes in sales are associated with state cigarette tax changes, implying the existence of tax avoidance. Additionally, the author emphasizes that legal border crossings play a minor role compared to other avoidance methods.

This body of literature shows that opportunities for tax avoidance can significantly influence the effectiveness of excise taxation. Cross-border purchases, smuggling, and other tax-minimizing strategies lead to variations in how taxes are passed on to consumers and complicate the optimal design of a tax system. Therefore, a thorough analysis of cigarette tax policies must take into account tax avoidance behaviors in order to accurately assess both public health outcomes and fiscal efficiency.

2.3 Data

We obtain historical data on U.S. state-level cigarette excise taxes from the Centers for Disease Control and Prevention (CDC). Excise tax rate data is available quarterly. The main advantage of using US data is that excise taxes are not uniform in the US and exhibit significant variability across states. This allows us to take into account not only changes in excise taxes over time but also state-level heterogeneity. Figure 12 displays the variation in cigarette excise taxes across US states as of June 2024.

In this study, we employ NielsenIQ Consumer Panel Data containing information about the purchase history of 40,000-60,000 households (varies by year) who continually provide information to NielsenIQ about their demographic characteristics, products they buy, and the timing and location where they make purchases, in a longitudinal study. Consumer panelists use in-home scanners to record all purchases intended for personal, in-home use. Panelists are geographically dispersed and demographically balanced (James M. Kilts Center for Marketing, NielsenIQ datasets, n.d.). The dataset includes 2,861,278 individual cigarette purchase transactions recorded by 35,672 unique households between 2004 and 2019. We exclude observations from the COVID-19 period to avoid potential biases introduced by pandemic-related restrictions on mobility and retail access, which may bias the estimation of the "border effect" related to cross-border purchases. We transform the transactional data into a panel format by aggregating purchases to the household-quarter level, aligning the frequency with the CDC excise tax dataset. We only include households that have a minimum of two quarterly observations. The resulting panel data comprises 327,534 quarterly observations. The data set covers the demographic characteristics of the households, including income range, size, gender composition, presence and age of children, marital status, type of residence, race, and Hispanic origin.

Additionally, it includes geographic characteristics, such as the panelist's ZIP code and product characteristics, which contain UPC code, description, brand, multi-pack, and size.

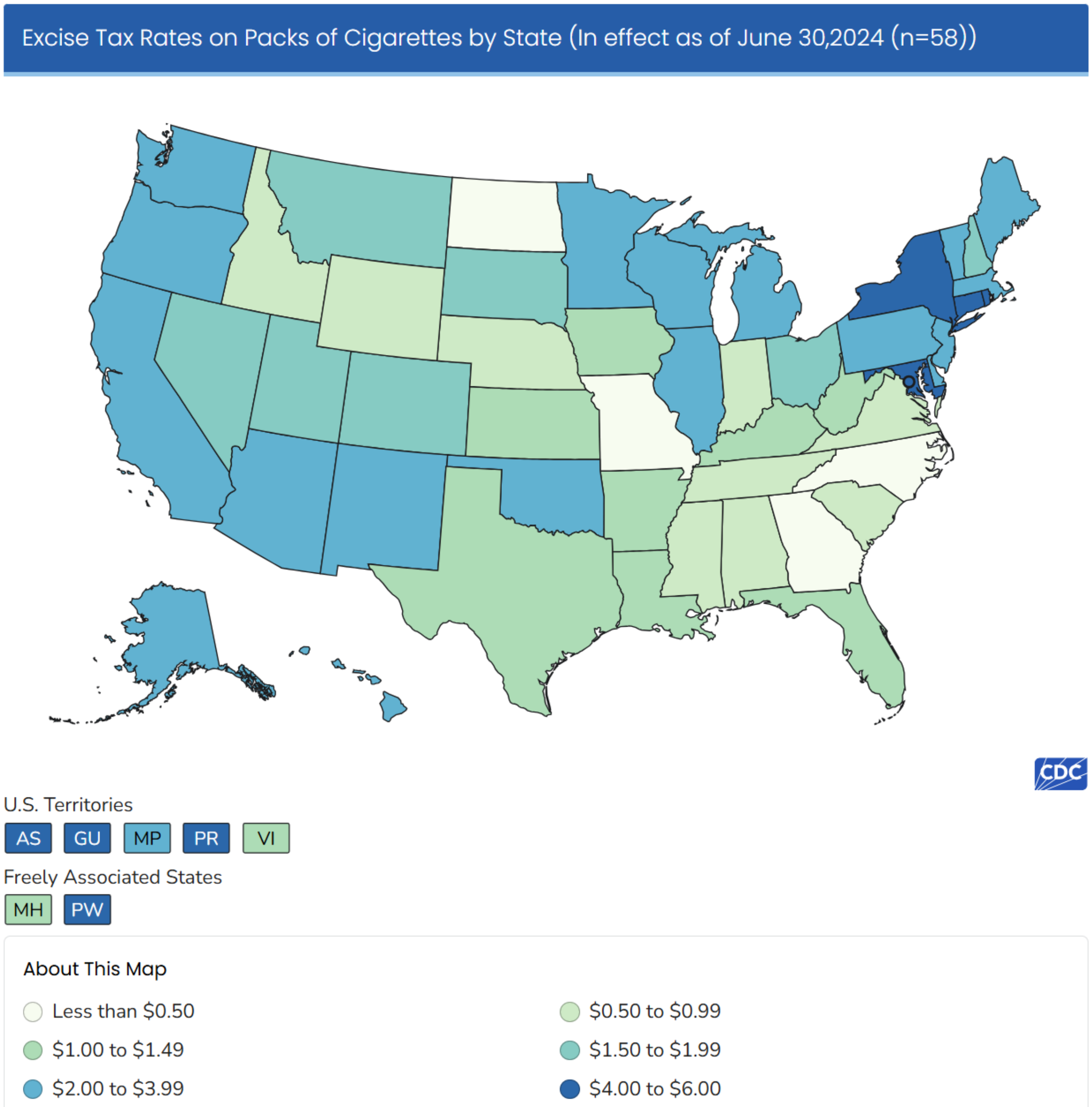


Figure 12: Excise Tax Rates on Packs of Cigarettes by State

Source: Centers for Disease Control and Prevention (CDC), 2024

The geographies of the data cover the entire United States (James M. Kilts Center for Marketing, NielsenIQ datasets, n.d.).

A major strength of the NielsenIQ dataset is the inclusion of panelists' residential addresses, which permits the incorporation of spatial controls in the empirical analysis. Figures 13 and 14 show the geographic distribution of household locations based on proximity to lower-tax state borders. Specifically, they depict panelist ZIP codes located within 50 kilometers and more than 100 kilometers of such borders, respectively.

We measure the distance to the nearest lower-tax state using Census TIGER/Line shape files provided by the United States Census Bureau.

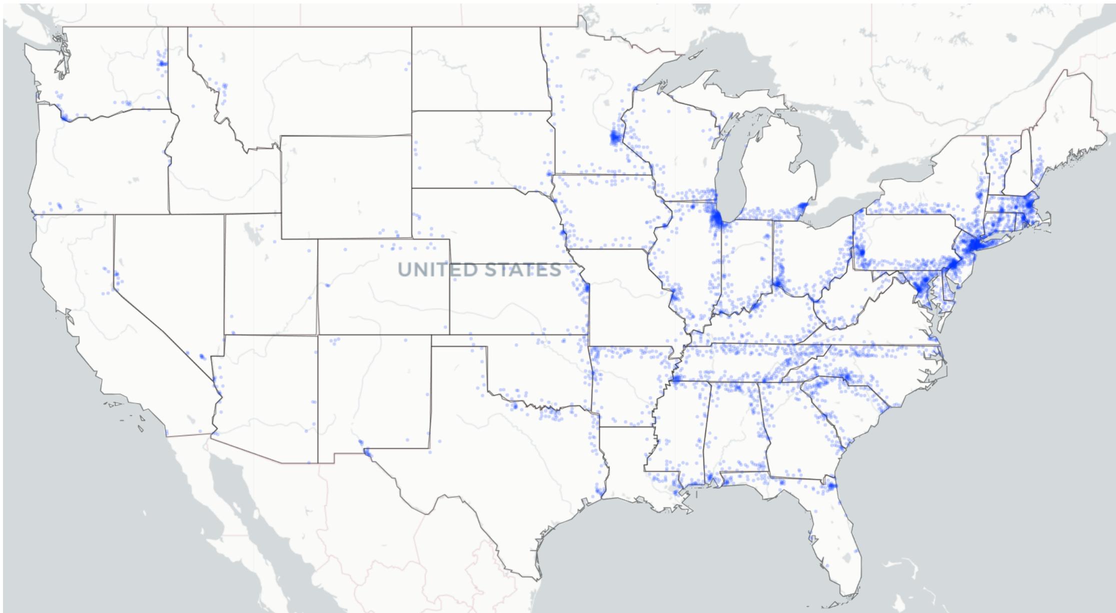


Figure 13: Distribution of Panelist ZIP Codes Located $\leq$ 50 km from a Lower-Tax State Border.

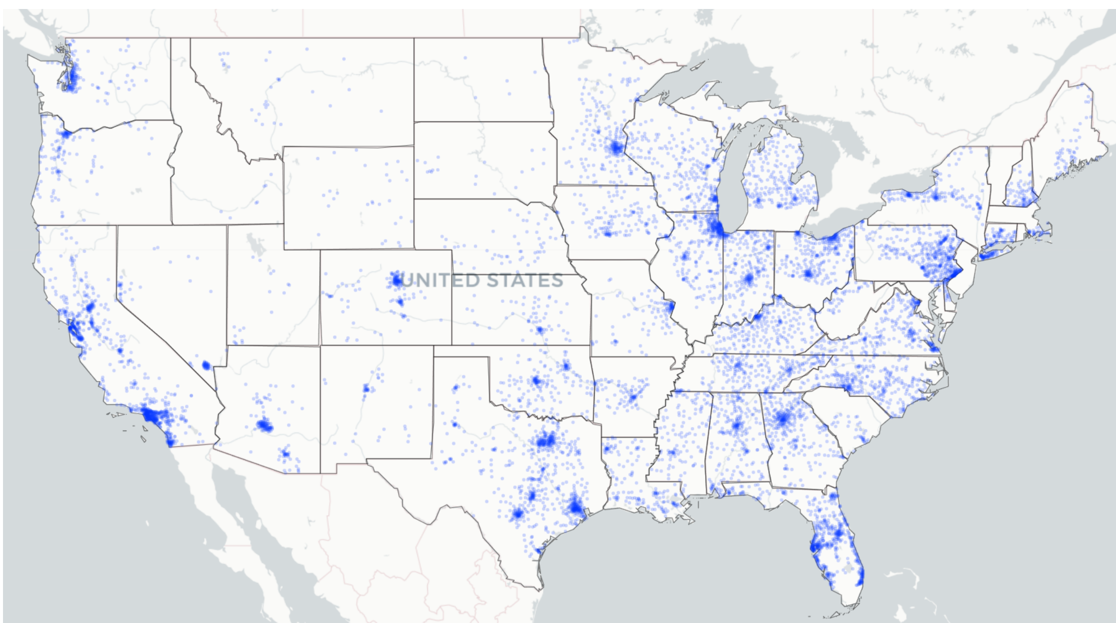


Figure 14: Distribution of Panelist ZIP Codes Located $>$ 100 km from a Lower-Tax State Border.

We estimate the distance between consumers and lower-tax borders as the distance from the household's place of residence provided in the data to the border of the closest lower tax state. The lower tax state does not need to be a border state. We focus on identifying the distance to the nearest state with a lower tax rate, regardless of whether it shares a common border with the state of residence. We identify the coordinates of boundaries for each US state and calculate the distance from each consumer ZIP code to the state boundaries of every US state. We estimate the distance to the lower tax state for each time period and consumer ZIP code as the closest distance to the border of the state with the lower state cigarette tax. Further, we match the tax rate with the corresponding lower tax state. Since we measure the distance to the lower tax state for each time period, we are able to properly capture the state and time level heterogeneity in cigarette taxes and the cost of cross-border purchasing.

Table 7 reports summary statistics for the key variables used in the analysis.

Table 7: Descriptive Statistics of Analysis Variables

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Total packs purchased	48	64	0.05	6	68	2,234
Price per pack [1]	6.3	2.6	0.011	4.9	7.2	99
Distance to the lower tax state (km)	144	118	0.015	50	210	500
Tax value [1]	1.8	1.1	0.065	0.91	2.4	5.6
Tax rate in the lower tax state	0.76	0.62	0.025	0.3	1	4.2
Tax difference	0.73	0.62	0	0.23	1.1	4.1
Smoking rate [2]	48	53	0.7	10	67	669
Time span	2003 Q4 - 2019 Q4					
Number of ZIP codes	13,431					
Number of states	49					
Number of panelists	35,672					
Number of observations	327,534					

Additionally, we outline the rules used to create demographic subgroups, with the distribution of variables across these characteristics shown in Table 8. It is important to note that demographic characteristics may change from year to year for the same household.

[1]The price per pack is adjusted for any applicable discounts and coupons. Cigarette prices and taxes are adjusted to inflation using the Consumer Price Index for tobacco and smoking products in the U.S. City Average provided by the U.S. Bureau of Labor Statistics and retrieved from the website of the Federal Reserve Bank of St. Louis (FRED). We used 2017 as the base year.

[2]The sample includes only those households in the NielsenIQ Homescan data sample that make at least one cigarette purchase. Smoking rate is calculated as the average number of cigarette packs consumed per household per quarter.

Table 8: Rules for Construction of Demographic Groups

Category	Panelists	N
Per capita income [3]		
High: >40,000$	$9,052$	$66,618$
Middle: 15.000$ - 40.000$	$21,315$	$157,598$
Low: $\leq$ 15.000$	$14,980$	$103,318$
Household size		
1: 1 member	$8,668$	$78,358$
2: 2 members	$17,085$	$138,461$
3: 3 members	$8,946$	$53,602$
4: 4 members	$6,375$	$33,769$
5: 5 members	$3,031$	$14,113$
6 plus: $\geq$ 6 members	$1,920$	$9,231$
Head Employment [4]		
$\leq$35 hours	$5,738$	$33,440$
35+ hours	$22,485$	$170,640$
Not employed for pay	$14,584$	$123,454$

2.4 Model Specification

We utilize panel data on consumers, which include continuous quarterly records of cigarette consumption. This study examines the impact of cigarette tax increases on consumption behavior. The baseline econometric model is as follows:

$$Y_{i,t} = \beta \cdot \tau_{s,t} + \gamma_s + \lambda_t + \epsilon_{i,t}, \tag{3}$$

where $Y_{i,t}$ denotes the number of cigarette packs consumed by a consumer i in time t, $\tau_{s,t}$ is a cigarette tax in the state s at time t, and γ_s, λ_t are state and time-fixed effects respectively.

The estimated tax sensitivity β is attenuated due to cross-border purchasing behavior, wherein consumers acquire cigarettes from neighboring states with lower tax rates. We refer to this phenomenon as the "border effect". To formally capture this mechanism, we analytically formulate tax sensitivity β as a function of a consumer's proximity to the nearest lower-tax state, denoted by $Dist_i$. Specifically, we model:

[3] We calculated per capita income by dividing annualized combined household income by household size. The income is adjusted to inflation using the Consumer Price Index for urban consumers in the U.S. City Average provided by the U.S. Bureau of Labor Statistics and retrieved from the website of the Federal Reserve Bank of St. Louis (FRED). We used 2017 as a base year.

[4] The sample includes only those households in the NielsenIQ Homescan data sample that make at least one cigarette purchase. "Head Employment", "Head Age", and "Head Education" refer to male household head if a male household head is present. In the cases in which no male household head is present, these variables refer to the female household head. This is in line with the study by the National Institute of Drug Abuse (April 2021) that finds men tend to use tobacco products at higher rates than women, and therefore men are more likely to be the primary buyers of cigarettes in grocery stores in a two-headed household.

$$\beta = \overline{\tau} + \tau(Dist_i), \tag{4}$$

where $\overline{\tau}$ reflects the true tax sensitivity of consumption, and $\tau(Dist_i)$ captures the magnitude of the border effect as a function of distance to the nearest lower-tax state. We reasonably assume that the "border effect" reaches a maximum at the border with the lower-tax state and then declines linearly with distance from the border of the lower-tax state, vanishing beyond a certain cutoff distance. Therefore, the function $\tau(Dist_i)$ is constructed such that $\tau(Dist_i) = 0$ for distances exceeding a threshold D_{cutoff} and attains its maximum value when $Dist_i = 0$.

Consequently, the estimator obtained from the regression that does not control for the "border effect" demonstrates a positive bias. In expectation, the observed coefficient can be expressed as:

$$\mathbb{E}[\beta] = \overline{\tau} + \mathbb{E}[\tau(Dist_i)], \tag{5}$$

where $\overline{\tau}$ represents the "true" tax sensitivity estimate, and $\mathbb{E}[\tau(Dist_i)]$ can be treated as the average bias introduced by cross-border tax evasion.

To parametrize the spatial heterogeneity in tax sensitivity, we impose a linear functional form on the $\tau(Dist_i)$ and estimate the following regression model:

$$Y_{i,t} = \overline{\tau} \cdot \tau_{s,t} + \mathbf{1}_{Dist_i \leq D_{\text{cutoff}}} \tau_{max}[1 - \frac{Dist_i}{D_{\text{cutoff}}}] \cdot \tau_{s,t} + \gamma_s + \lambda_t + \epsilon_{i,t}, \tag{6}$$

where $\mathbf{1}_{Dist_i \leq D_{\text{cutoff}}}$ is an indicator function which is equal to 1 if consumer i resides within D_{cutoff} kilometers of a lower-tax border and zero otherwise. This specification allows the "border effect" to vary linearly with proximity to the border, reaching its peak τ_{max} at the border and declining to zero at the cutoff threshold. Including this term enables us to recover both the unbiased tax sensitivity $\overline{\tau}$ and the magnitude of the "border effect", thus providing a more accurate estimation of the behavioral response to cigarette excise taxes in the presence of cross-border cigarette purchasing.

Further, we extend equation (6) by interacting the spatial effect with the difference in tax rates between the home state and the nearest lower-tax state. The extended regression is specified as follows:

$$Y_{i,t} = \overline{\tau} \cdot \tau_{s,t} + \mathbf{1}_{Dist_i \leq D_{\text{cutoff}}} \tau_{max}[1 - \frac{Dist_i}{D_{\text{cutoff}}}] \cdot [\tau_{s,t} - \tau_{l,t}] \cdot \tau_{s,t} + \gamma_s + \lambda_t + \epsilon_{i,t}, \tag{7}$$

where $\tau_{s,t}$ and $\tau_{l,t}$ denote the home state tax and the tax rate in the nearest lower tax state, respectively. In this case, the "border effect" is enhanced by the tax difference $[\tau_{s,t} - \tau_{l,t}]$, attaining its maximum at $\tau_{max} \cdot [\tau_{s,t} - \tau_{l,t}]$ when $Dist_i = 0$. This formulation allows the strength of the "border effect" to vary both with spatial proximity and the magnitude of the tax difference. This extension offers a more comprehensive understanding of the impact of cross-border cigarette purchasing on consumers' responses to cigarette taxation.

