THREE ESSAYS ON REGIONAL ECONOMIC IMPLICATIONS FOR HEALTH BEHAVIORS

By

JAMES MCKOWN BISHOP

Bachelor of Science in Economics
Oklahoma State University
Stillwater, OK
2011

Bachelor of Science in Mathematics
Oklahoma State University
Stillwater, OK
2011

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2016

THREE ESSAYS ON REGIONAL ECONOMIC
IMPLICATIONS FOR HEALTH BEHAVIORS

Dissertation Approved:

John Winters

Dissertation Adviser

Dan Rickman

Mehtabul Azam

Jayson Lusk

Name: JAMES MCKOWN BISHOP

Date of Degree: MAY 2016

Title of Study: THREE ESSAYS ON REGIONAL ECONOMIC IMPLICATIONS FOR HEALTH BEHAVIORS

Major Field: ECONOMICS

Abstract:

My dissertation comprises three chapters. The first two chapters examine how cigarette smuggling affects smoking rates, while the third chapter examines how commuting times and work hours affect sleep and sleep-related behaviors.

The first chapter measures how the effect of a state cigarette tax increase on smoking participation depends on the consumer's opportunity to avoid the tax by purchasing cigarettes in the nearest lower-tax state. I show theoretically that a home tax increase should deter more smokers the higher the nearest lower tax. I confirm the theory using health survey data from 1999 to 2012 with a standard fixed effects model, taking advantage of county-level identification to measure the distance to the nearest lower-tax state. I also find that tax increases diminish in effectiveness as a tax rises. I observe that nearest lower taxes rose just enough over the period to maintain the effectiveness of a home tax increase at the yearly means.

The second chapter tackles the same research question as the first, but with a novel methodology. Instead of using the standard fixed effects model, I compare counties on the high-tax side of a state border to counties on the other side of the same border and to counties that are not near a border before and after tax increases. This invites an empirical framework known as triple differences, or DDD. Compared to those in the first chapter, estimates using the DDD methodology imply an even stronger effect of cigarette smuggling opportunities on smoking participation.

The third chapter examines the relationships between commuting, working, and sleep. I complement health survey data on sleep with county-level averages of hours worked, commute time, and the prevalence of early and late commutes. As expected, I show that working and commuting for longer hours and at stranger times negatively affects the sleep behavior of workers, but I estimate no significant effects on non-workers. I also find stronger effects for women.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTERACTING EFFECTS OF STATE CIGARETTE TAXES ON SMOKING PARTICIPATION

1. Introduction

From a public health perspective, smoking deterrence is vital. Anti-smoking policies saved an estimated 157 million years of life from 1964 to 2012 (Holford et al., 2014). From an economic perspective, smoking deterrence is socially beneficial because it limits negative externalities such as secondhand smoke and litter. There may also be "internalities" (Herrnstein et al., 1993; Gruber and Koszegi, 2001) associated with smoking, whereby consumers, especially adolescents, are fundamentally unable to assess the full present and future costs of consuming an unhealthy and addictive product. Cigarette taxes are a pervasive method of correcting for these market failures as well as raising public revenue. Every state in the U.S. has collected a cigarette tax since 1970, and there have been over 120 increases in state cigarette taxes in the 21st century.

A common political argument against state cigarette tax increases is that consumers will react by traveling to another state with a lower tax and buying cigarettes there. Many states have outlawed such "casual smuggling" if it exceeds one or two cartons per trip. Larger-scale organized smuggling also arises from interstate tax differences. The Contraband Cigarette Trafficking Act of 1978 made smuggling more than 50 cartons (equivalent to 500 packs, or 10,000 cigarettes) a federal offense with a penalty of up to five years in prison.

The existence of both casual and organized smuggling implies that a consumer's decision to smoke may depend not only on the price of cigarettes in his home state, but the price in another state and the cost of travel. Thus, depending on the attractiveness of border crossing, a tax increase in one state may impact the effectiveness of a tax increase in another state.[1] Measuring these interacting effects of state cigarette taxes on smoking participation is the main purpose of this paper.

I model the decision to smoke as depending on the tax in the home state, the distance to the nearest state with a lower tax, and the value of the nearest lower tax. I use smoking microdata from the Behavioral Risk Factor Surveillance System (BRFSS) and take advantage of abundant variation across states in the timing and magnitude of cigarette tax increases from 1999 to 2012. In line with my theoretical predictions, I find that the effectiveness of a home state cigarette tax increase rises with the nearest lower tax, but that, keeping the nearest lower tax constant, successive tax increases diminish in effectiveness. I estimate that these two effects on the effectiveness of a home state tax increase cancel out when the nearest lower tax rises at about two-thirds the rate of the home state tax, which is roughly what occurred at the yearly means over the sample period. Lastly, I show my findings to be robust to various alternative specifications and sample restrictions.

My findings imply that, because it would become a destination for smuggled cigarettes, a state may not be able to acheive its smoking deterrence goals simply by raising its own tax to some prohibitive threshold. On the other side, a state that is a source of smuggled cigarettes exerts a positive externality on proximal states when it raises its cigarette tax. Thus, the power to reduce smoking in the U.S. is especially in the hands of states with low cigarette taxes. Therefore, even if it is certain that cigarette taxes will continue to rise, where they rise will be an important determinant of how much the national smoking rate responds.

[1] In this paper, the "effectiveness" of a cigarette tax increase in either the home state or another state refers to the drop in smoking participation it causes in the home state.

2. Literature

While policymakers are interested in cigarette taxation for the purposes of both smoking deterrence and revenue generation, the literature has focused on the latter. A vast set of theoretical and empirical works, rooted in Kanbur and Keen (1993), model the behavior of revenue-maximizing governments that share a crossable border. Leal et al. (2010) provide a survey of this literature.

Models of border crossing rely on the intuitive hypothesis that, all else equal, living closer to a lower-tax jurisdiction increases the likelihood of cigarette smuggling. Testing this hypothesis is complicated by the irregularity of state borders and population distributions. Some authors simplify the problem by focusing on one tax increase in one state. For example, Emery et al. (2002) investigate the extent of tax avoidance in California, while Hanson and Sullivan (2009) measure tax incidence in Wisconsin. Stehr (2005) measures cigarette tax avoidance without explicitly considering border geography by comparing cigarette consumption data from the BRFSS to state sales data and interpreting the difference as coming from cross-border purchases.

Few individual-level surveys provide the geographical identification necessary to estimate models of cross-border cigarette purchases. The Tobacco Use Supplements to the Current Population Survey (TUS-CPS) is the most prominent such survey in the literature (Lovenheim, 2008; Chiou and Muehlegger, 2008; and DeCicca et al., 2013a; 2013b). A key advantage of the TUS-CPS is that it asks the respondent whether their most recent cigarette pack purchase was in a state other than their state of residence. A key limitation is that the finest identified level of geography for most observations in the TUS-CPS is the metropolitan statistical area (MSA), which is generally a group of one or more counties economically linked with an urban core.

Nationwide data with greater detail than the TUS-CPS is rare, but available. Harding et al., (2012) use Nielsen Homescan data from 2006 and 2007 to measure the effects of proximity to lower-tax states on cigarette tax incidence. For consumers across the United

States, the Nielsen dataset provides their census tract of residence and the zip codes of the stores at which they shop. This is a vast improvement in geographical identification over the TUS-CPS. In their main specification, the authors interact the log of the distance from the consumer's home state to the closest lower-tax state with the difference in cigarette taxes between them. In an alternative specification, they generalize the marginal effect of distance by using interactions of distance range dummy variables and the tax difference. They find that the consumer burden of a cigarette tax rises with distance to the border, especially for households with annual income greater than $30,000.

While many authors use microdata to estimate price-participation elasticities for smoking (Gallet and List, 2003), few account for cross-border purchases. The first to do so is Lovenheim (2008), who uses the TUS-CPS waves from 1992 to 2002. Lovenheim estimates statistically insignificant price-participation mean elasticities between -0.02 and -0.06, depending on the use of a year trend and accounting for the possibility of the nearest lower-tax jurisdiction being an Indian reservation rather than another state. However, he finds that doubling the distance to a lower-tax state reduces the price-participation elasticity by about 0.2. Thus, he provides evidence that those with a greater opportunity to buy cigarettes across the border are less responsive to tax increases. Reinforcing that casual smuggling is an extensive problem, he estimates that between 13% and 25% of smokers buy cigarettes across a border.

To allow the dependent variable to be interpreted as either total consumption or the smoking rate, the theoretical model in Lovenheim (2008) relies on parametric assumptions for the cigarette demand function and the share of smokers that smuggle. In particular, Lovenheim assumes that increasing the home tax and the nearest lower tax by the same percentage has no effect on the smuggling share, which allows for estimation of the smuggling share as a function of the distance to the nearest lower tax. In this paper, I show that, for smoking participation in particular, the signs of the first- and second-order effects of the home and nearest lower taxes can be derived without parametric assumptions.

Another paper that considers cross-border purchases, though not chiefly, is Callison and Kaestner (2014) (hereafter C&K). Using the TUS-CPS waves from 1995 to 2007 with state and survey wave fixed effects, they estimate that a one dollar increase in the cigarette tax changes the probability that a person smokes by -0.7 percentage points. They estimate this effect to be -1.0 percentage point but statistically insignificant when they restrict the sample to only the 19 states with the largest tax increases over the period, and to be -0.3 percentage points when they use a paired difference-in-difference approach on the restricted sample.

To account for border crossing, C&K interact the home state tax with three dummy variables for intervals of distance to the nearest lower-tax state. They find that the home state tax actually has a stronger negative impact on an individual's probability of smoking when his home MSA is within 120 miles of a lower tax state. C&K do not account for the magnitude of the nearest lower tax. This is problematic because MSAs with higher taxes have a larger range of potential values for the corresponding nearest lower tax. This means that for MSAs with higher taxes, the nearest lower tax will tend to be both larger and closer. Therefore, distance and the nearest lower tax are likely to be negatively correlated. Indeed, in the sample used in this paper, this correlation is -0.22. Thus, if the magnitude of the nearest lower tax does impact the decision to smoke, estimates using only distance will be biased.

To the best of my knowledge, this paper makes three novel contributions to the literature. First, I show that the signs of the first-order and second-order effects of the home tax, nearest lower tax, and distance to the nearest lower tax on smoking participation can be theoretically derived without parametric assumptions. Second, and most importantly, I empirically confirm the theory with the use of a larger, more recent, and more geographically detailed dataset than that used in previous papers. Third, I document a curious observation: due to the mean nearest lower tax rising at about two thirds the rate of the mean home tax over the sample period, the effectiveness of an increase in the home tax at the means of the variables has remained roughly constant over time. At the means in each sample year, my estimates imply that a $1 increase in the home tax deterred 0.8% to 0.9% of the adult population from smoking.

3. Data

I use smoking data from the BRFSS, an annual nationwide health survey. This is not the first paper to use the BRFSS to measure the responsiveness of smoking participation to cigarette taxes, examples being Sloan and Trogdon (2004) and DeCicca and McLeod (2008). However, I am not aware of any that takes advantage of the county-level geographical identification the BRFSS offers. The BRFSS asks respondents if they have smoked 100 cigarettes in their lifetime, and if so, whether they currently smoke every day, some days, or not at all. I define a smoker as a person who smokes every day.

I pool all BRFSS surveys from 1999 to 2012 and exclude respondents that are pregnant or at least 75 years of age.[2] The period from 1999 to 2012 comprises every year following the Tobacco Master Settlement Agreement (TMSA) for which counties are identified in the BRFSS. The TMSA, adopted in November of 1998, fundamentally altered the U.S. tobacco industry by requiring tobacco companies to cease a variety of marketing practices and to make annual payments to fund state smoking-related medical costs and an anti-smoking advocacy group. Therefore I find it appropriate to restrict the sample to the post-TMSA era.

Information on federal and state cigarette taxes comes from *The Tax Burden on Tobacco* (Orzechowski and Walker, 2014), which I organize into a monthly panel. Most cigarette tax changes came into effect on the first of the month. For those that did not, I record the tax change as coming into effect beginning with the following month. Cigarette taxes in the U.S. are levied as a dollar amount per pack of cigarettes. I convert all taxes to January 2015 dollars using the monthly CPI.

The period from 1999 to 2012 contains a wealth of variation in cigarette taxes over

[2] I choose the same age cutoff as C&K. Other papers such as DeCicca and McLeod (2008) restrict the sample to those younger than 65. In Table A1 I show that the results are slightly stronger when those 65 and older are dropped from the sample, consistent with the recent literature showing that older smokers are relatively unresponsive to cigarette tax increases (Ma, 2015; MacLean et al., 2015).

states and time. Every year in this period saw multiple states raise their cigarette taxes, and every state except Missouri and North Dakota raised their tax at least once.[3] The highest state cigarette taxes are concentrated in the Northeast, while the lowest are concentrated in the South. As of January 2015, the highest state cigarette tax in the U.S. was $4.35 per pack in New York and the lowest was $0.17 in Missouri. The federal cigarette tax rose from $0.24 to $0.34 on January 1, 2000, to $0.39 on January 1, 2002, and to $1.01 on April 1, 2009.

I also take into account local cigarette taxes in the five counties that compose New York City, in Cuyahoga County, OH, and in Cook County, IL.[4] A few cities in Cook County have their own cigarette taxes, including Chicago, which contains over half the population of Cook County. Because I am unable to identify the respondent beyond the county level, I use the mean cigarette tax in Cook County weighted by population. The only other localities in the sample with their own cigarette taxes are in Alabama, Missouri, and Virginia. Each of these three states has many different cigarette taxes at the county and/or city levels, for which I do not account. In the results section, I check that the estimates are robust to the exclusion of these three states.

I measure distances between counties using 2010 centers of population from the U.S. Census. I measure the distance from a county to a state as the distance from the county's center of population to the closest center of population of all counties in that state. I exclude Alaska and Hawaii, leaving 48 states and the District of Columbia. For each month, I exclude the state that has the lowest cigarette tax of all states, as the nearest lower tax and the distance to it are undefined.[5]

Table 1 presents the means of the variables of interest for each survey year. After

[3]Cigarette taxes only decreased twice in the period, in each instance by 10 cents. The first was the expiration in 2004 of a temporary 10 cent increase in Oregon. The second was in New Hampshire in 2011 and was repealed in 2013 as initally planned.

[4]County-level taxes introduce the possibility that the nearest lower tax may be in the same state. In my preferred specification, I require the nearest lower tax to be in a different state. In Table A1 I show that the results are insensitive to allowing a county in the same state to be the nearest lower-tax jurisdiction. This is reasonable especially because my treatment of NYC is unaffected; the five NYC counties are all closer to a New Jersey county than any non-NYC county in New York.

[5]In Table A1, I alternatively include the lowest-tax state in each month by defining that state's "nearest lower tax" to be equal to its own tax and the log of distance to be 0. The estimates are practically identical.

excluding observations with missing values, the sample contains 3,366,814 observations.[6] The sample period is characterized by an upward trend in real cigarette taxes and a downward trend in smoking. However, Table 1 leaves unclear the relative and interacting roles of home and cross-border taxes in reducing smoking. The remainder of this paper theoretically and empirically untangles the relationships between these variables.

4. Theoretical Model

The following model generalizes the framework of Saba et al. (1995), in which the representative consumer chooses to either pay a higher price for cigarettes in his home state or pay the cost of traveling to a state with a lower price. I allow for consumer heterogeneity and focus on a third option: to not smoke at all. Essentially, the consumer compares his reservation price for the first pack of cigarettes to the price in his home state and to some function of the price in the other state and the distance to that state.

The home state is denoted as state 1 and the consumer is a distance d_2 from state 2. p_1 and p_2 are the prices of a pack of cigarettes in the respective states with $p_1 > p_2$. Let $x(p; \vec{\theta})$ be the consumer's demand function for cigarettes, where $\vec{\theta}$ is a vector of preference parameters. Let $p(x; \vec{\theta})$ be the inverse demand function. Define $p^0 = p(0; \vec{\theta})$, which represents the consumer's reservation price for the first pack of cigarettes. The consumer enjoys a surplus of $\int_{p_1}^{p^0} x(p; \vec{\theta})\, dp$ from buying cigarettes in the home state. He prefers buying cigarettes in the home state to buying no cigarettes at all if and only if the surplus is positive, which is true if and only if $p^0 > p_1$.

The consumer enjoys a surplus of $W(p_2, p^0; \vec{\theta}) = \int_{p_2}^{p^0} x(p; \vec{\theta})\, dp$ from buying cigarettes in the other state, but to do so must pay a transport cost $C(d_2)$ that increases with distance. He prefers buying cigarettes in the other state to buying no cigarettes at all if and only if the surplus exceeds the transport cost. Therefore there exists a "critical distance" $d^*(p_2, p^0; \vec{\theta}) = C^{-1}(W(p_2, p^0; \vec{\theta}))$ such that the net surplus $W - C$ is zero. Therefore the

[6] Missing values come primarily from non-identified counties, which are especially prominent in the earlier years of the BRFSS. The percentage of the sample for which counties are not identified falls from about 25% in 1999 to 7% in 2012. As a robustness check, I restrict the sample to only those counties that are identified in every year, and show that the results are very similar.

consumer prefers buying cigarettes in the other state to buying no cigarettes at all if and only if $d^* > d_2$. There also exists a "critical price" $p^*(d_2, p^0; \vec{\theta}) = W^{-1}(C(d_2), p^0; \vec{\theta})$ such that the consumer prefers buying cigarettes in the other state to buying no cigarettes at all if and only if $p^* > p_2$.

Consider $\vec{\theta}$ to be a vector of random variables, so that p^0 and p^* are dependent random variables with joint probability density function (pdf) $f(p^0, p^*; d_2)$ for a given value of d_2. Assume f is continuously differentiable and note that f is non-negative. Let S equal one if the agent smokes and zero otherwise. The probability that the agent smokes is the probability that at least one of the two states offers him a net surplus from purchasing cigarettes there:

$$\begin{aligned} P(S=1) = P(p^0 > p_1 \cup p^* > p_2) &= 1 - P(p^0 < p_1 \cap p^* < p_2) \\ &= 1 - \int_0^{p_1} \int_0^{p_2} f(p^0, p^*; d_2)\ dp^*\ dp^0 \end{aligned} \tag{1}$$

Differentiating shows how the probability of smoking and the effectiveness of price increases change with the prices in the two states:

$$\frac{\partial P(S=1)}{\partial p_1} = -\int_0^{p_2} f(p_1, p^*; d_2)\ dp^* \leq 0 \tag{2}$$

$$\frac{\partial P(S=1)}{\partial p_2} = -\int_0^{p_1} f(p^0, p_2; d_2)\ dp^0 \leq 0 \tag{3}$$

$$\frac{\partial^2 P(S=1)}{\partial p_1 \partial p_2} = -f(p_1, p_2; d_2) \leq 0 \tag{4}$$

$$\frac{\partial^2 P(S=1)}{\partial {p_1}^2} = -\int_0^{p_2} \frac{\partial f}{\partial p_1}(p_1, p^*; d_2)\ dp^* \tag{5}$$

$$\frac{\partial^2 P(S=1)}{\partial {p_2}^2} = -\int_0^{p_1} \frac{\partial f}{\partial p_2}(p^0, p_2; d_2)\ dp^0 \tag{6}$$

Equations 2 and 3 quantify the drop in the probability of smoking due to increases in the home price and cross-border price, respectively. Equation 4 quantifies the effect of an increase in one state's price on the effectiveness of an increase in the other state's price.

The cross-partial derivative in Equation 4 is non-positive, implying that, if anything, an increase in the home state price is made more effective by a higher price in the other state. Intuitively, this is because a higher price in the other state makes it less likely that the consumer will switch to buying there in response to an increase in the home price, and thus makes it more likely that he will quit smoking instead.

The signs of the pure partial derivatives in Equations 5 and 6 can be determined under the conditions described by the following proposition:

Proposition 1. *Suppose, for some finite m and possibly infinite n, that $f(p)$ is a non-negative continuously differentiable function such that $\frac{\partial f}{\partial p}(p) \neq 0 \ \forall p \in (m,n)$ and $\lim_{p\to n} f(p) = 0$. Then $\frac{\partial f}{\partial p}(p) < 0 \ \forall p \in (m,n)$.*

Proof. Because $\lim_{p\to n} f(p) = 0$, there exists $y \in (m,n)$ such that $0 \leq f(y) \leq f(m)$. By the mean value theorem, there exists $a \in (m,y)$ such that $\frac{\partial f}{\partial p}(a) = \frac{f(y)-f(m)}{y-m} \leq 0$. Suppose that $\frac{\partial f}{\partial p}(b) \geq 0$ for some $b \in (m,n)$. By the intermediate value theorem, there exists c in the interval bounded by a and b such that $\frac{\partial f}{\partial p}(c) = 0$. But $c \in (m,n)$, a contradiction. Therefore $\frac{\partial f}{\partial p}(p) < 0$ for all $p \in (m,n)$. □

To sign Equation 5, note that $f(p_1, p^*; d_2)$ is a non-negative continuously differentiable function of p_1, and that, by nature of being a pdf, $lim_{p_1\to\infty} f(p_1, p^*; d_2) = 0$. Then, if $\frac{\partial f}{\partial p_1}(p_1, p^*; d_2)$ is not zero for all p_1 greater than some m, it must be negative for $p_1 > m$.[7] $\frac{\partial f}{\partial p_1}(p_1, p^*; d_2)$ being non-positive $\forall p^* \in (0, p_2)$ and negative on some subinterval means that both sides of Equation 5 are positive. Therefore, at a high enough home price, successive increases in the home price will diminish in effectiveness.

To sign Equation 6, note that $f(p^0, p_2; d_2)$ is a non-negative continuously differentiable function of p_2, and that, because the consumer is never willing to pay as much for cigarettes that require travel than ones that don't, $lim_{p_2\to p^0} f(p^0, p_2; d_2) = 0$. Then, if $\frac{\partial f}{\partial p_2}(p^0, p_2; d_2)$ is not zero over some interval bounded above by p^0, it must be negative on that interval. $\frac{\partial f}{\partial p_2}(p^0, p_2; d_2)$ being non-positive $\forall p^0 \in (0, p_1)$ and negative on some

[7]What if there is no such m? Then either f has zero density for all $p_1 > m$, or the derivative of f changes sign an infinite number of times. In the first case the derivative is zero. For an example of the second case, consider $f(p) = \frac{2sin^2(p)}{\pi p^2}$.

subinterval means that both sides of Equation 6 are positive. This means that, at a high enough price in the other state, successive increases in that price will cause smaller and smaller drops in smoking in the home state.

