NBER WORKING PAPER SERIES

RELATIVE TAX RATES, PROXIMITY AND CIGARETTE TAX NONCOMPLIANCE: EVIDENCE FROM A NATIONAL SAMPLE OF LITTERED CIGARETTE PACKS

Shu Wang
David Merriman
Frank J. Chaloupka

Working Paper 22577
http://www.nber.org/papers/w22577

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2016

This research is supported by the National Cancer Institute of the National Institutes of Health under the State and Community Tobacco Control Initiative, grant number 5U01-CA154248 and the Robert Wood Johnson Foundation, under the Bridging the Gap: Research Informing Practice and Policy for Healthy Youth Behavior Initiative, grant numbers 64702 and 70175. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors have no financial arrangements that might give rise to conflicts of interest with respect to the research reported in this paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Shu Wang, David Merriman, and Frank J. Chaloupka. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Relative Tax Rates, Proximity and Cigarette Tax Noncompliance: Evidence from a National Sample of Littered Cigarette Packs
Shu Wang, David Merriman, and Frank J. Chaloupka
NBER Working Paper No. 22577
August 2016
JEL No. H26,I12

ABSTRACT

We analyze data about cigarette tax compliance from the first national scale littered cigarette packs collection. We code each pack based on whether an appropriate tax had been paid at the location where it was found. Noncompliance across our 132 sample communities ranges from zero to one hundred percent with an appropriately weighted mean of 21 percent. We provide evidence that noncompliance is due to both cross-border shopping and cigarette trafficking. OLS and binomial logit regressions demonstrate that the financial incentive for non-compliance is the most important explanatory variable and has a statistically and quantitatively significant impact on noncompliance.

Shu Wang
Department of Political Science
Michigan State University
368 Farm Lane, S344
East Lansing, MI 48824
swang24@msu.edu

David Merriman
Inst of Gov't and Public Affairs (MC191)
815 W Van Buren St, Suite 525
Chicago, Illinois 60607-3506
dmerrim@uic.edu

Frank J. Chaloupka
University of Illinois at Chicago
Department of Economics (m/c 144)
College of Liberal Arts and Sciences
601 S. Morgan Street, Room 713
Chicago, IL 60607-7121
and NBER
fjc@uic.edu

List of Tables

Table 1: Number of packs by tax differential and distance between purchase origin and sales destination
Table 2: Summary statistics
Panel 1: full sample
Panel 2: "Near the border" sample
Table 2: OLS regression: Effects of IFNC on tax noncompliance
Panel 1: full sample
Panel 2: "Near the border" sample
Table 4: Robustness analysis with alternative predictors
Panel 1: full sample
Panel 2: "Near the border" sample
Table 5: Binomial Logit parameter estimates
Table 6: Simulations of the relationship between five cent tax increase, predicted noncompliance and revenue at various levels of IFNC

List of Figures

Figure 1: Number of communities by IFNC
Panel 1: full sample
Panel 2: "Near the border" sample
Figure 2: Panel 1 – Population shares by IFNC at the block level (bottom 99% of blocks, Full sample)
Panel 2 - Population shares by IFNC at the block level (top 1% of blocks, Full sample)
Panel 3 - Population shares by IFNC at the block level (bottom 99% of the blocks, "Near the border" sample)
Panel 4 - Population shares by IFNC at the block level (top 1% of blocks, "Near the border" sample)
Figure 3: Change in IFNC from one cent change in tax (Full sample)

I. INTRODUCTION

"The central difficulty in estimating the parameters of … [the cigarette demand function] is... [that the propensity to cross-border shop] S_i is unobserved; location of purchase is not observed in the data. My solution to this problem is parameterize the S_i function and then incorporate these parameters into… [the cigarette demand function]." Lovenheim (p.16) 2008.

The problem Lovenheim identifies has plagued studies of cigarette demand for decades (Baltagi, Badi, and Levin, 1986; Chaloupka, 1991; Thursby, Jenssen, and Thursby, 1991; Beckman, Grossman, and Murphy, 1991; Saba et al, 1995; Galbraith and Kaiserman, 1997; Thursby and Thursby, 2000). Cigarette taxes, and hence prices, vary geographically, sometimes by a large amount in a small distance (Merriman 2010, Chernick and Merriman 2013). Cigarette purchasers may buy their cigarettes in a variety of places and may pay a variety of prices. Without knowing the location of purchase, the tax that is paid often cannot be inferred. The best solution available frequently has been as in Lovenheim (2008), to theorize, based on relatively abstract models about human behavior and to use econometric evidence to simultaneously infer both the propensity for tax noncompliance and the price elasticity of demand for cigarettes. Without empirical grounding for the tax noncompliance estimates however, estimates of the price elasticity of demand for cigarettes are likely to be biased since econometricians cannot be certain of the saliency of home and cross-border prices (Pesko et al 2016).

Fortunately, cigarette packs are physical objects and nearly all US states and many local governments require physical evidence of tax payment, in the form of a tax stamp displayed on the cigarette pack. Several recent studies (Fix et. al. 2014, Davis et. al. 2013, Kurti, von Lampe and Thompkins 2013, Chernick and Merriman 2013, Merriman 2010, and Lakhdar 2008) have exploited this information to determine the tax jurisdiction in which the pack was purchased. We extend this methodology by collecting the first national sample of littered cigarette packs. We use this sample to provide national estimates of tax noncompliance and to better understand its causes and implications.

II. LITERATURE REVIEW

Previous research has shown that the tax differential and distance between the home and a lower-rate jurisdiction are two major factors that drive tax noncompliance (e.g., Chiou and Muehlegger, 2008; Lovenheim, 2008; Harding, Leibtag, and Lovenheim, 2010; Merriman, 2010; Chernick and Merriman, 2013). It is the interactive effect of these two factors that determines the level of noncompliance. As Lovenheim (2008) points out, in many MSAs, there are multiple jurisdictions with lower prices than the home jurisdiction and border-crossers have a choice about where to shop.