2.5 Estimation Strategy and Results

We begin our analysis by estimating the baseline equation (6). This equation represents a continuous threshold regression model, where the threshold variable is the distance to the nearest lower-tax state $Dist_i$. The "border effect" variable decreases linearly with distance and eventually reaches zero at a certain cutoff distance, D_{cutoff}. Beyond this cutoff, the variable remains set to zero. As a result, the regression function is continuous, with no discontinuous jumps at the threshold point; however, there is a discontinuity in the slope at the threshold.

If D_{cutoff} were known, equation (6) can be estimated by the ordinary least squares method, and the estimates obtained will be consistent and asymptotically normal. In our problem setting, D_{cutoff} is unknown and needs to be estimated. Using the results of Hansen (2017) and Chan (1998), this model can be estimated by the method of conditional least squares (CLS). The regression parameters $(\overline{\tau}, \tau_{max}, \gamma_s, \lambda_t)$ are estimated by minimizing the sum of squared errors function $S_n(D_{\text{cutoff}})$ for a range of possible values of D_{cutoff}. CLS estimator of $\widehat{D}_{\text{cutoff}}$ is the value that minimizes $S_n(D_{\text{cutoff}})$:

$$\widehat{D}_{\text{cutoff}} = \min_{D_{\text{cutoff}} \in \Gamma} S_n(D_{\text{cutoff}}), \tag{8}$$

where Γ is a bounded set for possible values of the threshold parameter D_{cutoff}. Given a large number of unique distance values in the dataset, we approximated Γ by a grid of size N. More specifically, $\Gamma_N = \left\{D_{cutoff,(1)}, ..., D_{cutoff,(N)}\right\}$, which requires N functions evaluations. For our purpose, we tested 250 possible values of D_{cutoff}, spanning a grid ranging from 50 to 300 kilometers in 1-kilometer increments.

Chan (1998) shows that under suitable regularity conditions, the CLS estimator of the parameters, including the threshold parameter, is $n^{-\frac{1}{2}}$ consistent and asymptotically normally distributed. We construct approximate confidence intervals for the threshold parameter estimates following the methodology outlined by Hansen (2017). The proposed approach is based on inverting the F-type test statistic for the threshold parameter D_{cutoff}:

$$F_n(D_{\text{cutoff}}) = n \cdot \frac{S_n(D_{\text{cutoff}}) - S_n(\widehat{D}_{\text{cutoff}})}{S_n(\widehat{D}_{\text{cutoff}})}, \tag{9}$$

Assuming normally distributed errors and given asymptotic normality of the threshold regression estimates, this test statistic has an asymptotic χ^2_1 distribution. Specifically, the $1 - \alpha\%$ confidence interval for the threshold is given by the set $\hat{\Gamma}$ consisting of those values for which the null hypothesis is not rejected at significance level $1 - \alpha$:

$$\hat{\Gamma} = \{D_{\text{cutoff}} : \quad F_n(D_{\text{cutoff}}) \leq c_{1-\alpha}\}, \tag{10}$$

where $c_{1-\alpha}$ represents critical value for the F-type test statistic distribution.

Following Hansen (2017), confidence intervals for regression parameters $(\overline{\tau}, \tau_{\text{max}}, \gamma_s, \lambda_t)$ are obtained by computing confidence intervals for each $D_{\text{cutoff}} \in \hat{\Gamma}$ and taking their union.

This approach is particularly convenient for parameters obtained using the grid-search method, since confidence intervals can be directly obtained from the computation results of the least-square minimization.

We follow Hansen (2017) and calculate heteroskedasticity-robust asymptotic 95%-level confidence intervals for our threshold parameter D_{cutoff}. A graphical method to find the region $\hat{\Gamma}$ is to plot the test statistic $F_n(D_{\text{cutoff}})$ against D_{cutoff} values and draw a flat line at 95%-level critical value $c_{1-\alpha}$. Figure 15 illustrates confidence interval construction for the threshold parameter in the baseline model.

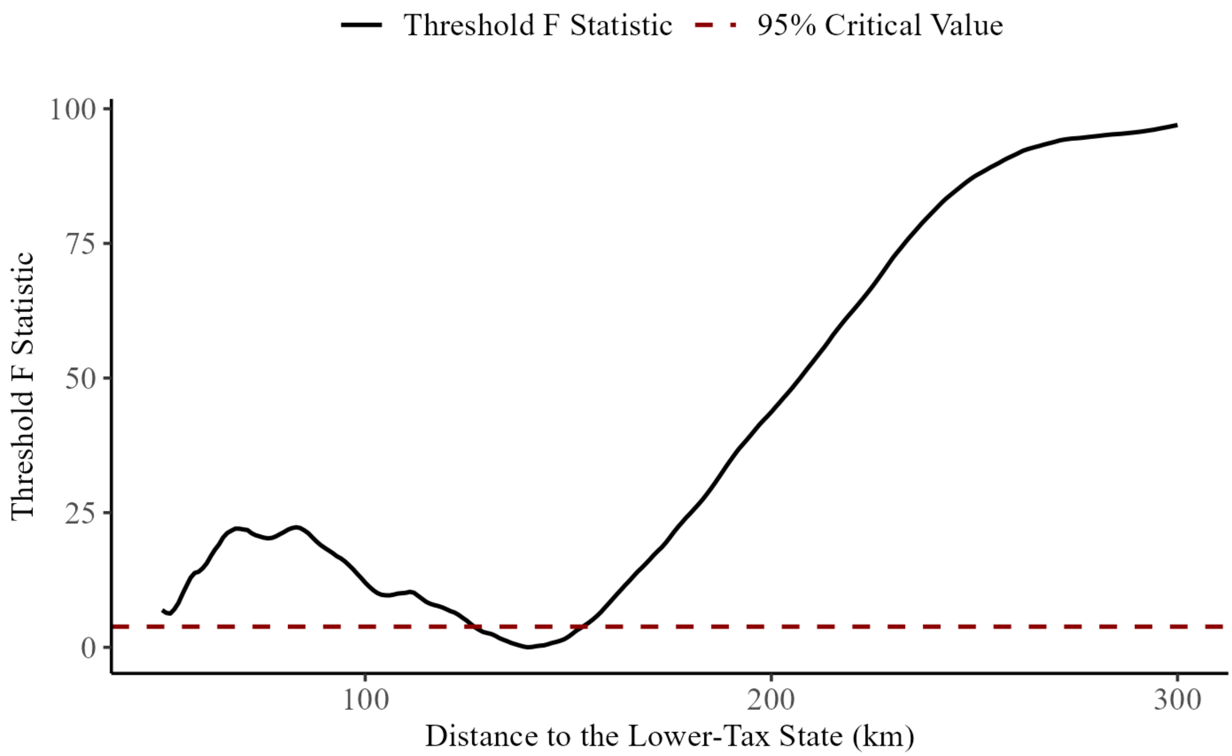


Figure 15: Baseline Model: Confidence Interval Construction for Threshold Parameter

Table 9 demonstrates estimation results of the equation (6), along with heteroskedasticity-robust asymptotic 95%-level confidence regions.

Table 9: Baseline Model: Border Effect and Tax Sensitivity Estimates

Variable	Estimate	95% CI	
D_{cutoff}	140	[127,	153]
τ_{max}	2.9272	[2.4944,	3.3505]
$\overline{\tau}$	-5.2329	[-5.8556,	-4.5990]

The results confirm the presence of spatial heterogeneity in the tax sensitivity. Specifically, consumers residing within 140 kilometers of the nearest lower-tax state are subject to the "border effect" and exhibit attenuated sensitivity to local cigarette excise tax increases. Our model allows us to recover both the true behavioral tax sensitivity $\overline{\tau}$ and the magnitude of the "border effect" bias τ_{max}.

We further proceed with our analysis by estimating the enhanced equation (7). We follow the same estimation procedure as described for the baseline equation (6). Figure 16 illustrates confidence interval construction for the threshold parameter in the extended model.

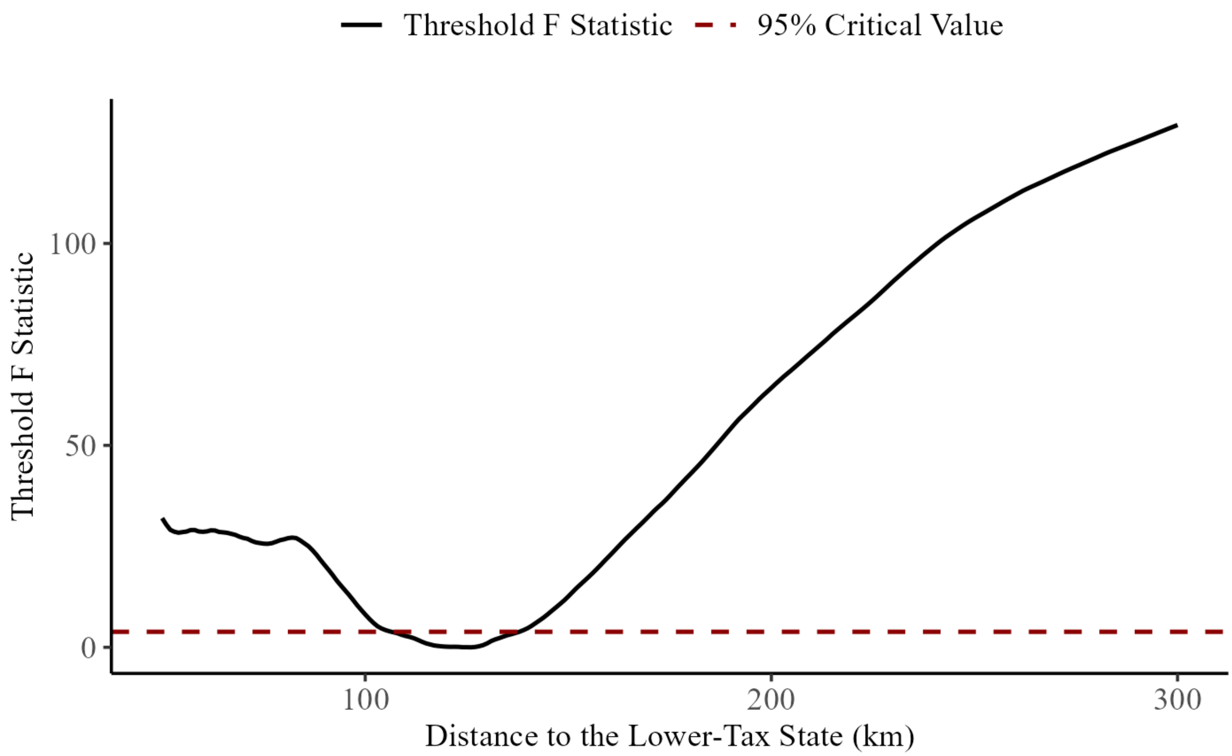


Figure 16: Extended Model: Confidence Interval Construction for Threshold Parameter

Table 10 demonstrates estimation results for the extended model, along with heteroskedasticity-robust asymptotic 95%-level confidence regions.

Table 10: Extended Model: Border Effect and Tax Sensitivity Estimates

Variable	Estimate	95% CI	
D_{cutoff}	126	[107,	137]
τ_{max}	1.7686	[1.5733,	1.9593]
$\overline{\tau}$	−5.6186	[−6.2403,	−4.9202]

The estimated threshold decreases slightly to 126 km, while the border effect parameter τ_{max} remains statistically significant. The enhanced model considers both geographic proximity and tax differentials to model the "border effect".

Our findings suggest that cross-border purchasing opportunities and differences in state taxes contribute to spatial heterogeneity in consumer sensitivity to cigarette taxes. This variation is an important factor to consider when developing an effective tobacco tax policy. Our model offers a more detailed understanding of how consumers respond to cigarette taxation in a system with heterogeneous tax regimes.

2.6 Robustness Analysis

To assess the robustness of our findings, we proceed with the robustness analysis of our models. Specifically, to confirm our assumptions, we examine how the responsiveness to cigarette excise taxes evolves with distance to the nearest lower-tax state without imposing a linear parametric structure as in equations (6) and (7) and then compare our results with the threshold regression estimates from Section 2.5.

We begin by estimating a segmented regression model corresponding to the baseline equation (6) in which the tax sensitivity is allowed to vary across equally sized discrete distance intervals. The specification is as follows:

$$Y_{i,t} = \sum_{g=1}^{G} \tilde{\tau}_g \cdot \mathbf{1}_{\left(Dist_i \geq D_{(g-1)}\right) \& \left(Dist_i < D_{(g)}\right)} \cdot \tau_{s,t} + \tilde{\gamma}_s + \tilde{\lambda}_t + \tilde{\epsilon}_{i,t}, \tag{11}$$

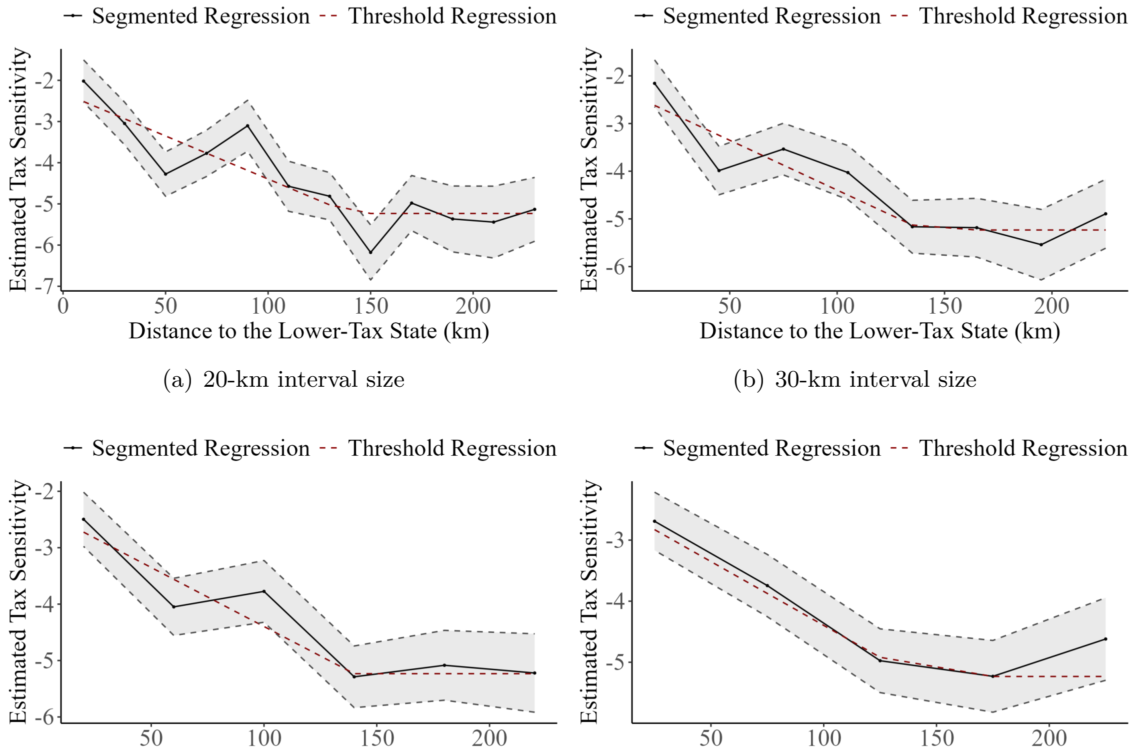
where $D_{(g-1)}$ and $D_{(g)}$ define cutoff values for the distance interval g, $D_{(0)}$ and $D_{(G)}$ correspond to the minimum and maximum values of the distance variable in the dataset.

Figures 17 display G estimates of tax sensitivity $\tilde{\tau}_g$ along with 95%-level confidence bands from equation (11). We plot the tax sensitivity estimates from the segmented regression against each midpoint of the distance interval. The red-dotted line overlays the fitted tax sensitivity function $\overline{\tau} + \mathbf{1}_{Dist_i \leq D_{\text{cutoff}}} \tau_{max}[1 - \frac{Dist_i}{D_{\text{cutoff}}}]$ from the threshold model (6). We observe that small interval sizes may produce volatile estimates, and too large intervals may absorb meaningful variation.

We proceed by estimating a segmented regression model corresponding to the extended equation (7). In this case, the tax sensitivity estimate varies across two dimensions: (1) distance to the nearest lower-tax state and (2) the difference in tax rates between the home state and the nearest lower-tax state. In order to evaluate a combined impact of these two factors on tax sensitivity, we estimated the segmented regression, in which the combined "border effect" defined as $\mathbf{1}_{Dist_i \leq \widehat{D}_{\text{cutoff}}}[1 - \frac{Dist_i}{\widehat{D}_{\text{cutoff}}}] \cdot [\tau_{s,t} - \tau_{l,t}]$ is allowed to vary across equally sized discrete intervals.

$$\begin{aligned} Y_{i,t} = & \bar{\tau}_{max,0} \cdot \mathbf{1}_{\text{BorderEffect}_i = 0} \cdot \tau_{s,t} + \\ & \sum_{g=1}^{G} \bar{\tau}_{max,g} \cdot \mathbf{1}_{\left(\text{BorderEffect}_i > B_{(g-1)}\right) \& \left(\text{BorderEffect}_i \leq B_{(g)}\right)} \cdot \tau_{s,t} + \bar{\gamma}_s + \bar{\lambda}_t + \bar{\epsilon}_{i,t}, \end{aligned} \tag{12}$$

where BorderEffect_i is defined as $\mathbf{1}_{Dist_i \leq \widehat{D}_{\text{cutoff}}}[1 - \frac{Dist_i}{\widehat{D}_{\text{cutoff}}}] \cdot [\tau_{s,t} - \tau_{l,t}]$ using estimate $\widehat{D}_{\text{cutoff}}$ from the threshold regression, $B_{(g-1)}$ and $B_{(g)}$ define cutoff values for the BorderEffect_i interval g, and $B_{(0)}$ and $B_{(G)}$ correspond to the minimum and maximum values of the "border effect" variable in the dataset.