The above logic merits some discussion. Proposition 1 states the conditions for a point of diminishing effectiveness, past which a function, such as a pdf, must decrease to zero. Up to that point, the second derivative with respect to a tax cannot be determined. This is because there may be disproportionately dense regions, or "spikes", in the pdf around certain reservation prices. Suppose, for illustration, that there happens to be a large share of people that would quit smoking if and only if the home price exceeded \$6.50. Raising the price from \$5 to \$6 is not very effective, but raising the price from \$6 to \$7 is super effective. Then, from \$5 to \$7, the second derivative of the probability of smoking with respect to the tax is negative. Proposition 1 ensures that, unless f is pathological, spikes in f must eventually be exhausted as the price increases. Therefore, at a high enough price, the second derivative cannot be negative.

To derive the marginal effects of distance, recall that the probability of smoking can be stated in terms of the critical distance rather than the critical price:

$$\begin{aligned} P(S=1) = P(p^0 > p_1 \cup d^* > d_2) &= 1 - P(p^0 < p_1 \cap d^* < d_2) \\ &= 1 - \int_0^{p_1} \int_0^{d_2} g(p^0, d^*; p_2)\ dd^*\ dp^0 \end{aligned} \tag{7}$$

where $g(p^0, d^*; p_2)$, assumed to be continuously differentiable, is the pdf of p^0 and d^* for a given value of p_2. This makes it simple to show how the probability of smoking and the effectiveness of price increases change with distance:

$$\frac{\partial P(S=1)}{\partial d_2} = -\int_0^{p_1} g(p^0, d_2; p_2)\ dp^0 \leq 0 \tag{8}$$

$$\frac{\partial^2 P(S=1)}{\partial d_2 \partial p_1} = -g(p_1, d_2; p_2) \leq 0 \tag{9}$$

$$\frac{\partial^2 P(S=1)}{\partial d_2 \partial p_2} = -\int_0^{p_1} \frac{\partial g}{\partial p_2}(p^0, d_2; p_2)\ dp^0 \tag{10}$$

$$\frac{\partial^2 P(S=1)}{\partial {d_2}^2} = -\int_0^{p_1} \frac{\partial g}{\partial d_2}(p^0, d_2; p_2)\, dp^0 \tag{11}$$

Equation 8 quantifies the drop in the probability of smoking due to an increase in distance. Equation 9 quantifies the effect of an increase in distance on the effectiveness of an increase in the home price. The cross-partial derivative in Equation 9 is non-positive, implying that, if anything, an increase in the home state price is made more effective by a greater distance to the other state. Intuitively, this is because a greater distance to the other state makes it less likely that the consumer will switch to buying there in response to an increase in the home price, and thus makes it more likely that he will quit smoking instead.

To sign Equation 10, recall Proposition 1. Note that $g(p^0, d_2; p_2)$ is a non-negative continuously differentiable function of p_2, and that, because the consumer is never willing to travel to pay the maximum price he is willing to pay at home, $lim_{p_2 \to p^0} g(p^0, d_2; p_2) = 0$. Then, as long as $\frac{\partial g(p^0, d_2; p_2)}{\partial p_2}$ is not zero over some interval bounded above by p^0, it must be negative on that interval. $\frac{\partial g}{\partial p_2}(p^0, d_2; p_2)$ being non-positive $\forall p^0 \in (0, p_1)$ and negative on some subinterval means that both sides of Equation 10 are positive. Therefore, for a high enough price in the other state, an increase in that price has a smaller effect when it is farther away.

To sign Equation 11, note that $g(p^0, d_2; p_2)$ is a non-negative continuously differentiable function of d_2, and that, by nature of being a pdf, $lim_{d_2 \to \infty} g(p^0, d_2; p_2) = 0$. Then, as long as $\frac{\partial g}{\partial d_2}(p^0, d_2; p_2)$ is not zero for all d_2 greater than some m, it must be negative for $d_2 > m$. $\frac{\partial g}{\partial d_2}(p^0, d_2; p_2)$ being non-positive $\forall p^0 \in (0, p_1)$ and negative on some subinterval means that both sides of Equation 11 are positive. Therefore, for a far enough distance, successive increases in distance will have smaller and smaller effects on the probability of smoking.

To relate the entire discussion of prices to taxes, the only necessary assumption is that prices and taxes are positively related. As long as pass-through rates are entirely positive, all of the derivatives with respect to taxes have the same signs as the corresponding derivatives with respect to prices, even if pass-through rates vary across states or time. Thus the theory presented here guides an examination of the reduced form relationship between taxes and smoking participation. From a public policy perspective, this is not a disadvantage, since policymakers are generally unable to skip the intermediate causal steps between these two variables.

The reduced form subsumes the responses of both buyers and sellers of cigarettes, while theoretical model in this paper considers only the buyer's response. Sellers' responses should similarly result in higher smoking rates where there are smuggling opportunities. Other papers, such as Harding et al., (2012) and Hanson and Sullivan (2009), show that cigarette taxes have a lower pass-through rate near borders. Tax differences also create an incentive for sellers to cater to smugglers by locating just across the border from a higher-tax state. Even if smuggling is not occuring across a given border, the possibility of it may induce cigarette sellers on the high-tax side to decrease their prices to compete. Policymakers have taken note: when Arkansas raised its cigarette tax in 2009, it included a provision that allowed sellers close to a lower-tax state to alternatively apply that state's tax plus 3 cents.[8] All of these supply-side effects make cigarettes cheaper where there are tax differences, which enables smoking and thus supports this paper's hypotheses.

5. Methods

5.1. Econometric Model

Following DeCicca and McLeod (2008), I estimate linear probability models with two-way fixed effects. Initially, I regress an indicator for being a smoker on the cigarette tax in the respondent's county and a set of controls:

$$P(S_{icst} = 1) = \beta_0 T_{ct} + \vec{\alpha} \cdot \vec{X}_i + \zeta u_{st} + \delta_s + \gamma_t \tag{12}$$

[8] See http://taxfoundation.org/article/border-zone-cigarette-taxation-arkansass-novel-solution-border-shopping-problem.

S_{icst} is equal to one if and only if respondent i in county c of state s at survey month t smokes every day and is otherwise equal to zero. T_{ct} is the sum of federal, state, and county cigarette taxes in county c at the start of month t in January 2001 dollars, centered at the sample mean. $\vec{X}_i$ is a set of individual characteristics comprised of dummy variables for marital status, employment status, education, sex, race, one, two, or three or more children in the household, and categories for income and age.[9] u_{st} is the unemployment rate in state s and month t and δ_s and γ_t are state and month-year fixed effects, respectively. β_0, the estimated effect of raising the tax by a dollar on an individual's probability of smoking, is the coefficient of interest. I expect β_0 to be negative by the law of demand. Alternatively, I allow the marginal effect to vary with the value of the tax by adding the square of the tax on the right hand side.

To take cross-border purchases into account, I include the cigarette tax in the nearest state to the respondent's county with a lower cigarette tax, which I denote ${T'}_{ct}$, the log of the distance in miles to that state, which I denote d_{ct}, and all second-order terms involving these two variables and the home state tax. I center the nearest lower tax and log of distance at their sample means.[10]

$$\begin{aligned} P(S_{icst} = 1) = \beta_0 T_{ct} + \beta_1 {T'}_{ct} + \beta_2 d_{ct} + \beta_3 (T_{ct} \times {T'}_{ct}) + \beta_4 (T_{ct} \times d_{ct}) \\ + \beta_5 ({T'}_{ct} \times d_{ct}) + \beta_6 T_{ct}^2 + \beta_7 {T'}_{ct}^2 + \beta_8 d_{ct}^2 + \vec{\alpha} \cdot \vec{X}_i + \zeta u_{st} + \delta_s + \gamma_t \end{aligned} \tag{13}$$

By Equations 4 and 9, I expect $\beta_3, \beta_4 < 0$. By Equations 2, 3, and 8, I also expect that the marginal effects of taxes and distance are also non-positive at all reasonable values of T, T', and d:

$$\frac{\partial P(S_{icst} = 1)}{\partial T_{ct}} = \beta_0 + \beta_3 \times {T'}_{ct} + \beta_4 \times d_{ct} + 2\beta_6 \times T_{ct} \leq 0 \tag{14}$$

[9] I include respondents that refused to answer, and, in the case of income, didn't know or were not sure (DK/NS). As shown in Table A2, I assign dummy variables to refusal and DK/NS entries as if they were simply another possible response outcome. In Table A1 I show that excluding all respondents with at least one refusal or DK/NS entry (about 13% of the sample) does not affect the results.

[10] Because centering only subtracts a constant (the sample mean) from each variable, it has no impact on the estimated marginal effects or their standard errors. However, it does allow the coefficients on the first-order terms to be interpreted as the marginal effects at the means of the variables.

$$\frac{\partial P(S_{icst}=1)}{\partial T'_{ct}} = \beta_1 + \beta_3 \times T_{ct} + \beta_5 \times d_{ct} + 2\beta_7 \times T'_{ct} \leq 0 \tag{15}$$

$$\frac{\partial P(S_{icst}=1)}{\partial d_{ct}} = \beta_2 + \beta_4 \times T_{ct} + \beta_5 \times T'_{ct} + 2\beta_8 \times d_{ct} \leq 0 \tag{16}$$

While Proposition 1 predicts that $\beta_5, \beta_6, \beta_7$ and β_8 are positive for large enough values of T, T', and d, whether the values in the sample are "large enough" is unknown. A negative estimate does not contradict the theory, and a positive estimate raises the possibility, but does not prove, that the point of diminishing effectiveness has been reached.

5.2. Identification

The two-way fixed effects model is identified by variation in cigarette taxes across both states and months. That is, the model relies on the fact that states raised their taxes by different amounts and at different times. State fixed effects control for all time-invariant state characteristics, which is important because states have inherent differences that affect both smoking rates and tax policies.[11] Differences in cultural attitudes toward smoking, for example, may cause cigarette taxes and smoking rates to be negatively correlated independent of the effect of taxes on quantity demanded. However, time-varying characteristics such as economic conditions may also affect both smoking rates (Ruhm, 2005; Xu, 2013; Kenkel et al., 2014) and state tax policies (Maag and Merriman, 2003). In particular, a economic downturn at the state level may simultaneously reduce the demand for cigarettes and motivate the state government to use a cigarette tax increase to combat budgetary concerns. I control for the state unemployment rate in the survey month to address this issue. In Table A1, I show that the results are largely robust to not controlling for the unemployment rate.

While DeCicca and McLeod (2008) maintain that state tax increases were motivated by budget shortfalls that are uncorrelated with smoking rates, other plausible motivations are similarly compatible with the identification strategy. Golden et al. (2014)

[11] A natural issue with using fixed effects is multicollinearity with the variables of interest. To test the level of multicollinearity, I regress each of the variables of interest on the other two in addition to state and month fixed effects. For the home tax, nearest lower tax, and log of distance, this results in variance inflation factors of 8.5, 6.29, and 1.21, respectively. Values below 10 are commonly considered acceptable.

consider a vast suite of explanatory variables and find that cigarette tax increases are most strongly associated with the political alignment of state governments. After the TMSA, Republican-controlled governments were less likely to raise cigarette taxes than mixed-party or Democrat-controlled governments. As the political alignment in a given state over time is unlikely to be correlated with the smoking rate, this explanation also suggests that the model is appropriate.

Ultimately, though, I cannot rule out the possibility that one or more omitted variables are correlated with both state cigarette taxes and smoking rates. One step to dealing with this problem is to include state-specific linear time trends. This amounts to adding $\delta_s \times t$, the interaction between the state fixed effect and the month-year of the survey, to the right-hand side of Equation 13. Without this interaction term, I assume that unobserved differences across states are time-invariant, hence the state effects being "fixed." By including linear trends, I generalize the state effects to be a linear function of time.

6. Results

In the following tables of estimates, the estimated coefficients $\hat{\beta}_0$ through $\hat{\beta}_8$ are listed in order. Estimated coefficients on the controls for the preferred specification are provided in Table A2. I multiply all of the estimated coefficients and standard errors by 100 to condense the tables and provide an intuitive interpretation of the estimates in terms of percentage points. This is equivalent to using a dependent variable equal to 100 if the respondent smokes and zero otherwise. Consistent with the previous literature, standard errors are clustered at the state level and observations are weighted by BRFSS sample weights. The results are not sensitive to the use of weights.

6.1. Main Results

Table 2 reports the main results. The first column estimates Equation 12, which ignores border crossing. I estimate that each \$1 increase in the cigarette tax per pack decreases a person's probability of smoking by about 0.74 percentage points. Based on a recent CDC report that 13.7% of Americans smoke every day (Jamal et al., 2014), this translates to a 5.4% decrease in smoking participation per dollar. Based on a November 2014 average after-tax retail price of $5.84 for a pack of cigarettes (Orzechowski and Walker, 2014) and a pass-through rate of 1.11 (Keeler et al., 1996), this translates to a price-participation elasticity of -0.28, reasonably consistent with surveys of the literature (Chaloupka and Warner, 2000; Gallet and List, 2003).

Columns 2 and 3 respectively show that adding a quadratic term or controlling linearly for the costs of border crossing does not reveal any new conclusive results. Column 2 shows that, without accounting for border crossing, the effect of a home tax increase on the probability of smoking does not appear to vary much with the level of the tax. The quadratic term is small, negative, and statistically insignificant. Column 3 introduces the nearest lower tax and the log of the distance to that tax on the right hand side. The coefficient on the home tax is very similar to that in column 1. While they are not statistically significant at conventional levels, the estimated effects of the nearest lower tax and the log of distance are negative as predicted by Equations 3 and 8.

Column 4 reports estimates of Equation 13, my preferred specification. In contrast to columns 2 and 3, including a full set of interaction terms reveals significant implications of border crossing. As predicted by Equation 4, $\hat{\beta}_3$ is negative and significant at the 1% level, indicating that a tax increase is more effective the higher the tax in the lower-tax state. The point estimate implies that each dollar of the nearest lower tax causes a $1 increase in the home state tax to decrease the probability of smoking by an additional 0.58 percentage points. $\hat{\beta}_6$ is positive and significant, indicating, in contrast to the result in column 2, that each successive increase in the home state tax causes a smaller drop in smoking than the last. $\hat{\beta}_7$ is also positive and significant, indicating that each successive increase in the nearest lower tax also causes a smaller drop in smoking than the last. Thus I report two main findings: 1) cigarette tax increases in the home state and the nearest lower-tax state complement each other as deterrents to smoking in the home state, and 2) the deterrence achieved by an increase in only one tax or the other diminishes as the tax rises.

$\hat{\beta}_4$ and $\hat{\beta}_5$ are not statistically significant and $\hat{\beta}_8$ is weakly significant. Therefore I do not observe compelling evidence of second-order effects involving distance. For the remainder of the paper, I focus on the implications and robustness of the second-order effects that do not involve distance.

$\hat{\beta}_0$, $\hat{\beta}_1$, and $\hat{\beta}_2$ in the preferred specification are the estimated marginal effects at the means of the variables over all years of the sample. For example, $\hat{\beta}_0$ implies that, at the means of the variables, a \$1 increase in the home tax decreases the smoking rate by 0.887 percentage points, which implies a price-participation elasticity of -0.34. However, the means of the tax variables generally increased over time as shown in Table 1, and so the estimated marginal effects may have changed as well. In Table 3, I use the year-specific means from Table 1, my preferred estimates from column 4 of Table 2, and Equations 14-16 to estimate marginal effects of increases in the home tax, nearest lower tax, and log of distance for each year of the sample. The values are all negative, as expected.

Home tax increases were remarkably stable in mean effectiveness over the period, only falling slightly after the federal tax increase in 2009. At each year's mean values, an increase in the home tax has a strongly significant negative effect on the probability of smoking, and the implied elasticities all fall between -0.315 and -0.345. These values are in accordance with the general cigarette tax literature (Chaloupka and Warner, 2000; Gallet and List, 2003), but are larger than the statistically insignificant elasticities reported by Lovenheim (2008). Many methodological differences could account for the differences in estimated elasticities at the means. Lovenheim uses an earlier dataset, restricts the sample to residents of MSAs, aggregates the data at the MSA level, uses a linear time trend rather than fixed effects, and uses a partial set of second-order terms justified by his stronger theoretical assumptions. However, the primary message of this paper is that cross-border purchasing opportunities impair the effectiveness of home tax increases. From this broader perspective, my results agree with those of Lovenheim. In any case, more interesting than the magnitudes of the mean elasticities is their apparent stability over time.

All else equal, $\beta_6 > 0$ implies that home tax increases would have diminished in effectiveness as they rose from 1999 to 2012. However, this was counteracted by the fact that cross-border taxes rose as well. From the first column of Table 3 it appears that these two effects approximately canceled out. I am not aware of a theoretical reason why the mean effectiveness of an increase in the home state tax should have necessarily been so stable as state taxes rose over time. Whether such stability persists will depend on the geographic pattern of future tax increases. Generally speaking, if the U.S. becomes more of a checkerboard of high and low taxes, home taxes will become less effective. If, on the other hand, taxes converge across states, home tax increases will become more effective. To quantify this, recall Equation 14, assume distance does not change, difference both sides, and insert the preferred estimates. The result is:

$$\Delta\left(\frac{\partial P(S_{icst}=1)}{\partial T_{ct}}\right) = -0.577 \times \Delta T'_{ct} + 0.384 \times \Delta T_{ct} \tag{17}$$

Then the effectiveness of a home tax increase is unchanged if $\frac{\Delta T'_{ct}}{\Delta T_{ct}} = \frac{0.384}{0.577} = 0.666$. Therefore, the point estimates imply that successive home tax increases do not decrease in effectiveness as long as each increase in the home tax is accompanied by an increase in the nearest lower tax of two-thirds the size.

Like the home tax, point estimates of the marginal effects of the nearest lower tax and the log of distance are quite stable at the means from year to year. However, they are not as strongly significant. In particular, the nearest lower tax has a small and statistically insignificant effect at the means in each year. This suggests that, for many counties, taxes have become high enough that an increase in the nearest lower tax alone will not decrease the smoking rate. For two reasons, this does not imply that cross-border taxes do not matter. First, these are only mean effects. If the home tax is high enough above the mean and nearest lower tax is low enough below the mean, the effect of an increase in the nearest lower tax will be significant. Second, the main result of the paper still holds.

Even if an increase in the nearest lower tax does not by itself have a significant effect on the home state smoking rate, it does have a significant effect when it accompanies an increase in the home tax.

6.2. Robustness Checks

Table 4 reports the results of various modifications to Equation 13, while Table 5 reports the results of various sample restrictions. My preferred estimates from column 4 of Table 2 are copied into the first column of both tables for comparison.

Column 2 of Table 4 replaces the state fixed effects with county fixed effects. In using state fixed effects, I assume that all unobserved characteristics that are correlated with the left- and right-hand side variables do not vary below the state level. Thus my preferred specification is vulnerable to unobserved differences across areas within the same state. For example, the northern part of a state may have different cultural, political, or personal health attitudes than the southern part. In column 2, I account for such differences at the county level, the finest identified level of geography in the BRFSS. The estimates are very similar to the baseline, confirming that the results are not driven by unobserved time-invariant differences across counties.

Column 3 includes state-specific linear trends in the probability of smoking. If linear changes in state characteristics over time are correlated with both cigarette taxes and smoking rates, then adding the trend term would affect the estimates. However, column 3 is very similar to the baseline, supporting the validity of the fixed effects model.

Column 4 removes federal taxes from the home and nearest lower taxes. In a linear specification, this would have no effect on the estimates because the federal tax is the same for all states at a given point in time, and is thus absorbed by the month fixed effects. However, because the results rely on the coefficients on the second-order terms, they may be sensitive to the exclusion of federal taxes. Column 4 confirms that the estimates are very similar whether or not federal taxes are excluded.

Column 5 adds three binary controls for state bans on indoor smoking. If such bans decrease smoking rates (Evans et al., 1999) and state governments that are more likely to enact them are also more likely to enact tax increases, the preferred estimates may confound the effects of taxes with the effects of bans. To address this, I use data on the effective dates of smoking ban legislation from the American Nonsmokers' Rights Foundation.[12] 34 states and the District of Columbia enacted bans on smoking in restaurants, bars, and/or non-hospitality workplaces over the sample period. I control separately for the legality of smoking in each of these three types of venues. Column 5 shows that doing so has practically no effect on the estimates.

Column 6 expands the definition of a smoker to include those who smoke some days as opposed to only those who smoke every day. Roughly 5% of respondents reported smoking some days but not every day in each year of the sample.[13] Including some-day smokers checks that the results are not driven by people who transition from everyday smoking to some-day smoking in response to tax increases. If this was the case, then β_3, β_6, and β_7 would be smaller in absolute value in column 6 than in the baseline. I show that this is not the case, as these estimates are all larger in absolute value in column 6. Therefore the results appear to be driven by people who are deciding not to smoke at all, and who are thus fully capturing the intended individual benefits of cigarette taxes.

Table 5 reports the results of various sample restrictions. Column 2 excludes the West census region, which is characterized by large counties that may exacerbate the measurement error in the county-based distance calculations. The estimates in columns 2 are muted compared to the baseline, but are qualitatively similar and remain statistically significant.

Column 3 excludes the Northeast, which is characterized by small, high-tax states, so that the nearest lower tax may be less likely to represent a resident's smuggling incentive. Consider, for example, Massachusetts, Rhode Island, and Connecticut, which on January 1, 2014 had per-pack taxes of \$3.51, \$3.50, and \$3.40, respectively. None of these states

[12] The data is available at http://www.no-smoke.org/pdf/EffectivePopulationList.pdf. I do not control for the thousands of local-level indoor smoking bans in the U.S.

[13] The annual share of some-day smokers was highest in 2001 at 5.9% and lowest in 2010 at 4.7%, with no clear pattern over time.

is likely to be a source of smuggled cigarettes, but the model does not consider the possibility of a farther state where the tax is even lower being the source instead.[14] This problem is more likely to arise when states are small and easily traveled across, as in the Northeast. As shown in column 3, removing the Northeast from the sample gives slightly stronger estimates but no qualitative difference in the results.

Column 4 excludes Alabama, Missouri, and Virginia from the sample. These three states are characterized by many different cigarette taxes at the county and city level that I do not account for in this paper. Excluding them has practically no effect on the estimates.[15]

In column 5, I restrict the sample to those counties that are observed in every year from 1999 to 2012, which removes about 27% of the observations.[16] County identification in the BRFSS improved over the sample period, so that some counties are only observed in the later years. By removing those counties that are absent in one or more years, I check that the results are not affected by changes in the composition of the sample over time. Indeed, the estimates in column 5 are very similar to the baseline.