Crosby, Merriman, Wang and Chaloupka (2014) compare the estimates of tax noncompliance using surveys and physical pack collections. They find that the estimates obtained from surveys are consistently lower than the ones from physical pack collections (Also see Institute of Medicine 2015 for review of different estimates of tax noncompliance using different methods.) In studies using surveys that ask smokers about the location of their purchase, national tax noncompliance is estimated to be between 3.6 percent (Current Population Survey Tobacco Use Supplement, or “CPS-TUS”, 2010-11) and 5 percent (DeCicca, Kenkel, and Liu, 2010). Estimates from studies using physical pack collections are much higher, including 19.7 percent (Fix et al, 2014, packs mailed in by consumers) and 21 percent (Barker et al forthcoming). In some localities physical packs provide evidence of higher rates of noncompliance. Davis et al (2013) used littered cigarette packs collected from random samples of Census tracts in Boston, New York City, Philadelphia, Providence, and Washington D.C. They found that overall nearly 59 percent of packs did not have a proper local tax stamp. Noncompliance ranged from about 27 percent in Philadelphia to 81 percent in Washington, D.C. Chernick and Merriman (2013) find evidence of noncompliance rates of more than 50 percent in NYC and Merriman (2010) puts noncompliance rates at 75 percent in the City of Chicago.

Chiou and Muehlegger (2008) use CPS-TUS data combined with respondents' county of residence to estimate consumers' choice of purchase locations as a function of differential taxation, travel costs and demographics. They estimate a consumer's tradeoff between distance and price using a term that interacts the tax-inclusive price differential between the state of purchase and the closest alternative location and the inverse of distance to the alternative location.

Similarly, Harding Leibtag, and Lovenheim (2010) use Nielsen Homescan data to investigate how tax incidence of state cigarette excise taxes varies with the excise tax rate, distance to nearest lower tax state, and smokers' socioeconomic status. They find that as excise taxes rise, the likelihood a consumer crosses state lines to purchase cigarettes increases significantly on the border, but this likelihood declines steeply as one moves farther from cross-state tax dollar saving opportunities. In particular, a one cent excise tax increase raises the probability of a cross-state purchase by nearly six percent for consumers on the border, but this probability declines by more than one percent for each one percent increase in distance as we move away from the border.

The probability of a consumer traveling cross-border to avoid higher cigarette taxes varies among studies. Harding et. al (2010) estimate the likelihood of cross-state purchasing to be 5 percent whereas Lovenheim (2008) estimate it to be 13 percent. Harding and co-authors attribute the difference to the fact that the average distance to the border is closer in the sample (i.e., 102.7 miles in CPS-TUS) used by Lovenheim (2008) than the one (Nielsen homescan data) they used (126.5 miles). Both studies find a significant relationship between tax noncompliance and the cigarette tax rate differential interacted with distance to the border of a lower tax locality.

Neither Chiou and Muehlegger (2008) nor Harding et al (2010) directly observe cross-border travel. Instead these authors use a low self-reported price paid to infer that cross-border travel occurred. Merriman (2010) extends the research on tax noncompliance by using the cigarette tax stamp affixed onto the pack as tangible evidence of tax noncompliance. Using a random sample of littered cigarette packs in Chicago, he finds that nearly 75 percent of the packs did not have a Chicago tax stamp, and that both the tax differential and the distance to the border with a lower-rate locality affected the probability of a non-compliant stamp. The $2.68 in 2008 difference between the tax in Chicago and surrounding counties decreased the probability of a local stamp by close to 60 percent overall. However the probability of tax noncompliance decreased about one percent with each one mile increase in distance to the lower-tax state (i.e. Indiana) border.

Extending the methodology of Merriman (2010) to New York City (NYC), Chernick and Merriman (2013) collect four waves of littered cigarette pack data and examine the change in tax noncompliance in response to the New York State (NYS) cigarette tax rate increase in 2008 (from $1.50 to $2.75). Unlike previous research that depends on cross-sectional variation, this study uses a natural quasi-experiment to study changes in tax noncompliance before and after a change in the cigarette tax rate. Chernick and Merriman (2013) find that the share of littered packs with a NYC tax stamp fell from 55 percent to 49 percent immediately after the tax increase, and this was unchanged after three months, and also after one year and three months. They also find evidence that distance to low tax sources of cigarettes, including NYS border, New Jersey state border, and Poospatuck Indian Reservation influences tax noncompliance.

In summary, variations in estimates of tax noncompliance are partly due to variations in methodologies used in the studies. Studies using surveys of smokers have the advantage of a large nationally representative sample, but may be subject to the weakness of under-reporting of potentially anti-social behavior and also only capture individual tax avoidance/evasion but generally will not capture activities like buying bootlegged cigarettes from legitimate retailers. Using tax stamps to determine the location of purchase addresses this weakness; however, both Merriman (2010) and Chernick and Merriman (2013) conduct their studies in the cities with high cigarette tax rates adjacent to several lower-rate jurisdictions, and thus the findings may not be generalizable to the rest of the US. Our study applies the methodology used in Merriman (2010) to a national sample and therefore is able to obtain findings that may more plausibly generalize to the entire nation[1]. We acknowledge that the littered pack methodology also has significant limitations since we are unable to demonstrate the representativeness of our sample using conventional measures (such as checking the demographic characteristics of respondents) and we caution that some non-compliant packs are the results of tourism or normal interjurisdictional commuting patterns (See Merriman 2010 for attempts to address these potential weaknesses).

III. CONCEPTUAL FRAMEWORK, DATA AND METHODS

A. Conceptual Framework

Our primary objective is to collect and analyze tangible, direct and representative estimates of cigarette tax noncompliance in a large number of areas around the US. We combine this evidence with data about hypothesized determinants of tax noncompliance to estimate parameters, including the tax noncompliance gradient (i.e. the change in tax noncompliance by distance to lower tax cigarettes) that are fundamental to this literature and relevant to the policy

[1] While our sample is national in scope it is not necessarily nationally representative. The sample was designed to be representative of communities with school children in certain age ranges. Additional detail about sampling is discussed below in the section on data collection strategy.

discussion. We begin by discussing the relationship between the phenomena that we observe (i.e. the share of littered packs collected without a tax stamp or with a stamp from a lower tax jurisdiction) and the population-wide cigarette tax noncompliance rate.

Let $S_i = 1$ if the ith cigarette pack consumed in a particular location was purchased in a lower tax jurisdiction and $S_i = 0$ otherwise. Define the tax noncompliance rate as $\overline{S} = \left(\frac{\sum_{i=1}^{N} S_i}{N} \right)$ where N is the total number of cigarette packs consumed at that location. Using our littered pack collection we observe a sample of the packs consumed and calculate $\hat{\overline{S}} \equiv \left(\frac{\sum_{i=1}^{N_L} S_i}{N_L} \right)$ with $N_L < N$.