(a) 20-km interval size

(b) 30-km interval size

(c) 40-km interval size

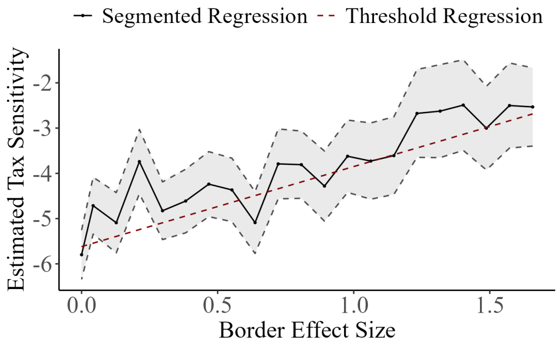
(d) 50-km interval size

Figure 17: Segmented regression estimates of tax sensitivity $\tilde{\tau}_g$ along with 95%-level confidence bands from the baseline model for 20, 30, 40 and 50-kilometer interval sizes. Red-dotted line overlays the fitted tax sensitivity function $\overline{\tau} + \mathbf{1}_{Dist_i \leq D_{\text{cutoff}}} \tau_{max}[1 - \frac{Dist_i}{D_{\text{cutoff}}}]$ from the threshold model.

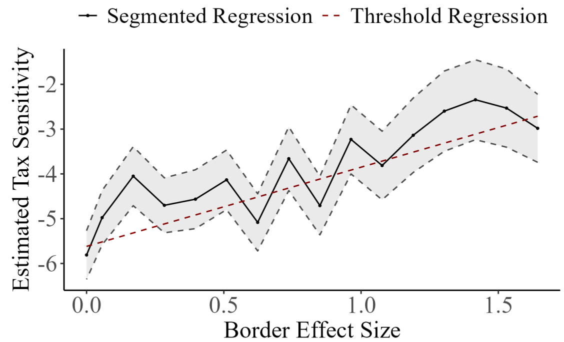
Figures 18 display G estimates of tax sensitivity $\bar{\tau}_{max,g}$ along with 95%-level confidence bands from equation (12) for $G = 5$, $G = 10$, $G = 15$, and $G = 20$. We plot the tax sensitivity estimates from the segmented regression against each midpoint of the "border effect" interval. The red-dotted line overlays the fitted tax sensitivity function $\overline{\tau} + \mathbf{1}_{Dist_i \leq D_{\text{cutoff}}} \tau_{max}[1 - \frac{Dist_i}{D_{\text{cutoff}}}] \cdot [\tau_{s,t} - \tau_{l,t}]$ from the threshold model (7).

The results provide strong evidence of a positive and increasing relationship between tax sensitivity and the size of the composite "border effect" that is jointly dependent on both distance to the nearest lower tax state and tax differentials.

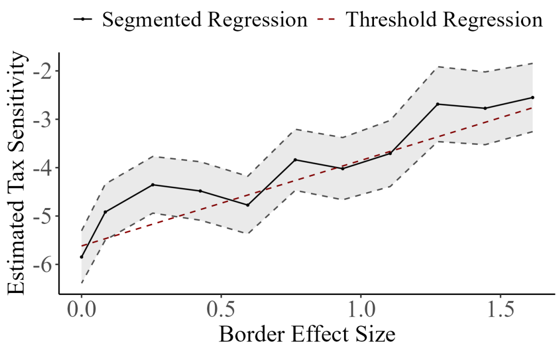
We can conclude that the robustness analysis results validate our parametric specification: the observed tax sensitivities from the segmented regressions follow the analytical tax sensitivity imposed in the threshold models.



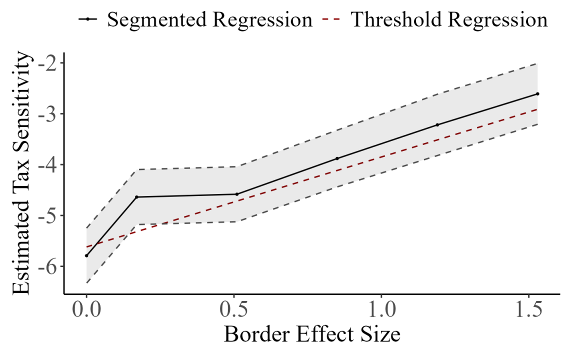
(a) Number of intervals used is 20



(b) Number of intervals used is 15



(c) Number of intervals used is 10



(d) Number of intervals used is 5

Figure 18: Segmented regression estimates of the tax sensitivity estimate $\bar{\tau}_{max,g}$ along with 95%-level confidence bands from the extended model for $G = 5$, $G = 10$, $G = 15$, and $G = 20$. The red-dotted line overlays the fitted tax sensitivity function $\overline{\tau}+\mathbf{1}_{Dist_i \leq D_{\text{cutoff}}}\tau_{max}[1-\frac{Dist_i}{D_{\text{cutoff}}}]\cdot[\tau_{s,t}-\tau_{l,t}]$ from the threshold model.

2.7 Demographic Heterogeneity

In this section, we analyze demographic heterogeneity in tax sensitivities across demographic groups. We focus on two key dimensions: annual per capita income and employment status. Both dimensions are directly related to the financial situation of consumers and, therefore, likely to shape their behavioral responses to taxation.

2.7.1 Income Groups

We begin by analyzing tax sensitivities across households with different levels of annual per capita income. Per capita income is calculated by dividing the combined annual household income by the household size. Income values are adjusted for inflation, using 2017 as the base year. Households are classified into three income groups, as shown in Table 11.

Table 11: Distribution of Households by Income Group

Category	Panelists	N
Per capita income[1]		
High: >40,000$	$9,052$	$66,618$
Middle: 15.000$ - 40.000$	$21,315$	$157,598$
Low: $\leq$ 15.000$	$14,980$	$103,318$

For each income group, we estimate the extended model defined in equation (7). Table 12 demonstrates estimation results, along with heteroskedasticity-robust asymptotic 95%-level confidence intervals.

Table 12: Extended Model: Border Effect and Tax Sensitivity Estimates by Income Groups

	High Income		Middle Income		Low Income	
Variable	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
D_{cutoff}	62	$[52, 72]$	145	$[111, 190]$	129	$[113, 156]$
τ_{max}	2.5446	$[2.0077, 3.0633]$	1.6905	$[1.3952, 1.9698]$	1.8463	$[1.5047, 2.1631]$
$\overline{\tau}$	-5.1909	$[-6.3830, -3.9293]$	-5.5759	$[-6.7187, -4.4845]$	-6.2732	$[-7.5094, -5.1200]$

For each observation in the dataset, we calculate the predicted tax sensitivity of cigarette consumption. Figure 19 shows the average predicted tax sensitivity along the distance to the nearest lower-tax state.

The results indicate that tax sensitivity decreases with income levels. High-income consumers show the lowest sensitivity to excise tax increases, while low-income consumers exhibit the highest sensitivity. The tax sensitivity of middle-income consumers falls in between these two groups, which aligns with economic expectations.

All income groups are influenced by the “border effect”. However, for high-income consumers, this effect is present only for distances up to 62 kilometers from the lower-tax state. For middle- and low-income consumers, the shape of the “border effect” functions is similar; however, low-income consumers demonstrate significantly higher sensitivity to excise tax increases compared to middle-income consumers.

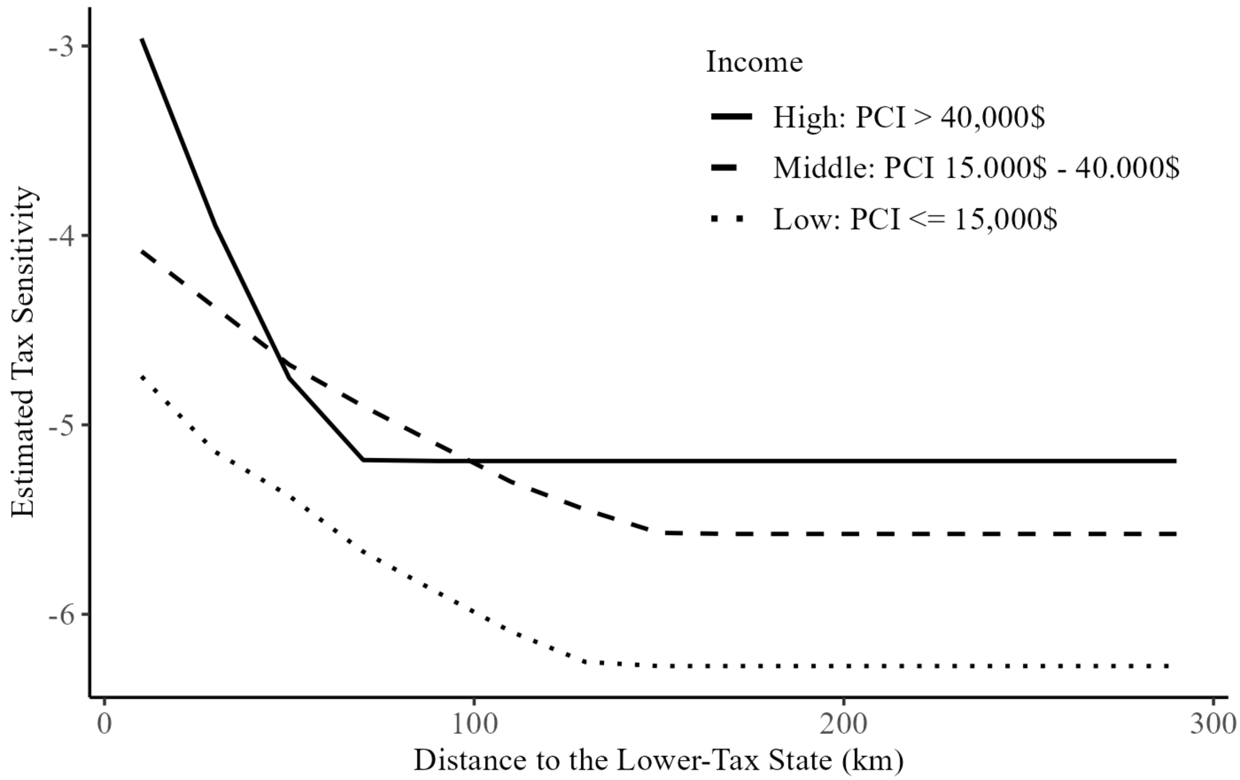


Figure 19: Predicted Tax Sensitivity by Income Group.

2.7.2 Employment Status

Next, we examine demographic heterogeneity by employment status. Households are divided into two categories based on the employment status of the male household head: "Employed" and "Not Employed for Pay". If there is no male head in the household, we will consider the female head of the household. The distribution of households across these categories is summarized in Table 13.

Table 13: Distribution of Households by Employment Status

Category	Panelists	N
Head Employment[2]		
Employed	25,813	204,080
Not employed for pay	14,584	123,454

As before, we estimate the extended model outlined in equation (7) for each demographic subgroup. Table 14 shows estimation results, along with heteroskedasticity-robust asymptotic 95%-level confidence intervals.

Table 14: Extended Model: Border Effect and Tax Sensitivity Estimates by Employment Status

	Employed		Not Employed for Pay	
Variable	Estimate	95% CI	Estimate	95% CI
D_{cutoff}	127	$[113, 143]$	116	$[92, 143]$
τ_{max}	1.7669	$[1.5185, 2.0048]$	1.8061	$[1.4744, 2.1160]$
$\overline{\tau}$	-4.9924	$[-5.7799, -4.2228]$	-6.6409	$[-7.7842, -5.5035]$

For every observation in the dataset, we calculate the predicted tax sensitivity of cigarette consumption. Figure 20 illustrates the predicted tax sensitivities by employment status along the distance to the nearest lower-tax state.

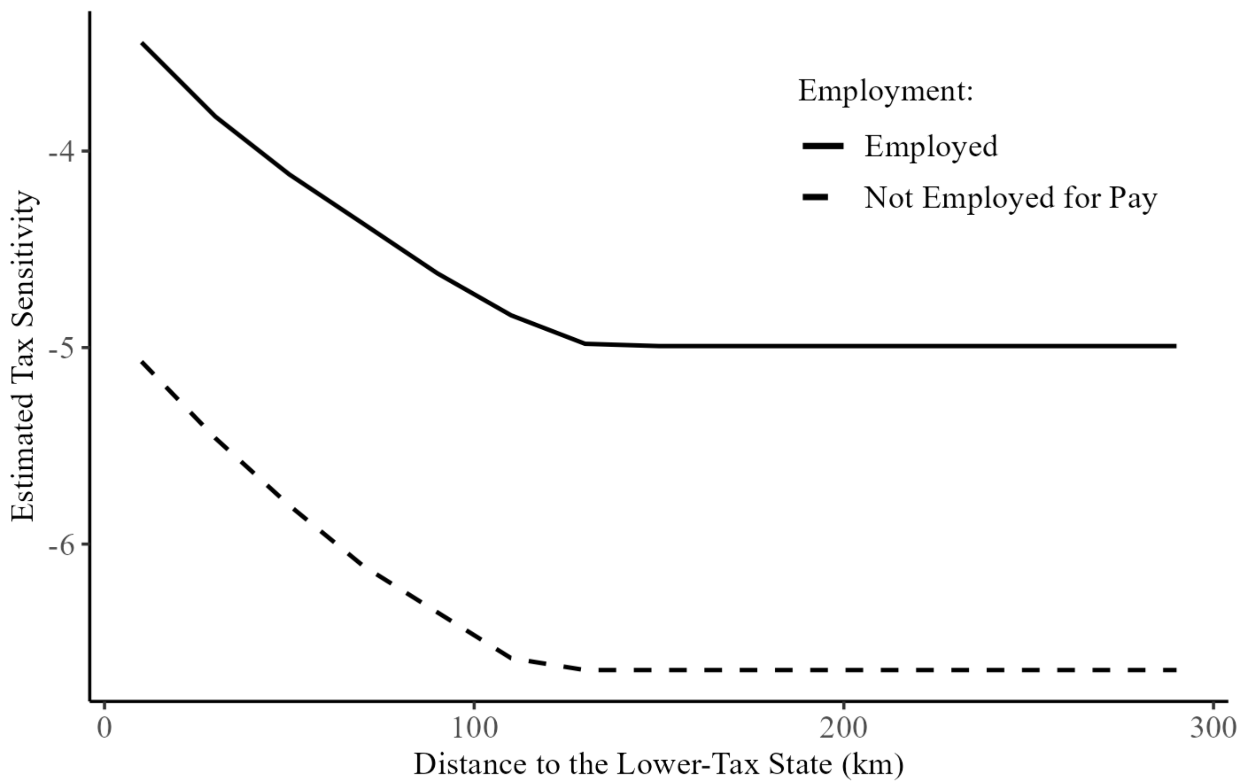


Figure 20: Predicted Tax Sensitivity by Employment Status.

For both employed and non-employed consumers, the shape of the "border effect" function is similar. However, non-employed consumers demonstrate significantly higher sensitivity to excise tax increases compared to employed consumers.

2.8 Conclusion

This study provides empirical evidence of spatial variation in consumer responses to cigarette excise taxes, which is influenced by cross-border purchasing behavior. We developed and estimated a threshold regression model to capture the "border effect", which represents a systematic decrease in tax sensitivity among consumers living near states with lower taxes. Our results indicate that this effect diminishes with distance and disappears beyond a critical threshold estimated to be approximately 126–140 kilometers, depending on the model specification. Furthermore, we found that the strength of the border effect increases with the size of the tax differential between a consumer's home state and the nearest lower-tax state. Our robustness analysis confirms the validity of the parametric assumptions underlying our main model and supports the conclusion that spatial variation in tax sensitivity is both statistically and economically significant.

Additionally, our analysis reveals that this heterogeneity extends across different demographic subgroups. Tax sensitivity decreases with income levels. High-income consumers exhibit the lowest sensitivity to excise tax increases, while low-income consumers show the highest sensitivity. The tax sensitivity of middle-income consumers falls between these two groups, which aligns with economic expectations. All income groups are influenced by the "border effect", but for high-income consumers, this effect is only present at distances up to 62 kilometers from the lower-tax state. While the shape of the "border effect" functions for middle- and low-income consumers is similar, low-income consumers demonstrate significantly higher sensitivity to excise tax increases compared to middle-income consumers. Moreover, our analysis based on employment status indicates that both employed and non-employed consumers display a similar shape in the "border effect" function; however, non-employed consumers show significantly higher sensitivity to excise tax increases than employed consumers.

These findings have important implications for the design of tobacco tax policy within heterogeneous tax regimes such as the United States. Future research could build on these results by exploring interstate cross-border purchasing with Mexico and investigating the additional impact of variability in sales tax.

3 How State Borders Shape the Impact of Cigarette Taxes on Prices

This study was published as CERGE-EI Working Paper Series No 804.

Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

3.1 Introduction

In the United States, excise taxes on cigarettes serve as both a significant source of government revenue and a widely used public health policy tool. Raising cigarette prices through higher excise taxes effectively reduces smoking rates and thereby improves public health outcomes. However, potential unintended consequences of cigarette taxes should also be considered. Darden et al. (2025) demonstrated that, although increases in cigarette taxes are associated with health improvements, higher cigarette taxes may reduce human capital spending among low-income smokers.

Furthermore, the effectiveness of these tax measures can be considerably undermined by various forms of tax avoidance, such as cross-border purchases made in states with lower taxes, buying cigarettes from Indian reservations, smuggling, and making purchases over the Internet. These actions affect how much of the excise tax is reflected in cigarette prices. If the full tax amount is passed on to consumers, they face higher costs and may reduce their cigarette consumption. Conversely, if only a portion of the tax is passed on to consumers, the intended effects of the tax policy measure may not be fully realized.

In the United States, where we can track the variability of state excise taxes across states, geographic proximity to lower-tax jurisdictions creates opportunities for tax avoidance through cross-border purchases. Consumers living near state boundaries can partially offset the impact of local tax increases by buying cigarettes in neighboring states with lower tax rates. Moreover, due to profit motives, shops near borders may adjust their prices to mitigate the unfavorable tax difference to a certain extent. As a result, these "border effects" lead to spatial variation in the pass-through of taxes to prices, which tends to be systematically lower in areas close to the neighboring lower-tax states. This effect diminishes as the distance from the lower-tax state border increases.

This paper develops a novel analytical framework to model the "border effect"—the reduction in the pass-through of cigarette excise taxes to prices that arises from cross-border purchases. To quantify the magnitude of the border effect, we estimate an exponential growth model with UPC, location, and time-fixed effects. The border effect is specified as an exponentially declining function of distance to the nearest lower-tax border, with the strongest influence at the boundary and diminishing impact as distance increases. The analysis draws on comprehensive NielsenIQ Consumer Panel data from 2004 to 2019, which provide detailed information on household-level cigarette purchases over time and the geographic location of panelists.

As a robustness check, we also estimate a segmented regression using separate tax sensitivity estimates for a range of distance intervals. We verify that the segmented estimates are consistent with the pattern derived from the exponential model. Further, we refine the "border effect" specification by incorporating the difference between the home state's tax and the tax in the closest lower-tax state as an additional factor. We then compare the estimation results from both specifications.