In sum, my results are robust to many alternative methods. The estimated coefficients on the square of the home tax, the square of the nearest lower tax, and their interaction, which form my main results, all remain statistically significant at conventional levels throughout. The estimates on the square of the nearest lower tax and the interaction term are especially robust, being significant at 1% in most specifications. Thus I provide evidence that border crossing can limit a state's ability to use taxes to deter its residents from smoking.

In the next two tables, I consider additional threats to the use of the nearest lower tax as the relevant out-of-state option. First, I consider that the nearest lower-tax county

[14] Table A4 includes the nearest tax that is lower than the nearest lower tax and the distance to that "even-lower tax" (ELT). This results in 20 terms involving the five variables of interest in the second-order approximation. Most of the newly introduced terms, including the interaction between the home tax and the ELT, do not have statistically significant coefficients and the original coefficients of interest are not significantly different from the baseline.

[15] The number of observations only falls by about 115,000. This is partly because Virginia was the lowest-tax state for the first 68 months of the sample and Missouri was the lowest-tax state for the last 30 months, so those states were already excluded in those periods, respectively.

[16] Each county-year with a positive number of observations has at least 21 observations.

may be sparsely populated such that cigarettes are not available for purchase there. To address this, columns 2 and 3 of Table 6 require the nearest lower-tax county to have at least 10,000 or 50,000 residents, respectively, as recorded by the 2010 Census. The results are very similar, and if anything are stronger with a stricter population cutoff.

Second, I consider the availability of untaxed cigarettes from Indian reservations. Lovenheim (2008) obtains similar estimates with and without accounting for Indian reservations. In an earlier working paper version (Lovenheim, 2007), he details the many difficulties he had to overcome to do so. In short, states vary widely in their enforcement of tribal sales of cigarettes to non-members, and tribal taxes and market power may undermine the assumption that tribes sell cigarettes at the non-tax price. To work around this, I use the most recent TUS-CPS, January 2011, to estimate the share of smokers from each state that bought their most recent pack of cigarettes from an Indian reservation. In column 4 I exclude the states for which that share is 10% or greater: Oklahoma, Arizona, New Mexico, Washington, Nevada, New York, Maryland, and Montana. In column 5, I set the cutoff at 5%, which further excludes South Dakota, Nebraska, Kansas, Wisconsin, Vermont, Idaho, and Wyoming. Though the estimates lose some precision, they remain qualitatively similar to the baseline. It does not appear that the results are significantly affected by purchases from Indian reservations.

Third, I conduct a placebo analysis in column 6 by restricting the nearest lower-tax state to be in a different census region. Finding significant coefficients on any of the terms in this regression involving the nearest lower tax or distance to it would be a signal that the model is not capturing something other than plausible smuggling options. As expected, the second-order effects disappear in column 6.

Fourth, I consider in Table 7 the possibility of cigarette smuggling from Mexico, as shown by Connelly et al. (2009).[17] Close proximity to the Mexican border may cause the U.S.-exclusive nearest lower tax and distance measures to overstate the costs of buying cigarettes from outside the state of residence. Column 2 excludes a liberally-defined set of

[17] Connelly et al. (2009) find no association between state cigarette sales and sharing a border with Canada. This is sensible as cigarette taxes in Canada are higher than those in the U.S.

southwestern states: California, Nevada, Arizona, Utah, New Mexico, Texas, Oklahoma, and Colorado. Column 3 excludes only the four states that share a border with Mexico: California, Arizona, New Mexico, and Texas. The estimates in columns 2 and 3 are muted compared to the baseline, but are qualitatively similar and remain statistically significant.

In columns 4 and 5, the sample is restricted to those states and counties, respectively, that share a border with Mexico. This removes a great deal of variation from the sample and therefore the estimates are imprecise. In the Mexican border state sample, the point estimates for β_3, β_6, and β_7 are much stronger and in the same direction as the baseline, but are not statistically significant. In the border county sample, the estimates flip sign, but the sample is not large enough to make any inferences. Overall, while cigarette smuggling from Mexico may occur, I find no evidence that it is a state-wide issue or that it is driving the results.

6.3. Results by Income Group

Differential effects by income groups is a long-studied topic in the cigarette tax literature. Most studies have found that lower-income individuals are more responsive to cigarette tax increases (Colman and Remler, 2008; DeCicca and McLeod, 2008; Farrelly et al., 2001; Townsend et al., 1994; Siahpush et al., 2009), though there is some contention (Franks et al., 2007). In Table 8 I consider that the previous results of this paper may vary by income.

There are two primary issues with using household income from the BRFSS. The first is that household income may not be an appropriate indicator of purchasing power or socioeconomic status for students or retirees. To alleviate this issue, I restrict the sample to respondents aged 25 to 64. The results of this sample restriction for respondents of all incomes is presented in column 1. The interacting and diminishing effects documented above for the full sample appear to be stronger for working age respondents.

The second issue with the BRFSS income data is that it is reported in categories which are the same in each survey year and therefore do not allow for inflation adjustment. With this in mind, I choose the income value among the survey-defined cutoffs that minimizes differences in the share of respondents above the cutoff from year to year. [18] In this way, an annual household income of $25,000 is the optimal cutoff. The share of respondents above this cutoff in a given year ranges from a high of 79.6% in 2008 to a low of 75.5% in 2011.

In columns 2 and 3, I use the baseline specification with the two income subsamples, once again restricting both samples to working age respondents. In line with the majority of the previous literature, it appears from comparing β_0 from the two regressions that low-income respondents are more responsive to home tax increases. However, none of the reported coefficients are significantly different.[19] I therefore do not find evidence that smuggling opportunities vary in their impact on smoking propensity by income.

7. Conclusion

The existence of both casual and organized cigarette smuggling in the United States suggests that a consumer's decision to smoke does not only depend on the tax in his own state. In particular, it implies an interacting effect between state taxes, such that a tax increase is more effective when smuggling is less attractive. Intuitively, the marginal home state cigarette consumer, when faced with a home state tax increase, will either switch to buying cigarettes from a lower-tax state or quit smoking. The higher the tax in the lower-tax state, the smaller the surplus from buying there, and thus the more likely he is to quit instead.

Using BRFSS data from 1999 to 2012, I estimate that a $1 increase in a state's cigarette tax reduces a resident's likelihood of smoking by an additional 0.58 percentage points for every dollar of the nearest lower tax. This is a large impact: at the sample

[18] In Table A3 I split the sample by BRFSS income categories. Though home tax increases appear to be most effective for respondents with incomes between $10,000 and $20,000, no clear patterns appear in the second-order coefficients.

[19] The z-score used to test the equality of coefficients β_{iA} and β_{iB} from regressions on separate samples A and B is $(\beta_{iA} - \beta_{iB}) / \sqrt{SE(\beta_{iA})^2 + SE(\beta_{iB})^2}$ (Clogg et al., 1995).

means, about a \$1.50 increase in the nearest lower tax doubles the effectiveness of an increase in the home state tax. Such an increase in the nearest lower tax is not unheard of; the mean nearest lower tax in the sample rose by \$1.23 (in January 2015 dollars) from 1999 to 2012.

I also find that successive tax increases, whether in the home state or the nearest lower-tax state, cause smaller and smaller drops in the home state smoking rate if the other tax does not change. I estimate that, especially after the federal cigarette tax increase in 2009, nearest lower taxes were high enough that an increase in the nearest lower tax alone would not decrease the mean survey respondent's probability of smoking. Home state tax increases, on the other hand, did not decrease in effectiveness over time at the yearly means. I estimate that maintaining the effectiveness of home state tax increases requires the nearest lower tax to rise at two-thirds the rate of the home tax.

My results suggest that one state's decision to raise its tax may make it worthwhile for a neighboring state to do the same. For example, state legislatures in Ohio and West Virginia have each considered \$1 cigarette tax increases in 2015. Based on my estimates, a \$1 tax increase in West Virginia would cause 0.58% of the adult population (or roughly 3.5% of everyday smokers) in eastern Ohio to respond to a \$1 Ohio tax increase by quitting smoking instead of buying cigarettes from West Virginia. Ohio would lose no revenue from this group's decision to quit, since the group's alternative to quitting would be to buy across the border. In addition, some Ohio smokers who were already buying cigarettes in West Virginia may switch to buying in Ohio if the West Virginia tax rises. Thus a tax increase in West Virginia would be a gain for Ohio in terms of both revenue generation and smoking deterrence. This example also highlights that a cigarette tax increase will have differential effects across counties within the same state. For a given home state tax increase, those counties near a state with very low taxes will not experience as great a drop in smoking as those for which border crossing is not as attractive an option.

Thus interstate politics are entangled in the major public health hazard that is smoking. While smoking rates have fallen over the past decades, industry influence, innovations such as electronic cigarettes, and strong preferences of those who still smoke mean the issue will not go away quietly. States that wish to further curb smoking through higher taxes are to some extent at the mercy of the states that surround them. Greater disincentives to smoking on the part of either the most smoking-friendly states or the federal government may be required for smoking rates to continue to fall.

CHAPTER II

CROSS-BORDER EFFECTS OF STATE CIGARETTE TAXES: A TRIPLE DIFFERENCE APPROACH

1. Introduction

Cigarette taxes vary widely across states in the U.S., creating opportunities both for arbitrageurs and price-conscious individuals who are willing to transport cigarettes across state lines. These acts of cigarette smuggling deny state governments the ability to use taxes to convince their residents to stop smoking. Measuring the extent of this problem is the aim of this paper. That is, I ask: how much does the deterrence effect of a cigarette tax increase depend on the lowest tax nearby? I find that nearby lower taxes significantly impair the ability of a state to use a cigarette tax increase to decrease its smoking rate.

Previous works have studied the extent of cigarette smuggling (Emery et al., 2002; Stehr, 2005; Chiou and Muehlegger, 2008), as well as its implications with respect to tax incidence (Harding et al., 2012; Hanson and Sullivan, 2009; DeCicca et al., 2013b), the optimal tax rate (Kanbur and Keen, 1993; Leal et al., 2010; DeCicca et al., 2013a) and smoking participation (Lovenheim, 2008; Callison and Kaestner, 2014; Bishop, 2016). This paper contributes to the literature on smoking participation by exploring a novel empirical methodology which takes care of multiple concerns with the standard methods. This methodology is commonly known as triple differences (DDD).

In topic, this paper is most similar to Bishop (2016), who uses a standard fixed effects model to show that a cigarette tax increase is more effective the higher the nearest lower tax. In methodology, it is most similar among the smoking literature to Callison and Kaestner (2014) and Kaestner and Callison (2015), who use a paired differences-in-differences (DiD) methodology to estimate cigarette demand. The paired DiD matches each state that increased its tax in a given period to a set of states that had similar smoking rates prior to the tax increase but did not increase their taxes over the same period. Using a standard fixed effects model, Callison and Kaestner (2014) estimate that a one dollar increase in the cigarette tax decreases the probability that a person smokes by 0.7 to 1.0 percentage points. This value is only 0.3 percentage points per dollar and statistically insignificant in the paired DiD model.

Whereas the paired DiD matches treatment and control based on the pre-treatment outcome, the DDD can be thought of as matching treatment and control based on geography. That is, I compare smoking participation before and after a tax increase in counties close to a lower-tax border, counties in the same state, and counties on the other side of the border. Previous papers have used a nationwide set of counties on opposite sides of state borders to study policies such as bank-branching deregulations (Huang, 2008) and minimum wages (Dube et al., 2010). However, to the best of my knowledge, cigarette taxation and the resulting smuggling incentives have not yet been studied in this way.

2. Data

Studying the topic of neighboring cigarette taxes requires geographically identified smoking data. While the Tobacco Use Supplements to the Current Population Survey (TUS-CPS) provide the greatest detail as to tobacco use behaviors, most observations in the TUS-CPS are identified at no finer than the metropolitan statistical area (MSA) level. MSAs are groups of economically and geographically linked counties. I instead use the Behavioral Risk Factor Surveillance System (BRFSS), which offers the county of residence for most respondents in the 2012 survey and in each year prior.[20]

[20] County of residence is identified for 87.7% of observations from 1999 to 2012.

While other papers have used the BRFSS to associate smoking participation and cigarette taxes (Sloan and Trogdon, 2004; DeCicca and McLeod, 2008), I am not aware of any besides Bishop (2016) that takes advantage of this county-level geographical identification. Each respondent in the BRFSS reports whether they currently smoke every day, some days, or not at all. I define a person to be a smoker if and only if they smoke every day.

I begin the sample with the year 1999, the first year following the Tobacco Master Settlement Agreement (TMSA), which spawned a new era of anti-tobacco programs and legislation. I exclude respondents that are pregnant or at least 75 years of age. The resulting sample contains 3,388,181 valid observations.

Like the previous literature, this paper relies on differences in the timing and magnitude of state cigarette tax increases. The post-TMSA era offers over 100 such increases and at least one in all but two states, with data coming from *The Tax Burden on Tobacco* (Orzechowski and Walker, 2014). Conveniently, cigarette taxes in the sample are universally levied as a dollar amount per pack of cigarettes. As of January 2015, state cigarette taxes ranged from $0.17 per pack in Missouri to $4.35 per pack in New York. The federal cigarette tax rose from $0.24 per pack to $1.01 per pack due to three separate increases over the period.

I also take into account local cigarette taxes in the five counties that compose New York City, in Cuyahoga County, OH, and in Cook County, IL. Chicago and a few other cities in Cook County have their own cigarette taxes. To account for these, I use the mean cigarette tax in Cook County weighted by population. I do not account for the many different, mostly small cigarette taxes at the county and/or city levels in Alabama, Missouri, and Virginia. No other cigarette taxes existed in the U.S. during the sample period.

3. Methods

The primary contribution of the this paper is to employ a DDD model to measure how the effectiveness of a cigarette tax increase depends on cross-border taxes. The standard empirical method of measuring the effects of cigarette taxes in general is to use two-way fixed effects (DeCicca and McLeod, 2008; Bishop, 2016). Econometrically, this resembles the following estimation equation:

$$P(S_{ist} = 1) = \beta_0 T_{st} + \vec{\alpha} \cdot \vec{X}_i + \zeta \vec{U}_{st} + \delta_s + \gamma_t \tag{18}$$

where S_{ist} is equal to one if and only if the respondent smokes, T_{st} is the tax in the respondent's state at the time of the survey, $\vec{X}_i$ and $\vec{U}_{st}$ are individual- and state-level controls, and δ_s and γ_t are state and time fixed effects.[21] Because cigarette taxes in the United States vary both over states and within states over time, the use of state and time fixed effects leaves ample identifying variation. This variation arises from differences in the timing and magnitude of cigarette tax increases across states. In other words, the standard method of identification relies on all states that don't increase their taxes in a given period as controls for states that do.

Bishop (2016) expands Equation 18 to include the distance to the nearest county with a lower tax, the value of the nearest lower tax, and all second-order terms involving these two variables and/or the local tax. His primary hypothesis is that the coefficient on the interaction between the local tax and the nearest lower tax is negative. The intuition behind the hypothesis is that a higher tax across the border makes smokers less likely to turn to smuggling in response to a cigarette tax increase, and thus more likely to quit instead.

The primary issue in testing this hypothesis concerns how to measure spatial variation in cigarette taxes while controlling for other spatial variation that could confound the estimates. The assumption in the standard model with state and time fixed effects is

[21] The time fixed effect is usually either at the year or the month-year level. I refer to the month-year fixed effect as a month fixed effect for short.

that any unobserved state-level variation that is correlated with the variables of interest is constant over time. The DDD methodology uses state-month fixed effects to instead absorb all of this unobserved variation. Variation remains due to differences in border proximity among counties in the same state.

3.1. Group Formation

This section describes how counties are assigned to groups based on border proximity. Distances from county population centroids to state borders come from Holmes (1996)[22] Counties within a specified proximity, or "distance cutoff" of the closest state border are assigned to the corresponding border group. Counties that are not within the distance cutoff of a state border are assigned to the "interior group" of their respective states.

I note a few observations on the grouping method. First, every county is assigned to a well-defined group. That is, no counties are left unassigned and groups do not overlap. Second, the county compositions of groups, like states, do not vary over time. Third, every group is comprised of counties from exactly one or two states, which classifies them as interior groups and border groups, respectively. Fourth, the groupings do not depend whatsoever on cigarette taxes, only geography, and thus could be generally applied in future research.

Each observation is assigned two values: the tax in the respondent's county ("local tax"), and the lowest tax among all the counties in the respondent's group ("group tax"). For residents of interior groups or those on the low-tax side of a border group, these two values are the same, except when the local tax includes a county-level tax. Those on the high-tax side of a border group have a group tax that is lower than their local tax. Each respondent is assigned the values of the two taxes on the first day of the survey month, adjusted to January 2015 dollars using the monthly CPI.

Figure 1 illustrates how counties are assigned to groups, using the example of Indiana and a 40-mile distance cutoff. The lightly-colored center region is the interior group, while

[22]The dataset is easily accessed at http://www.econ.umn.edu/ holmes/data/BorderData.html.

a distinct border group exists for each of the four neighboring states. Many counties are within 40 miles of more than one state border. For example, Steuben County is the most northeastern county in Indiana and borders both Michigan and Ohio, but its population centroid is closer to Michigan and therefore assigned to the Indiana-Michigan group. The group tax for a resident of Steuben County in a given month is the minimum of the taxes in Indiana and Michigan in that month.

I consider an alternative method of group formation, which is to use MSAs as groups. I use the 1999 MSA definitions from the U.S. Census Bureau.[23] 37 of these MSAs span multiple states, making them analagous to the border groups described above. The remaining 237 MSAs are entirely in one state and are analagous to the interior groups. The group tax is defined as the lowest tax among the counties in the respondent's MSA. Some counties are not part of an MSA, and so are removed from the sample, leaving 2,444,347 valid observations.

3.2. Econometric Models

The DDD model includes the interaction of the local tax and the group tax as well as state-group, state-month, and group-month fixed effects. The fixed effect interactions absorb practically all of the independent variation in local and group taxes but allows me to estimate the coefficient on the interaction of the two.[24] The following represents the DDD framework (Wooldridge, pp. 150-151):

$$P(S_{ist} = 1) = \beta_0 T_{st} * T'_{gt} + \vec{\alpha} \cdot \vec{X}_i + \delta_s * \zeta_g + \delta_s * \gamma_t + \zeta_g * \gamma_t \tag{19}$$

where T'_{gt} is the group tax and ζ_g is a border group fixed effect. The coefficient of interest, β_0, represents the increase in the effectiveness of a \$1 local cigarette tax increase due to a \$1 increase in the group tax, keeping the local tax constant.[25] My hypothesis is

[23] Accessed at https://www.census.gov/population/estimates/metro-city/99mfips.txt.

[24] There is still minor variation in local taxes despite the fixed effects interactions due to county- and city-level taxes, the same local-level taxes that are accounted for in Bishop (2016).

[25] For illustration, suppose the tax in county A is \$3.00 and the tax in a neighboring county B is \$1.00. Some share of

that β_0 is negative.

In Equation 20, the state-month and group-month interactions are replaced by the local tax and lowest in-group tax, respectively.[26] In this way, they are continuous variables rather than indicator variables, but represent variation over the same dimensions. This allows me to estimate the independent effects of the two tax measures. At the very least, this allows me to check that I obtain reasonable estimates for the marginal effects of a local or group tax increase, all else equal.

Replacing the fixed effect interactions with the continuous tax variables imposes a functional form assumption on the treatment effects of the two taxes. If no other terms are added to the regression, the assumption is that taxes have a linear effect on smoking outcomes. There are likely second-order effects of taxes (Bishop, 2016), so I also include squares of the tax variables on the right hand side. In the absence of state-month fixed effects, we also control for state-level characteristics that vary over time. The result is the following empirical model, which I term "pseudo-DDD":

$$\begin{aligned} P(S_{ist}=1) = \beta_0 T_{st} * T'_{gt} + \beta_1 T_{st} + \beta_2 T'_{gt} + \beta_3 T_{st}^2 + \beta_4 {T'}_{gt}^2 + \\ \vec{\alpha} \cdot \vec{X}_i + \zeta \vec{U}_{st} + \delta_s * \zeta_g + \gamma_t \end{aligned} \tag{20}$$

where $\vec{U}_{st}$ contains state-level controls for the unemployment rate and bans on smoking in workplaces, restaurants, and bars.

3.3. Identification

An assumption critical to the DDD methodology is that smokers only buy cigarettes within their own border group. If smokers buy outside their border group, then the outside price they enjoy is an important omitted variable. This is especially relevant for counties that are close to more than one border. A very similar issue arises with using

the population of county A would quit due to an increase in the tax from \$3.00 to \$4.00. I seek to measure how much larger that share would be if the neighboring tax was \$2.00 instead.

[26] A month fixed effect remains, as in the standard fixed effects model.

the nearest lower tax, which is standard in the literature (Lovenheim, 2008). If a much lower tax is slightly farther away, then the nearest lower tax is unlikely to represent the relevant cigarette smuggling incentives. Therefore, the border group assumption offers not necessarily a better or worse, but a different approach to modeling smuggling behavior compared to the previous literature.

One improvement due to the DDD methodology is that state-group fixed effects provide finer geographical controls than state fixed effects. State fixed effects imply the assumption that unobservable variables correlated with smoking behavior are the same over all regions of a state. For example, after controlling for observable characteristics, a standard fixed effects model would attribute higher smoking rates in southern Indiana to the lower cigarette tax in Kentucky. However, it may be that cultural differences are driving the difference in smoking rates. By including state-group fixed effects, I control for differences in unobservable characteristics between eastern, western, northern, southern, and central Indiana, as well as those counties on the other side of the corresponding borders.

DDD also provides a more concise treatment of the effect of distance. Including distance as a continuous regressor, as in Lovenheim (2008) and Bishop (2016), requires a parametric assumption and complicates the right-hand-side. DDD uses a distance cutoff to form the groups, then uses indicator variables for the group assignments. The task is to determine an appropriate value for the distance cutoff. Intuitively, it should be the maximum distance a person would be willing to travel to buy cigarettes in another jurisdiction. Lovenheim (2008) estimates that smuggling occurs up to a distance of 77 miles, which he notes is a large result compared to conjectures in the previous literature. Most notably, Lewit et al. (1981) and Lewit and Coate (1982) assume that all smuggling occurs within 20 miles, while Chaloupka (1991) assumes a 25-mile maximum distance.

A distance cutoff that is too low will assign counties with significant smuggling incentives to the interior group. This will bias the estimates towards zero since some of the "treatment" will be considered as the control. A distance cutoff that is too high will assign counties with insignificant smuggling incentives to border groups. This will also bias the estimates toward zero since some of the control will be considered as "treated." Either way, an improper choice of distance cutoff will attenuate the estimates because it will blur the distinction between treatment and control. In the results below, I consider a range of distance cutoffs from 20 miles to 100 miles. I also report results where I interact the local tax-group tax interaction with indicator variables representing distance from a border in 20 mile increments up to 100 miles. In this "distance tier" regression, the groups are formed with a 100-mile cutoff for the purpose of the fixed effects, but the coefficient of interest is allowed to vary based on how close a given county is to the closest border.