If our sampled packs are an independent and identically distributed random sample of all packs then the law of large numbers implies that $E\left(\hat{\overline{S}}\right) = \overline{S}$. Of course, a non-trivial concern is the extent to which the littered packs that we recovered were, in fact, a representative sample of all packs consumed in the area. Merriman (2008) and Chernick and Merriman (2012) discuss the general issue of the representativeness of littered cigarette packs and provide evidence that they are reasonably representative. In the research design used here the places that were sampled within each community were intended to be representative of the community area. The evidence provided in earlier literature and the consistency of our results with the findings in Fix et. al. (2014) gives us confidence that our sample is not biased either for, or against, areas with particularly high levels of tax noncompliance.

B. Data Collection Strategy

Barker et al (2016) provides detailed documentation of the protocol that was used to select collection sites and to collect and code data. We provide a brief summary of the most important points here. Our data on littered cigarette packs was collected and analyzed as a part of a large project funded jointly by the Robert Woods Johnson Foundation and the National Cancer Institute. The project was designed to examine how recent increase in tobacco taxes affect tobacco use and related behaviors among U.S. adults and youth[2]. Data for this project came from a large variety of primary and secondary sources. The sample communities in which we attempted to collect littered packs represented school enrollment areas for nationally representative samples of 8th, 10th, and 12th grade public school students in the continental U.S.[3].

As a part of this primary data collection, two-person teams of data collectors were trained to follow the data collection protocol, and traveled to 160 communities to collect littered cigarette packs between May and July 2012. The collectors looked for and collected littered packs from a representative sample of streets, business entrances, and parks, and marked the location where the pack was found.

We coded the communities in which packs were found, whether a stamp was affixed, and the taxing authority reflected on the stamp. Out of a total number of 3,867 packs collected, there were 2,116 packs with cellophane[4]. Because the tax stamps are affixed to the cellophane, only these packs with cellophane were used to analyze tax noncompliance behavior. After excluding the packs with no cellophane, there were 132 communities from 38 states represented in the analysis.

---

[2] See http://www.ihrp.uic.edu/study/monitoring-and-assessing-impact-tax-and-price-policies-us-tobacco-use for more information about the larger project.

[3] A community, is one in which most of the students who attend the school live. It is identified either through maps produced and available from the school district, or in a few cases (when we cannot obtain the maps and no one at the district can/will tell us) there is an algorithm that looks at location of other schools against the school district boundary area (available from census) and population data. [Bachman et al 2011]

[4] To simplify data collection and avoid arbitrary judgments in the field, collectors were told to pick up every littered pack on the route.

In most of our analyses we follow Merriman (2010) and weight each community (rather than each pack) equally. Because communities that represent a larger share of the population have a higher probability of being drawn equal weighting of communities gives estimates of population parameters (e.g. the probability of non-compliance). See Merriman (2010) for more detail.

C. Construction of Dependent variable: Tax Noncompliance

Since the tax stamp found on a littered pack shows the location of its purchase, a discrepancy between the jurisdiction identified by the tax stamp and the jurisdiction where the littered pack was found suggests tax noncompliance. We coded a pack as having tax noncompliance if the pack had no tax stamp or had a stamp from a lower-rate jurisdiction. At the community level, we measured tax noncompliance by the share of packs that avoided tax.

Both cross-border shopping by individual consumers and organized smuggling contribute to tax noncompliance. A discrepancy between the location identified by the tax stamp on a littered pack and the actual location where the pack was collected indicates noncompliance, but does not reveal how noncompliance is achieved. Inappropriate tax stamps could be found because tourists or consumers in the normal course of affairs made purchases in another (possibly distant) location; or because consumers crossed lower tax borders in order to purchase cigarettes (cross border shopping); or because someone purchased the cigarettes in a distant location and brought them to the communities in which we collected data for resale (i.e. trafficking). Merriman (2010) uses data from Chicago to demonstrate that the first explanation for inappropriate stamps can plausibly explain only a small fraction of such stamps.

We use the distance between the two locations to infer the extent to which non-compliance was due to trafficking. Because it would not be economically efficient for individual consumers to travel long distances in order to cross-border shop it is reasonable to assume that the longer the distance, the noncompliance is more likely to result from trafficking. We also expect larger tax differentials on packs with longer distance between the location of purchase and the location of consumption.

To better understand the difference between cross-border shopping and trafficking in our sample, we measured the distance from the centroid of the community where a pack was found to the nearest border of the state that issued the tax stamp on the pack. We assigned a zero distance if the pack had a stamp that indicated the location where it was found. We did not assign a distance if the pack had no stamp or an unidentifiable stamp.

We also calculated the tax differential between the jurisdiction where a pack was found and the jurisdiction that issued the tax stamp. A positive tax differential indicates monetary savings from tax noncompliance, while a negative tax differential indicates that the consumer paid higher cigarette taxes. We did not code the latter packs as being tax noncompliant. For the packs with no stamp, we calculated a tax differential based on the assumption that no state or local taxes are paid for these packs. However, it is possible that these are packs from North Carolina, South Carolina, or North Dakota, where state cigarette taxes are levied but no tax stamps are affixed on the pack. We could not calculate the differential if the pack had an unidentifiable stamp.

Table 1 reports a cross-tabulation of the number of packs by distances and tax differentials. Readers are cautioned that the data in Table 1 weights each pack (rather than each community) equally. This is necessitated by the fact that, within a community, noncompliant packs may have inappropriate stamps from a variety of origins. Because each pack is equally weighted the analyses will not necessarily be representative of the population. Nevertheless we believe the analyses are informative.

1,719 packs out of 2,116 had appropriate tax stamps that indicate the locations where they were found. 303 packs (15 percent, excluding the 61 packs that have unidentifiable stamps) indicated tax noncompliance by having no stamp or an inappropriate and low tax stamp compared to the location where they were found.

255 packs (12 percent) had an inappropriate stamp that had been transported from locations beyond 50 miles or had no stamp. That is, cross-border shopping seems to be an unlikely explanation for noncompliant stamps on 84 percent of the noncompliant packs (255 out of 303 packs)[5]. Seventy-two of the 120 packs that had noncompliant stamps (60 percent) were transported more than 50 miles and 61 (51 percent) had stamps issued by states that are more than 150 miles away. Packs with higher tax differentials also tend to have longer distances or no stamps, indicating more incentives for tax noncompliance.

D. Construction of: “Incentive for tax noncompliance”

Because Table 1 suggests that both cross-border shopping and cigarette trafficking may be important avenues for noncompliance our empirical analyses, especially Table 4, will examine both. We hypothesize that the most important variable determining cigarette trafficking will be the total tax rate in a community because transport costs will be relatively insignificant when large quantities of cigarettes are transported.