Beyond geographic variation, we also examine how the pass-through of cigarette taxes differs across demographic groups. By linking household-level purchase data with demographic characteristics from the NielsenIQ Consumer Panel, we assess whether the level and functional form of tax pass-through differ across population subgroups. The analysis reveals important heterogeneity across households. High-income households face the highest tax pass-through and are largely unaffected by border proximity, while middle-income households are affected by the "border effects" only when the distance from the lower-tax state does not exceed 90 kilometers. Low-income households remain sensitive to "border effects" at greater distances, though their responsiveness declines beyond 200 kilometers. Moreover, consumers who are not engaged in paid work exhibit significantly lower pass-through, suggesting greater scope for tax avoidance through flexible shopping behavior. Therefore, the welfare implications of cigarette taxation vary across demographic groups: high-income households experience a higher tax pass-through and thus bear a larger tax burden, while low-income and non-employed households face a lower effective tax burden due to reduced tax pass-through from cross-border cigarette purchases and other tax avoidance actions.

Our findings contribute to the literature in several important respects. First, we provide empirical evidence of spatial differences in how excise taxes are passed through to cigarette prices in the United States. Second, we quantify how these differences vary based on the distance to lower-tax state borders and differences in state tax rates. Third, we examine how tax pass-through varies across different demographic groups. Together, these results highlight the importance of considering spatial factors in the evaluation and design of excise tax policy, particularly in a tax system with heterogeneous tax regimes such as the US.

3.2 Data

We use historical data on state-level cigarette excise taxes in the United States, obtained from the Centers for Disease Control and Prevention (CDC). The data reports the excise tax per pack in US dollars and is available on a quarterly basis. Excise taxes are not uniform in the US and exhibit significant variability across states. This allows us to take into account both changes in excise taxes over time and also state-level heterogeneity. Figure 21 displays the variation in cigarette excise taxes across US states as of June 2024.

In this study, we employ transaction-level data from NielsenIQ Consumer Panel containing information about the purchase histories of 40,000-60,000 households (numbers vary by year) who continually provide information to NielsenIQ about their demographic characteristics, products they buy, and the timing and location where they make purchases in a longitudinal study. Consumer panelists use in-home scanners to record all purchases intended for personal, in-home use. Panelists are geographically dispersed and demographically balanced (James M. Kilts Center for Marketing, NielsenIQ datasets, n.d.). The dataset includes 2,847,908 individual cigarette purchase transactions recorded by 47,457 unique households between 2004 and 2019. We exclude observations from the COVID-19 period to avoid potential biases introduced by pandemic-related restrictions on mobility and retail access, which may bias the estimation of the "border effect" related to cross-border purchases. The data set includes the demographic characteristics of the households, including income range, size, gender composition, presence and age of children, marital status, type of residence, race, and Hispanic origin.

Additionally, it includes geographic characteristics, such as the panelist's ZIP code and product characteristics, which contain UPC code, description, brand, multi-pack, and size. The geographies of the data cover the entire US (James M. Kilts Center for Marketing, NielsenIQ datasets, n.d.).

A major strength of the NielsenIQ dataset is the inclusion of panelists' residential addresses, which permits incorporation of spatial controls in empirical analyses. Figures 22 and 23 show the geographic distribution of household locations based on proximity to the borders of states with lower cigarette taxes. Specifically, they depict panelist ZIP codes located within 50 kilometers and more than 100 kilometers of such borders.

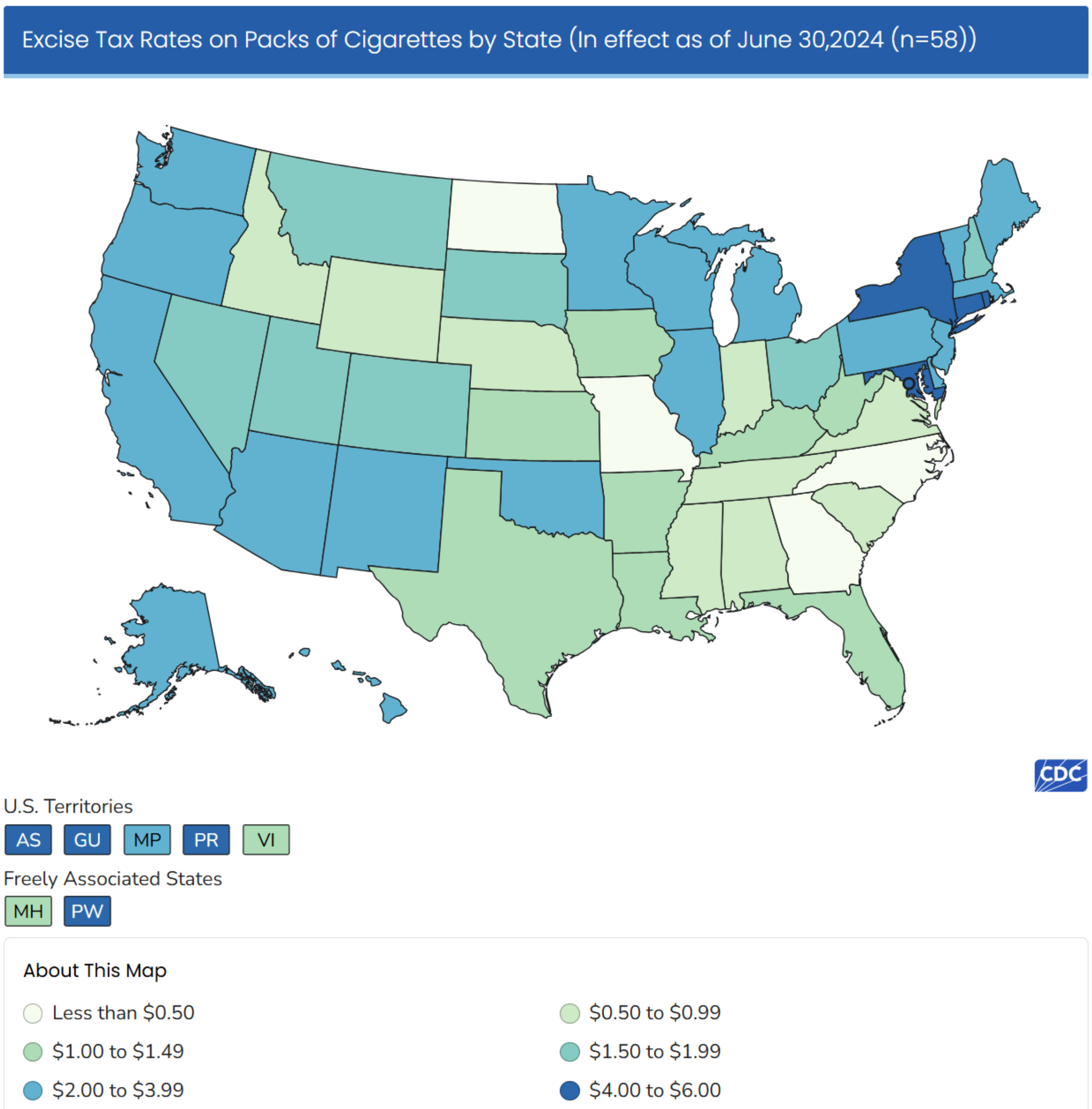


Figure 21: Excise Tax Rates on Packs of Cigarettes by State

Source: Centers for Disease Control and Prevention (CDC), 2024

We measure the distance to the nearest lower-tax state using Census TIGER/Line shape files provided by United States Census Bureau. We estimate the distance between consumers and lower-tax borders as the distance from the household's place of residence provided in the data to the border of the closest lower tax state. The lower tax state may not be a bordering state.

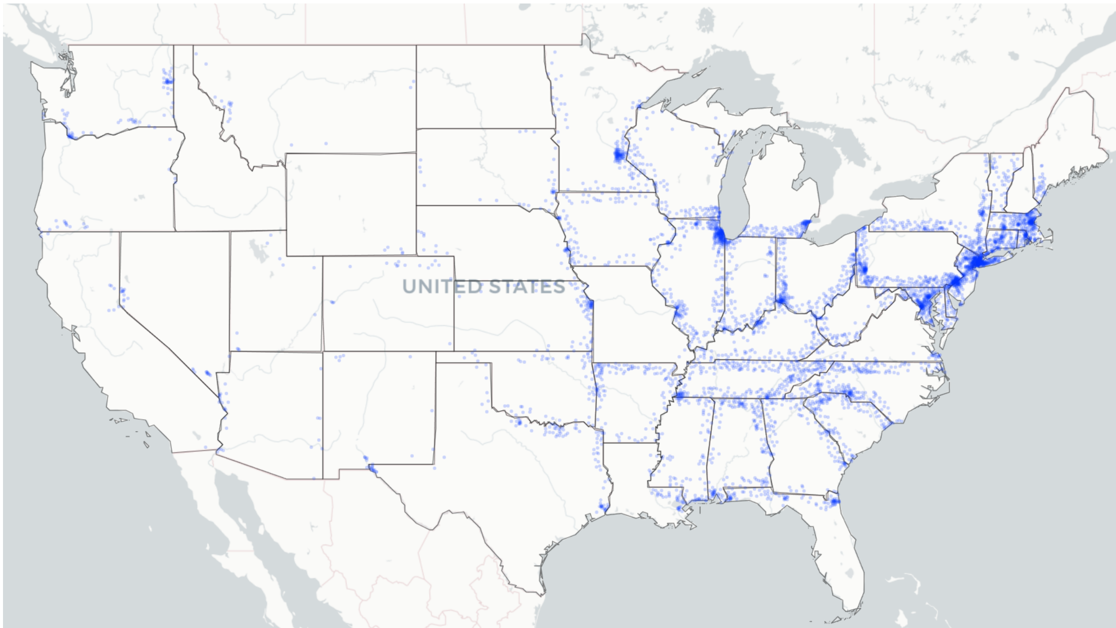


Figure 22: Distribution of Panelist ZIP Codes Located ≤ 50 km from a Lower-Tax State Border.

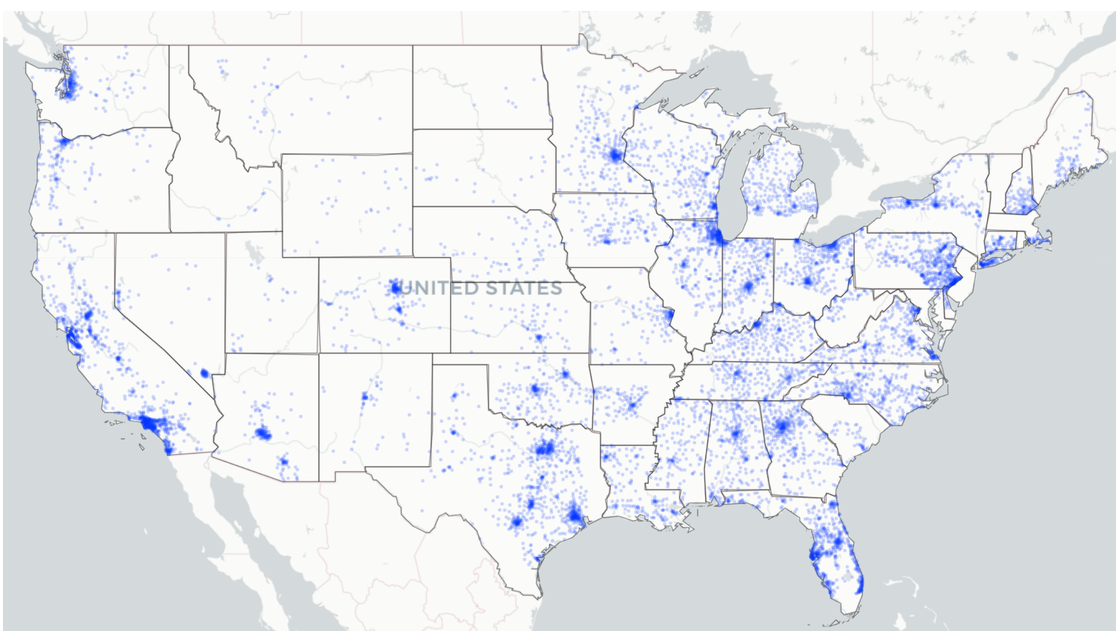


Figure 23: Distribution of Panelist ZIP Codes Located > 100 km from a Lower-Tax State Border.

We identify the coordinates of boundaries for each US state and calculate the distance from each consumer ZIP code to the state boundaries of every US state. We estimate the distance to the lower tax state for each time period and consumer ZIP code as the closest distance to the border of the state with the lower cigarette tax. Further, we match the tax rate in the corresponding lower tax state. Because we measure the distance to the lower tax state for each time period, we are able to correctly capture the state and time level heterogeneity in cigarette taxes and the cost of cross-border purchasing.

As a target variable, we use the final price paid per cigarette pack adjusted for possible discounts and coupons. For cigarette multi-packs, we normalize the price paid to the number of units in a multi-pack. We adjust prices and taxes for inflation using the Consumer Price Index for tobacco and smoking products in the US City Average provided by the US Bureau of Labor Statistics and retrieved from the website of the Federal Reserve Bank of St. Louis (FRED). We use 2017 as a base year.

Table 15 reports summary statistics for the key variables used in the analysis.

Table 15: Descriptive Statistics of Analysis Variables

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Final price per pack	6.4	2.1	1	5	7.4	35
Total packs purchased per trip	5.5	7.3	1	1	10	430
Distance to the lower tax state (km)	141	117	0.015	49	207	500
Tax value	1.8	1.1	0.065	0.94	2.4	5.6
Tax rate in the lower tax state	0.77	0.63	0.025	0.3	1	4.2
Tax difference	0.73	0.61	0.0046	0.24	1.2	4.1
Time span	2003 Q4 - 2019 Q4					
Number of ZIP codes	14,831					
Number of states	49					
Number of panelists	47,457					
Number of observations	2,847,908					
Number of UPC	6,418					

Finally, we define the rules used to construct demographic subgroups, with the distribution of variables across these characteristics reported in Table 16.

Table 16: Rules for Construction of Demographic Groups

Category	Panelists	N
Per capita income [3]		
High: Per capita income>40,000$	$11,656$	$503,466$
Middle: Per capita income 15.000$ - 40.000$	$26,827$	$1,375,787$
Low: Per capita income $\leq$ 15.000$	$18,388$	$968,655$
Household size		
1: 1 member	$10,794$	$689,838$
2: 2 members	$21,436$	$1,251,405$
3: 3 members	$10,950$	$465,653$
4: 4 members	$8,130$	$263,690$
5: 5 members	$3,874$	$106,241$
6 plus: $\geq$ 6 members	$2,434$	$71,081$
Head Employment [2]		
$\leq$35 hours	$6,899$	$316,162$
35+ hours	$29,551$	$1,410,850$
Not employed	$18,017$	$1,120,896$

3.3 Model Specification

To assess the impact of excise taxes on cigarette prices, we estimate the following baseline econometric model using transaction-level data on cigarette purchases:

$$P_{u,i,s,t} = \beta \cdot \tau_{s,t} + \gamma_s + \lambda_t + \alpha_u + \epsilon_{u,i,s,t}, \tag{13}$$

where $P_{u,i,s,t}$ denotes the price of a cigarette pack purchased for UPC u by consumer i in state s at time t; $\tau_{s,t}$ represents the cigarette tax in state s at time t; and γ_s, λ_t, and α_u denote state, time, and UPC fixed effects, respectively.

It is important to note that the use of UPC fixed effects enables us to control for potential brand-substitution effects following tax increases. UPC fixed effects account for product-specific characteristics that may confound the relationship between cigarette taxes and cigarette prices. Since different cigarette UPCs can vary in nicotine content, flavor, and other attributes, price comparisons across UPCs may reflect differences in product quality rather than the impact of cigarette taxes. For example, in Cotti et al. (2022) study, they also estimate tax pass-through of e-cigarette

[3] We calculate per capita income by dividing annualized combined household income by household size. The income is adjusted to inflation using the Consumer Price Index for urban consumers in the US City Average provided by the US Bureau of Labor Statistics and retrieved from the website of the Federal Reserve Bank of St. Louis (FRED). We use 2017 as a base year.

[2] The sample includes only those households in the NielsenIQ Homescan data sample that make at least one cigarette purchase. "Head Employment", "Head Age", and "Head Education" refer to male household head if a male household head is present. In the cases in which no male household head is present, these variables refer to the female household head. This is in line with the study by the National Institute of Drug Abuse (April 2021) that finds men tend to use tobacco products at higher rates than women, and therefore men are more likely to be the primary buyers of cigarettes in a household.

prices, incorporating UPC-by-state (locality) and UPC-by-period fixed effects in the regression model. In addition, state fixed effects account for unobserved interstate price differentials, while time fixed effects capture common trends in cigarette price dynamics over time.

We assume that the estimated pass-through of taxes to prices β is attenuated due to cross-border purchasing, wherein consumers purchase cigarettes from neighboring states with lower tax rates. We refer to this phenomenon as the "border effect". We analytically formulate tax pass-through β as a function of a consumer's proximity to the nearest lower-tax state, denoted by $Dist_i$:

$$\beta = \overline{\tau} + \tau(Dist_i), \tag{14}$$

where $\overline{\tau}$ reflects the "true" pass-through rate of cigarette taxes to prices, and $\tau(Dist_i)$ captures the "border effect" as a function of distance to the nearest lower-tax state. We assume that the "border effect" reaches a maximum at the border with the lower-tax state and then gradually declines with increasing distance from the border.

Consequently, the estimator of tax pass-through rate obtained from equation (13), which does not account for the "border effect", demonstrates a negative bias. The expected value of the tax pass-through rate can be expressed as:

$$\mathbb{E}[\beta] = \overline{\tau} + \mathbb{E}[\tau(Dist_i)], \tag{15}$$

where $\overline{\tau}$ represents the "true" tax pass-through estimate, and $\mathbb{E}[\tau(Dist_i)]$ can be treated as the average bias introduced by cross-border tax evasion.

To parametrize the spatial heterogeneity in the pass-through of taxes to prices, we impose an exponential functional form on the $\tau(Dist_i)$ and estimate the following regression model:

$$P_{u,i,s,t} = \overline{\tau} \cdot \tau_{s,t} + \tau_{max} e^{-\phi \cdot Dist_i} \cdot \tau_{s,t} + \gamma_s + \lambda_t + \alpha_u + \epsilon_{u,i,s,t}, \tag{16}$$

where parameter ϕ controls the concavity of the distance function. Figure 24 illustrates the shape of $e^{-\phi \cdot Dist_i}$ for a range of ϕ values.

This specification allows the "border effect" to vary with proximity to the border, reaching its peak τ_{max} at the border and declining towards zero at large distances. Including this term enables us to recover both the unbiased tax pass-through $\overline{\tau}$ estimate and the magnitude of the border effect, thus providing a more accurate estimate of the tax pass-through in the presence of cross-border cigarette purchasing.