To get an idea of MSA size compared to border and interior groups, note that the median MSA covered 1637 square miles in according to the 2005-2009 American Community Survey. A circular MSA of this area would have a radius of about 23 miles, while the distance from the center of a square MSA of the same area would be 20 miles to one of its edges and 29 miles to a corner. The 90th percentile MSA in terms of area covered 5368 square miles, corresponding to a circle of radius 41 miles. So MSAs are generally small compared to the border groups I form using my chosen distance cutoffs.

4. Pre-Intervention Trends

The DDD model compares deviations from state mean smoking rates for counties on either side of a border before and after a tax. The identifying assumption is that deviations from state mean smoking rates have the same trends over time on each side of a border in the absence of a tax change. In Figure 2, I aggregate the deviations from state mean smoking rates for the three types of state-group-month cells: state-groups on the high-tax side of a border group, state-groups on the low-tax side of a border group, and interior groups. I use all of the observations in the first five years of the sample for which taxes did not change anywhere in the assigned group. Table 9 reports the number of these observations for each state-group type.

Table 9 and Figure 2 confirm the preliminary expectations for the data. First, residents on the high-tax side of a border (who thus enjoy the greatest smuggling opportunity) are more likely to smoke than other residents in the same state, by about 0.438 percentage points. Second, this is not the case for residents on the low-tax side of a border or not close to a border. Third, there does not appear to be a difference in trends among the three types prior to a tax change. That is, deviations from the state mean do not appear to systematically rise or fall over the period for any of the three types.

Figure 3 displays the differences in deviations from state mean smoking rates between high-tax and low-tax sides of each border group. Only months for which both sides of the border have not yet had a tax change and are observed in that month are shown. Breaking the sample in this way leads to greater variance in trends, but the main takeaway is that the differences across state-groups do not tend or either rise or fall over time prior to tax changes.

5. Results

The remainder of this paper reports and discusses the empirical results. In the following tables, I multiply all of the estimated coefficients and standard errors by 100 to condense the columns and provide an intuitive interpretation of the estimates in terms of percentage points. Standard errors are clustered at the group level and observations are weighted by BRFSS sample weights.

The first five columns of Table 10 display the results from estimating Equation 20, the pseudo-DDD, on the 1999-2012 BRFSS sample at border distance cutoffs in 20 mile intervals from 20 to 100 miles. The coefficient on the interaction of the local tax and group tax is negative in all five columns as expected, and significant at the 5% level for all distance cutoffs greater than 20 miles. From using distance cutoffs of 40 to 100 miles, I estimate that a $1 increase in the local cigarette tax, keeping the group tax constant, decreases an individual's probability of smoking by an additional 0.31 to 0.37 percentage points for each dollar of the group tax.

In the first five columns, the estimated effects of a local tax increase at the means are all significant at 5% and the point estimates are insensitive to the choice of distance cutoff. I estimate that a $1 increase in the local cigarette tax decreases the mean individual's probability of smoking by about 0.5 percentage points. This is somewhat lower than the estimated effect in Bishop (2016). At the means of the variables, the estimated effect of an increase in the lowest in-group tax, keeping the local tax constant, is negative but insignificant at conventional levels. This is consistent with Bishop (2016), who finds a similar result for the nearest lower tax. The squared terms are all positive, with much stronger second-order effects associated with the lowest in-group tax than the local tax. This also is consistent with Bishop (2016).

The sixth and final column of Table 10 displays the results of using MSA boundaries instead of border distances to form groups. The estimated second-order coefficients disagree with the other results and the intuitive hypothesis that lower nearby taxes make a tax increase less effective. This is possibly because MSAs are too small, such that cigarettes are smuggled not just across states within MSAs but across MSA borders. In addition, a resident on the edge of an MSA may be very close to another state that is outside the MSA but has a lower tax than the lowest tax inside the MSA. In this case, the lowest tax inside the MSA is likely to be greater than the relevant cost of smuggling, meaning that there will be positive measurement error in the group tax variable.

Table 11 displays the results from estimating Equation 19 on the same sample with the same grouping specifications as in the previous table.[27] The quadratic functions of the two tax measures are replaced by state-month and group-month interactions. The estimated coefficents from this "full DDD" model are much stronger than those for the pseudo-DDD. The estimate using the 20-mile cutoff is now negative and significant, though still smaller than the estimates with larger cutoffs. For distance cutoffs of 40 to 100 miles, the estimates are approximately twice as large. I estimate that a $1 increase in the local cigarette tax, keeping the group tax constant, decreases an individual's probability of

[27] The stricter fixed effect interactions exclude a small number of observations due to some states or groups being observed only in a single month.

smoking by about 1 percentage point for each dollar of the group tax.

Econometrically, it is logical that the more restrictive model gives stronger estimates. The pseudo-DDD model includes variation in local and group taxes over time, while the more restrictive, "true" DDD absorbs it. Since the tax measures are naturally positively correlated with their interaction and negatively correlated with the decision to smoke, it is plausible that misspecifying the functional forms of the taxes causes the pseudo-DDD model to underestimate the coefficient on the interaction term. This is evidenced by the absurd results that arise from applying the pseudo-DDD model to the MSA dataset. I therefore consider the estimates from the full DDD specification for border groups with distance cutoffs of 40 miles and greater to be the main results of this paper.

The results for the MSA grouping method are not so absurd with the full DDD model as with the pseudo-DDD. The estimated coefficient on the interaction of the local and group taxes is negative but not significant at conventional levels, and is smaller than those for border groups with distance cutoffs of at least 40 miles. This is consistent with the notion that the MSA estimates are upwardly biased due to measurement error in the tax variables, especially so in the pseudo-DDD.[28]

In Table 12, I form groups with a 100-mile distance cutoff and interact the interaction of the local and group taxes with indicator variables representing distance from the respondent's county to the nearest border in 20 mile increments. That is, I use the same set of fixed effect interactions as in column 5 of Table 11, but I allow the reported coefficient to vary with tiers of distance. Each reported coefficient is the estimated effect of the cross-border tax on the effectiveness of a tax increase for residents of the corresponding distance tier relative to residents that live more than 100 miles from the nearest border. The closer a respondent lives to a border, the stronger the expected effect of having a cheaper tax on the other side.

I estimate that there is a significant effect of nearby lower taxes up to, but not further

[28]In result not shown, I find very similar estimates for both the DDD and pseudo-DDD specifications at each distance cutoff when I restrict the sample to residents of MSAs. Therefore, it does not appear that the weak results using the MSA grouping method are driven by sample selection.

than a distance of 80 miles. However, the estimated effect between 60 and 80 miles away is no weaker than it is at any closer distance tier. While the point estimate dips between 20 and 40 miles, none of the coefficients up to 80 miles are significantly different from each other at the 10% level. However, the estimated effect at the 60-to-80-mile tier is significantly different from that at 80-to-100 miles at the 5% level. So while I do find evidence that cigarette smuggling occurs only within some distance to a border, it appears that it may not decay steadily with distance as assumed in Lovenheim (2008).

6. Conclusion

I employ a DDD methodology to measure the impact of smuggling incentives on the effectiveness of a cigarette tax increase. My results largely agree with previous work despite the difference in methodology. The most restrictive model gives especially strong evidence that proximity to a lower cigarette tax makes a cigarette tax increase less effective. A $1 tax increase on the high-tax side of a border deters about an additional 1% of that side's population from smoking when the tax on the low-tax side is $1 higher. In other words, cigarette smuggling props up smoking rates. There are people who smoke who would not smoke if they were forced to buy cigarettes from their state of residence. These people do not only live on state borders. My results suggest that cigarette smuggling is a significant problem up to a distance of 80 miles from a lower-tax state.

This paper joins the literature that points to cigarette smuggling as a public health hazard. It also points to the promise of a federal cigarette tax increase relative to the smatterings of state tax increases in the past years. Further, perhaps the more effective anti-smoking policies are those people seek rather than seek to avoid, such as quit-help and education programs. Unlike taxes, however, these programs generate no revenue. Thus state and local governments face a tradeoff between revenue collection and public health. A government that focuses on the latter requires the cooperation of its neighbors to maximize the effectiveness of its cigarette taxes.

CHAPTER III

COMMUTING, HOURS WORKED, AND SLEEP

1. Introduction

What are the relationships between time spent commuting, working, and sleeping? To explore this question, this paper combines health survey data on sleep behavior with commuting and working data at the county level. The results indicate that longer commutes, longer workweeks, and commutes that begin at more unusual times of the day raise the risk of short sleep and unintentionally falling asleep, with considerable differences between men and women.

My results are summarized as follows. Counties with residents that spend more time commuting and working are more sleep-deprived. So are counties with a greater share of residents that commute outside the range of 5 a.m. to 9 a.m. In accordance with intuition, these relationships appear to be driven by employed people. They also appear to be driven disproportionately by women.

A well-developed medical literature has established that short sleep is associated with a long list of health risks, including obesity, diabetes, hypertension, and cardiovascular disease (Buxton and Marcelli 2010). Despite this, the incidence of short sleep among full-time workers in the U.S. has increased over previous decades (Knutson et al. 2006; Luckhaupt et al. 2010). Knutson et al. also find that, after adjusting for socioeconomic factors, there has been no such significant change for those that don't work full-time. Fundamental economic theory states that consumption of sleep, like any other good, should fall as its cost rises. It appears that critical opportunity costs of sleep arise particularly from work and related activities such as commuting.

Based on the American Community Survey (ACS) from 2006 to 2011, over 95% of workers commute in some fashion, and the average commute in the United States is around 25.5 minutes (McKenzie 2014). According to the National Household Travel Survey, commuting accounted for 28% of all vehicle miles traveled in the U.S. in 2009 (Santos et al. 2011). There is evidence that increased commuting has negative consequences for mental health (Turcotte 2011; Roberts et al. 2011) and physical health (Sugiyama et al. 2013; Ding et al. 2014), though how much of this occurs through loss of sleep as opposed to other channels is unclear. Further, these health effects do not necessarily translate into effects on overall well-being (Dickerson et al. 2014).

The relationship between commuting and sleep as competing uses of time is especially important because of the dangers of driving while sleep-deprived. A decrease in driving performance due to a decrease in sleep time has been documented by laboratory experiments (Philip, et al. 2005) and occupational surveys (Scott, et al. 2007). MacLean et al. (2003) provide a review of the literature on sleep deprivation and driving safety. They note that sleepiness may be factor in up to 20% of motor vehicle accidents, and that prolonged wakefulness can have an effect on driving performance comparable to illegal levels of alcohol intoxication.

Independent of the amount of time devoted to sleep, driving ability has been shown to depend on time of day (Williamson et al. 2011). In particular, driving performance is impaired at early morning hours and in the early afternoon (Lenne et al. 1997), and crashes attributed to falling asleep occur primarily during these periods (Pack et al. 1995). Work schedules that move commute times to these periods can therefore be expected to increase the risk of motor vehicle accidents. "Strange hours" of work are especially prevalent in the U.S. relative to other countries, even after controlling for the number of hours worked (Hamermesh and Stancanelli 2014).

A commuter's poor driving performance is a negative externality to other drivers.

Accidents lead to both injury and traffic congestion for those that did not cause them. Urban sprawl leads to longer commutes, which leads to less sleep and more accidents. The trading away of sleep is therefore a public issue.

I estimate that 7 to 13 seconds of each additional minute spent working and 13-18 seconds of each additional minute spent commuting replace sleep. Keeping working and commuting time constant, commuting before 5 a.m. and after 9 a.m. each increase the likelihood of sleeping fewer than 6.5 hours per day. I also find that time spent commuting and working and the prevalence of these strange commute times each contribute to unintentionally falling asleep at some time during the day, while early commuting in particular increases the likelihood of falling asleep while driving. I find evidence that decreased sleep is not the only biological channel through which commuting duration and timing impact the likelihood of falling asleep. None of these effects appear for non-workers as opposed to the employed, supporting the validity of the results. Overall, most of the effects are stronger for women than for men, though the prevalence of early commutes is particularly associated with less sleep among men.

2. Previous Literature

Becker (1965) famously developed a theory of time use, with implications for theories of commuting (Nelson 1977) and sleep (Biddle and Hamermesh 1990). Intuitively, an agent maximizes utility under a time constraint, such as 24 hours in a day, in the same way he does under a budget constraint. In fact, these two types of constraints can be represented by one equation as long as the time spent working is a choice variable. Thus, the fundamental economic theory on consumption and income is directly analogous to activity and time endowment. From this point of view, authors have investigated how commuting times affect choices between work and leisure (Ross and Zenou 2008; Van Ommeren and Gutierrez-i-Puigarnau 2011). Whether sleep is considered "leisure" depends on the definition (Aguiar and Hurst 2007), but it nonetheless represents a vital, universal, and extensive use of time.

Regarding time tradeoffs, this paper is most similar to Christian (2012), who estimates that 28-35% of additional commuting time would have otherwise been spent sleeping. On the topic of early and late commutes, the most similar work is Basner et al. (2014), who show that starting work or educational training one hour earlier is associated with 20 fewer minutes of sleep. These and other studies of sleep time such as Knutson et al. (2006) and Basner et al. (2007) primarily use the American Time Use Survey (ATUS).[29] Studies of non-market activities such as sleep rely on the level of detail unique to time use surveys such as the ATUS (Aguiar and Hurst 2007). In particular, the ATUS is the largest sample that contains individual-level observations of both sleeping and commuting time. A subsample of the Current Population Survey has participated in the ATUS each year since 2003. Participants are asked to report what they did each minute for 24 hours starting at 4 a.m. of the day prior to the survey. Each activity is recorded in one of 431 categories, including working, traveling to and from work, and sleeping. However, there are multiple difficulties associated with using the ATUS or other time use surveys.

Most importantly, time use surveys have small samples relative to other nationwide surveys, and thus allow only coarse regional distinctions. Aguiar et al. (2013) merge state-level unemployment data with the ATUS to measure how work hours lost during the Great Recession were replaced by a comprehensive list of activities, including sleep. Antillon et al. (2014) focus specifically on the relationship between state unemployment and ATUS-observed sleep time. Brochu et al. (2012) perform a similar analysis with Canadian time use data and unemployment at the province level. Each paper provides evidence that sleep time is countercyclical, but each is restricted by the nature of time use surveys to a highly aggregated regional measure of business cycles. In this paper, I match county-level variables to individually reported sleep durations. While commuting and work variables are the focus of this paper, I am also able to include unemployment rates among other controls at the county level.

[29]Time use surveys from other nations have been used, but none are as common as the ATUS. For example, Chatzitheochari and Arber (2009) use the 2000 United Kingdom Time Use Survey to examine relationships between working time, shift work, and sleeping time, and Szalontai (2006) uses a time use survey from South Africa to show that sleeping time has an inverse relationship with economic opportunities.

A problem with the ATUS specific to the topic of sleep concerns the start and end times of the diary day. Because the time diary begins and ends at 4 a.m., few continuous sleep durations are actually observed.[30] The number of minutes of sleep observed over the diary is usually the sum of time slept since 4 a.m. in the morning and time slept from some bedtime until 4 a.m. the following night. In other words, most diary days are bookended by separate periods of sleep. If individuals have very regular sleep schedules, then this sum is representative of daily sleep time, but this is a strong assumption to make.

Another issue is that the variables of interest may be endogenous. This is true for commuting variables if those with a weaker preference for sleep or sleep regularity sort into areas or occupations with longer or more irregular commutes. Similarly, light sleepers may disproportionately choose jobs with longer hours. Since preference for sleep cannot be observed in the cross-sectional data, the results may be biased.

Since few other surveys ask about both driving and sleeping time, these issues have remained in the literature. One exception is Ding et al. (2014), who use a health survey conducted in New South Wales, Australia to examine relationships between driving time and health behaviors, including sleep. Their analysis is limited to a binary variable for insufficient sleep (defined to be fewer than seven hours) and four intervals (0-30 minutes, 31-60 minutes, 61-120 minutes, and 120 minutes or more) for the answer to the question "About how many hours in each 24 hour day do you usually spend driving?" Monaco et al. (2005) use a survey of truck drivers to show that longer hours of driving are associated with decreased sleep time and increased risk of falling asleep while driving. By nature, their sample is restricted to a special case where driving and working are one and the same.

The Behavioral Risk Factor Surveillance System (BRFSS) in the United States offers suitable information on sleep for the general population, but lacks information on commuting and working. This paper introduces the solution of complementing the BRFSS

[30] This is identically the case for the time use subsample of the Canadian General Social Survey and the 2000 UK Time Use Survey.

data with county-level data from the ACS. This novel approach provides many advantages over time use data. First, it removes bias due to within-county sorting based on individual preferences. Second, it provides sleep-related outcomes besides time spent sleeping. Third, it allows for county-level controls. Fourth, it provides a larger sample. Aggregation does not introduce any bias to the results, but has the disadvantage of decreased precision. This is counteracted by a larger sample size and does not prevent me from obtaining statistically significant results.

3. Data

3.1. BRFSS individual responses

The BRFSS is an annual nationwide telephone survey overseen by the Centers for Disease Control and Prevention. The BRFSS sleep data come from the optional "inadequate sleep" survey module that was conducted from 2009 to 2012. BRFSS optional modules are sets of questions on specific topics that states (and Washington D.C. and Puerto Rico) may choose to include in their surveys in a given year. Table 13 summarizes regional participation from 2009 to 2012, unlisted states having not participated.[31] No state participated in the module for all four years, and most only participated for one year. New York, California, Pennsylvania, Florida, and Texas, and thus some of the longest mean commutes in the U.S., are absent from the sample.

I consider four questions from the sleep module. The first asks "On average, how many hours of sleep do you get in a 24-hour period? Think about the time you actually spend sleeping or napping, not just the amount of sleep you think you should get."[32] The distribution of responses is presented in Figure 4. For this question, the respondent is not instructed to give a whole number of hours, but the survey proctor is instructed to round the response to the nearest hour. Still, reported sleep time compares reasonably

[31] Puerto Rico participated in the module in 2012, with 1777 employed respondents, but they are not included in the sample for this paper.

[32] This question became a core survey question in 2013. However, 2013 is the first year of the BRFSS for which sub-state geography is not identified for the sake of confidentiality.

with a normal distribution. 36.1% were recorded as sleeping 6 or fewer hours per day. Due to rounding, this is equivalent to the share of respondents that reported sleeping fewer than 6.5 hours per day, which I define as "short sleep."

The second question asks "During the past 30 days, for about how many days did you find yourself unintentionally falling asleep during the day?" For employees, 71.3% of the responses are zeroes and only 8.8% of responses exceed five days. The lack of variation in positive responses diminishes the potential value of a count model for this variable. I use a binary variable equal to one if the reported number of days is greater than zero.

A third question from the module asks "During the past 30 days, have you ever nodded off or fallen asleep, even just for a brief moment, while driving?" 56,962 employed respondents responded either yes or no, with 2,153 of these respondents, or about 3.9%, admitting to falling asleep while driving in the past 30 days.[33] This statistic should be interpreted with caution. Drivers that are judged to have fallen asleep just prior to a crash often deny having done so, and laboratory studies show that most people who fall asleep for no longer than a few minutes honestly fail to acknowledge it (Horne and Reyner, 1999). Not only should 3.9% be interpreted as a lower bound, but relatively weak estimated effects on this variable should not be taken as a lack of an effect on road safety, especially with contrasting evidence for unintentionally falling asleep some time during the day.

The last question I consider asks "During the past 30 days, for about how many days have you felt you did not get enough rest or sleep?" Besides being part of the sleep module in 2011 and 2012, this question was also part of the core survey in 2008, 2009, and 2010. These three years of the BRFSS offer about ten times the sample size compared to the four years of the sleep module. 26.8% of employed respondents reported getting enough sleep every night in the past month. As with my measure of falling asleep during the day, I use a binary variable equal to one if the reported number of days is greater

[33] This is after removing from the sample the 584 respondents who reported not having a driving license or otherwise not being a driver.

than zero.[34]

The advantage of larger and geographically broader sample size is diminished by the subjective nature of this question. This question relies on the respondent's definition of "enough sleep." Thus, of the four questions described, it is likely the least informative about sleep and sleep-related safety.

3.2. ACS county-level variables

The BRFSS does not provide data on commuting. To solve this issue, I use county-level means from the 5-year American Community Survey sample from 2008-2012. Each county is assigned the mean commute time among residents of the county that are age 16 or older and work somewhere other than at home. Assuming that a day of commuting involves one trip to work and another equally long trip back home, the commute time measure is doubled to represent daily travel time.

I also obtain county-level population sixteen years and older, population density, labor force participation, and unemployment from the same sample.[35] These county-level controls are a luxury not offered by the ATUS. In addition to mean commuting time, I include mean weekly hours worked and two measures of "strange hours" as variables of interest. For strange hours, I use the percentage of commuters that depart home after or at midnight and before 5 a.m. ("early commuters"), and the percentage that depart after or at 9 a.m. and before midnight ("late commuters").

Table 14 summarizes the ACS data separately for the two BRFSS samples for employees and non-workers. The employee and non-worker subsamples are conveniently similar in size. Non-workers tend to live in less populated areas with lower labor force participation and higher unemployment rates, but the differences are small. The sleep module and core sample means are practically the same; though it excludes some of the largest cities in the U.S., the sleep module sample is reasonably representative of the nation.

[34]Though in general one could consider a count model for an answer in this format, there is a strong response bias toward multiples of five days, suggesting that this would be inappropriate.

[35]Weekly hours worked are averaged over all workers, non-workers excluded.

4. Methodology

4.1. Specifications

I regress the BRFSS sleep question responses on the ACS county-level variables and a set of individual level controls that may influence sleep behavior. Specifically, I include dummy variables for the month of the survey, marital status, education, sex, race, the presence of one, two, or three or more children in the household, and intervals for income and age. Respondents to the BRFSS are at least 18 years old. I make no further age restrictions. Each observation in the sample is given the same weight.

Each of the four variables of interest (commute time, share of early commuters, share of late commuters, and hours worked) are expected to negatively affect sleeping time and increase the likelihood and frequency of unintentionally falling asleep. Commuting time and weekly work hours represent opportunity costs of sleep, while strange departure times interfere with the body's circadian rhythm. There is a large medical literature on the biological reasons for sleep difficulty as a result of irregular work hours, such that the existence of "Shift Work Sleep Disorder" is well-established and associated a wide range of health risks (Schwartz and Roth 2006).