To control for cross border shopping we follow Lovenheim (2008) and hypothesize that the incentive to avoid taxes rises with the tax difference between home and nearby jurisdictions’

[5] Davis et. al. 2013 use a similar approach to estimate cigarette trafficking. In their study of littered packs in five U.S. east coast cities they find that 58.7 percent are noncompliant. They estimate that between 30.5 and 42.1 percent of all packs were the due to trafficking suggesting that between 52 and 72 percent of noncompliant packs were due to trafficking in their data.

cigarette tax and falls with distance to lower tax jurisdictions. Starting with census block level locations and tax rates, we constructed an index of the "incentive for tax noncompliance" (IFNC) for each community in our sample. IFNC measures the maximum per mile reduction in the tax paid on a pack of cigarettes by traveling to adjacent jurisdictions. For each block (i=1…N) within each community, let

$BTax_i$ = cigarette tax rate in Census block i

$ATax_j$ =cigarette tax rate in state j that is adjacent to the state that houses Census block i

d_{ij} = distance (in miles) from Census block i to the border of adjacent state j.

The cigarette tax rate includes the tax levied by state, county, and city governments, if applicable. Distance is the straight line as-the-crow-flies distance from the block centroid to the state border. For each Census block, the potential reduction in cigarette taxes paid per mile of travel is:

$$S_{ij} = \frac{\max(BTax_i - ATax_j, 0)}{d_{ij}},$$

The potential reduction in cigarette taxes paid per mile as the result of border crossing is $S_i = \underset{j}{Max}\left(S_{ij}\right)$. We use the proportion of the community population in each block, $\frac{Pop_i}{Pop_c}$, to weight S_i when we aggregate to the community level. The implication is that a block with a lower proportion of a community's population would have a lower impact on that community's potential for tax noncompliance. IFNC is measured for each community as:

$IFNC = \sum_{i=1}^{n}\left(\frac{Popi}{Popc)}\right) Si$, where n = the number of blocks in community c.

If a smoker resides too far from the state border, tax noncompliance may not take place at all because traveling to another state is too costly and time-consuming and therefore infeasible. Merriman (2010) found that a one mile increase in the distance from Chicago to the Indiana border reduced the probability of noncompliance by 0.026 ($p<.01$). In other words, the probability of tax noncompliance decreased from almost one at the Chicago/Indiana border to approximately 0 when the Indiana border was beyond 38 miles (1/0.026). To take into account this behavioral response to the distance to the state border, some of our analysis is restricted to the communities that are within 38 miles of a lower tax border.

In the vast majority of communities the average tax savings per mile traveled is less than ten cents per mile traveled as shown in panels 1 and 2 of Figure 1. Out of the entire sample of 132 communities 115 have an IFNC of ten cents or less. Only a few have an IFNC greater than twenty cents and only two communities (located quite close to a lower tax border) have an IFNC greater than 70 cents. The story is only a little different if we restrict the sample to communities within 38 miles of a tax border (Figure 1 panel 2). 51 of the 67 communities (78 percent) have an IFNC less than 10 cents.

While the absolute value of these IFNCs might seem small, it is premature to conclude that they will have no effect on behavior. An IFNC of 10 cents implies that someone living five miles from the border would save 50 cents per pack (or $5 per carton) by crossing the border. It is plausible that this differential might be enough to induce behavioral changes. Furthermore, IFNC will vary within a community so that some individuals close to a tax border will have a greater than average inducement to cross border shop. Merriman (2010) has presented evidence that small differences in proximity to lower tax cigarettes can result in relatively large differences in noncompliance. Figure 2 panels 1 and 3 show the share of the population (analyzed at a block level) by IFNC for all of our sampled communities. Panel 1 shows that about 60 percent of the population has an IFNC of only a few cents and about 99 percent of the population has an IFNC under 80 cents per mile. As shown in panel 3, roughly one percent of the population has a much higher IFNC. For some individuals that live very close to a tax border with a large discontinuity IFNC can be as high as $40 per mile. Panels 2 and 4 restrict the sample blocks to those in communities within 38 miles of a tax border. Even in these relatively close-to-the-border-communities, about 40 percent of the population faces a relatively small IFNC and almost 99 percent of the population faces an IFNC of $1 per mile or less.

These descriptive statistics make it clear that the tax policy issue most relevant to the bulk of the population is the extent to which a change in IFNC from a very low level (say under five cents per mile) to a moderate level (say ten or 20 cents per mile) will change tax noncompliance. Our empirical analyses are designed to shed light on this issue.

E. Control variables

In some of our specifications we included covariates to control for factors other than tax differentials and distance that may affect tax noncompliance. We include median household income (natural log transformed) of the community, because smokers' motivation to avoid taxes may vary with income.

We also include a measure of population density. Chernick and Merriman (2013) hypothesize that the greater density of population in New York City (NYC) may explain why poor residents of NYC seem to evade cigarette taxes at a higher rate than poor residents of Chicago. NYC's high population density brings scale economies and lowers smugglers' cost to distribute untaxed cigarettes in poor neighborhoods, and makes it more profitable to sell illegal cigarettes at the street level.

As a measure of mobility, or the cost of traveling for tax noncompliance, we include the percentage of households with cars. We also include the percentage of land used for retail or service as a measure of the local availability of cigarettes. In addition, we control for the geographic region in which the community is located to account for numerous factors such as climate and ease of auto travel that may affect propensity to avoid cigarette taxes[6]. Seven binary regional variables were included following the US Bureau of Economic Analysis classification scheme. The regions are Great Lakes, Midwest, New England, Plains, Rocky Mountain, Southeast, or Southwest. The reference group is the far west region.

IV. RESULTS

A. Basic Tables

Table 2 reports summary statistics of the variables used in the analysis, with Panel 1 and Panel 2 reporting the statistics for the full sample and "near the border" subset, respectively. The highest IFNC occurs in a large city with a very large population. The city government levies a cigarette tax in addition to the state tax rate, making the rate among the highest in the U.S. The fact that the city is close to several states facilitates noncompliance[7].

The mean of tax noncompliance is higher (.21 versus .28) in the "near the border" sample. As expected, IFNC has a higher mean (7.4 cents versus 13.8 cents) in the near the border sample. The proportion of communities in the Far West and Southwest decreases in the near the border sample, whereas the proportion of communities in New England increases.