The tax pass-through Π, following equation (16), has the following analytical formulation:

$$\Pi = \overline{\tau} + \tau_{max} e^{-\phi \cdot Dist_i}, \tag{17}$$

where $\overline{\tau}$ term represents the "true" unbiased tax pass-through estimate and the term $\tau_{max} e^{-\phi \cdot Dist_i}$ represents the "border effect" function.

Further, we extend equation (16) by interacting the spatial component with the difference in tax rates between the home state and the nearest lower-tax state. The extended regression specification is as follows:

$$P_{u,i,s,t} = \overline{\tau} \cdot \tau_{s,t} + \tau_{max} e^{-\phi \cdot Dist_i} \cdot [\tau_{s,t} - \tau_{l,t}] + \gamma_s + \lambda_t + \alpha_u + \epsilon_{u,i,s,t}, \tag{18}$$

where $\tau_{s,t}$ and $\tau_{l,t}$ denote the home state tax and the tax rate in the nearest lower tax state, respectively. In this case, the "border effect" formulation is enhanced by the tax difference $[\tau_{s,t} - \tau_{l,t}]$, reaching its maximum at $\tau_{max} \cdot [\tau_{s,t} - \tau_{l,t}]$ at the border when $Dist_i = 0$. This formulation allows the strength of the "border effect" to vary both with spatial proximity and the magnitude of the inter-state tax differential.

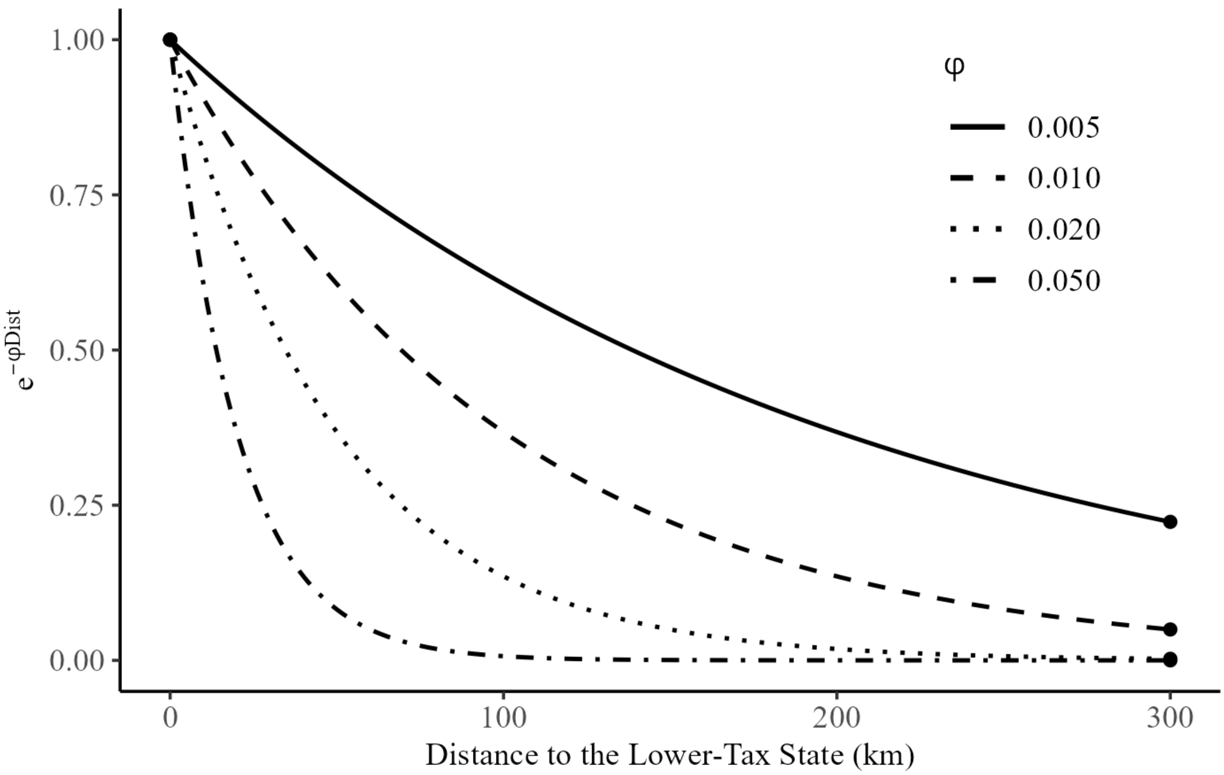


Figure 24: $e^{-\phi \cdot Dist_i}$ for a Range of ϕ Parameters

The tax difference between the home state tax and the nearest lower tax state determines the slope of the distance function. Figure 25 illustrates the values for $\Delta\tau \cdot e^{-\phi \cdot Dist_i}$ for a range of different $\Delta\tau$ parameters with $\phi = 0.01$.

Note that equation (18) can be re-written in terms of relative tax difference $\Delta_{s,t} = \frac{[\tau_{s,t} - \tau_{l,t}]}{\tau_{s,t}}$, which allows us to recover tax pass-through rate estimate:

$$P_{u,i,s,t} = \overline{\tau} \cdot \tau_{s,t} + \tau_{max} e^{-\phi \cdot Dist_i} \cdot \Delta_{s,t} \tau_{s,t} + \gamma_s + \lambda_t + \alpha_u + \epsilon_{u,i,s,t}, \tag{19}$$

where $\Delta_{s,t}$ is defined as the tax difference between the home state and the lower-tax state relative to the home state tax rate. Consequently, the tax pass-through Π for the extended model can be expressed using the following equation:

$$\Pi = \overline{\tau} + \tau_{max} e^{-\phi \cdot Dist_i} \cdot \Delta_{s,t}, \tag{20}$$

in which the term $\overline{\tau}$ term represents "true" unbiased tax pass-through estimate and $\tau_{max} e^{-\phi \cdot Dist_i} \cdot \Delta_{s,t}$ term represents the "border effect" component.

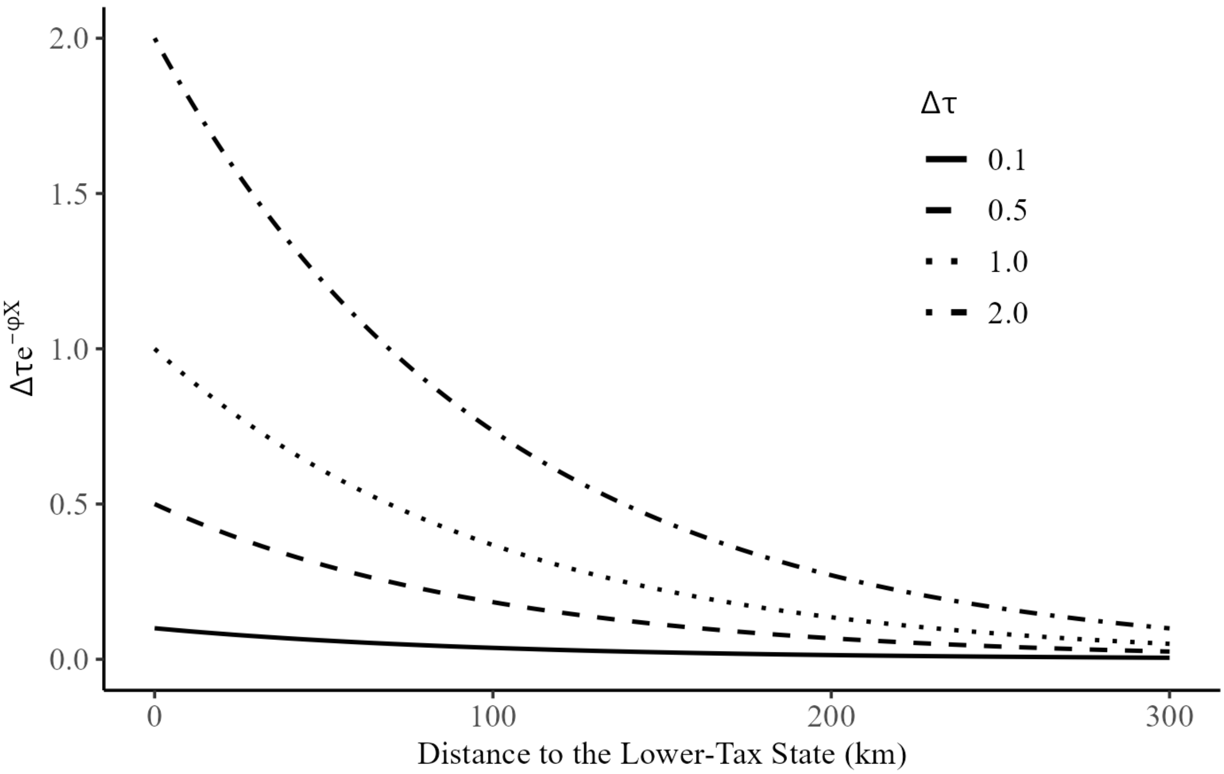


Figure 25: $\Delta\tau \cdot e^{-\phi \cdot Dist_i}$ for a Range of Different $\Delta\tau$ Parameters and $\phi = 0.01$

3.4 Estimation Strategy and Results

We begin our empirical analysis by estimating the baseline equation (16). If ϕ were known, the equation (16) can be estimated by the ordinary least squares (OLS) method, and the estimates obtained will be consistent and asymptotically normal. In this problem setting, ϕ is unknown and needs to be estimated. We estimate the exponential growth model represented by equation (16) using the non-linear least squares (NLS) method, as the model is non-linear in parameter ϕ. Following the pioneering work of Becsey et al. (1968), we employ a direct grid search approach. Grid search is a brute-force optimization technique that systematically evaluates a range of possible parameter values and selects the one that minimizes the sum of squared errors, $S_n(\phi)$. Formally, the NLS estimator of ϕ is defined as:

$$\hat{\phi} = \min_{\phi \in \Gamma} S_n(\phi), \tag{21}$$

where Γ is a bounded set for possible values of the parameter ϕ. For our purpose, we tried 200 possible values of ϕ for the grid ranging from 0.0005 to 0.1 in 0.0005 increments.

We selected this method due to its simplicity, robustness, and a large number of fixed effects. Our specification includes 6,418 UPC fixed effects, 49 state fixed effects, and 65 time fixed effects. Estimating a non-linear model with such high dimensionality would be computationally intensive. Because the fixed-effect estimates themselves are not focus of interest, we reduce the dimensionality of the problem by applying the Frisch–Waugh–Lovell (FWL) theorem. Using this theorem, we regress the residualized target variable on the residualized components of the predictors of interest.

The residuals are obtained from models in which both the target and the predictors are regressed on the complete set of fixed effects. In our setting, the target variable is the price of a cigarette pack, while the key predictors are the state cigarette tax and the border effect variable.

The residualized version of each variable is obtained in two steps. In the first stage, we regress the variables of interest (including state and time fixed effects) on UPC fixed effects and compute the residuals. This step is equivalent to subtracting UPC-specific averages from each variable. In the second stage, we regress these first-stage residuals on the residualized state and time fixed effects to obtain the second-stage residuals, which are then used in the final reduced-form regression. In this problem setting, the residualized versions of the target variable and the tax rate are computed only once, and only the residualized version of the "border effect" variable is recalculated at each iteration of the grid search. This approach significantly simplifies computational complexity.

Under the assumption that the model falls into the category of moderately non-linear models with normally distributed errors, we can construct an approximate confidence interval for the parameter ϕ following the methodology proposed by Beale (1960). In this framework, confidence regions for non-linear estimators are derived using the likelihood ratio criterion. Specifically, a given parameter ϕ is included in the $1-\alpha$ confidence region if the corresponding residual sum of squares, $S_n(\phi)$, lies within the margin defined by the following equation:

$$S_n(\phi) = S_n(\hat{\phi}) \cdot [1 + \frac{p}{n-p} F_{1-\alpha}(p, n-p)], \tag{22}$$

where p is number of parametes, $F_{1-\alpha}(p, n-p)$ is an appropriate quantile of F-distribution with $(p, n-p)$ degrees of freedom. It is important to note that when the number of observations n is large, the scaled $p \cdot F$-distribution converges to a χ^2_p distribution.

Given the large sample size, and following equation (22), the F-type statistic for the non-linear parameter ϕ has an asymptotic χ^2_1 distribution:

$$F_n(\phi) = n \cdot \frac{S_n(\phi) - S_n(\hat{\phi})}{S_n(\hat{\phi})}, \tag{23}$$

Thus $1-\alpha$ asymptotic confidence interval for ϕ can be defied as the set of ϕ for which $F_n(\phi)$ is smaller than critical value $c_{1-\alpha}$ of χ^2_1 distribution:

$$C_\phi = \{\phi : F_n(\phi) \leq c_{1-\alpha}\}, \tag{24}$$

This approach is particularly convenient for parameters obtained using the grid-search method, because confidence intervals can be directly obtained from the computation results of the least-square minimization. For example, the same approach was used by Hansen (2017) to construct confidence intervals for the threshold parameter in the regression kink model using the grid search method.

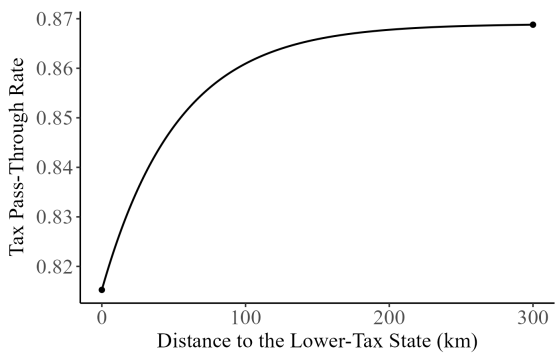
Confidence intervals for the regression coefficients $\overline{\tau}$ and τ_{max} are obtained by computing confidence intervals for each $\phi \in C_\phi$ and taking their union.

Table 17 demonstrates estimation results of the equation (16), along with heteroskedasticity-robust asymptotic 95%-level confidence regions.

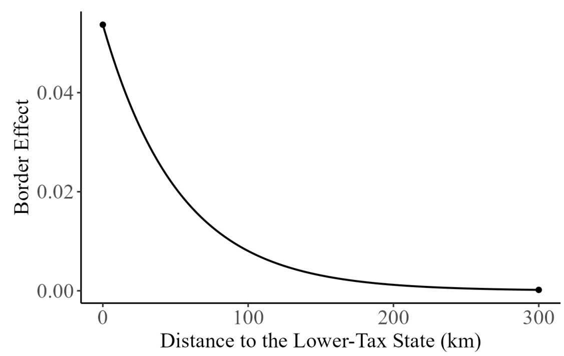
Table 17: Baseline Model: Border Effect and Tax Pass-Through Estimates

Variable	Estimate	95% CI	
ϕ	0.019	[0.0165,	0.0215]
τ_{max}	-0.0537	[-0.0581,	-0.0495]
$\overline{\tau}$	0.8690	[0.8630,	0.8754]

Figure 26 illustrates how the estimated pass-through rate and border effect vary with distance to the nearest lower-tax state.



(a) Tax Pass-Through



(b) Border Effect

Figure 26: Baseline Model: Exponential Growth Model.

The results confirm the presence of spatial heterogeneity in the pass-through of state cigarette taxes to prices. The model allows us to recover both the "true" behavioral tax pass-through estimate $\overline{\tau}$ and the magnitude of the "border effect" bias τ_{max}.

The analysis proceeded with estimation of the enhanced equation (18). We follow the same estimation procedure as described for the baseline equation (16). For this purpose, we tried 191 possible values of ϕ for the grid ranging from 0.001 to 0.02 in 0.0001 increments. Table 18 demonstrates estimation results of the equation (18) together with heteroskedasticity-robust asymptotic 95%-level confidence regions.

Table 18: Extended Model: Border Effect and Tax Pass-Through Estimates

Variable	Estimate	95% CI	
ϕ	0.0076	[0.0072,	0.0079]
τ_{max}	-0.3036	[-0.3105,	-0.2964]
$\overline{\tau}$	0.9573	[0.9516,	0.9634]

The two-dimensional Figure 27 illustrates the evolution of the pass-through rate across distance to the nearest lower tax state, as well as the relative difference between the home-state tax and the lower tax state.

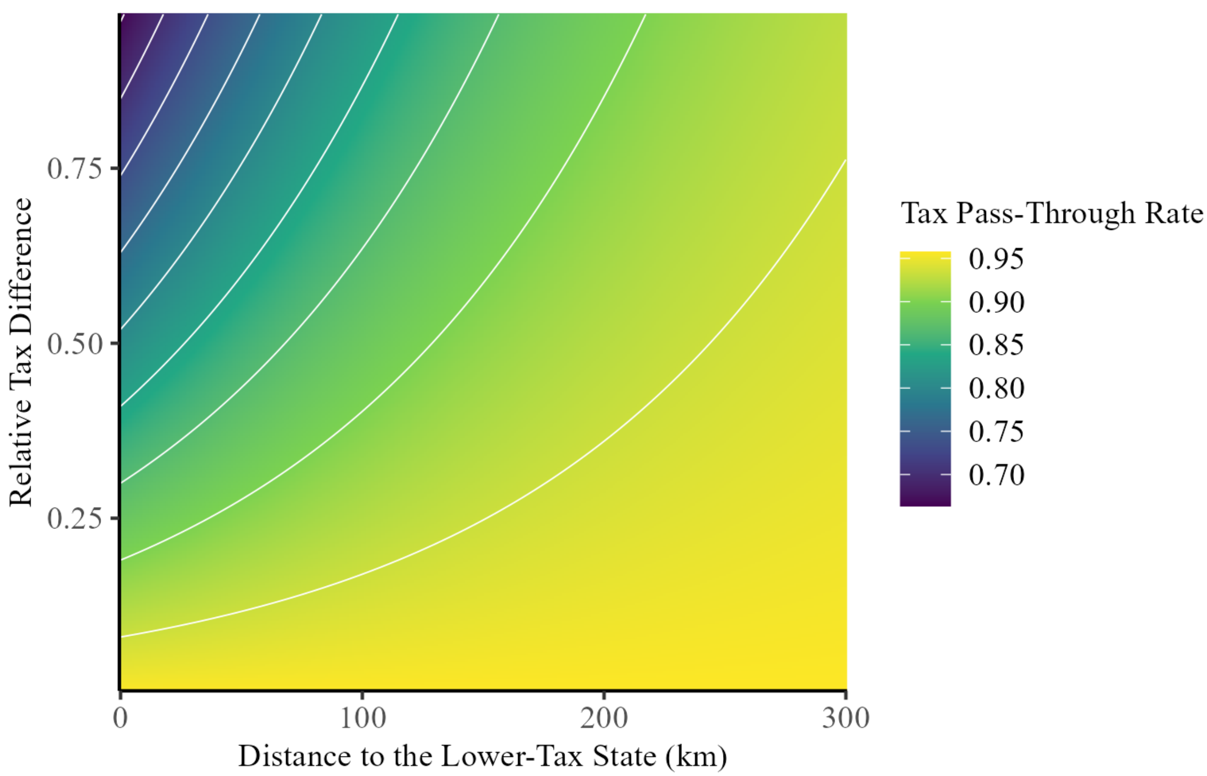


Figure 27: Extended Model: Exponential Growth Model.