I restrict the sample to workers that are not self-employed.[36] As a falsification test, I perform the same regressions for the subsample of non-workers. Non-workers in the sample are primarily retirees, but also include those out of work, homemakers, and students.[37] The expectation is that county-level commuting variables have weak, if any, effects on sleep time for non-workers relative to employees. The effects may conceivably be non-zero due to longer non-commuting travel time because of increased traffic congestion, or because sleep patterns formed while working may persist. Still, the effects for employees should be larger in absolute value.

[36] The self-employed comprise about 8.5% of all respondents for both the three-year core survey sample and the sleep module sample. Self-employment represents a unique situation with regard to commuting and work scheduling. The results are insensitive to including the self-employed in the sample of workers.

[37] The BRFSS only distinguishes between individuals that have been "out of work" for greater than and less than one year, without regard as to whether the individual is unemployed or has exited the labor force.

I report results for employees with and without the inclusion of state fixed effects. Including state fixed effects eliminates variation in the variables of interest across states. This is valuable to the extent that state fixed effects are picking up on inter-state differences in sleep preferences, but it is counterproductive to the extent that they are picking up on differences in inter-state means of the variables of interest. Based on this, the results without state fixed effects are tentatively preferred.

There is a well-developed literature on the different responses to commute times for men and women (White 1986). Authors have found evidence of sex-specific effects for such variables as labor force participation (Black et al. 2014) and psychological health (Roberts et al. 2011). Recognizing this, I perform the same regressions as described above on male and female subsamples with sex-specific versions of the variables of interest.[38]

4.2. Econometric Concerns

The estimation is vulnerable to bias because individual preferences for sleep are unobservable. A greater preference for sleep is positively correlated with sleep time and negatively correlated with those variables that interfere with sleep, including all of the variables of interest. This will downwardly bias the already negative estimates for sleep time and upwardly bias the already positive estimates related to lack of sleep. The use of county-level rather than individually observed variables reduces this bias, because the county-level variables are unrelated to any choice of workplace location, method of commuting, or any other determinant of the variables of interest at the individual level, except those that involve a change in the county of residence.

The bias is not wholly eliminated because individuals may still sort into counties based on their preferences for sleep. For example, more career-motivated individuals may sort into larger cities (Rosenthal and Strange 2008), and simultaneously be more willing to sacrifice sleep. Alternatively, people who are more vulnerable to sleepiness may locate where they can be safer by spending less time driving. There are many more imaginable

[38] Using sex-specific county means is necessary to avoid bias from aggregation.

examples, but they all share the effect of exaggerating the estimates.

The drawback to using county-level means is measurement error. Because this measurement error is not correlated with an individual's actual unobserved commute time, it does not introduce any bias into estimation of the effects of the county-level variables. Only the precision of the estimates suffers, which is remedied by the larger sample size relative to previous studies. There is also measurement error in the sleep time regression due to rounding of the dependent variable to the nearest hour, but this also does not bias the coefficients (Schneeweis et al. 2010).

Another concern is that the county-level variables may be highly collinear. To examine this possibility, Table 15 provides correlations between all of the county-level variables, weighted by the number of observations per county in the sleep module sample.[39] More populated counties, as expected, tend to be more population-dense, have longer average commutes, have a greater percentage of late commutes, and have a smaller percentage of early commutes. Longer weekly hours are correlated with a larger percentage of early commuters and a smaller percentage of late commuters. The latter correlation of -0.533 is the only one in the table above 0.5 in absolute value, indicating that collinearity among the county-level variables is not a major concern.

5. Results

The remainder of this paper presents and discusses the results. In all tables, standard errors are in parentheses and clustered by county. One, two, and three stars show statistical significance at p-values of .10, .05, and .01, respectively. All logit coefficients are reported as odds ratios for ease of interpretation.

[39]The correlations are similar in the core sample.

5.1. Sleep Time

Panel A of Table 16 reports OLS estimates with time spent sleeping per 24 hours as the dependent variable.[40] The first column reports estimated effects for employed people, the second includes state fixed effects, and the last two columns do the same for the sample of non-workers instead of the employed. Below the four estimated coefficients in each column, I test their joint significance.[41] As expected, none of the estimates for non-workers are individually or jointly statistically significant at conventional levels. In addition, I report χ^2 tests of the equality of the four coefficients across the worker and non-worker sample. I reject the null hypothesis that the coefficients are the same across samples both with and without state fixed effects. In sum, all of the results for non-workers support the claim that estimates for the employee sample are measuring the actual effects of the variables of interest as opposed to some omitted correlates. With this in hand, I turn to the estimates for the employee sample.

As expected, the estimated coefficients for all variables of interest are negative. Coefficients on daily commute time, percentage of late commutes, and weekly hours are statistically significant at the 1%, 5%, and 10% levels, respectively, and this is robust to the inclusion of state fixed effects. The coefficient on daily commute time of -0.304 implies that 30% of extra time devoted to commuting would have otherwise been devoted to sleep. The results imply an analogous percentage for time devoted to work of about 11%.[42] These percentages are significantly different at the 10% level. However, including state fixed effects doubles the coefficient on work hours and reduces the coefficient on commute time, causing the percentages to be practically equal. This is not enough evidence to infer whether sleep substitution differs for time spent commuting versus working. The inclusion of state fixed effects has otherwise little consequence.

A one percentage point increase in late commutes is associated with 27.4 fewer seconds

[40] Count models (treating the number of sleep hours as discrete) give analogous results. OLS results are preferred for interpretation.

[41] Under the null hypothesis of no effects, the test statistic for OLS is distributed as F(4, c - 1), where c is the number of county-level clusters. For the employee sample, c = 624. For the non-worker sample, c = 622.

[42] This is the coefficient on working hours per week times the number of days in a week divided by the number of minutes in an hour: $-0.923 * 7/60 \approx -10.8\%$.

of sleep (45.7% of a minute), keeping commuting and working time constant. This implies a disruptive effect of irregular working hours on sleep time as opposed to a time tradeoff. Because this variable cannot distinguish between late morning, afternoon, and evening commuters, it is unclear if the results are being driven by one of these types or a combination. The estimated coefficient for early commuters is negative and larger in absolute value than that for the percentage of late commutes, but it is not statistically significant at conventional levels.

Panel B of Table 16 reports results of logit regressions using a binary sleep time measure as the dependent variable. The dependent variable is equal to one if the respondent reported sleeping 6.5 or fewer hours per day. For employees, estimates for all variables of interest are greater than one, as expected, and significant at the 5% level. A one-minute increase in daily commute time increases the likelihood of short sleep by 0.6%. A one-hour increase in mean weekly hours worked increases the likelihood of short sleep by 2.7%. One percentage point increases in early and late commutes are associated with 3.6% and 1.1% increases in the likelihood of short sleep, respectively. Reassuringly, none of the estimates are individually or jointly statistically significant at conventional levels for non-workers.[43] Further, the estimates are significantly different between workers and non-workers both with and without state fixed effects.

The estimates for commuting and working time are not significant when state fixed effects are included. However, the estimated effect of early commutes, the only variable of interest that did not show a significant effect on sleep time in Panel A, remains positive and significant with respect to short sleep. Table 16 as a whole thus shows that all four variables of interest play a role in decreasing the amount of sleep Americans consume.

[43]Under the null hypothesis of no effects of any of the four variables of interest, the test statistic for logit estimation is distributed as χ^2 with four degrees of freedom.

5.2. Falling Asleep

Panel A of Table 17 reports results of logit regressions using a dependent variable equal to one if the respondent reported unintentionally falling asleep in the past 30 days.[44] The first column reports the baseline results for employees and the second, third and fourth columns experiment with additional controls. The last two columns report results for non-workers. Commute time, working time, and early commuting all increase the likelihood of unintentionally falling asleep for employed people. These estimates are largely robust to the inclusion of state fixed effects and are not statistically significant at conventional levels for non-workers. Including state fixed effects reduces precision in the estimation for hours worked but does little to affect the point estimate. Late commuting is only estimated to have a significant effect with the inclusion of state fixed effects, for both the worker and non-worker samples. The evidence that commuting after 9 a.m. affects the likelihood of unintentionally falling asleep is therefore unreliable.

In the third and fourth columns, I include the individual's reported sleep time as a control. The fourth column shows that state fixed effects have a similar effect on the estimates with and without the sleep time control. The estimates of the variables of interest are robust to the inclusion of sleep time. This means that, though commuting and working reduce sleep time as evidenced in Table 16, there appear to be one or more other channels through which commuting and working raise the likelihood of unintentionally falling asleep. The simplest explanation is fatigue, such that, among two similar people who sleep the same amount each night, the one who spends more time working and commuting has greater difficulty staying awake.

Panel B of Table 17 reports results of logit regressions using a dependent variable equal to one if the respondent reported falling asleep while driving in the past 30 days. For employees, a one-minute increase in daily commute time is associated with a 1.7% increase in the likelihood of falling asleep while driving. A one percentage point increase

[44] In results unreported, I experimented with a zero-inflated negative binomial model, which supplements the logit model with a count model for the number of days reported. Other than commute time acting through the count effect rather than the probability of a nonzero response, the results are similar.

in early commutes is associated with a 4.3% increase in the likelihood of falling asleep while driving in the past month. The percentage of late commutes and weekly work hours do not have a statisitically significant effect.

The result for early commutes is consistent with the previous literature on driving performance by time of day (Pack et al. 1995; Lenne et al. 1997). The result for early commutes is also robust to controlling for sleep time, suggesting that this effect is working through some channel other than simply lack of sleep, such as circadian disturbance. The lack of a result for late commutes is not too surprising given the 15-hour width of the interval.

Overall, the results in Panel B are weaker than those in Panel A. Once again, none of the estimates are statistically significant at conventional levels for non-workers and the estimates are significantly different across samples. However, the four coefficients are jointly significant at the 5% level for non-workers both with and without state fixed effects. In addition, none of the estimates are robust to the inclusion of state fixed effects. These weak results are consistent with some response bias concerning falling asleep while driving, as discussed in the data section. Though it is less specific, falling asleep some time during the day may be a more reliable indicator of fatigue, which translates to unsafe driving. All together, the results in Table 17 provide evidence that commuting and working times that are longer and especially early in the morning have a negative impact on road safety.

5.3. Not Enough Sleep

Table 18 reports results of logit regressions using a dependent variable equal to one if the respondent reported not getting enough sleep at least once in the past 30 days. Longer commutes and a greater prevalence of late commutes are associated with a higher likelihood of not getting enough sleep some time in the past month. Strangely, longer work hours are associated with a decreased likelihood of not getting enough sleep. Also, the result for late commutes is not robust to the inclusion of state fixed effects. However, these issues are overshadowed by the similarity in the estimates for employees and non-workers. From results not shown, this puzzling result is robust to many specifications.[45] I interpret this as evidence that the question on getting “enough sleep” is too subjective to be a valuable measure of sleep behavior. Consequently, Table 18 allows little inference regarding the relationship between commuting, work hours, and sleep.

5.4. Sex-Specific Results

In Table 19 I divide the sleep module sample by sex and estimate separate effects on males and females. The variables of interest are sex-specific in these regressions. That is, county means are averaged over only workers of the sex to which the sample is restricted. I compare coefficients across samples as in Clogg et al. (1995).

Longer commute times and especially working hours have a greater marginal effect on women than men. For all four of the dependent variables, the difference in the estimated effect of longer working hours on women is statistically greater than that on men at the 1% level. It appears that women are fully driving the effects of weekly work hours on sleep time and falling asleep. Women also appear to be driving the effect of commute time on unintentionally falling asleep. While the time tradeoff between sleep and commuting for men (-0.17) is statistically significant at the 5% level, it is also statistically smaller than the estimated tradeoff for females (-0.447) at the 10% level.

Early commuting appears to have a larger effect on sleep time for men, while late commuting appears to have a larger effect on sleep time for women.[46] This is consistent with differences in departure times for men and women in the U.S. 2.6% of women in the 2008-2012 ACS had early commutes compared to 5.6% of men, while 25.9% of women had late commutes compared to 22.5% of men. However, for women, the prevalence of early commuting has especially strong association with unintentionally falling asleep. Each percentage point increase in the share of female workers that commute before 5

[45]These include restricting the sample to counties below specified population density percentiles, restricting the sample to states that participated in the sleep module, removing the county-level controls, alternative binary dependent variables such as at least 10 or 30 days of not enough sleep, count models with and without zero-inflation, and interval regression. Results using only the sleep module sample are insignificant at conventional levels for both employees and non-workers.

[46]This interpretation should be taken with some caution, as the early commute estimates are particularly imprecise.

a.m. is associated with a 9.7% increase in the likelihood of falling asleep while driving in the past month, compared to an analogous 3.6% for men, with the difference being significant at the 10% level.

Table 20 reports results by sex for non-workers. As expected, the coefficients of interest are not jointly significant in any of the regressions for the male sample. However, they are jointly significant for short sleep and falling asleep while driving for the female sample. For both sexes, I fail to reject the hypothesis that the effects on falling asleep while driving are the same for workers and non-workers. This reinforces the notion that this dependent variable is less reliable than the others.

The results suggest that commuting time decreases sleep time and increases the likelihood of short sleep for female non-workers. As discussed in the methods section, this is plausible, seeing as non-workers can be expected to travel the same roads as commuters. All other coefficients in the first three columns for both sexes are insignificant at conventional levels. In all, the results in Table 20 are consistent with intuition and support the validity of the sex-specific results for workers.

Why do women generally show greater marginal sacrifices of sleep? Perhaps a greater likelihood of having a working spouse or children at home restrict women's locational choices for residence and/or workplace (White 1986). In this way, women may be more likely to trade away sleep than to adjust their commuting behavior to accomodate it. Women may also be more involved in household production, such that men are more likely than women to sacrifice such activities rather than sacrifice sleep.

6. Conclusion

This paper presents an array of evidence that commuting and working erode sleep time and an individual's ability to stay awake throughout the day. For each additional minute spent commuting and working, I estimate that 13-18 and 7-13 seconds of sleep are lost, respectively. I also provide evidence that strange commute times are associated with less sleep, and that early commute times in particular are associated with a higher likelihood of falling asleep while driving. Further, I show that the results differ significantly by sex: commuting time, working time, and late commuting have greater marginal effects on women, while early commuting is especially detrimental to men's sleep time and women's ability to stay awake during the day. While I also present estimated effects of the same variables on the likelihood of not getting enough sleep, their validity is more questionable. The survey question is too subjective be meaningful, despite the tenfold larger sample size. This leaves the BRFSS sleep module as the preferred source of information on sleep behavior in this paper.

Because many of the largest and densest urban areas are absent from the sleep module sample, this analysis is most applicable to areas of low or moderate density. Another reason is that mean commute times in the largest urban areas are curiously stable over time (Gordon et al. 1991). This is an empirical regularity that authors have attempted for decades to explain (Levinson and Kumar 1994; Van Ommeren and Rietveld 2005; Anas 2014). It is not clear how stable urban commute times will be in the future. Using data on a comprehensive set of U.S. Census tracts in 1990 and 2000, Kirby and LeSage (2009) show that the prevalence of long commutes rose between those years, the rise being strongly associated with sex-specific changes in employment and age distribution. Unlike other studies, they observe the entire range of regional densities in the United States. Their results indicate that as demographics change, so will commute times.

Even without a change in commute times, the results suggest that reducing the share of early commutes (starting between midnight and 5 a.m.) would enhance road safety. This is supported by the economics literature on the relationship between mental clarity and time of day (Dickinson and McElroy 2010; Dickinson and Whitehead 2014). The potential benefit is limited by the small share of early commutes as a percentage of all commutes in the U.S. (2.6% for women and 5.6% for men). Further, the burden of risk due to early morning commuting is likely to be predominantly on the part of the commuter himself, as collisions with other vehicles are less likely at those times of the day when there are fewer vehicles on the road. In a free market, this private risk is capitalized into wages and rents, so that an early commuter is no worse off than an otherwise equivalent counterpart.

The negative consequences of a commuter's lack of sleep go beyond the externality imposed on other drivers. These range from effects on productivity, to work safety, to physical, mental, and social health. The results here show that the timing and duration of commuting and working influence these various outcomes through tradeoffs in terms of sleep and sleepiness.

REFERENCES

Aguiar, Mark, and Erik Hurst. "Measuring trends in leisure: The allocation of time over five decades." *The Quarterly Journal of Economics* 122.3 (2007): 969-1006.

Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis. "Time use during the great recession." *The American Economic Review* 103.5 (2013): 1664-1696.

Anas, Alex. "Why are urban travel times so stable?." *Journal of Regional Science* (2014) doi: 10.1111/jors.12142.

Antillon, Marina, Diane S. Lauderdale, and John Mullahy. "Sleep behavior and unemployment conditions." *Economics & Human Biology* 14 (2014): 22-32.

Basner, Mathias, et al. "American Time Use Survey: Sleep time and its relationship to waking activities." *Sleep* 30.9 (2007): 1085-1095.

Basner, Mathias, Andrea M. Spaeth, and David F. Dinges. "Sociodemographic characteristics and waking activities and their role in the timing and duration of sleep." *Sleep* 37.12 (2014): 1889-1906.

Becker, Gary S. "A theory of the allocation of time." *The Economic Journal* (1965): 493-517.

Biddle, Jeff E., and Daniel S. Hamermesh. "Sleep and the Allocation of Time." *The Journal of Political Economy* 98.5 Part 1 (1990): 922-943.

Bishop, James. "Interacting effects of state cigarette taxes on smoking participation." Working Paper (2016). https://mpra.ub.uni-muenchen.de/66609/1/MPRA_paper_66609.pdf.

Black, Dan A., Natalia Kolesnikova, and Lowell J. Taylor. "Why do so few women work in New York (and so many in Minneapolis)? Labor supply of married women across US cities." *Journal of Urban Economics* 79 (2014): 59-71.

Brochu, Pierre, Catherine Deri Armstrong, and Louis-Philippe Morin. "The trendiness of sleep: an empirical investigation into the cyclical nature of sleep time." *Empirical Economics* 43.2 (2012): 891-913.

Buxton, Orfeu M., and Enrico Marcelli. "Short and long sleep are positively associated with obesity, diabetes, hypertension, and cardiovascular disease among adults in the United States." *Social Science & Medicine* 71.5 (2010): 1027-1036.

Callison, Kevin, and Robert Kaestner. "Do higher tobacco taxes reduce adult smoking? New evidence of the effect of recent cigarette tax increases on adult smoking." *Economic Inquiry* 52.1 (2014): 155-172.

Chaloupka, Frank J. "Rational Addictive Behavior and Cigarette Smoking." Journal of Political Economy 99.4 (1991): 72242.

Chaloupka, Frank J., and Kenneth E. Warner. "The economics of smoking." *Handbook of Health Economics* 1.B (2000): 1539-1627.
Chatzitheochari, Stella, and Sara Arber. "Lack of sleep, work and the long hours culture: evidence from the UK Time Use Survey." *Work, Employment & Society* 23.1 (2009): 30-48.
Chiou, Lesley, and Erich Muehlegger. "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 (2008): 1-41.
Christian, Thomas J. "Trade-offs between commuting time and health-related activities." *Journal of Urban Health* 89.5 (2012): 746-757.
Clogg, Clifford C., Eva Petkova, and Adamantios Haritou. "Statistical methods for comparing regression coefficients between models." *American Journal of Sociology* 100.5 (1995): 1261-1293.
Colman G, Remler D. "Vertical equity consequences of very high cigarette tax increases: If the poor are the ones smoking, how could cigarette tax increases be progressive?" *Journal of Policy Analysis and Management* 27.2 (2008): 376400.
Connelly, Richard T., Rajeev K. Goel, and Rati Ram. "Demand for cigarettes in the United States: effects of prices in bordering states and contiguity with Mexico and Canada." *Applied Economics* 41.18 (2009): 2255-2260.
DeCicca, Philip, Donald Kenkel, and Feng Liu. "Excise tax avoidance: the case of state cigarette taxes." *Journal of Health Economics* 32.6 (2013a): 1130-1141.
DeCicca, Philip, Donald Kenkel, and Feng Liu. "Who pays cigarette taxes? The impact of consumer price search." *Review of Economics and Statistics* 95.2 (2013b): 516-529.
DeCicca, Philip, and Logan McLeod. "Cigarette taxes and older adult smoking: evidence from recent large tax increases." *Journal of Health Economics* 27.4 (2008): 918-929.
Dickerson, Andy, Arne Risa Hole, and Luke A. Munford. "The relationship between well-being and commuting revisited: Does the choice of methodology matter?." *Regional Science and Urban Economics* 49 (2014): 321-329.
Dickinson, David L., and Todd McElroy. "Rationality around the clock: Sleep and time-of-day effects on guessing game responses." *Economics Letters* 108.2 (2010): 245-248.
Dickinson, David L., and John C. Whitehead. "Dubious and dubiouser: Contingent valuation and the time of day." *Economic Inquiry* 53.2 (2015): 1396-1400.
Ding, Ding, Klaus Gebel, Philayrath Phongsavan, Adrian E. Bauman, and Dafna Merom. "Driving: a road to unhealthy lifestyles and poor health outcomes." *PLOS ONE* 9.6 (2014): e94602.
Dube, Arindrajit, T. William Lester, and Michael Reich. "Minimum wage effects across state borders: Estimates using contiguous counties." *The Review of Economics and Statistics* 92.4 (2010): 945-964.
Emery, Sherry, Martha M. White, Elizabeth A. Gilpin, and John P. Pierce. "Was there significant tax evasion after the 1999 50 cent per pack cigarette tax increase in California?" *Tobacco Control* 11.2 (2002): 130-134.
Evans, William N., Matthew C. Farrelly, and Edward Montgomery. "Do workplace smoking bans reduce smoking?." *American Economic Review* 89.4 (1999): 728-747.
Farrelly, Matthew C., Jeremy W. Bray, Terry Pechacek, and Trevor Woollery. "Response by adults to increases in cigarette prices by sociodemographic characteristics." *Southern Economic Journal* 68.1 (2001): 156-165.
Franks, Peter, Anthony F. Jerant, J. Paul Leigh, Dennis Lee, Alan Chiem, Ilene

Lewis, and Sandy Lee. “Cigarette prices, smoking, and the poor: Implications of recent trends.” *American Journal of Public Health* 97.10 (2007): 1873-1877.
Gallet, Craig A., and John A. List. “Cigarette demand: a meta-analysis of elasticities.” *Health Economics* 12.10 (2003): 821-835.
Golden, Shelley D., Kurt M. Ribisl, and Krista M. Perreira. “Economic and political influence on tobacco tax rates: a nationwide analysis of 31 years of state data.” *American Journal of Public Health* 104.2 (2014): 350-357.
Gordon, Peter, Harry W. Richardson, and Myung-Jin Jun. “The commuting paradox evidence from the top twenty.” *Journal of the American Planning Association* 57.4 (1991): 416-420.
Gruber, Jonathan, and Botond Koszegi. ”Is Addiction Rational? Theory and Evidence.” *Quarterly Journal of Economics* 116.4 (2001): 1261-1303.
Hamermesh, Daniel S., and Elena Stancanelli. Long Workweeks and Strange Hours. Working Paper No. w20449, National Bureau of Economic Research (2014).
Hanson, Andrew, and Ryan Sullivan. “The incidence of tobacco taxation: evidence from geographic micro-level data.” *National Tax Journal* 62.4 (2009): 677-698.
Harding, Matthew, Ephraim Leibtag, and Michael F. Lovenheim. “The heterogeneous geographic and socioeconomic incidence of cigarette taxes: evidence from Nielsen Homescan Data.” *American Economic Journal: Economic Policy* 4.4 (2012): 169-198.
Herrnstein, Richard J., George F. Loewenstein, Drazen Prelec, and William Vaughan. “Utility maximization and melioration: Internalities in individual choice.” *Journal of Behavioral Decision Making* 6.3 (1993): 149-185.
Holford, Theodore R., et al.“Tobacco control and the reduction in smoking-related premature deaths in the United States, 1964-2012.” *Journal of the American Medical Association* 311.2 (2014): 164-171.
Holmes, Thomas J. “The effects of state policies on the location of industry: evidence from state borders.” No. 205. Federal Reserve Bank of Minneapolis (1996).
Horne, Jim, and Louise Reyner. “Vehicle accidents related to sleep: a review.” *Occupational and Environmental Medicine* 56.5 (1999): 289-294.
Huang, Rocco R. “Evaluating the real effect of bank branching deregulation: Comparing contiguous counties across US state borders.” *Journal of Financial Economics* 87.3 (2008): 678-705.
Jamal, Ahmed, et al. “Current cigarette smoking among adults—United States, 2005-2013.” *Morbidity and Mortality Weekly Report* 63.47 (2014): 1108-1112.
Kaestner, Robert, and Kevin Callison. “Are Cigarettes like Apples? An Assessment of the Forward Looking Hypothesis of the Demand for Cigarettes.” Working Paper (2015). http://tippie.uiowa.edu/economics/midwest-health-economics/papers/kaestner-callison-cigarettetaxes.pdf
Kanbur, Ravi, and Michael Keen. “Jeux sans frontieres: Tax competition and tax coordination.” *American Economic Review* 83.4 (1993): 877.
Keeler, Theodore E., Teh-wei Hu, Paul G. Barnett, Willard G. Manning, and Hai-Yen Sung. “Do cigarette producers price-discriminate by state? An empirical analysis of local cigarette pricing and taxation.” *Journal of Health Economics* 15.4 (1996): 499-512.
Kenkel, Donald S., Maximilian D. Schmeiser, and Carly Urban. “Is smoking inferior? Evidence from variation in the earned income tax credit.” *Journal of Human Resources* 49.4 (2014): 1094-1120.