B. Regression Analysis

In Table 3 we report estimates from a set of ordinary least squares (OLS) regressions in which the dependent variable is the proportion of packs without a tax stamp or a tax stamp from a lower-rate jurisdiction. Panel 1 presents the estimates using the full sample whereas Panel 2 presents estimates using the near the border sample. Independent variables include IFNC, median household income (log-transformed), the proportion of households with cars, population

[6] Regional dummies may also control for ease of cigarette trafficking since sources of and distance to low tax cigarettes differs by region.
[7] Confidentiality agreements prevent us from explicitly identifying the communities in which we collected data.

density (log-transformed), proportion of land used for retail and service, and a set of regional dummies. Robust standard errors are reported for all regressions.

In the first column of Panel 1 in Table 3, the coefficient on IFNC is statistically significant and indicates that a ten cent per mile increase in the tax differential will increase tax noncompliance by 3.1 percentage points. The relatively small standard error suggests that, conditional on the model assumptions; we can be 95% confident that the actual increase in tax noncompliance will be between about 2.5 and 3.7 percentage points. In Model 2, we add a variable to measure median household income and its coefficient has the expected sign—suggesting that tax noncompliance falls as median income rises—but the coefficient is measured imprecisely and is not statistically significant. The coefficient on IFNC is essentially unchanged.

Model 3 adds the share of households with cars but this variable is not a statistically significant predictor of tax noncompliance. Median household income has a statistically significant effect on tax noncompliance in this specification, suggesting that a ten percent increase in median household income is associated with a 1.2 percentage point decrease in tax noncompliance. Model 4 includes population density, which has a statistically significant negative effect on tax noncompliance. A ten percent increase in population density is associated with a four tenths of one percentage point decrease in tax noncompliance. In this specification, the effects of median household income and car ownership are statistically insignificant.

In Model 5 we also include the proportion of land used for retail or service. The coefficient on this variable is significant and has a large magnitude; a one percentage point increase in the share of land used for retail or service would result in a 3.7 percentage point decrease in tax noncompliance. None of the other control variables is associated with a statistically significant change in tax noncompliance. Model 6 adds a set of regional dummies. In this specification only median household income and IFNC are statistically significant predictors of tax noncompliance. The magnitude of the coefficient on median household income is larger compared to Model 3—a ten percent increase in median household income is associated with a 1.7 percentage point decrease in tax noncompliance.

The control variables do not have robust significant effects on the proportion of packs with no tax stamp or a lower-rate tax stamp across different models. On the other hand, IFNC is a statistically significant and quantitatively important predictor of tax noncompliance in every specification.

Panel 2 in Table 3 reports analogous estimates from applying the same models to the subset of communities within 38 miles of a state border. Specification 1 (column 1) reports the results using all 67 such communities while specifications 2 to 7 drop two communities that are "outliers" because they have very high IFNCs[8].

In these specifications IFNC is again the only consistently statistically significant independent variable. Its magnitude is consistently two to three times as large in columns (2) to (7) of panel 2 as it is in the full sample results shown in Panel 1. The explanatory power of the regressions is essentially the same in the two panels.

What do these results suggest about the impact of tax differentials on tax noncompliance behavior? In the near the border sample the mean tax noncompliance rate is about 28 percent. The coefficient on IFNC in column (7) of panel 2 of Table 3 suggests that a 10 cent increase in IFNC would cause the dependent variable—tax noncompliance—to increase by eight percentage

[8] The IFNCs in one of the omitted communities is $1.13 per mile. The IFNC in the other omitted community is $3.12 per mile. If we exclude these two outliers in OLS model using the full sample, the coefficient on IFNC is larger and about the same size as reported in Columns 2 to 7 in Panel 2 of Table 3. If we exclude these outliers from the full sample regressions reported in panel 1 of Table 3 we also get larger (and still statistically significant) coefficients on IFNC. Thus, the major difference between the coefficients on IFNC reported in panel 1 and 2 of Table 3 is the omission of the two outliers in panel 2 columns 2 to 7.

points to a noncompliance rate of about 36 percent. While this may sound quite large, readers should understand that a 50 cent increase in the cigarette tax differential would be required to generate a 10 cent increase in the IFNC of an individual living five miles from the border. An individual living 10 miles from the border would have their IFNC increase by only five cents.

C. Robustness

In this section we discuss various checks on the robustness of our results. Some of our analyses are displayed in Table 4 panel 1. Column (1) modifies the regression in column (6) of Table 3 (panel 1) by replacing median income with the poverty rate. This causes a slight decline in explanatory power (as measured by adjusted r-squared) but has almost no impact on the magnitude or significance of the coefficient on IFNC. Columns 2 and 3 of Table 4 panel 1 experiment by substituting the total tax rate and distance to a lower tax border for IFNC. This specification is, in part, motivated by the evidence presented in Table 1 which suggests that a substantial portion of noncompliance may result from trafficking. If this is the case, then the total tax may be more important than the tax relative to neighboring states.

Here we measure distance by the average of straight line as-the-crow-flies distance from the block centroid to the nearest state border with a lower tax weighted by population in each block that constitutes a community. This is in some ways a more general specification of incentives for tax noncompliance than the one used in Table 3 since it allows both distance to a tax border and tax rate to separately affect noncompliance. In columns (2) and (3) we show that the total tax rate has a positive and statistically significant association with noncompliance—a $1 increase in the tax rate will raise the noncompliance rate by eight percentage points. Distance to the border has a negative and (in column 2) statistically significant association with noncompliance but the coefficient is surprisingly small suggesting that, holding the total tax rate constant, a 100 mile increase in distance to the border would lower the noncompliance rate by only two percentage points. In these specifications the explanatory power of the model is reduced compared to specifications including IFNC which suggests that incentives for cross-border shopping are an important determinant of noncompliance.

In column (4) of Table 4 panel 1 we include IFNC in addition to the total tax rate and distance. Even with these two collinear (by construction) variables the coefficient on IFNC is statistically and economically significant and is, in fact, quite similar to the coefficient in column (1). Column (5) is identical to column (4) except that we drop the two communities with outlier (that is very high) values for IFNC. The magnitude of the coefficient on IFNC rises but the standard error also increases so that the coefficient is statistically insignificant. The coefficient on total tax rate also becomes insignificant.

Column (6) of Table 4 panel (1) reverts to the specification of Table 3 panel 1 column (6) but excludes the two outliers and includes only the 56 communities in which we found ten or more littered packs for which we could determine noncompliance status. For this sample, in which we might have the most confidence about our measures of noncompliance the coefficient on IFNC is statistically significant and its coefficient is 2.1 times as large as the coefficient in the full sample.