The results suggest that incorporating tax differentials into the model substantially increases the variation of the estimated tax pass-through, with values ranging from 0.67 to 0.96. In contrast, the baseline model produces a significantly narrower range of estimates, between 0.81 and 0.87. This greater variation observed in the extended model underscores the importance of including inter-state tax differentials into the model equation.

3.5 Robustness Check

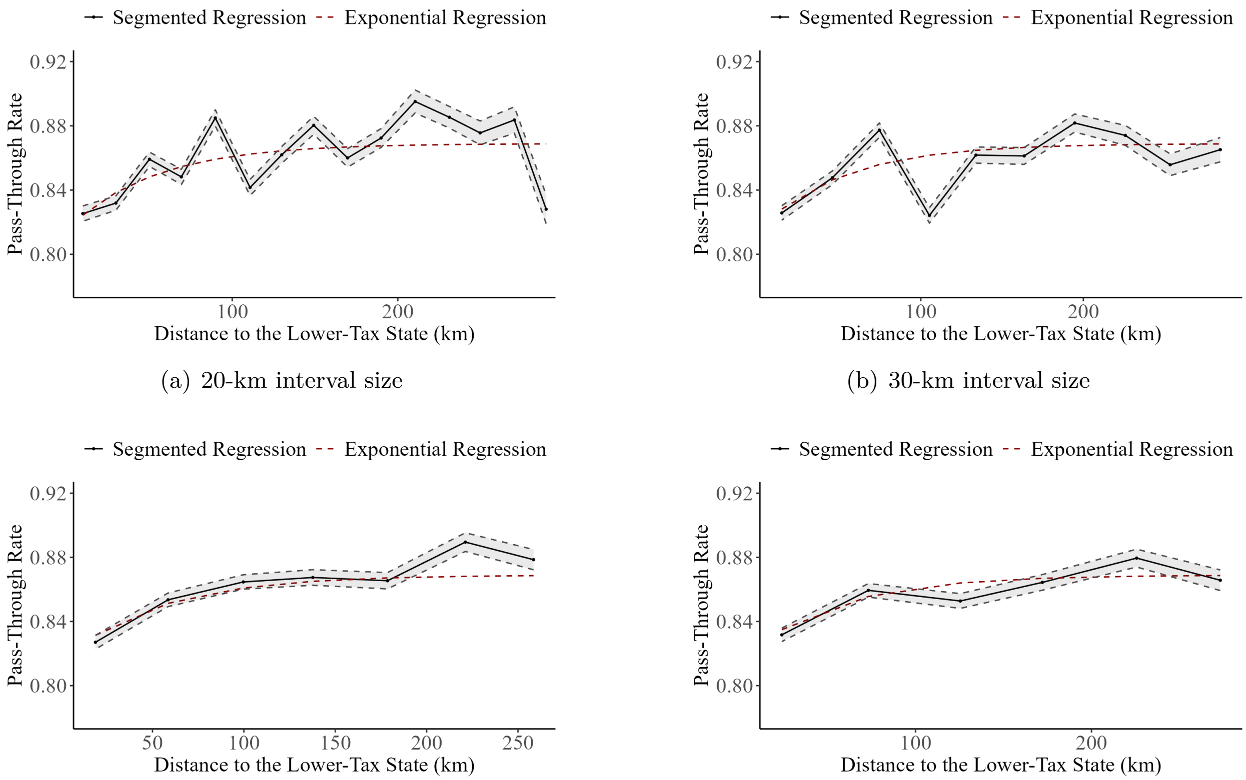
To assess the robustness of the results obtained, we perform additional analysis of the models. Specifically, to validate the model assumptions, we examine how the pass-through of state excise taxes to cigarette prices varies with distance to the nearest lower-tax state without imposing the parametric structure specified in equations (16) and (18). We then compare these findings with the non-linear regression estimates reported in Section 3.4.

We begin by estimating a segmented regression model corresponding to the baseline equation (16) in which the tax pass-through is allowed to vary across equally sized discrete distance intervals. The specification is as follows:

$$P_{u,i,s,t} = \sum_{g=1}^{G} \tilde{\tau}_g \cdot \mathbf{1}_{\left(Dist_i \geq D_{(g-1)}\right) \& \left(Dist_i < D_{(g)}\right)} \cdot \tau_{s,t} + \gamma_s + \lambda_t + \alpha_u + \epsilon_{u,i,s,t}, \tag{25}$$

where $D_{(g-1)}$ and $D_{(g)}$ define cutoff values for the distance interval g, $D_{(0)}$ and $D_{(G)}$ correspond to the minimum and maximum values of the distance variable in the dataset.

The figures below display G estimates of tax pass-through $\tilde{\tau}_g$ along with 95%-level confidence bands from the equation (25). We plot the tax pass-through estimates from the segmented regression against each midpoint of the distance interval. The red-dotted line overlays the fitted tax pass-through function $\Pi = \bar{\tau} + \tau_{max} e^{-\widehat{\phi} \cdot Dist_i}$ from the exponential model (16). We observe that small interval sizes may produce volatile estimates, and too-large intervals may absorb meaningful variation.



(a) 20-km interval size

(b) 30-km interval size

(c) 40-km interval size

(d) 50-km interval size

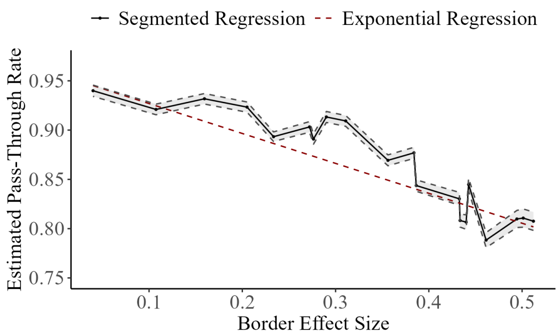
Figure 28: Segmented regression estimates of tax pass-through $\tilde{\tau}_g$ along with 95%-level confidence bands from the baseline model for 20, 30, 40, and 50-kilometer interval sizes. The red-dotted line overlays the fitted tax pass-through function $\Pi = \bar{\tau} + \tau_{\text{max}} e^{-\widehat{\phi} \cdot \text{Dist}_i}$ from the exponential model.

We next estimate a segmented regression model corresponding to the extended equation (18). In this case, the tax pass-through estimate varies across two dimensions: (1) distance to the nearest lower-tax state and (2) the difference in tax rates between the home state and the nearest lower-tax state. In order to evaluate the combined impact of these two factors on pass-through rate, we estimate the segmented regression, in which the combined "border effect" defined as $e^{-\widehat{\phi} \cdot Dist_i} \cdot [\tau_{s,t} - \tau_{l,t}]$ is allowed to vary across equally sized discrete intervals.

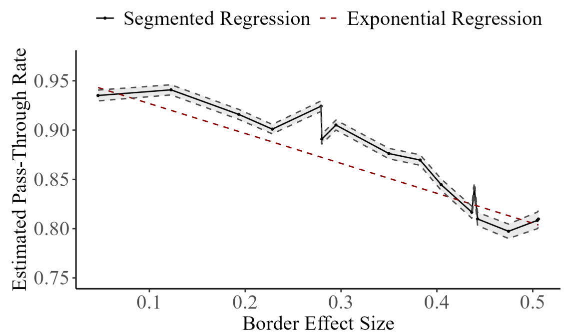
$$\begin{aligned} P_{u,i,s,t} = &\bar{\tau}_{max,0} \cdot \mathbf{1}_{\text{BorderEffect}_i = 0} \cdot \tau_{s,t} + \\ &\sum_{g=1}^{G} \bar{\tau}_{max,g} \cdot \mathbf{1}_{\left(\text{BorderEffect}_i > B_{(g-1)}\right) \& \left(\text{BorderEffect}_i \leq B_{(g)}\right)} \cdot \tau_{s,t} + \gamma_s + \lambda_t + \alpha_u + \epsilon_{u,i,s,t}, \end{aligned} \tag{26}$$

where BorderEffect_i is defined as $e^{-\widehat{\phi} \cdot Dist_i} \cdot [\tau_{s,t} - \tau_{l,t}]$ using estimate $\widehat{\phi}$ from the exponential regression, $B_{(g-1)}$ and $B_{(g)}$ define cutoff values for the BorderEffect_i interval g, and $B_{(0)}$ and $B_{(G)}$ correspond to the minimum and maximum values of the "border effect" variable in the dataset.

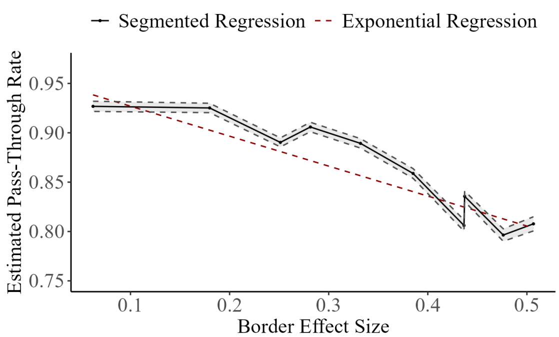
The figures below display G estimates of tax pass-through $\overline{\tau}_{max,g}$ along with 95%-level confidence bands from the equation (26) for $G = 6$, $G = 10$, $G = 16$, and $G = 20$. We plot the tax pass-through estimates from the segmented regression against mean values of $e^{-\widehat{\phi} \cdot Dist_i} \cdot \Delta_{s,t}$ at each "border effect" interval. The red-dotted line overlays the fitted tax pass-through function $\Pi = \overline{\tau} + \tau_{max} e^{-\widehat{\phi} \cdot Dist_i} \cdot \Delta_{s,t}$ from the exponential model (19).



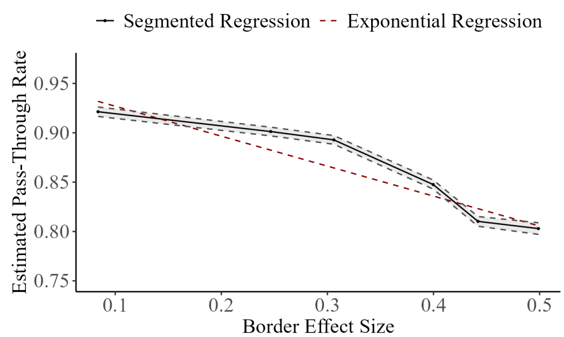
(a) Number of intervals used is 20



(b) Number of intervals used is 16



(c) Number of intervals used is 10



(d) Number of intervals used is 6

Figure 29: Segmented regression estimates of the tax pass-through estimate $\overline{\tau}_{\max,g}$ along with 95% confidence bands from the extended model for $G = 6$, $G = 10$, $G = 16$, and $G = 20$. The red-dotted line overlays the fitted tax pass-through function $\Pi = \overline{\tau} + \tau_{\max} e^{-\widehat{\phi} \cdot Dist_i} \cdot \Delta_{s,t}$ from the exponential model.

The results provide strong evidence of a negative relationship between tax pass-through rates and the extent of the composite "border effect", which depends on both the distance to the nearest lower-tax state and the differences in tax rates between the home state and the nearest lower-tax state.

Furthermore, the robustness analysis supports our parametric specifications: the tax pass-through observed in the segmented regressions is consistent with the tax pass-through predicted by the exponential models.

3.6 Demographic Heterogeneity

In this section, we analyze how tax pass-through rates differ among various demographic sub-groups. We concentrate on two important aspects of demographic heterogeneity: annual per capita income and employment status. Both of these factors are closely linked to consumers' financial situation and, therefore, are likely to shape their behavioral responses to taxation.

3.6.1 Income Groups

We begin by analyzing pass-through rates across households with different levels of annual per capita income. Per capita income is computed by dividing the combined annualized household income by household size. All income values are adjusted for inflation using 2017 as the base year. Households are classified into three income groups, as shown in Table 19.

Table 19: Distribution of Households by Income Group

Category	Panelists	N
Per capita income[1]		
High: $>$ \$40,000	11,656	503,466
Middle: \$15,000 – \$40,000	26,827	1,375,787
Low: $\leq$ \$15,000	18,388	968,655

For each income group, we estimate the extended model defined in equation (18). Table 20 demonstrates estimation results, along with heteroskedasticity-robust asymptotic 95%-level confidence intervals.

Table 20: Extended Model: Border Effect and Tax Pass-Through Estimates by Income Groups

	High Income		Middle Income		Low Income	
Variable	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
ϕ	0.0045	$[0.004, 0.005]$	0.0375	$[0.034, 0.0415]$	0.0065	$[0.006, 0.007]$
τ_{max}	-0.2216	$[-0.2381, -0.2046]$	-0.3960	$[-0.4208, -0.3733]$	-0.2618	$[-0.2739, -0.2495]$
$\overline{\tau}$	0.9522	$[0.9401, 0.9641]$	0.8985	$[0.8901, 0.9069]$	0.9280	$[0.9188, 0.9372]$

For each observation in the dataset, we calculate the predicted pass-through rate. Figure 30 shows the average predicted pass-through rate along the distance to the nearest lower-tax state.

The results indicate that high-income households exhibit the highest tax pass-through rates and are largely unaffected by cross-border shopping opportunities. Middle-income consumers display sensitivity to the "border effect" only when they reside within 90 kilometers of a lower-tax state border. Beyond this distance, higher travel costs discourage cross-border shopping. In contrast, low-income consumers remain highly responsive to the "border effect", even at distances greater than 90 kilometers. However, beyond 200 kilometers, low-income consumers demonstrate higher tax pass-through compared to middle-income consumers.

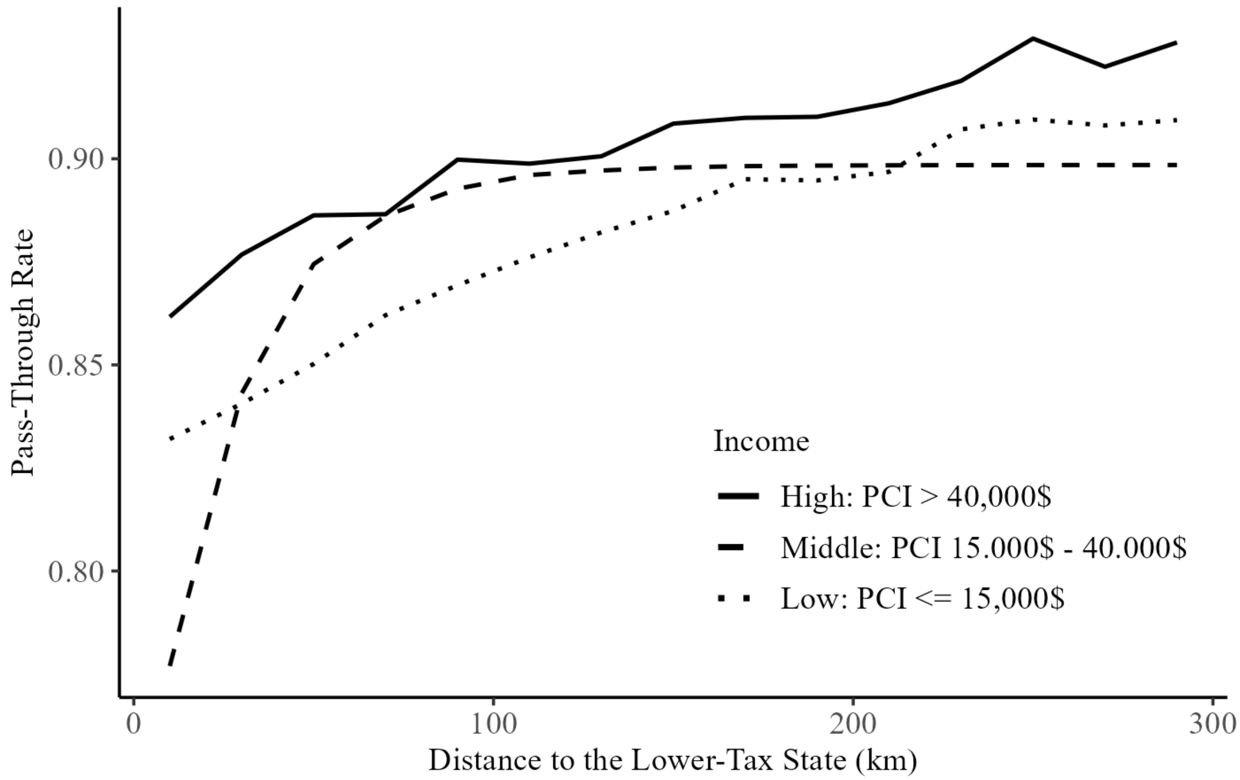


Figure 30: Predicted Tax Pass-Through by Income Group.

3.6.2 Employment Status

Next, we examine demographic heterogeneity by employment status. We divide households into two categories based on the employment status of the male household head: "Employed" and "Not Employed for Pay". If there is no male head in the household, we consider the female head of the household. The distribution of households across these categories is summarized in Table 21.

Table 21: Distribution of Households by Employment Status

Category	Panelists	N
Head Employment[2]		
Employed	36,450	1,727,012
Not employed for pay	18,017	1,120,896

As before, we estimate the extended model defined in equation (18) for each demographic subgroup. Table 22 shows estimation results, along with heteroskedasticity-robust asymptotic 95%-level confidence intervals.

Table 22: Extended Model: Border Effect and Tax Sensitivity Estimates by Employment Status

	Employed		Not Employed for Pay	
Variable	Estimate	95% CI	Estimate	95% CI
ϕ	0.016	$[0.015, 0.0165]$	0.0031	$[0.0029, 0.0034]$
τ_{max}	-0.3205	$[-0.3305, -0.3095]$	-0.2784	$[-0.2893, -0.2673]$
$\overline{\tau}$	0.9427	$[0.9362, 0.9504]$	0.9416	$[0.9334, 0.9497]$

For every observation in the dataset, we calculate the predicted pass-through rate. Figure 31 displays the predicted pass-through rates based on employment status along the distance to the nearest lower-tax state.

Figure 31: Predicted Tax Pass-Through by Employment Status.

The findings show that households in which the head is not employed for pay demonstrate significantly lower pass-through rates compared to households with an employed head. This suggests that non-employed consumers are more likely to avoid taxes, possibly due to greater flexibility in shopping behavior or stronger incentives to engage in cross-border purchases.

3.7 Conclusion

This study provides empirical evidence of spatial heterogeneity in the pass-through of cigarette taxes to prices driven by cross-border purchasing behavior. Using detailed NielsenIQ Consumer Panel data from 2004 to 2019, we develop and estimate an exponential growth model with UPC, location, and time-fixed effects to assess the "border effect"—a systematic attenuation in tax pass-through among consumers residing near lower-tax states. Our results show that this effect is a declining function of distance to the nearest lower-tax border, with the strongest influence at the boundary and diminishing impact as distance increases. Additionally, we find that the border effect's strength increases with the magnitude of the tax difference between a consumer's home state and the nearest lower-tax state. Our robustness analysis validates the parametric assumptions of our model and supports the conclusion that spatial variation in tax pass-through rate is statistically and economically meaningful.