Kirby, Dustin K., and James P. LeSage. “Changes in commuting to work times over the 1990 to 2000 period.” *Regional Science and Urban Economics* 39.4 (2009): 460-471.
Knutson, Kristen L., Eve Van Cauter, Paul J. Rathouz, Thomas DeLeire, and Diane S. Lauderdale. “Trends in the prevalence of short sleepers in the USA: 19752006.” *Sleep* 33.1 (2010): 37-45.
Leal, Andres, Julio Lopez-Laborda, and Fernando Rodrigo. “Cross-border shopping: A survey.” *International Advances in Economic Research* 16.2 (2010): 135-148.
Lenne, Michael G., Thomas J. Triggs, and Jennifer R. Redman. “Time of day variations in driving performance.” *Accident Analysis & Prevention* 29.4 (1997): 431-437.
Levinson, David M., and Ajay Kumar. “The rational locator: why travel times have remained stable.” *Journal of the American Planning Association* 60.3 (1994): 319-332.
Lewit, Eugene, Douglas Coate, and Michael Grossman. “The Effects of Government Regulation on Teenage Smoking.” *Journal of Law and Economics* 24.3 (1981): 54573.
Lewit, Eugene M., and Douglas Coate. “The Potential for Using Excise Taxes to Reduce Smoking. *Journal of Health Economics* 1.2 (1982): 12145.
Lovenheim, Michael F. “How far to the border?: The extent and impact of cross-border casual cigarette smuggling.” Stanford Institute for Economic Policy Research Policy Paper No. 06040. Stanford, CA: Stanford Institute for Economic Policy Research (2007).
Lovenheim, Michael F. “How far to the border?: The extent and impact of cross-border casual cigarette smuggling.” *National Tax Journal* 61.1 (2008): 7-33.
Luckhaupt, Sara E., Sangwoo Tak, and Geoffrey M. Calvert. “The prevalence of short sleep duration by industry and occupation in the National Health Interview Survey.” *Sleep* 33.2 (2010): 149-159.
Ma, Zhen. “The effects of increases in cigarette prices on cigarette consumption among smokers after the Master Settlement Agreement.” *Review of Economics of the Household* (2015) doi:10.1007/s11150-015-9290-0.
Maag, E., Merriman, D. “Tax policy responses to revenue shortfalls.” *State Tax Notes* (2003).
MacLean, Alistair W., David R.T. Davies, and Kris Thiele. “The hazards and prevention of driving while sleepy.” *Sleep Medicine Reviews* 7.6 (2003): 507-521.
MacLean, Johanna Catherine, Asia Sikora Kessler, and Donald S. Kenkel. “Cigarette taxes and older adult smoking: evidence from the Health and Retirement Study.” *Health Economics* (2015) doi:10.1002/hec.3161.
McKenzie, Brian. Modes less traveled bicycling and walking to work in the United States: 20082012. Technical Report ACS-25, U.S. Census Bureau (2014).
Monaco, Kristen, Lindy Olsson, and Justin Hentges. “Hours of sleep and fatigue in motor carriage.” *Contemporary Economic Policy* 23.4 (2005): 615-624.
Nelson, Jon P. “Accessibility and the Value of Time in Commuting.” *Southern Economic Journal* 43.3 (1977): 1321-1329.
Orzechowski and Walker. “The tax burden on tobacco: historical compilation.” 49 (2014).
Pack, Allan I., et al. “Characteristics of crashes attributed to the driver having fallen asleep.” *Accident Analysis & Prevention* 27.6 (1995): 769-775.
Philip, Pierre, et al. “Fatigue, sleep restriction and driving performance.”

Accident Analysis and Prevention 37.3 (2005): 473-478.
Roberts, Jennifer, Robert Hodgson, and Paul Dolan. "'It's driving her mad': Gender differences in the effects of commuting on psychological health." *Journal of Health Economics* 30.5 (2011): 1064-1076.
Rosenthal, Stuart S., and William C. Strange. "Agglomeration and hours worked." *The Review of Economics and Statistics* 90.1 (2008): 105-118.
Ross, Stephen L., and Yves Zenou. "Are shirking and leisure substitutable? An empirical test of efficiency wages based on urban economic theory." *Regional Science and Urban Economics* 38.5 (2008): 498-517.
Ruhm, Christopher J. "Healthy living in hard times." *Journal of Health Economics* 24.2 (2005): 341-363.
Saba, Richard P., T. Randolph Beard, Robert B. Ekelund, Jr., and Rand W. Ressler. "The demand for cigarette smuggling." *Economic Inquiry* 33.2 (1995): 189.
Santos, Adella, Nancy McGuckin, Hikari Yukiko Nakamoto, Danielle Gray, and Susan Liss. "Summary of travel trends: 2009 National Household Travel Survey." U.S. Department of Transportation Report No. FHWA-PL-ll-022 (2011).
Schneeweiss, Hans, John Komlos, and Amar S. Ahmad. "Symmetric and asymmetric rounding: a review and some new results." *AStA Advances in Statistical Analysis* 94.3 (2010): 247-271.
Scott, Linda D., et al. "The relationship between nurse work schedules, sleep duration, and drowsy driving." *Sleep* 30.12 (2007): 1801-1807.
Schwartz, Jonathan RL, and Thomas Roth. "Shift Work Sleep Disorder." *Drugs* 66.18 (2006): 2357-2370.
Siahpush, Mohammad, Melanie A. Wakefield, Matt J. Spittal, Sarah J. Durkin, and Michelle M. Scollo. "Taxation reduces social disparities in adult smoking prevalence." *American Journal of Preventive Medicine* 36.4 (2009): 285-291.
Sloan, Frank A., and Justin G. Trogdon. "The impact of the Master Settlement Agreement on cigarette consumption." *Journal of Policy Analysis and Management* 23.4 (2004): 843-855.
Stehr, Mark. "Cigarette tax avoidance and evasion." *Journal of Health Economics* 24.2 (2005): 277-297.
Sugiyama, Takemi, Ding Ding, and Neville Owen. "Commuting by car: weight gain among physically active adults." *American Journal of Preventive Medicine* 44.2 (2013): 169-173.
Szalontai, Gabor. "The demand for sleep: a South African study." *Economic Modelling* 23.5 (2006): 854-874.
Townsend, Joy, Paul Roderick, and Jacqueline Cooper. "Cigarette smoking by socioeconomic group, sex, and age: effects of price, income, and health publicity." *British Medical Journal* 309.6959 (1994): 923-928.
Turcotte, Martin. "Commuting to work: Results of the 2010 General Social Survey." *Canadian Social Trends* 92 (2011): 25-36.
Van Ommeren, Jos N., and Eva Gutirrez-i-Puigarnau. "Are workers with a long commute less productive? An empirical analysis of absenteeism." *Regional Science and Urban Economics* 41.1 (2011): 1-8.
Van Ommeren, Jos, and Piet Rietveld. "The commuting time paradox." *Journal of Urban Economics* 58.3 (2005): 437-454.
White, Michelle J. "Sex differences in urban commuting patterns." *The American Economic Review* 76.2 (1986): 368-372.
Williamson, Ann, et al. "The link between fatigue and safety." *Accident Analysis*

& Prevention 43.2 (2011): 498-515.
Wooldridge, Jeffrey M. *Econometric analysis of cross section and panel data.* Second edition. MIT press (2010).
Xu, Xin. "The business cycle and health behaviors." *Social Science & Medicine* 77 (2013): 126-136.

Figure 1: Example Border Groups: Indiana

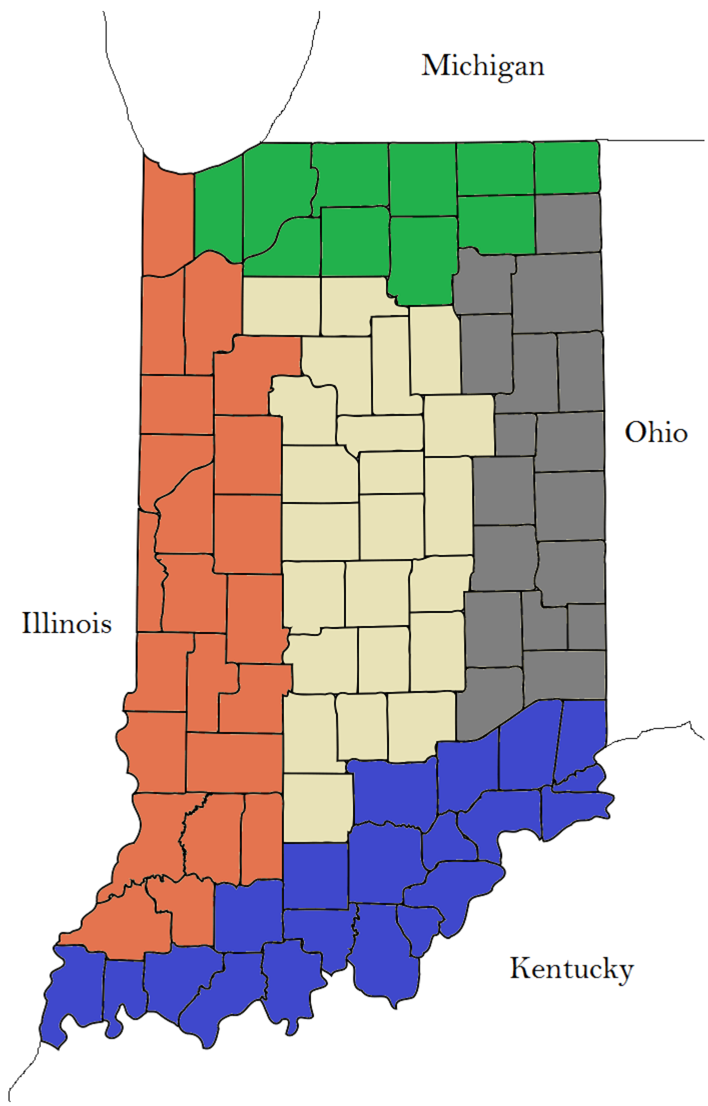
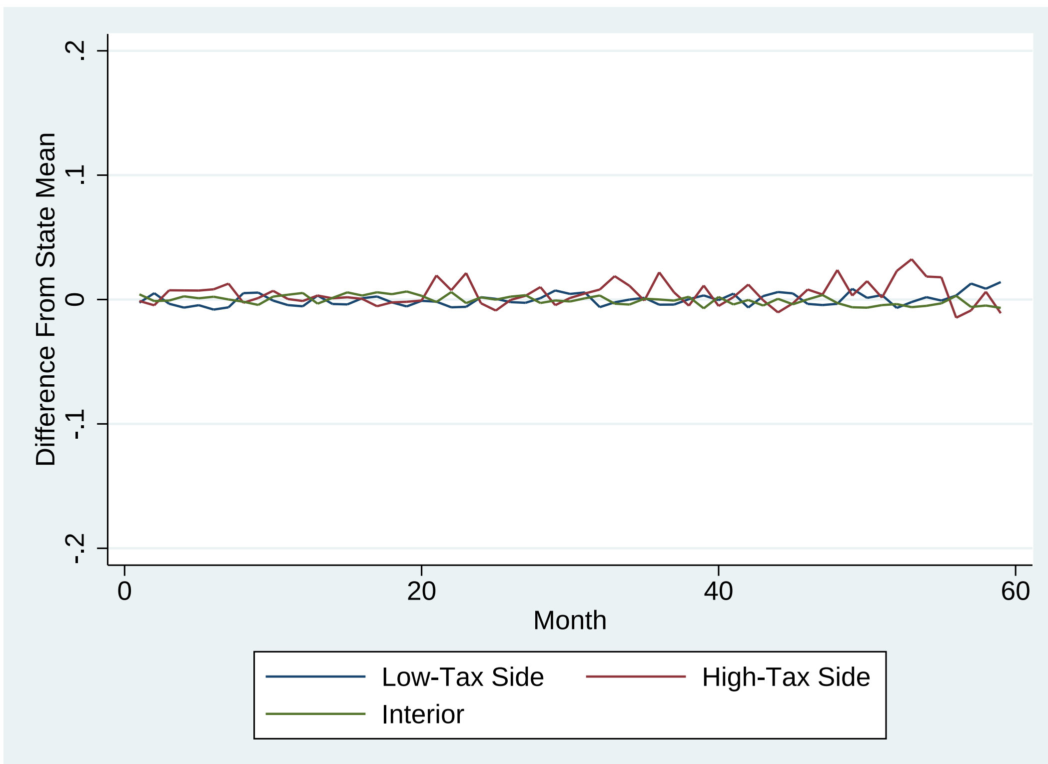


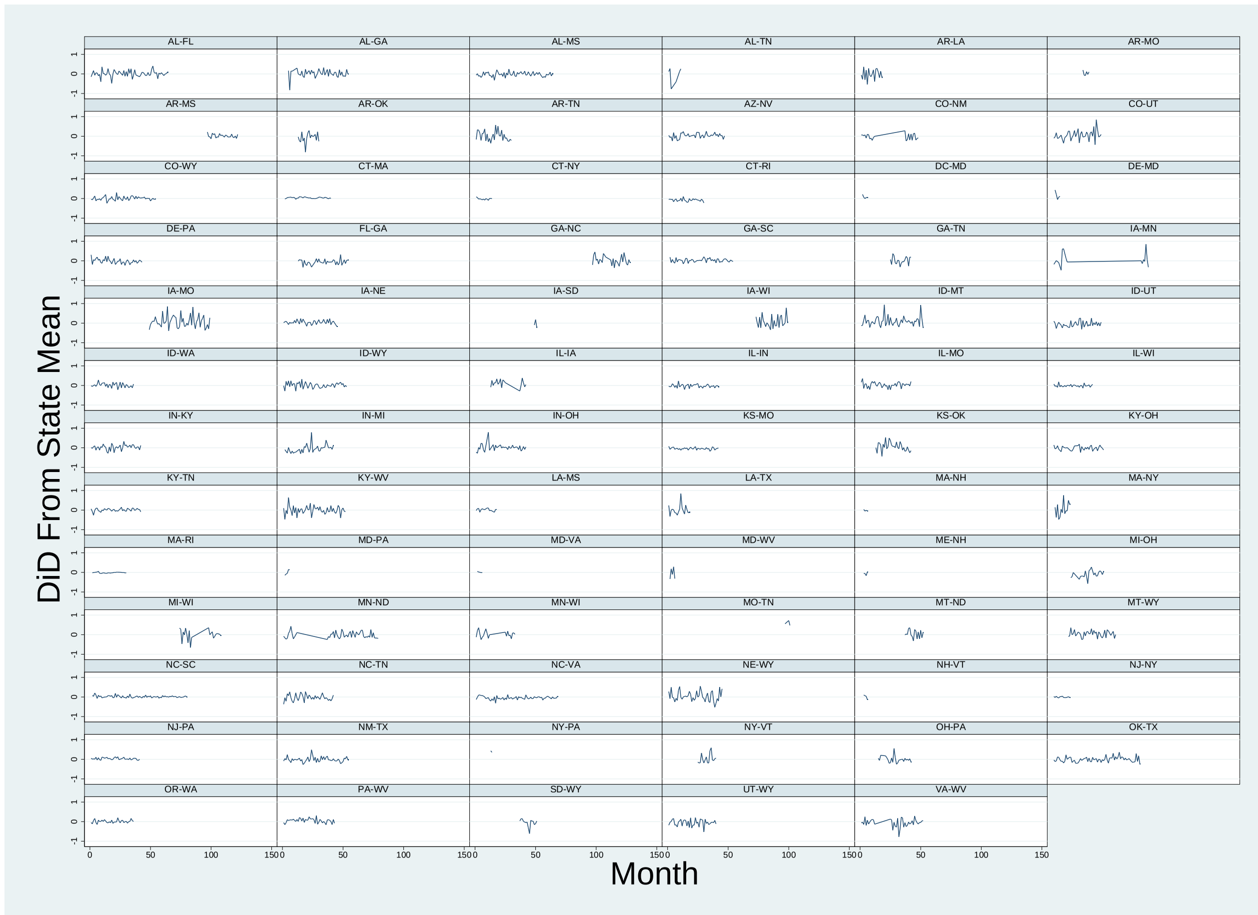
Illustration of state-groups formed using to a 40-mile distance cutoff using Indiana as an example. The assignment of counties to Illinois (orange), Michigan (green), Ohio (gray), Kentucky (blue), and the Indiana interior group (beige) result in five state-groups.

Figure 2: Pre-Intervention State-Adjusted Trends by State-Group Type



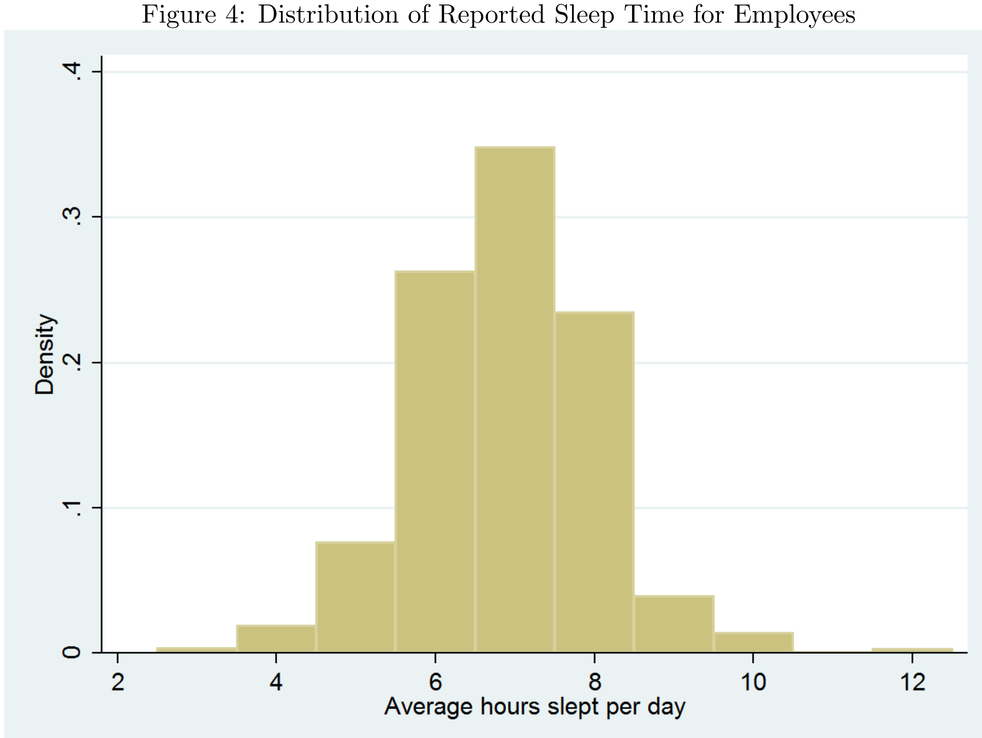
Each line connects the monthly mean differences between the state-group smoking rate and the state smoking rate for residents of border groups that did not experience a tax increase in the observed or a prior sample month. Values are aggregated over state-group types and weighted by the number of observations per state-group. Values are depicted for each month from January 1999 to December 2004.

Figure 3: Differences in Pre-Intervention State-Adjusted Trends by Border Group



Each line connects the monthly mean differences in deviations from the state smoking rate for state-groups on opposite sides of the given border in those months prior to any tax changes in that border group in the sample period. Values are depicted for each month from January 1999 to December 2004 prior to any tax changes in the border group for which data is available for both state-groups.

Figure 4: Distribution of Reported Sleep Time for Employees



Distribution of responses of employed persons to the survey question "On average, how many hours of sleep do you get in a 24-hour period? Think about the time you actually spend sleeping or napping, not just the amount of sleep you think you should get." N = 56150 respondents, with sample mean of 6.9 hours and Sheppard-corrected standard deviation of 1.2 hours.