Table 4 panel 2 reports regressions analogous to Table 4 panel (1) (columns 1 to 4) for our close to the border sample (dropping our two outliers). In this sample replacing median income with the poverty rate causes the coefficient on IFNC to fall and become statistically insignificant. The coefficient on the poverty rate is very large and has the intuitive sign (suggesting that increases in poverty are associated with more noncompliance) but it is not statistically significant. In columns (2) and (3) of Table 4 panel 2 we replace IFNC with separate variables for the total tax rate and distance to the border. Again, this fits the data less well than including IFNC and total tax rate has the intuitive sign although it is not statistically significant in column (3). Distance to the border is statistically insignificant and has a (non-intuitive) positive sign in both specifications (2) and (3). In column (4) we include IFNC along with total tax rate and distance to the border and find that none of the individual variables has a statistically significant effect.

Overall, the results presented in Table 3 suggest caution because conclusions drawn from regressions like those presented in Table 3 may be sensitive to functional form and sample. Despite this there seems to be strong evidence that noncompliance rates vary consistently with our measure of incentives for noncompliance.

V. NON-LINEARITY BETWEEN TAX NONCOMPLIANCE AND IFNC

Although the OLS regressions in the previous two sections provide strong evidence that financial incentives to avoid cigarette taxation are a very important determinant of behavior they have limited direct relevance for policymakers in any particular jurisdiction. Policymakers must assess the impact of a cigarette tax increase, or neighboring jurisdictions' tax increases on noncompliance, consumption and revenue in their home jurisdiction. The OLS results suggest that there will be some reaction but in any particular jurisdiction magnitude depends on relative current tax rates and proximity to lower tax jurisdictions. For example, if a jurisdiction already has much higher taxes than its neighbors a tax increase might cause little additional noncompliance because those who can easily avoid will already have done so.

The appropriate econometric treatment is to allow our estimates of the relationship between IFNC and noncompliance to vary non-linearly with the level of IFNC. In Table 5 we report the results of binomial logit estimation that relates our dependent variable (number of packs that either had no stamp or a stamp from a lower tax jurisdiction) to the independent variables used in the last column Table 3 controlling for the number of packs found in each community. Coefficients are estimated by maximum likelihood. This specification conforms to the a priori constraint that predicted noncompliance is always between 0 and 1 and is flexible enough to allow the marginal effect of IFNC on tax noncompliance to vary with its level. The parameter estimates in Table 5 are consistent with those in Tables 3 and 4—IFNC is a statistically significant determinant of noncompliance.

Because of the nonlinear specification, the magnitudes of the estimated coefficients in Table 5 are not very informative but we can use the coefficients to produce enlightening simulations that relate tax policy to tax revenue and noncompliance. In particular, we use the estimated coefficients in Table 5 to estimate how changes in noncompliance and tax revenue vary with tax increases at various levels of tax and proximity to lower tax borders.

Recall that the incentive to avoid depends upon the savings per mile traveled to lower tax borders as well as the distribution of the population within the community, i.e. $IFNC = \sum_{i=1}^{n} \left(\frac{Pop_i}{Pop_c} \right) S_i$. Taking the derivative of IFNC with respect to the local tax rate implies that $\frac{\partial IFNC_C}{\partial t_i} = \sum_{i=1}^{N} \frac{Pop_i}{Pop_c} \frac{\partial S_i}{\partial t_i} = \sum_{i=1}^{N} \frac{Pop_i}{Pop_c} \left(\frac{1}{d_{ij}} \right)$ so the change in noncompliance incentives as a result of a tax increase will depend upon the distance to neighboring jurisdictions as well as the distribution of the population. We calculate this derivative for each of the 160 communities visited by data collectors and display the results in figure 3 below. The figure shows that only ten of the 160 communities have a derivative greater than 0.3. That is in all but ten communities a one cent tax increase will cause IFNC to increase by less than 0.3 cents. This is because, in these communities the average person lives more than 3.3 miles (=1/.3) miles from a lower tax border. This also explains why, as shown in figure 2, the vast majority of the population has an IFNC of less than ten cents per mile.

Using this information and the parameter estimates reported in column 2 of Table 5 we simulate the impact of tax increases on noncompliance and tax revenue. Our results are reported in Table 6. For illustrative purposes the table lists IFNC's ranging from 15 to 25 cents and 45 to 55 cents. Recall from table 3 that the mean IFNC in our close to the border sample was just under 14 cents and the median was much lower (a little less than 1 penny). We choose to illustrate our findings with much higher IFNCs because these are the communities that are most at risk of increased noncompliance when tax rates go up. Clearly, the vast majority of communities with lower levels of IFNC will face an even lower threat of noncompliance if they raise their taxes.

For each level of IFNC we use the parameter estimates reported in Table 5 evaluated at the mean level of all variables except IFNC to compute the predicted share and standard error of noncompliant packs. The simulated share of noncompliant packs varies from 15.3 percent at an IFNC of 15 cents to 59.3 percent at an IFNC of 55 cents. Using this prediction we calculate the share of packs that have paid the tax (column 4) and the predicted percentage decline in tax paid packs due to noncompliance (column 5) when IFNC rises by one cent per mile.

The impact of increased noncompliance on tax revenue will depend on the amount that taxes must rise to cause a one cent increase in IFNC. This, in turn will depend upon the geographic distribution of the population within the jurisdiction. As we have shown in Figure 3, in the vast majority of areas that we sampled IFNC will rise little when taxes increase because most of the population lives relatively far from the lower tax border.

How will increases in taxes and hence IFNC affect tax revenues? This depends upon the current tax rate and price/tax elasticity of demand. As shown in column (5) marginal noncompliance rates rise with IFNC. To better understand the implied relationship between cigarette tax increases and tax revenue we used our estimated results to calculate the maximum tax rate at which tax revenue would increase[9] at each level of IFNC. The higher the starting tax rate the more revenue is lost when noncompliance increases due to an increase in IFNC.

For illustrative purposes, we make the conservative assumption that a five cent tax increase is required to raise IFNC by a penny. This is equivalent to assuming that the average person in the community lives five miles from the lower tax border. Using this assumption we calculate the highest tax at which a five cent tax increase (i.e. a one cent per mile increase in IFNC) will generate an increase in revenue despite the predicted increase in noncompliance. This is displayed in column (6)[10]. So, for example, we predict that a community with an IFNC of 15 cents per mile would increase revenue when it increased its tax rate by 5 cents per pack as long as its starting tax rate was under $7.72. Since no community in the US has a tax rate this high our predictions imply that any community with an IFNC under 15 will increase revenue by increasing the cigarette tax by five cents.

As IFNC rises, the percentage decline in fully tax paid packs increases with increases in IFNC so that the maximum tax rate at which revenue rises with a tax increases falls. For communities with an IFNC of 55 cents (i.e. for communities that have a tax rate much higher than their very close neighbors) we predict that revenue will fall with a five cent tax increase for any tax rate above $2.06.