In addition to geographic variation, we investigate differences in cigarette tax pass-through across demographic groups. High-income households face the highest tax pass-through and are largely unaffected by border proximity, while middle-income households are affected by the "border effect" only when the distance from the lower-tax state does not exceed 90 kilometers. Low-income households remain sensitive to "border effects" at greater distances, though their responsiveness declines beyond 200 kilometers. Moreover, consumers not employed for pay exhibit significantly lower pass-through, suggesting greater scope for tax avoidance through flexible shopping behavior. Therefore, the welfare implications of cigarette taxation differ across demographic groups. High-income households experience a higher tax pass-through, meaning they bear a larger tax burden. In contrast, low-income and non-employed households face a lower effective tax burden because they benefit from reduced tax pass-through resulting from cross-border cigarette purchasing and other tax avoidance actions.

The previous chapter analyzed the influence of cross-border purchasing on the tax sensitivity of cigarette consumption. The current chapter extends this analysis by investigating how cross-border purchasing opportunities in the lower-tax states affect the pass-through of cigarette taxes to prices. When cigarette taxes increase in the home state, consumers residing close to a lower-tax state border may engage in cross-border shopping. Consequently, the prices these consumers effectively pay for cigarettes do not increase by the full amount of the tax, and their cigarette consumption responds only partially to higher taxes. Empirical findings in both chapters support this mechanism. Moreover, retailers near borders with lower-tax states may refrain from fully passing tax increases onto cigarette prices to avoid losing customers to cross-border shopping. These profit-driven pricing incentives further influence both the tax sensitivity of cigarette consumption and the degree of tax pass-through for border residents.

These findings have important implications for the design of tobacco tax policy. We show that cigarette excise taxes, while effective on average, may operate unevenly across geographies and demographic groups. As a result, achieving uniform public health outcomes and accurately forecasting tax revenues becomes challenging when neighboring jurisdictions maintain large tax differentials. For example, when a neighboring state maintains a relatively low cigarette tax rate, a substantial increase in the home-state cigarette tax may induce consumers residing near the border to engage in cross-border purchasing. In response, retailers near the border may refrain from fully passing the tax increase on to prices to retain customers. Conversely, when a neighboring state increases its cigarette tax while the home state maintains a relatively low tax rate, the home state may experience increased cigarette demand due to cross-border purchases by consumers from the higher-tax jurisdiction. These results emphasize the need to incorporate spatial considerations into the evaluation and design of excise tax policy, particularly in a tax system with heterogeneous tax regimes such as the US.

Future research could build on these results by exploring interstate cross-border purchases in Mexico and investigating the additional impact of the variability in sales taxes.

4 A Comparison of Threshold and Exponential Regression Model Specifications

This section compares threshold and exponential regression model specifications. The first subsection measures the "border effect" on the tax sensitivity of cigarette consumption using an exponential regression specification. The estimated border-effect function is then compared with those obtained from the threshold regression and segmented regression models discussed in Section 2. The analysis is conducted for both baseline and extended model specifications.

The second subsection examines the border-effect function for the pass-through of cigarette taxes to prices using a threshold regression specification. Estimates from the threshold regression model are compared with those from the exponential and segmented regression models discussed in Section 3, using both baseline and extended model formulations.

4.1 Spatial Heterogeneity in Tax Sensitivity of Cigarette Consumption

In this sub-section, we analytically define tax sensitivity as a function of a consumer's proximity to the nearest lower-tax state, employing an exponential regression specification. We start with the baseline model and estimate the following regression equation:

$$Y_{i,t} = \overline{\tau} \cdot \tau_{s,t} + \tau_{max} e^{-\phi \cdot Dist_i} \cdot \tau_{s,t} + \gamma_s + \lambda_t + \epsilon_{i,t}, \tag{27}$$

where the parameter ϕ determines the concavity of the distance function. Consistent with prior assumptions, the "border effect" attains its maximum at the border with the lower-tax state and gradually diminishes as the distance from the border increases.

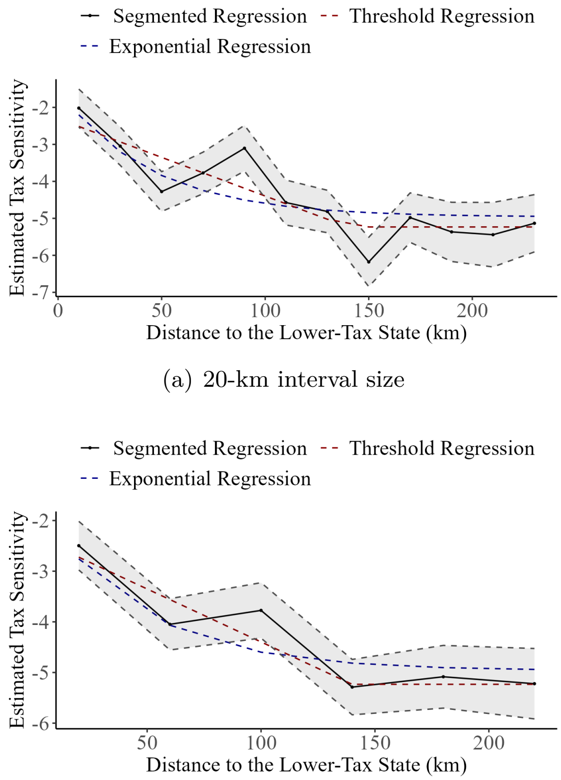
This specification allows the "border effect" to vary with proximity to the border, reaching its peak τ_{max} at the border and declining towards zero at large distances. Including this term enables us to recover both the unbiased tax sensitivity estimate, $\overline{\tau}$, and the magnitude of the border effect, thereby providing a more accurate estimate of tax sensitivity in the presence of cross-border cigarette purchasing.

As before, we estimate the exponential regression using the non-linear least squares (NLS) method, given the model's non-linearity in parameter ϕ. A direct grid search approach is employed. Table 23 presents the estimation results for equation 27, along with heteroskedasticity-robust asymptotic 95% confidence intervals.

Table 23: Baseline Model: Border Effect and Tax Sensitivity Estimates

Variable	Estimate	95% CI	
ϕ	0.0225	[0.0165,	0.0315]
τ_{max}	3.4585	[3.0419,	3.9395]
$\overline{\tau}$	−4.9637	[−5.7389,	−4.2198]

Figures 32 compare the estimated tax sensitivity function from exponential regression with those derived from the threshold and segmented regression models discussed in Section 2.



(a) 20-km interval size

(b) 30-km interval size

(c) 40-km interval size

(d) 50-km interval size

Figure 32: Segmented regression estimates of tax sensitivity along with 95%-level confidence bands from the baseline model for 20, 30, 40, and 50-kilometer interval sizes. The blue-dotted line overlays the fitted tax sensitivity function $\overline{\tau} + \tau_{\max} e^{-\widehat{\phi}\cdot \text{Dist}_i}$ from the exponential model. Red-dotted line represents the fitted tax sensitivity function $\overline{\tau} + \mathbf{1}_{Dist_i \leq D_{\text{cutoff}}} \tau_{max}[1 - \frac{Dist_i}{D_{\text{cutoff}}}]$ from the threshold model.

We proceed by estimating the extended model using an exponential regression specification. In this case, the tax sensitivity estimate varies across two dimensions: (1) distance to the nearest lower-tax state and (2) the difference in tax rates between the home state and the nearest lower-tax state. The model is estimated using the following equation:

$$Y_{i,t} = \overline{\tau} \cdot \tau_{s,t} + \tau_{max} e^{-\phi \cdot Dist_i} \cdot [\tau_{s,t} - \tau_{l,t}] \cdot \tau_{s,t} + \gamma_s + \lambda_t + \epsilon_{i,t}, \tag{28}$$

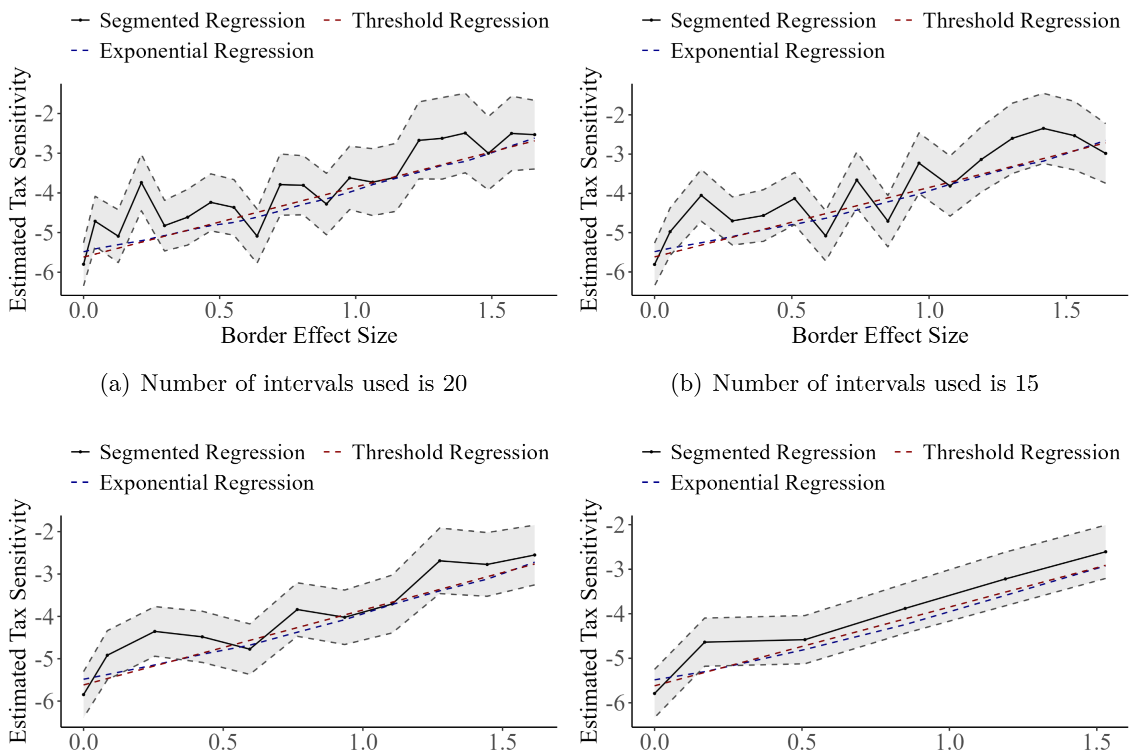
where $\tau_{s,t}$ and $\tau_{l,t}$ denote the home state tax and the tax rate in the nearest lower tax state, respectively. In this case, the "border effect" is enhanced by the tax difference $[\tau_{s,t} - \tau_{l,t}]$, attaining its maximum at $\tau_{max} \cdot [\tau_{s,t} - \tau_{l,t}]$ when $Dist_i = 0$. This formulation allows the strength of the "border effect" to vary both with spatial proximity and the magnitude of the tax difference. This extension provides a more comprehensive understanding of how cross-border cigarette purchasing influences consumer responses to cigarette taxation.

Table 24 demonstrates the estimation results for equation 28, along with heteroskedasticity-robust asymptotic 95% confidence intervals.

Table 24: Extended Model: Border Effect and Tax Sensitivity Estimates

Variable	Estimate	95% CI
ϕ	0.0205	[0.016, 0.026]
τ_{max}	2.0968	[1.8206, 2.3781]
$\overline{\tau}$	-5.4843	[-6.2480, -4.7517]

Figures 33 present a comparison of the estimated tax sensitivity function obtained from exponential regression with those derived from the threshold and segmented regression models discussed in Section 2.



(a) Number of intervals used is 20

(b) Number of intervals used is 15

(c) Number of intervals used is 10

(d) Number of intervals used is 5

Figure 33: Segmented regression estimates of the tax sensitivity along with 95%-level confidence bands from the extended model for $G = 5$, $G = 10$, $G = 15$, and $G = 20$. The blue-dotted line overlays the fitted tax sensitivity function $\overline{\tau} + \tau_{\max} e^{-\widehat{\phi}\cdot\text{Dist}_i} \cdot [\tau_{s,t} - \tau_{l,t}]$ from the exponential model. The red-dotted line represents the fitted tax sensitivity function $\overline{\tau} + \mathbf{1}_{Dist_i \leq D_{\text{cutoff}}} \tau_{max}[1 - \frac{Dist_i}{D_{\text{cutoff}}}] \cdot [\tau_{s,t} - \tau_{l,t}]$ from the threshold model.

We find that the exponential and threshold regression specifications produce similar tax-sensitivity functions. However, for the baseline model, the threshold regression specification more closely aligns with the segmented regression estimates than the exponential regression specification does.

4.2 Spatial Heterogeneity in Tax Pass-Through of Cigarette Taxes to Prices

In this sub-section, we provide an analytical formulation of the pass-through of cigarette taxes to prices as a function of consumer proximity to the nearest lower-tax state, using a threshold regression specification. We begin by estimating the baseline model with the following regression equation:

$$P_{u,i,s,t} = \overline{\tau} \cdot \tau_{s,t} + \mathbf{1}_{Dist_i \leq D_{\text{cutoff}}} \tau_{max}[1 - \frac{Dist_i}{D_{\text{cutoff}}}] \cdot \tau_{s,t} + \gamma_s + \lambda_t + \alpha_u + \epsilon_{u,i,s,t}, \tag{29}$$

where $\mathbf{1}_{Dist_i \leq D_{\text{cutoff}}}$ is an indicator function which is equal to 1 if consumer i resides within D_{cutoff} kilometers of a lower-tax border and zero otherwise. This specification allows the "border effect" to vary linearly with proximity to the border, reaching its peak τ_{max} at the border and declining to zero at the cutoff threshold. Including this term enables us to recover both the unbiased tax pass-through $\overline{\tau}$ and the magnitude of the "border effect".

Consistent with the previous analysis, the threshold regression is estimated using the conditional least squares (CLS) method. A direct grid search approach is used to estimate the parameter D_{cutoff}. Table 25 presents the estimation results for equation 29, together with heteroskedasticity-robust asymptotic 95% confidence intervals.

Table 25: Baseline Model: Border Effect and Tax Pass-Through Estimates

Variable	Estimate	95% CI	
D_{cutoff}	83	[81,	84]
τ_{max}	−0.0455	[−0.0487,	−0.0421]
$\overline{\tau}$	0.8650	[0.8605,	0.8692]

Figures 34 compare the estimated tax pass-through function from threshold regression with those derived from the exponential and segmented regression models discussed in Section 3.

The analysis proceeded with estimation of the extended model using threshold regression specification. In this framework, the tax pass-through estimate varies across two dimensions: (1) distance to the nearest lower-tax state and (2) the difference in tax rates between the home state and the nearest lower-tax state. The model is estimated using the following equation

$$P_{u,i,s,t} = \overline{\tau} \cdot \tau_{s,t} + \mathbf{1}_{Dist_i \leq D_{\text{cutoff}}} \tau_{max}[1 - \frac{Dist_i}{D_{\text{cutoff}}}] \cdot [\tau_{s,t} - \tau_{l,t}] + \gamma_s + \lambda_t + \alpha_u + \epsilon_{u,i,s,t}, \tag{30}$$

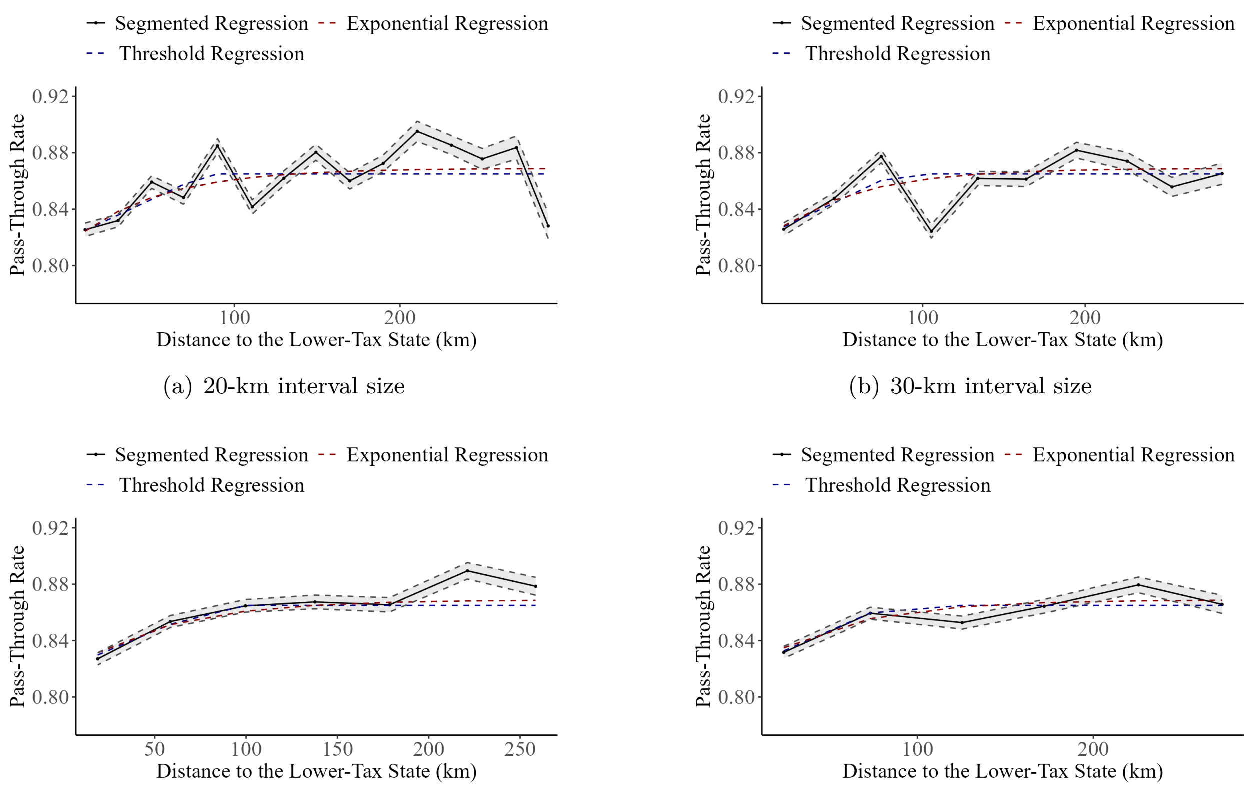
where $\tau_{s,t}$ and $\tau_{l,t}$ denote the home state tax and the tax rate in the nearest lower tax state, respectively. In this case, the "border effect" formulation is enhanced by the tax difference $[\tau_{s,t} - \tau_{l,t}]$, reaching its maximum at $\tau_{max} \cdot [\tau_{s,t} - \tau_{l,t}]$ at the border when $Dist_i = 0$. This formulation allows the strength of the "border effect" to vary both with spatial proximity and the magnitude of the inter-state tax differential.