Table 1: Means of Variables

Year	Everyday Smoker	Home Tax	Nearest Lower Tax	Log(Distance)	N
1999	0.191	0.914	0.679	4.365	101,276
2000	0.187	1.059	0.819	4.350	117,786
2001	0.184	1.029	0.792	4.345	137,294
2002	0.185	1.218	0.893	4.337	165,580
2003	0.177	1.484	1.081	4.312	181,605
2004	0.163	1.513	1.076	4.335	219,549
2005	0.159	1.709	1.180	4.307	252,032
2006	0.154	1.699	1.168	4.267	238,523
2007	0.156	1.656	1.134	4.414	315,441
2008	0.141	1.786	1.241	4.334	300,636
2009	0.134	2.349	1.788	4.333	310,872
2010	0.130	2.683	1.988	4.354	326,121
2011	0.138	2.585	1.938	4.360	364,551
2012	0.132	2.580	1.913	4.342	335,548
Full Sample	0.152	1.922	1.400	4.341	3,366,814

Home and nearest lower taxes are the sum of federal, state, and county cigarette taxes in the respondent's county measured in January 2015 dollars. Distance is the number of miles to the nearest county in a lower-tax state.

Table 2: Estimated Effects of Cigarette Taxes on Smoking Participation

	(1)	(2)	(3)	(4)
Home Tax	-0.741***	-0.669***	-0.702***	-0.887***
	(0.187)	(0.240)	(0.178)	(0.247)
Lower Tax			-0.202	-0.129
			(0.142)	(0.228)
Log(Distance)			-0.084	-0.183*
			(0.136)	(0.096)
Home Tax × Lower Tax				-0.577***
				(0.188)
Home Tax × Log(Distance)				-0.004
				(0.089)
Lower Tax × Log(Distance)				0.187
				(0.150)
(Home Tax)^2		-0.020		0.192**
		(0.034)		(0.080)
(Lower Tax)^2				0.589***
				(0.137)
(Log(Distance))^2				-0.155*
				(0.091)
N	3,366,814	3,366,814	3,366,814	3,366,814

The dependent variable in all regressions is an indicator equal to one only if the respondent smokes every day. Tax and distance variables are centered at their sample means and all reported coefficients and standard errors are multiplied by 100 to aid the presentation of the results. All regressions include state fixed effects, month-year fixed effects, the state unemployment rate in the survey month, and individual-level controls for marital status, employment status, education, sex, race, the presence of one, two, or three or more children in the household, income category, and age category. Taxes are in January 2015 dollars and distance is in miles. All regressions are weighted by BRFSS sample weights and all standard errors are clustered by state.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 3: Marginal Effects at Year-Specific Means

Year	Home Tax	Lower Tax	Log(Distance)
1999	-0.905***	-0.188	-0.219**
	(0.309)	(0.305)	(0.104)
2000	-0.901***	-0.197	-0.211**
	(0.307)	(0.306)	(0.104)
2001	-0.900***	-0.204	-0.210**
	(0.308)	(0.310)	(0.104)
2002	-0.879***	-0.206	-0.211**
	(0.289)	(0.285)	(0.101)
2003	-0.881***	-0.142	-0.208**
	(0.270)	(0.242)	(0.101)
2004	-0.901***	-0.082	-0.208**
	(0.268)	(0.223)	(0.105)
2005	-0.898***	-0.067	-0.202*
	(0.260)	(0.209)	(0.108)
2006	-0.894***	-0.072	-0.200*
	(0.257)	(0.208)	(0.108)
2007	-0.881***	-0.080	-0.178
	(0.246)	(0.198)	(0.109)
2008	-0.886***	-0.052	-0.178
	(0.241)	(0.186)	(0.115)
2009	-0.876***	-0.055	-0.170
	(0.233)	(0.178)	(0.118)
2010	-0.839***	-0.093	-0.166
	(0.212)	(0.167)	(0.124)
2011	-0.830***	-0.106	-0.173
	(0.209)	(0.168)	(0.125)
2012	-0.838***	-0.096	-0.164
	(0.212)	(0.168)	(0.123)

Each value is calculated using the year-specific means from Table 1, the estimates in column 4 of Table 2, and Equations 14-16. Standard errors are calculated using the delta method.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 4: Robustness Checks: Alternative Specifications

	(1) Baseline	(2) County FEs	(3) State Trends	(4) No Fed. Taxes	(5) Control for Smoking Bans	(6) Some-day Smokers
Home Tax	-0.887***	-0.671**	-0.668***	-0.900***	-0.871***	-0.596**
	(0.247)	(0.322)	(0.246)	(0.274)	(0.233)	(0.254)
Lower Tax	-0.129	-0.128	-0.282	-0.089	-0.165	-0.187
	(0.228)	(0.303)	(0.218)	(0.232)	(0.234)	(0.248)
Log(Distance)	-0.183*	-0.057	-0.256***	-0.186*	-0.173*	-0.098
	(0.096)	(0.073)	(0.088)	(0.101)	(0.096)	(0.090)
Home Tax × Lower Tax	-0.577***	-0.501**	-0.570***	-0.501***	-0.555***	-0.944***
	(0.188)	(0.223)	(0.164)	(0.154)	(0.177)	(0.222)
Home Tax × Log(Distance)	-0.004	-0.044	0.015	0.019	0.007	0.117
	(0.089)	(0.086)	(0.090)	(0.088)	(0.084)	(0.115)
Lower Tax × Log(Distance)	0.187	0.165	0.075	0.156	0.171	-0.123
	(0.150)	(0.171)	(0.183)	(0.200)	(0.144)	(0.196)
(Home Tax)^2	0.192**	0.200*	0.119*	0.176**	0.186**	0.238***
	(0.080)	(0.101)	(0.069)	(0.082)	(0.075)	(0.076)
(Lower Tax)^2	0.589***	0.417**	0.510***	0.525***	0.577***	0.848***
	(0.137)	(0.170)	(0.133)	(0.166)	(0.129)	(0.212)
(Log(Distance))^2	-0.155*	-0.005	-0.176*	-0.154	-0.152	-0.077
	(0.091)	(0.062)	(0.097)	(0.093)	(0.093)	(0.091)
N	3,366,814	3,366,814	3,366,814	3,366,814	3,366,814	3,366,814

The dependent variable in all columns except column 6 is an indicator equal to one only if the respondent smokes every day. In column 6, the dependent variable is also equal to one for those that smoke some days. Tax and distance variables are centered at their sample means and all reported coefficients and standard errors are multiplied by 100 to aid the presentation of the results. All regressions include state fixed effects (or county fixed effects in column 2), month-year fixed effects, the state unemployment rate in the survey month, and individual-level controls for marital status, employment status, education, sex, race, the presence of one, two, or three or more children in the household, income category, and age category. Taxes are in January 2015 dollars and distance is in miles. All regressions are weighted by BRFSS sample weights and all standard errors are clustered by state.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 5: Robustness Checks: Sample Restrictions

	(1) Baseline	(2) No West Region	(3) No Northeast	(4) No AL, MO, or VA	(5) Yearly-Sampled Counties Only
Home Tax	-0.887***	-0.722***	-0.928***	-0.880***	-0.851***
	(0.247)	(0.243)	(0.254)	(0.254)	(0.277)
Lower Tax	-0.129	-0.119	-0.123	-0.151	-0.057
	(0.228)	(0.224)	(0.248)	(0.231)	(0.306)
Log(Distance)	-0.183*	-0.119	-0.191*	-0.228**	-0.206*
	(0.096)	(0.084)	(0.098)	(0.099)	(0.108)
Home Tax × Lower Tax	-0.577***	-0.395**	-0.652***	-0.561***	-0.599***
	(0.188)	(0.161)	(0.200)	(0.189)	(0.194)
Home Tax × Log(Distance)	-0.004	0.050	-0.026	0.025	0.009
	(0.089)	(0.088)	(0.091)	(0.088)	(0.082)
Lower Tax × Log(Distance)	0.187	0.174	0.222	0.163	0.050
	(0.150)	(0.163)	(0.167)	(0.153)	(0.163)
$(\text{Home Tax})^2$	0.192**	0.127*	0.212**	0.194**	0.190*
	(0.080)	(0.068)	(0.083)	(0.082)	(0.097)
$(\text{Lower Tax})^2$	0.589***	0.472***	0.675***	0.592***	0.613***
	(0.137)	(0.161)	(0.135)	(0.139)	(0.162)
$(\text{Log(Distance)})^2$	-0.155*	-0.115	-0.157	-0.175*	-0.196*
	(0.091)	(0.096)	(0.094)	(0.099)	(0.098)
N	3,366,814	2,559,092	2,892,812	3,251,058	2,453,890

The dependent variable in all regressions is an indicator equal to one only if the respondent smokes every day. Tax and distance variables are centered at their sample means and all reported coefficients and standard errors are multiplied by 100 to aid the presentation of the results. All regressions include state fixed effects, month-year fixed effects, the state unemployment rate in the survey month, and individual-level controls for marital status, employment status, education, sex, race, the presence of one, two, or three or more children in the household, income category, and age category. Taxes are in January 2015 dollars and distance is in miles. All regressions are weighted by BRFSS sample weights and all standard errors are clustered by state.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 6: Robustness Checks: Smuggling Incentives

	(1) Baseline	(2) Pop. Min. 10k	(3) Pop. Min. 50k	(4) Under 10% From Res.	(5) Under 5% From Res.	(6) Regional Placebo
Home Tax	-0.887***	-0.822***	-0.873***	-0.979***	-0.969***	-0.562**
	(0.247)	(0.226)	(0.236)	(0.306)	(0.315)	(0.263)
Lower Tax	-0.129	-0.251	-0.082	0.148	0.100	-0.159
	(0.228)	(0.219)	(0.222)	(0.178)	(0.187)	(0.304)
Log(Distance)	-0.183*	-0.182*	-0.152	-0.215**	-0.224*	-0.108
	(0.096)	(0.100)	(0.093)	(0.103)	(0.113)	(0.325)
Home Tax $\times$ Lower Tax	-0.577***	-0.581***	-0.675***	-0.510*	-0.514*	0.287
	(0.188)	(0.185)	(0.215)	(0.262)	(0.278)	(0.360)
Home Tax $\times$ Log(Distance)	-0.004	0.007	0.063	-0.075	-0.080	0.155
	(0.089)	(0.080)	(0.089)	(0.121)	(0.123)	(0.253)
Lower Tax $\times$ Log(Distance)	0.187	0.168	0.104	0.194	0.190	-0.078
	(0.150)	(0.133)	(0.158)	(0.169)	(0.178)	(0.280)
$(\text{Home Tax})^2$	0.192**	0.192**	0.248**	0.279***	0.275***	-0.095
	(0.080)	(0.080)	(0.095)	(0.099)	(0.099)	(0.120)
$(\text{Lower Tax})^2$	0.589***	0.598***	0.608***	0.377	0.397	-0.164
	(0.137)	(0.134)	(0.135)	(0.241)	(0.261)	(0.301)
$(\text{Log(Distance)})^2$	-0.155*	-0.152*	-0.144	-0.071	-0.087	0.046
	(0.091)	(0.090)	(0.089)	(0.088)	(0.095)	(0.111)
N	3,366,814	3,332,462	3,332,462	2,808,220	2,371,017	3,231,171

The dependent variable in all regressions is an indicator equal to one only if the respondent smokes every day. Tax and distance variables are centered at their sample means and all reported coefficients and standard errors are multiplied by 100 to aid the presentation of the results. All regressions include state fixed effects, month-year fixed effects, the state unemployment rate in the survey month, and individual-level controls for marital status, employment status, education, sex, race, the presence of one, two, or three or more children in the household, income category, and age category. Taxes are in January 2015 dollars and distance is in miles. All regressions are weighted by BRFSS sample weights and all standard errors are clustered by state.

*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 7: Robustness Checks: Mexican Border

	(1) Baseline	(2) No Southwest	(3) No Border States	(4) Border States	(5) Border Counties
Home Tax	-0.887***	-0.859***	-0.792***	0.696	2.189
	(0.247)	(0.224)	(0.237)	(0.802)	(2.828)
Lower Tax	-0.129	-0.193	-0.236	-1.087	-4.440
	(0.228)	(0.222)	(0.235)	(0.737)	(3.610)
Log(Distance)	-0.183*	-0.054	-0.046	-0.679*	-0.756
	(0.096)	(0.082)	(0.078)	(0.234)	(0.584)
Home Tax × Lower Tax	-0.577***	-0.496***	-0.487**	-2.170	4.103
	(0.188)	(0.179)	(0.182)	(2.514)	(6.931)
Home Tax × Log(Distance)	-0.004	0.050	0.029	-1.387*	-1.058
	(0.089)	(0.080)	(0.083)	(0.535)	(0.956)
Lower Tax × Log(Distance)	0.187	0.063	0.084	1.901	2.655
	(0.150)	(0.145)	(0.145)	(1.097)	(1.894)
$(\text{Home Tax})^2$	0.192**	0.153*	0.138*	0.345	-1.147
	(0.080)	(0.077)	(0.078)	(1.114)	(3.588)
$(\text{Lower Tax})^2$	0.589***	0.583***	0.579***	1.568	-0.923
	(0.137)	(0.147)	(0.150)	(1.768)	(3.068)
$(\text{Log(Distance)})^2$	-0.155*	-0.107	-0.097	-0.346	-0.342
	(0.091)	(0.100)	(0.098)	(0.410)	(0.698)
N	3,366,814	2,782,455	3,056,734	310,080	50,437

The dependent variable in all regressions is an indicator equal to one only if the respondent smokes every day. Tax and distance variables are centered at their sample means and all reported coefficients and standard errors are multiplied by 100 to aid the presentation of the results. All regressions include state fixed effects, month-year fixed effects, the state unemployment rate in the survey month, and individual-level controls for marital status, employment status, education, sex, race, the presence of one, two, or three or more children in the household, income category, and age category. Taxes are in January 2015 dollars and distance is in miles. All regressions are weighted by BRFSS sample weights and all standard errors are clustered by state.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 8: Results for Low and Non-Low Income

	(1) Working Age	(2) Under 25k / year	(3) Over 25k / year
Home Tax	-1.099***	-1.502**	-0.827***
	(0.296)	(0.576)	(0.236)
Lower Tax	0.087	0.127	0.059
	(0.279)	(0.650)	(0.271)
Log(Distance)	-0.229*	-0.217	-0.064
	(0.119)	(0.274)	(0.087)
Home Tax × Lower Tax	-0.990***	-0.858	-0.754***
	(0.229)	(0.702)	(0.184)
Home Tax × Log(Distance)	-0.026	0.018	-0.104
	(0.121)	(0.320)	(0.121)
Lower Tax × Log(Distance)	0.299	0.472	0.313
	(0.189)	(0.417)	(0.208)
(Home Tax)^2	0.317***	0.414*	0.197***
	(0.092)	(0.226)	(0.072)
(Lower Tax)^2	0.936***	0.571	0.802***
	(0.194)	(0.586)	(0.210)
(Log(Distance))^2	-0.230**	-0.380	-0.120
	(0.102)	(0.244)	(0.074)
N	1,822,601	364,497	1,305,666

The dependent variable in all regressions is an indicator equal to one only if the respondent smokes every day. Tax and distance variables are centered at their sample means and all reported coefficients and standard errors are multiplied by 100 to aid the presentation of the results. All regressions include state fixed effects, month-year fixed effects, the state unemployment rate in the survey month, and individual-level controls for marital status, employment status, education, sex, race, the presence of one, two, or three or more children in the household, income category, and age category. Samples for this table are restricted to respondents age 25 to 64. Taxes are in January 2015 dollars and distance is in miles. All regressions are weighted by BRFSS sample weights and all standard errors are clustered by state.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 9: Pre-Intervention State-Adjusted Trends by State-Group Type

State-Group Type	Deviation from State Mean Smoking Rate	N
Low-Tax Side	-0.033	174,848
High-Tax Side	0.438	86,443
Interior Group	-0.022	203,987

Differences between the state-group smoking rate and the state smoking rate for respondents in the first five years of the sample residing in a border group that did not experience a tax increase in the observed or a prior sample month. Groups are formed using a 40-mile cutoff. Values are aggregated over state-group types and weighted by the number of observations per state-group.

Table 10: Pseudo-DDD With Various Border Group Bounds

	(1) 20 Miles	(2) 40 Miles	(3) 60 Miles	(4) 80 Miles	(5) 100 Miles	(6) MSA
Local Tax	-0.516**	-0.507**	-0.511***	-0.515***	-0.458**	-0.968***
	(0.259)	(0.249)	(0.191)	(0.161)	(0.178)	(0.244)
Group Tax	-0.147	-0.165	-0.203	-0.255	-0.374	0.316
	(0.332)	(0.334)	(0.288)	(0.280)	(0.268)	(0.242)
Local Tax × Group Tax	-0.065	-0.355***	-0.314**	-0.371**	-0.362**	0.766***
	(0.160)	(0.126)	(0.130)	(0.172)	(0.157)	(0.171)
$(\text{Local Tax})^2$	0.024	0.075	0.064	0.085	0.072	-0.234***
	(0.088)	(0.066)	(0.061)	(0.066)	(0.063)	(0.055)
$(\text{Group Tax})^2$	0.094	0.374***	0.358***	0.399**	0.444**	-0.437***
	(0.102)	(0.100)	(0.101)	(0.170)	(0.170)	(0.113)
N	3,388,181	3,388,181	3,388,181	3,388,181	3,388,181	2,444,347

The dependent variable in all regressions is an indicator equal to one only if the respondent smokes every day. Tax variables are centered at their sample means and measured in January 2015 dollars. All reported coefficients and standard errors are multiplied by 100 to aid the presentation of the results. All regressions include state-group fixed effects, month fixed effects, the state unemployment rate in the survey month, indicators for state smoking bans, and individual-level controls for marital status, employment status, education, sex, race, the presence of one, two, or three or more children in the household, income category, and age category. All regressions are weighted by BRFSS sample weights and all standard errors are clustered by border group.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 11: Full DDD With Various Border Group Bounds

	(1) 20 miles	(2) 40 miles	(3) 60 miles	(4) 80 miles	(5) 100 miles	(6) MSA
Local Tax × Group Tax	-0.520***	-1.073***	-0.991***	-0.937***	-0.914***	-0.208
	(0.080)	(0.150)	(0.186)	(0.203)	(0.209)	(0.311)
N	3,388,077	3,388,086	3,388,076	3,388,073	3,388,056	2,444,347

The dependent variable in all regressions is an indicator equal to one only if the respondent smokes every day. Tax variables are centered at their sample means and measured in January 2015 dollars. All reported coefficients and standard errors are multiplied by 100 to aid the presentation of the results. All regressions include state-group fixed effects, state-month fixed effects, group-month fixed effects, the state unemployment rate in the survey month, indicators for state smoking bans, and individual-level controls for marital status, employment status, education, sex, race, the presence of one, two, or three or more children in the household, income category, and age category. All regressions are weighted by BRFSS sample weights and all standard errors are clustered by border group.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 12: Distance Tier Interactions

Distance Tier	Local Tax × Group Tax Coefficient
80-100 miles	-0.225
	(0.453)
60-80 miles	-0.960**
	(0.460)
40-60 miles	-0.817**
	(0.372)
20-40 miles	-0.439**
	(0.192)
0-20 miles	-0.813***
	(0.163)
N	3,388,056

Same regression as in column 5 of Table 11, except that the interaction of the two tax variables is interacted with an indicator for distance tier, with distance of greater than 100 miles to the nearest border being the omitted category.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 13: State Participation in the BRFSS Sleep Module

	Number of Employed Respondents				
State	2009	2010	2011	2012	Total
Alaska	0	0	1623	2105	3728
Arkansas	0	1275	0	0	1275
Connecticut	0	2919	0	0	2919
Delaware	0	1837	0	0	1837
D.C.	0	1728	0	0	1728
Georgia	2360	0	0	0	2360
Hawaii	2735	2605	0	0	5340
Illinois	2462	0	0	0	2462
Kansas	0	0	0	2537	2537
Louisiana	3402	0	0	0	3402
Minnesota	2577	4457	7319	0	14353
Missouri	0	1966	0	0	1966
Nevada	0	1453	0	1732	3185
Oregon	0	1617	0	1685	3302
Tennessee	0	0	1832	0	1832
Wyoming	2507	0	0	0	2507
Total	16043	19857	10774	8059	54733

Number of respondents to the BRFSS sleep module questions that reported being employed for wages at the time of the survey, by state and year. Zero responses reflect years for which states opted not to participate in the sleep module. States not listed did not participate in the module in any of the years. Puerto Rico participated in 2012, but is not included in the analysis.

Table 14: Summary Statistics for County-Level Variables

	Employees		Non-Workers	
	Mean	Std. Dev.	Mean	Std. Dev.
Sleep Module, 2009-2012				
Daily commute time (mins.)	47.21	8.11	46.91	8.09
% Early commuters	4.45	2.21	4.60	2.19
% Late commuters	23.55	3.43	23.61	3.52
Weekly work hours	38.42	1.31	38.39	1.32
Population over age 16	403484.61	589012.90	377421.41	582398.23
Population density (per mile^2)	945.66	1679.34	817.45	1552.24
% Labor force participation	67.08	5.41	65.73	5.72
% Unemployment	8.49	2.39	8.85	2.51
N	47868		47165	
Core Sample, 2008-2010				
Daily commute time (mins.)	48.00	9.58	47.77	9.22
% Early commuters	4.30	1.98	4.45	2.02
% Late commuters	23.49	3.51	23.54	3.51
Weekly work hours	38.46	1.25	38.43	1.24
Population over age 16	395640.77	734652.08	391640.39	769228.82
Population density (per mile^2)	943.95	3199.75	848.52	3081.98
% Labor force participation	65.10	6.10	63.80	6.56
% Unemployment	8.78	2.72	9.16	2.81
N	486305		493115	

All variables are from the 2008-2012 American Community Survey. Each county is weighted by the number of corresponding observations in the BRFSS sample.

Table 15: Correlations of County-Level Variables

	Commute	% Early	% Late	Workweek	Pop.>16	Unemp.	LFP
Daily Commute Time	1.000						
% Early Commuters	0.227	1.000					
% Late Commuters	-0.164	-0.242	1.000				
Weekly Work Hours	0.227	0.274	-0.533	1.000			
Pop. Over 16	0.338	-0.109	0.319	-0.027	1.000		
Unemployment	0.201	0.070	0.354	-0.200	0.157	1.000	
LFP	0.022	-0.313	0.152	-0.005	0.225	-0.421	1.000
Pop. Density	0.384	-0.341	0.176	0.122	0.413	0.146	0.201

Each county is weighted by the number of corresponding observations in the BRFSS sleep module sample. Weighting by core sample observations gives similar results.