---

[9] This is also the minimum tax rate at which revenue would be predicted to fall.

[10] Calculations in column (6) of Table 5 ignore changes in tax revenue due to changes in consumption. This makes little difference to the reported results since the simulated 5 cent price increase will be very small relative to the total price of cigarettes. This is demonstrated more formally in appendix 1.

VI. CONCLUSION

Tobacco control policies are designed to reduce consumption not just tax paid sales. Although data on tax paid cigarette sales is often readily available it may be a poor proxy for cigarette consumption. Empirical estimates of cigarette demand functions have been important tools in the discussion of tobacco control policies but have been plagued by technical problems associated with unobserved consumption of low or no tax cigarettes. Our data on a national sample of littered cigarette packs provides information about tax noncompliance in a large variety of areas and circumstances. Our statistical analyses provide strong evidence that tax noncompliance increases with tax differentials and proximity to low tax cigarettes. Empirically our proxy for incentives for tax noncompliance—cents per mile saved by the average person in the community—explains the variance in measured noncompliance rates better than any of the other potential explanatory variables we examine. We show however that for the vast majority of people in our sample communities potential tax savings from cross border shopping is only a few cents per mile. We also show that while increases in cigarette tax rates may result in increased noncompliance, the increase in noncompliance is low enough that tax revenues will rise in almost all cases. Taken together, this evidence suggests that even in communities with quite high cigarette tax rates and close proximity to adjacent areas with relatively low tax rates cigarette tax, increases are a viable tool for increasing revenues.

APPENDIX

Equations underlying calculations in table 5

R = revenue

t = per pack tax in cents

P = pre tax price of pack of cigarettes (assumed to be $3.50)

Q = tax paid sales

$C =$ consumption

$\varepsilon =$ price elasticity of cigarette consumption

$\sigma(t) = \dfrac{Q(t)}{C(t)}$ is based on simulations reported in column 4 of table 6

(1) $$\varepsilon = \frac{\partial C}{C}\frac{(P+t)}{\partial(P+t)} \approx \frac{\%\Delta C}{\%\Delta(P+t)} \approx \frac{\%\Delta C}{\%\Delta(3.50+t)} \approx \frac{\%\Delta C}{\left(\dfrac{.05}{(3.50+t)}\right)}$$

(2) $$R(t) = Q*t = C(t)*\sigma(t)*t$$

(3) $$R(t+5) = Q(t+5)*(t+5) = C(t+5)*\sigma(t+5)*(t+5)$$

(4) $$\Delta R \equiv R(t) - R(t+5) = \left[C(t)*\sigma(t)*(t)\right] - \left[C(t+5)*\sigma(t+5)*(t+5)\right]$$

(5) $$\Delta R = \left[C(t)*\sigma(t)*(t)\right] - \left[C(t)*(1+\%\Delta C)*\sigma(t+5)*(t+5)\right]$$

(6) $$\Delta R = \left[C(t)*\sigma(t)*(t)\right] - \left[C(t)*\left(1+\left(\varepsilon*\left(\frac{.05}{3.50+t}\right)\right)\right)*\sigma(t+5)*(t+5)\right]$$

(7) $$\Delta R = C(t)\left[\sigma(t)*(t)\right] - \left[\left(1+\left(\varepsilon*\left(\frac{.05}{3.50+t}\right)\right)\right)*\sigma(t+5)*(t+5)\right]$$

Case 1: Assume $\varepsilon = 0$ then equation (7) implies that tax revenue is maximized at the t that satisfies

(8) $$\Delta R = 0 = C(t)*\left[\sigma(t)*(t) - \left[\sigma(t+5)*(t+5)\right]\right]$$

(9) $$0 = \left[\sigma(t)*(t) - \left[\sigma(t+5)*(t+5)\right]\right] \Rightarrow \sigma(t)*(t) = \sigma(t+5)*(t+5)$$

(10) $$\left[\sigma(t) - \sigma(t+5)\right]*(t) = \sigma(t+5)*5 \Rightarrow t = \frac{\sigma(t+5)*5}{\left[\sigma(t) - \sigma(t+5)\right]}$$ (this is column 6)

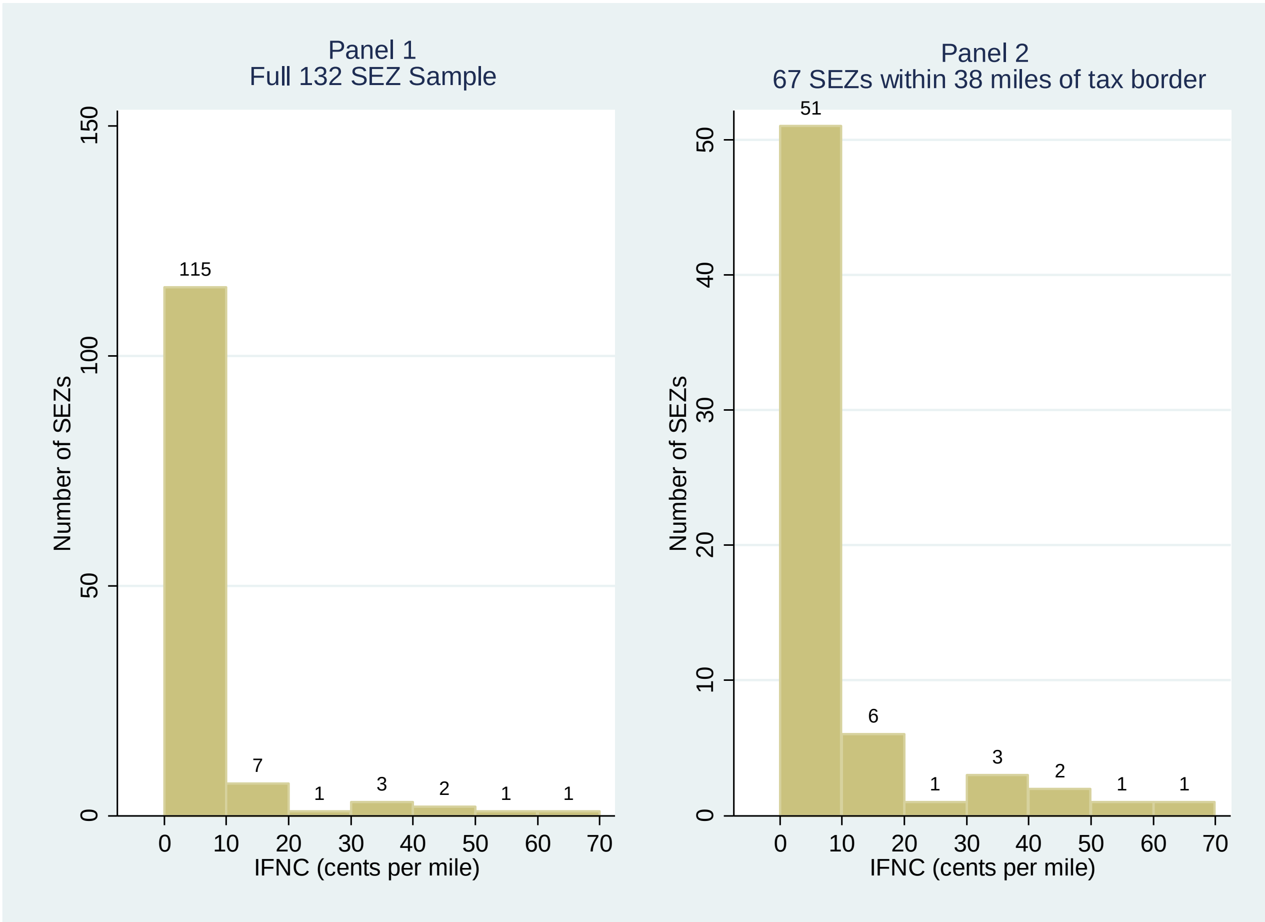
Case 2: Assume $\varepsilon \neq 0$ then equation (7) implies that tax revenue is maximized at the t that satisfies