Table 26 demonstrates the estimation results for equation 30, along with heteroskedasticity-robust asymptotic 95% confidence intervals.

Table 26: Extended Model: Border Effect and Tax Pass-Through Estimates

Variable	Estimate	95% CI	
D_{cutoff}	207	[204,	210]
τ_{max}	-0.2559	[-0.2615,	-0.2502]
$\overline{\tau}$	0.9431	[0.9381,	0.9481]

Figures 35 present a comparison of the estimated tax pass-through function obtained from threshold regression with those derived from the exponential and segmented regression models discussed in Section 3.



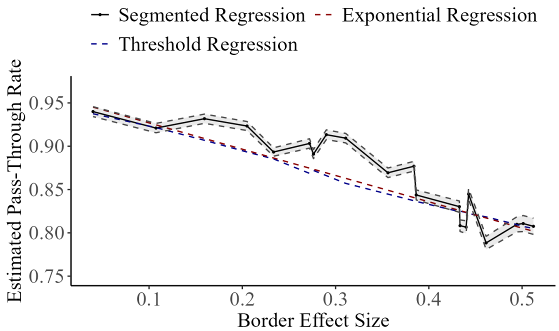
(a) 20-km interval size

(b) 30-km interval size

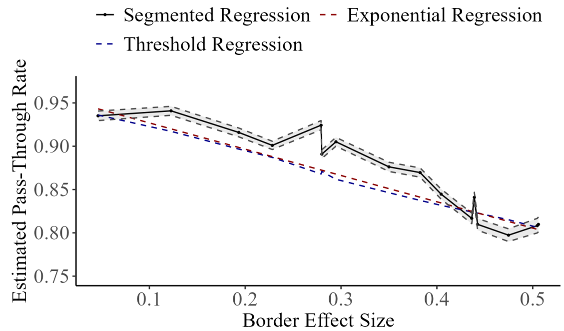
(c) 40-km interval size

(d) 50-km interval size

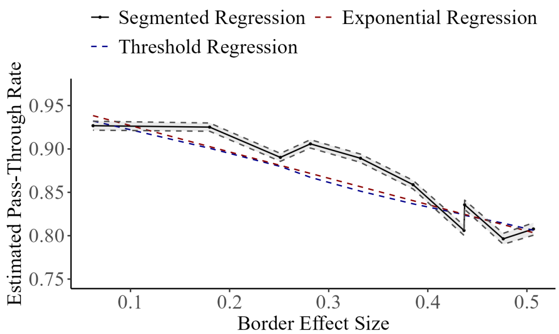
Figure 34: Segmented regression estimates of tax pass-through along with 95%-level confidence bands from the baseline model for 20, 30, 40, and 50-kilometer interval sizes. The red-dotted line overlays the fitted tax sensitivity function $\overline{\tau}+\tau_{\max}e^{-\widehat{\phi}\cdot\text{Dist}_i}$ from the exponential model. Blue-dotted line represents the fitted tax sensitivity function $\overline{\tau}+\mathbf{1}_{Dist_i\leq D_{\text{cutoff}}}\tau_{max}[1-\frac{Dist_i}{D_{\text{cutoff}}}]$ from the threshold model.



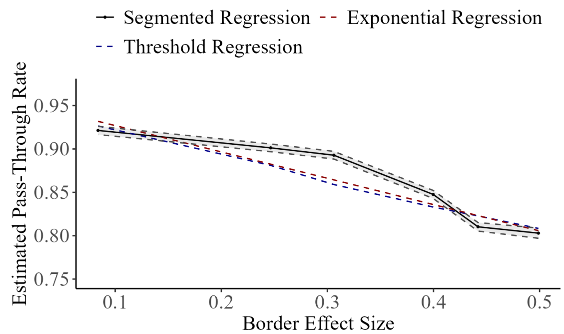
(a) Number of intervals used is 20



(b) Number of intervals used is 16



(c) Number of intervals used is 10



(d) Number of intervals used is 6

Figure 35: Segmented regression estimates of the tax pass-through estimate along with 95% confidence bands from the extended model for $G = 6$, $G = 10$, $G = 16$, and $G = 20$. The red-dotted line overlays the fitted tax pass-through function $\overline{\tau} + \tau_{\max} e^{-\widehat{\phi} \cdot Dist_i} \cdot \Delta_{s,t}$ from the exponential model. Blue-dotted line represents the fitted tax sensitivity function $\overline{\tau} + \mathbf{1}_{Dist_i \leq D_{\text{cutoff}}} \tau_{max} [1 - \frac{Dist_i}{D_{\text{cutoff}}}] \cdot \Delta_{s,t}$ from the threshold model. $\Delta_{s,t}$ is defined as the tax difference between the home state and the lower-tax state relative to the home state tax rate.

The results demonstrate that both exponential and threshold regression specifications produce similar tax pass-through functions in the baseline and extended model specifications. Moreover, the unbiased tax pass-through estimate $\overline{\tau}$ and the "border effect" parameter τ_{max} yield very similar values across threshold regression and exponential model specifications. However, within the threshold regression framework, the estimated threshold parameter D_{cutoff} varies considerably between the baseline model (83 kilometers) and the extended model (207 kilometers). This variability limits the interpretability of the threshold parameter D_{cutoff}. Therefore, we prefer the exponential regression specification, in which the nonlinear parameter ϕ directly characterizes the concavity of the pass-through function and is more intuitive in this context.

Appendix

Appendix A

Table 27: Estimation of the Baseline Model for Households That Do Not Relocate.

	Dependent variable:	
	Total packs purchased	
	(1)	(2)
Tax difference	5.589***	
	(0.539)	
Lower tax state distance	−0.004***	
	(0.001)	
Tax distance interaction	−0.027***	
	(0.002)	
Tax value	−14.045***	−12.166***
	(0.332)	(0.244)
Factor: Low income	2.017***	2.146***
	(0.529)	(0.529)
Factor: Middle Income	1.598***	1.666***
	(0.396)	(0.396)
Factor: Household size 2	2.070***	2.222***
	(0.520)	(0.520)
Factor: Household size 3	5.422***	5.646***
	(0.625)	(0.625)
Factor: Household size 4	4.573***	4.918***
	(0.749)	(0.749)
Factor: Household size 5	8.892***	9.221***
	(0.934)	(0.934)
Factor: Household size 6 plus	6.663***	6.953***
	(1.117)	(1.117)
Factor: Head employment 35+ hours	3.795***	3.877***
	(0.432)	(0.432)
Factor: Head employment Not employed	−0.779*	−0.929**
	(0.444)	(0.444)
Factor: Head education HS graduate or lower	−0.984	−1.023*
	(0.600)	(0.600)
Factor: Head education Some college	−1.263***	−1.228**
	(0.483)	(0.484)
Factor: Head age ≥50	−10.144***	−10.746***
	(1.038)	(1.038)
Factor: Head age 35-49	−3.187***	−3.401***
	(0.983)	(0.983)
Factor: Presence of children = yes	−1.320***	−1.281***
	(0.491)	(0.491)
Factor: Gender composition Female head only	−7.810***	−7.837***
	(0.645)	(0.645)
Factor: Gender composition Male head only	−4.164***	−4.211***
	(0.987)	(0.988)
Consumer fixed effects:	*yes*	*yes*
State fixed effects:	*no*	*no*
Observations	323,545	323,545
R^2	0.017	0.016
F Statistic	236.850*** (df = 20; 274149)	263.636*** (df = 17; 274152)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 27 presents the estimation results from a regression that was estimated only on households that did not relocate. Consequently, this regression specification does not include state fixed effects. We observe that the regression results have not changed significantly.

Appendix B

Table 28: Estimation of Baseline Model Excluding Household-Level Fixed Effects.

	Dependent variable:	
	Total packs purchased	
	(1)	(2)
Tax difference	7.331^{***}	
	(0.436)	
Lower tax state distance	-0.004^{***}	
	(0.001)	
Tax distance interaction	-0.027^{***}	
	(0.002)	
Tax value	-14.527^{***}	-12.063^{***}
	(0.277)	(0.215)
Factor: Low income	3.009^{***}	3.335^{***}
	(0.332)	(0.332)
Factor: Middle Income	3.420^{***}	3.618^{***}
	(0.272)	(0.272)
Factor: Household size 2	4.279^{***}	4.260^{***}
	(0.354)	(0.354)
Factor: Household size 3	-0.445	-0.459
	(0.433)	(0.433)
Factor: Household size 4	-2.726^{***}	-2.805^{***}
	(0.531)	(0.531)
Factor: Household size 5	-2.364^{***}	-2.461^{***}
	(0.669)	(0.670)
Factor: Household size 6 plus	-3.270^{***}	-3.448^{***}
	(0.774)	(0.774)
Factor: Head employment 35+ hours	1.048^{***}	1.165^{***}
	(0.357)	(0.357)
Factor: Head employment Not employed	3.941^{***}	4.023^{***}
	(0.359)	(0.359)
Factor: Head education HS graduate or lower	3.267^{***}	3.323^{***}
	(0.271)	(0.272)
Factor: Head education Some college	1.670^{***}	1.717^{***}
	(0.269)	(0.270)
Factor: Head age $\geq$50	22.883^{***}	22.861^{***}
	(0.494)	(0.494)
Factor: Head age 35-49	13.403^{***}	13.465^{***}
	(0.498)	(0.499)
Factor: Presence of children = yes	-9.191^{***}	-9.155^{***}
	(0.369)	(0.369)
Factor: Gender composition Female head only	-8.325^{***}	-8.466^{***}
	(0.317)	(0.318)
Factor: Gender composition Male head only	0.670^{*}	0.560
	(0.405)	(0.405)
Constant	38.601^{***}	37.344^{***}
	(1.015)	(1.009)
Consumer fixed effects:	*no*	*no*
State fixed effects:	*yes*	*yes*
Observations	378,101	378,101
R^2	0.044	0.042
F Statistic	255.316^{***} (df = 68; 378032)	257.801^{***} (df = 65; 378035)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 28 presents the estimation results of an alternative regression specification without household-level fixed effects and shows that the coefficient estimates predict that cigarette consumption on average decreases with household size, which is inconsistent with economic theory assumptions. The presence of heterogeneity bias determines the choice of the “within” fixed effects model. Nevertheless, estimated tax sensitivity and variables related to border effects are within a similar range.

References

Azagba, S., & Sharaf, M. (2011). Cigarette taxes and smoking participation: evidence from recent tax increases in Canada. *International journal of environmental research and public health*, *8*(5), 1583–1600.

Beale, E. M. L. (1960). Confidence regions in non-linear estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, *22*(1), 41–88.

Becsey, J. C., Berke, L., & Callan, J. R. (1968). Nonlinear least squares methods: A direct grid search approach. *Journal of Chemical Education*, *45*(11), 728.

Bureau, U. C. (2023). *TIGER/Line Shapefiles — census.gov.* `https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html`. (Accessed: 2023-11-30)

CDC STATE System Tobacco Legislation - Tax. (2021). `https://data.cdc.gov/Legislation/CDC-STATE-System-Tobacco-Legislation-Tax/2dwv-vfam/data_preview`. (Accessed: 2021-09-30)

Cebula, R. (2012). A new form of federal or state-level cigarette excise taxation in the US. *International Journal of Economics and Finance*, *4*(2), 16.

Cebula, R. J., Angjellari-Dajci, F., & Kashian, R. (2016). Are there interregional differences in the response of cigarette smoking to state cigarette excise taxes in the USA? exploratory analysis. *Journal of Applied Statistics*, *43*(8), 1494–1507.

Cebula, R. J., Foley, M., & Houmes, R. (2014). Empirical analysis of the impact of cigarette excise taxes on cigarette consumption: estimates from recent state-level data. *Journal of Economics and Finance*, *38*(1), 164–180.

Chaiyasong, S., Limwattananon, S., Limwattananon, C., Thamarangsi, T., Tangchareonsathien, V., & Schommer, J. (2011). Impacts of excise tax raise on illegal and total alcohol consumption: A Thai experience. *Drugs: education, prevention and policy*, *18*(2), 90–99.

Chan, K. (1998). Limiting properties of the least squares estimator of a continuous threshold autoregressive model. *Biometrika*, *85*(2), 413–426.

Chiou, L., & Muehlegger, E. (2008). Crossing the line: Direct estimation of cross-border cigarette sales and the effect on tax revenue. *The B.E. Journal of Economic Analysis & Policy*, *8*(1), 1–41.

Cotti, C., Courtemanche, C., Maclean, J. C., Nesson, E., Pesko, M. F., & Tefft, N. W. (2022). The effects of e-cigarette taxes on e-cigarette prices and tobacco product sales: Evidence from retail panel data. *Journal of Health Economics*, *86*, 102676.

Cotti, C., Nesson, E., & Tefft, N. (2018). The relationship between cigarettes and electronic cigarettes: Evidence from household panel data. *Journal of health economics*, *61*, 205–219.

Darden, M., Hebert, R., Pesko, M., & Sturm, S. (2025). *Cigarette Taxes and the Household Budget* (Tech. Rep.). National Bureau of Economic Research.

DeCicca, P., Kenkel, D., & Liu, F. (2013a). Excise tax avoidance: The case of state cigarette taxes. *Journal of Health Economics*, *32*(6), 1130–1141.

DeCicca, P., Kenkel, D., & Liu, F. (2013b). Who pays cigarette taxes? the impact of consumer price search. *Review of Economics and Statistics*, *95*(2), 516–529.

Evans, W. N., & Farrelly, M. C. (1998). The compensating behavior of smokers: taxes, tar, and nicotine. *The Rand journal of economics*, 578–595.

Farrelly, M. C., Nimsch, C. T., Hyland, A., & Cummings, M. (2004). The effects of higher cigarette prices on tar and nicotine consumption in a cohort of adult smokers. *Health Economics*, *13*(1), 49–58.

Gruber, J., Sen, A., & Stabile, M. (2003). Estimating price elasticities when there is smuggling: the sensitivity of smoking to price in canada. *Journal of Health Economics*, *22*(5), 821–842.

Gruenewald, P. J., Ponicki, W. R., Holder, H. D., & Romelsjö, A. (2006). Alcohol prices, beverage quality, and the demand for alcohol: quality substitutions and price elasticities. *Alcoholism: Clinical and Experimental Research*, *30*(1), 96–105.

Hansen, B. E. (2017). Regression Kink With an Unknown Threshold. *Journal of Business & Economic Statistics*, *35*(2), 228–240.

Hanson, A., & Sullivan, R. (2009). The incidence of tobacco taxation: Evidence from geographic micro-level data. *National Tax Journal*, *62*(4), 677–698.

Harding, M., Leibtag, E., & Lovenheim, M. F. (2012). The heterogeneous geographic and socioeconomic incidence of cigarette taxes: Evidence from nielsen homescan data. *American Economic Journal: Economic Policy*, *4*(4), 169–198.

Heckley, G., Jarl, J., & Gerdtham, U.-G. (2017). Frequency and intensity of alcohol consumption: new evidence from Sweden. *The European Journal of Health Economics*, *18*(4), 495–517.

James M. Kilts Center for Marketing, NielsenIQ Datasets. (2021). `https://www.chicagobooth.edu/research/kilts/datasets/nielsen`. (Accessed: 2021-09-30)

Kanny, D., Naimi, T. S., Liu, Y., Lu, H., & Brewer, R. D. (2018). Annual total binge drinks consumed by US adults, 2015. *American journal of preventive medicine*, *54*(4), 486–496.

Kim, H., & Lee, D. (2020). Racial demographics and cigarette tax shifting: evidence from scanner data. *Empirical Economics*, *61*(2), 1011–1037.

Lee, J.-M. (2008). Effect of a large increase in cigarette tax on cigarette consumption: an empirical analysis of cross-sectional survey data. *Public health (London)*, *122*(10), 1061-1067.

Lewit, E. M., & Coate, D. (1982). The potential for using excise taxes to reduce smoking. *Journal of Health Economics*, *1*(2), 121–145.

Lin H., L. Z. e. a., Chang C. (2020). The effect of the presence of children on adult smoking behaviour: empirical evidence based on china family panel studies. *BMC Public Health*, *20*(1448), 1471–2458.

Lovenheim, M. F. (2008). How far to the border?: The extent and impact of cross-border casual cigarette smuggling. *National Tax Journal*, *61*(1), 7–33.

Merriman, D. (2010). The micro-geography of tax avoidance: Evidence from littered cigarette packs in chicago. *American Economic Journal: Economic Policy*, *2*(2), 61–84.

Moore, S. C. (2010). Substitution and complementarity in the face of alcohol-specific policy interventions. *Alcohol and Alcoholism*, *45*(5), 403–408.

National Institute on Drug Abuse. Are There Gender Differences in Tobacco Smoking? (2021). https://nida.nih.gov/publications/research-reports/. (Accessed: 2022-02-21)

Nicholson, A., Turner, T. M., & Alvarado, E. (2014). Cigarette Taxes and Cross-border Revenue Effects. *Public Finance Review*, *44*(3), 311–343.

Pesko, M. F., Courtemanche, C. J., & Maclean, J. C. (2020). The effects of traditional cigarette and e-cigarette tax rates on adult tobacco product use. *Journal of risk and uncertainty*, *60*(3), 229-258.

Stehr, M. (2005). Cigarette tax avoidance and evasion. *Journal of Health Economics*, *24*(2), 277–297.

U.S. Bureau of Labor Statistics. (2023a). *Consumer Price Index for All Urban Consumers: All Items in U.S. City Average, retrieved from FRED, Federal Reserve Bank of St. Louis.* https://fred.stlouisfed.org/series/CPIAUCSL. ([Accessed: 2023-11-30])

U.S. Bureau of Labor Statistics. (2023b). *Consumer Price Index for All Urban Consumers: Tobacco and Smoking Products in U.S. City Average, retrieved from FRED, Federal Reserve Bank of St. Louis.* https://fred.stlouisfed.org/series/CUSR0000SEGA. ([Accessed: 2023-11-30])

White, J. S., & Ross, H. (2015). Smokers' strategic responses to sin taxes: Evidence from panel data in Thailand. *Health economics*, *24*(2), 127–141.

Xu, X., Malarcher, A., O'Halloran, A., & Kruger, J. (2014). Does every us smoker bear the same cigarette tax? *Addiction*, *109*(10), 1741–1749.

Xuan, Z., Chaloupka, F. J., Blanchette, J. G., Nguyen, T. H., Heeren, T. C., Nelson, T. F., & Naimi, T. S. (2015). The relationship between alcohol taxes and binge drinking: evaluating new tax measures incorporating multiple tax and beverage types. *Addiction*, *110*(3), 441–450.