Table 16: Effects on Sleep Time and Short Sleep

	Employees		Non-Workers	
Panel A: OLS Regression Dependent Variable: Minutes of Sleep Per Day				
Daily Commute Time	-0.304***	-0.223***	-0.154	-0.047
	(0.094)	(0.083)	(0.108)	(0.072)
% Early Commuters	-0.756	-0.402	-0.685	-0.073
	(0.511)	(0.251)	(0.753)	(0.341)
% Late Commuters	-0.457**	-0.491**	0.156	-0.196
	(0.213)	(0.194)	(0.211)	(0.185)
Weekly Work Hours	-0.923*	-1.809***	0.963	-0.514
	(0.498)	(0.566)	(0.657)	(0.623)
State FE		✓		✓
F(4, c - 1) (H_0: no effects)	4.12***	6.37***	1.14	0.44
$\chi^2(4)$ (H_0: same as workers)			20.17***	11.00**
N	47868	47868	47165	47165
Panel B: Logit Regression Dependent Variable: Fewer Than 6.5 Hours of Sleep Per Day				
Daily Commute Time	1.006**	1.002	1.004	1.001
	(0.003)	(0.002)	(0.003)	(0.002)
% Early Commuters	1.036**	1.028***	1.024	1.003
	(0.014)	(0.007)	(0.020)	(0.009)
% Late Commuters	1.011**	1.010*	0.998	1.007
	(0.005)	(0.005)	(0.006)	(0.005)
Weekly Work Hours	1.027**	1.025	0.984	1.026
	(0.014)	(0.017)	(0.021)	(0.017)
State FE		✓		✓
$\chi^2(4)$ (H_0: no effects)	18.48***	26.81***	3.10	4.33
$\chi^2(4)$ (H_0: same as workers)			17.34***	11.45**
N	47868	47868	47165	47165

Sample is pooled over all observations from the BRFSS sleep module from 2009 to 2012. All regressions include individual-level controls for month of the survey, marital status, education, sex, race, the presence of one, two, or three or more children in the household, and intervals for income and age, as well as county-level controls for population, population density, labor force participation, and unemployment. "Employees" were employed for wages and "Non-Workers" were not employed at the time of the survey. Standard errors are clustered by county. Coefficients in Panel B are reported as odds ratios for ease of interpretation.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 17: Effects on Falling Asleep During the Day and While Driving

	Employees				Non-Workers	
Panel A: Logit Regression Dependent Variable: Unintentionally Fell Asleep In Past Month						
Daily Commute Time	1.006***	1.005**	1.005***	1.004**	1.000	1.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
% Early Commuters	1.023**	1.015**	1.020**	1.013**	0.999	1.010
	(0.011)	(0.006)	(0.010)	(0.006)	(0.006)	(0.007)
% Late Commuters	1.006	1.013***	1.004	1.011**	1.005	1.015***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
Weekly Work Hours	1.031**	1.034*	1.027**	1.027	1.012	1.011
	(0.014)	(0.018)	(0.013)	(0.017)	(0.011)	(0.015)
Daily Sleep Time			0.996***	0.996***		
			(0.000)	(0.000)		
State FE		✓		✓		✓
$\chi^2(4)$ (H_0: no effects)	24.72***	27.90***	25.13***	21.54***	1.69	16.56***
$\chi^2(4)$ (H_0: same as workers)					23.77***	12.62**
N	47868	47868	47868	47868	47165	47165
Panel B: Logit Regression Dependent Variable: Fell Asleep While Driving In Past Month						
Daily Commute Time	1.009*	1.007	1.007	1.006	1.007	1.005
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.008)
% Early Commuters	1.043**	1.017	1.038*	1.013	1.029	1.038
	(0.022)	(0.017)	(0.020)	(0.017)	(0.022)	(0.027)
% Late Commuters	0.993	1.005	0.989	1.001	1.009	1.026
	(0.011)	(0.013)	(0.011)	(0.013)	(0.017)	(0.018)
Weekly Work Hours	1.003	1.022	0.996	1.011	1.059	1.072
	(0.031)	(0.034)	(0.030)	(0.034)	(0.042)	(0.048)
Daily Sleep Time			0.994***	0.994***		
			(0.000)	(0.000)		
State FE		✓		✓		✓
$\chi^2(4)$ (H_0: no effects)	14.40***	7.47	13.03**	4.48	10.20**	11.63**
$\chi^2(4)$ (H_0: same as workers)					10.47**	9.33*
N	47323	47323	47323	47323	44040	44040

Sample is pooled over all observations from the BRFSS sleep module from 2009 to 2012. All regressions include individual-level controls for month of the survey, marital status, education, sex, race, the presence of one, two, or three or more children in the household, and intervals for income and age, as well as county-level controls for population, population density, labor force participation, and unemployment. "Employees" were employed for wages and "Non-Workers" were not employed at the time of the survey. Standard errors are clustered by county. Coefficients are reported as odds ratios for ease of interpretation. *: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 18: Effects on Not Getting Enough Sleep

Logit Regression Dependent Variable: > 0 Days of Not Enough Sleep in Past Month				
	Employees		Non-Workers	
Daily Commute Time	1.002***	1.004***	1.002***	1.002***
	(0.001)	(0.001)	(0.001)	(0.001)
% Early Commuters	1.002	0.997	0.995	0.994***
	(0.004)	(0.003)	(0.003)	(0.002)
% Late Commuters	1.010***	1.003	1.007***	1.005***
	(0.002)	(0.002)	(0.002)	(0.002)
Weekly Work Hours	0.984**	0.978***	0.984***	0.994
	(0.006)	(0.005)	(0.005)	(0.004)
State FE		✓		✓
$\chi^2(4)$ (H_0: no effects)	10.33**	23.94***	12.98**	4.63
$\chi^2(4)$ (H_0: same as workers)			8.32*	14.18***
N	481174	481174	482131	482130

Sample is pooled over all observations from the BRFSS core survey from 2008 to 2010. All regressions include individual-level controls for month of the survey, marital status, education, sex, race, the presence of one, two, or three or more children in the household, and intervals for income and age, as well as county-level controls for population, population density, labor force participation, and unemployment. "Employees" were employed for wages and "Non-Workers" were not employed at the time of the survey. Standard errors are clustered by county. Coefficients are reported as odds ratios for ease of interpretation.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 19: Sex-Specific Effects for Workers

Estimation:	OLS	Logit	Logit	Logit
Dependent Variable:	Minutes of Sleep Per Day	<6.5 Hours of Sleep Per Day	Fell Asleep During Day	Fell Asleep While Driving
Panel A: Male Sample				
Daily Commute Time	-0.170**	1.005*	1.003	1.004
	(0.084)	(0.002)	(0.003)	(0.005)
% Early Commuters	-1.073***	1.040***	1.018*	1.036**
	(0.409)	(0.013)	(0.011)	(0.016)
% Late Commuters	-0.042	1.000	0.997	0.981
	(0.216)	(0.007)	(0.006)	(0.012)
Weekly Work Hours	0.229	0.991	1.006	0.967
	(0.527)	(0.015)	(0.015)	(0.029)
N	19643	19643	19643	19591
Panel B: Female Sample				
Daily Commute Time	-0.447***	1.009**	1.009***	1.025***
	(0.157)	(0.004)	(0.003)	(0.009)
% Early Commuters	-0.888	1.034*	1.046***	1.097***
	(0.747)	(0.020)	(0.014)	(0.039)
% Late Commuters	-0.624***	1.019***	1.012*	1.008
	(0.240)	(0.007)	(0.007)	(0.014)
Weekly Work Hours	-2.127***	1.088***	1.071***	1.107**
	(0.563)	(0.017)	(0.020)	(0.045)
N	28225	28225	28225	28200

Male and female samples are pooled over all observations from the BRFSS sleep module from 2009 to 2012 of males and females, respectively. For each regression, all four reported variables are specific to the sex to which the sample is restricted. All regressions include individual-level controls for month of the survey, marital status, education, race, the presence of one, two, or three or more children in the household, and intervals for income and age, as well as county-level controls for population, population density, labor force participation, and unemployment. All samples are restricted to respondents who were employed for wages at the time of the survey. Standard errors are clustered by county. Logit coefficients are reported as odds ratios for ease of interpretation. *: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 20: Sex-Specific Effects for Non-Workers

Estimation:	OLS	Logit	Logit	Logit
Dependent Variable:	Minutes of Sleep Per Day	<6.5 Hours of Sleep Per Day	Fell Asleep During Day	Fell Asleep While Driving
Panel A: Male Sample				
Daily Commute Time	0.081	0.994*	0.996	1.002
	(0.111)	(0.003)	(0.002)	(0.005)
% Early Commuters	-0.481	1.029	0.999	1.017
	(0.690)	(0.019)	(0.007)	(0.018)
% Late Commuters	0.321	0.993	0.998	1.017
	(0.236)	(0.006)	(0.005)	(0.020)
Weekly Work Hours	0.602	0.988	1.021*	1.017
	(0.704)	(0.020)	(0.012)	(0.037)
F(4, c - 1) or $\chi^2(4)$ (H_0: no effects)	0.52	4.99	5.97	2.38
$\chi^2(4)$ (H_0: same as workers)	10.94**	20.60***	15.09***	6.76
N	16087	16087	16087	16021
Panel B: Female Sample				
Daily Commute Time	-0.464**	1.013***	1.003	1.020***
	(0.183)	(0.005)	(0.002)	(0.007)
% Early Commuters	-1.017	1.024	1.007	1.043
	(0.945)	(0.026)	(0.011)	(0.033)
% Late Commuters	-0.291	1.006	1.005	1.021*
	(0.272)	(0.007)	(0.005)	(0.012)
Weekly Work Hours	-0.345	1.024	0.995	1.137***
	(0.745)	(0.022)	(0.013)	(0.046)
F(4, c - 1) or $\chi^2(4)$ (H_0: no effects)	1.67	9.90**	3.34	17.53***
$\chi^2(4)$ (H_0: same as workers)	6.11	18.62***	42.90***	5.55
N	31078	31078	31078	31063

Male and female samples are pooled over all observations from the BRFSS sleep module from 2009 to 2012 of males and females, respectively. For each regression, all four reported variables are specific to the sex to which the sample is restricted. All regressions include individual-level controls for month of the survey, marital status, education, race, the presence of one, two, or three or more children in the household, and intervals for income and age, as well as county-level controls for population, population density, labor force participation, and unemployment. All samples are restricted to respondents who were not employed at the time of the survey. Standard errors are clustered by county. Logit coefficients are reported as odds ratios for ease of interpretation.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

APPENDIX

Table A1: Further Robustness Checks

	(1) Baseline	(2) Age <65	(3) Same State Lower Tax	(4) Add Lowest Tax State	(5) No Unemp.	6 Exclude Refusers
Home Tax	-0.887***	-1.008***	-0.778***	-0.905***	-0.688**	-0.926***
	(0.247)	(0.290)	(0.271)	(0.251)	(0.284)	(0.249)
Lower Tax	-0.129	-0.079	-0.154	0.102	-0.223	-0.147
	(0.228)	(0.268)	(0.242)	(0.227)	(0.229)	(0.270)
Log(Distance)	-0.183*	-0.184*	-0.094	-0.124	-0.157	-0.171*
	(0.096)	(0.104)	(0.091)	(0.108)	(0.095)	(0.097)
Home Tax × Lower Tax	-0.577***	-0.670***	-0.617*	-0.574***	-0.465**	-0.595***
	(0.188)	(0.218)	(0.347)	(0.189)	(0.204)	(0.168)
Home Tax × Log(Distance)	-0.004	0.010	0.066	-0.005	-0.001	-0.039
	(0.089)	(0.105)	(0.086)	(0.086)	(0.090)	(0.083)
Lower Tax × Log(Distance)	0.187	0.219	0.120	0.181	0.154	0.208
	(0.150)	(0.164)	(0.166)	(0.136)	(0.149)	(0.155)
$(\text{Home Tax})^2$	0.192**	0.228**	0.197	0.201**	0.131	0.184**
	(0.080)	(0.096)	(0.124)	(0.082)	(0.087)	(0.080)
$(\text{Lower Tax})^2$	0.589***	0.700***	0.632***	0.499***	0.547***	0.594***
	(0.137)	(0.160)	(0.223)	(0.137)	(0.167)	(0.137)
$(\text{Log(Distance)})^2$	-0.155*	-0.175*	-0.075	-0.118	-0.152	-0.143
	(0.091)	(0.098)	(0.062)	(0.092)	(0.093)	(0.090)
N	3,366,814	2,757,785	3,366,814	3,434,559	3,366,814	2,923,448

The dependent variable in all regressions is an indicator equal to one only if the respondent smokes every day. Tax and distance variables are centered at their sample means and all reported coefficients and standard errors are multiplied by 100 to aid the presentation of the results. All regressions include state fixed effects, month-year fixed effects, the state unemployment rate in the survey month (except in column 5), and individual-level controls for marital status, employment status, education, sex, race, the presence of one, two, or three or more children in the household, income category, and age category. Taxes are in January 2015 dollars and distance is in miles. All regressions are weighted by BRFSS sample weights and all standard errors are clustered by state.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table A2: Expanded Estimates From Preferred Specification

Taxes and Distance			**Race**		
Home Tax	-0.887***	(0.247)	White, non-Hispanic	Omitted	
Lower Tax	-0.129	(0.228)	Black, non-Hispanic	-7.574***	(0.850)
Log(Distance)	-0.183*	(0.096)	Other, non-Hispanic	-5.199***	(0.467)
Home Tax $\times$ Lower Tax	-0.577***	(0.188)	Multiracial, non-Hispanic	0.964*	(0.490)
Home Tax $\times$ Log(Distance)	-0.004	(0.089)	Hispanic	-13.174***	(0.498)
Lower Tax $\times$ Log(Distance)	0.187	(0.150)	Refused	-2.755***	(0.558)
(Home Tax)^2	0.192**	(0.080)			
(Lower Tax)^2	0.589***	(0.137)	**Children in Household**		
(Log(Distance))^2	-0.155*	(0.091)	No Children	Omitted	
			One Child	-0.416*	(0.213)
Marital Status			Two Children	-1.833***	(0.229)
Married	Omitted		$>$ Two Children	-1.937***	(0.357)
Divorced	9.157***	(0.323)	Refused	-2.926***	(0.713)
Widowed	4.976***	(0.279)			
Separated	9.135***	(0.418)	**Income**		
Never Married	3.286***	(0.232)	$<$ \$10,000	Omitted	
In Unmarried Couple	8.107***	(0.948)	\$10,000-\$15,000	0.171	(0.383)
Refused	3.947***	(0.883)	\$15,000-\$20,000	0.887**	(0.389)
			\$20,000-\$25,000	0.354	(0.341)
Employment Status			\$25,000-\$35,000	-1.196***	(0.424)
Employed For Wages	Omitted		\$35,000-\$50,000	-2.493***	(0.496)
Self-Employed	0.353*	(0.189)	\$50,000-\$75,000	-4.704***	(0.631)
Out of Work $\geq$ 1 Year	6.547***	(0.352)	$\geq$ \$75,000	-7.154***	(0.704)
Out of Work $<$ 1 Year	6.241***	(0.465)	Don't Know/Not Sure	-3.401***	(0.653)
Homemaker	-1.088***	(0.240)	Refused	-6.922***	(0.496)
Student	-7.443***	(0.664)			
Retired	0.349**	(0.154)	**Age**		
Unable to Work	6.055***	(0.400)	18-24	Omitted	
Refused	-0.183	(1.012)	25-29	3.989***	(0.334)
			30-34	3.518***	(0.255)
Completed Education			35-39	3.930***	(0.259)
No School	Omitted		40-44	4.305***	(0.258)
Grade 1-8	1.641*	(0.848)	45-49	3.892***	(0.287)
Grade 9-11	11.905***	(1.151)	50-54	2.182***	(0.263)
High School	3.463***	(0.929)	55-59	-0.312	(0.328)
Some College	-1.223	(1.046)	60-64	-3.525***	(0.454)
College Degree	-9.345***	(1.205)	65-69	-6.841***	(0.706)
Refused	0.111	(1.495)	70-74	-10.276***	(0.836)
			Refused	-3.065***	(0.304)
Sex					
Male	Omitted		State Unemployment	-0.430***	(0.068)
Female	-2.603***	(0.252)			
r^2	0.090		N	3,366,814	

The dependent variable is an indicator equal to one only if the respondent smokes every day. Tax and distance variables are centered at their sample means and all reported coefficients and standard errors are multiplied by 100 to aid the presentation of the results. Coefficients on state and month-year fixed effects are not reported. Taxes are in January 2015 dollars and distance is in miles. The regression is weighted by BRFSS sample weights and standard errors are clustered by state.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table A3: Results By Income Category

	less than $10,000	$10,000-$15,000	$15,000-$20,000	$20,000-$25,000	$25,000-$35,000	$35,000-$50,000	$50,000-$75,000	$75,000 and above
Home Tax	0.638	-2.088***	-3.592***	-0.841	-1.244**	-1.533**	0.324	-0.512*
	(0.702)	(0.700)	(1.100)	(0.746)	(0.607)	(0.573)	(0.271)	(0.277)
Lower Tax	-3.135**	0.187	2.763**	-0.396	-0.929	0.738	-0.390	0.199
	(1.170)	(1.076)	(1.272)	(1.085)	(0.854)	(0.729)	(0.495)	(0.244)
Log(Distance)	-0.402	-0.454	0.442	-0.514	-0.359	-0.173	0.364**	-0.078
	(0.401)	(0.621)	(0.413)	(0.345)	(0.309)	(0.135)	(0.150)	(0.147)
Home Tax × Lower Tax	-0.596	-0.972	-1.751*	-0.396	-0.774	-1.364**	0.246	-0.654**
	(0.958)	(1.540)	(0.971)	(0.878)	(0.601)	(0.602)	(0.331)	(0.279)
Home Tax × Log(Distance)	-1.392***	1.205**	0.838	-0.511	-0.252	-0.357	0.238*	-0.167
	(0.313)	(0.514)	(0.570)	(0.420)	(0.278)	(0.241)	(0.133)	(0.165)
Lower Tax × Log(Distance)	1.872***	-0.636	-0.578	1.061*	0.412	0.703*	-0.112	0.278
	(0.587)	(1.082)	(0.675)	(0.608)	(0.447)	(0.399)	(0.205)	(0.269)
$(\text{Home Tax})^2$	-0.151	0.628	1.389***	-0.146	0.229	0.303	-0.276**	0.182*
	(0.297)	(0.514)	(0.374)	(0.302)	(0.197)	(0.210)	(0.119)	(0.104)
$(\text{Lower Tax})^2$	0.502	0.715	0.321	0.865	0.906	1.566**	-0.089	0.583*
	(0.963)	(0.950)	(0.947)	(0.870)	(0.574)	(0.648)	(0.158)	(0.299)
$(\text{Log(Distance)})^2$	-0.933**	-0.108	-0.224	-0.350	-0.625**	-0.029	0.125	-0.140
	(0.428)	(0.423)	(0.217)	(0.332)	(0.261)	(0.124)	(0.110)	(0.087)
N	75,048	66,938	96,489	126,022	184,558	271,783	322,538	526,787

The dependent variable in all regressions is an indicator equal to one only if the respondent smokes every day. Tax and distance variables are centered at their sample means and all reported coefficients and standard errors are multiplied by 100 to aid the presentation of the results. All regressions include state fixed effects, month-year fixed effects, the state unemployment rate in the survey month, and individual-level controls for marital status, employment status, education, sex, race, the presence of one, two, or three or more children in the household, income category, and age category. Samples for this table are restricted to respondents age 25 to 64. Taxes are in January 2015 dollars and distance is in miles. All regressions are weighted by BRFSS sample weights and all standard errors are clustered by state.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table A4: Second-Nearest Lower Tax

Home Tax	-0.832***
	(0.223)
Lower Tax	-0.566
	(0.492)
Log(Distance)	-0.225*
	(0.120)
Home Tax × Lower Tax	-0.863**
	(0.381)
Home Tax × Log(Distance)	-0.004
	(0.104)
Lower Tax × Log(Distance)	-0.039
	(0.291)
$(\text{Home Tax})^2$	0.165**
	(0.078)
$(\text{Lower Tax})^2$	0.729
	(0.618)
$(\text{Log(Distance)})^2$	-0.159
	(0.102)
Even Lower Tax	0.391
	(0.437)
Log(dist to ELT)	-0.163
	(0.130)
Even Lower Tax × Log(dist to ELT)	0.718*
	(0.366)
Home Tax × Even Lower Tax	0.459
	(0.414)
Home Tax × Log(dist to ELT)	0.085
	(0.175)
Lower Tax × Even Lower Tax	-0.322
	(1.118)
Lower Tax × Log(dist to ELT)	-0.805*
	(0.420)
Log(Distance) × Even Lower Tax	0.336
	(0.289)
Log(Distance) × Log(dist to ELT)	0.126
	(0.136)
Even Lower Tax × Even Lower Tax	0.105
	(0.701)
Log(dist to ELT) × Log(dist to ELT)	-0.400***
	(0.110)
N	3,010,867

The dependent variable in all regressions is an indicator equal to one only if the respondent smokes every day. Tax and distance variables are centered at their sample means and all reported coefficients and standard errors are multiplied by 100 to aid the presentation of the results. All regressions include state fixed effects, month-year fixed effects, the state unemployment rate in the survey month, and individual-level controls for marital status, employment status, education, sex, race, the presence of one, two, or three or more children in the household, income category, and age category. Taxes are in January 2015 dollars and distance is in miles. All regressions are weighted by BRFSS sample weights and all standard errors are clustered by state.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

VITA

James McKown Bishop

Candidate for the Degree of

Doctor of Philosophy

Thesis: THREE ESSAYS ON REGIONAL ECONOMIC IMPLICATIONS FOR HEALTH BEHAVIORS

Major Field: Economics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Economics at Oklahoma State University, Stillwater, Oklahoma in May, 2016.

Completed the requirements for the Bachelor of Science in Economics at Oklahoma State University, Stillwater, Oklahoma in May, 2011.

Completed the requirements for the Bachelor of Science in Mathematics at Oklahoma State University, Stillwater, Oklahoma in May, 2011.

Experience:

Graduate Student Instructor, Introduction to Microeconomics
August 2013 - December 2015

Graduate Teaching Assistant
January 2012 - August 2013

Graduate and Professional Student Government Association (GPSGA) Representative August 2011 - May 2016

GPSGA Social Chair August 2012 - May 2013