(11) $$\Delta R = 0 = C(t)\left[\sigma(t)*(t)\right] - \left[\left(1+\left(\varepsilon*\left(\frac{.05}{3.50+t}\right)\right)\right)*\sigma(t+5)*(t+5)\right]$$

Let $\lambda(t) = \left(1+\left(\varepsilon*\left(\dfrac{.05}{3.50+t}\right)\right)\right)$. Note that $\lambda(0) = 1+(\varepsilon*0.014) \approx 1$ if $-1 \le \varepsilon \le 0$.

Since $\lambda(t)$ goes to 1 as t goes to ∞. For all practical purposes we can assume that $\lambda(t) \approx 1$ and equation (11) is equivalent to equation (8). Bottom line is that we can essentially ignore changes in consumption as is done when calculating column 6 of table 5.

Figure 1: Number of catchments by IFNC



All communities have IFNC lower than 70 cents, except two communities that have IFNC of 113 cents and 312 cents, respectively. The graphs here exclude these two outliers.

Figure 2: IFNC by Cumulative Share of Population

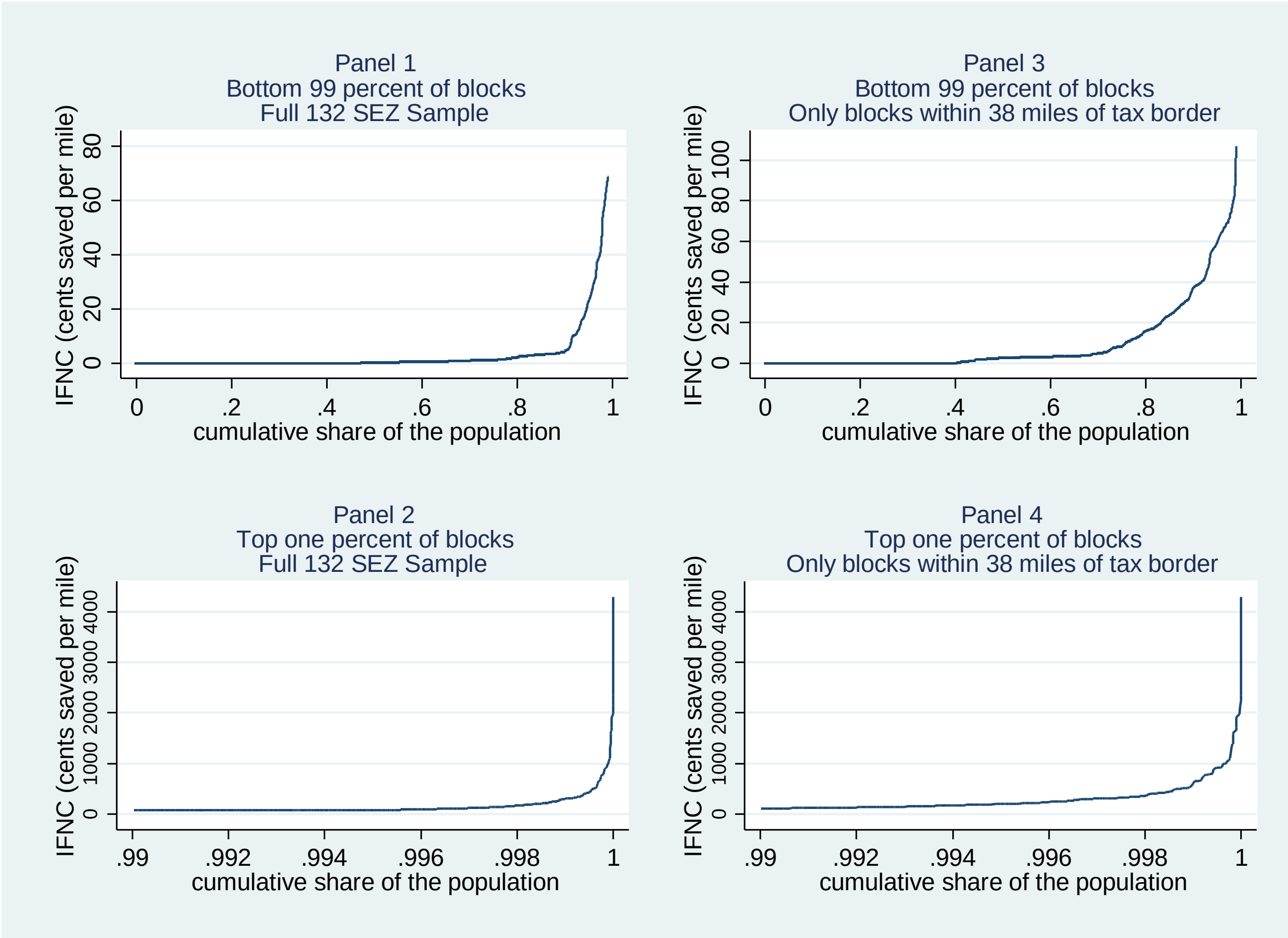
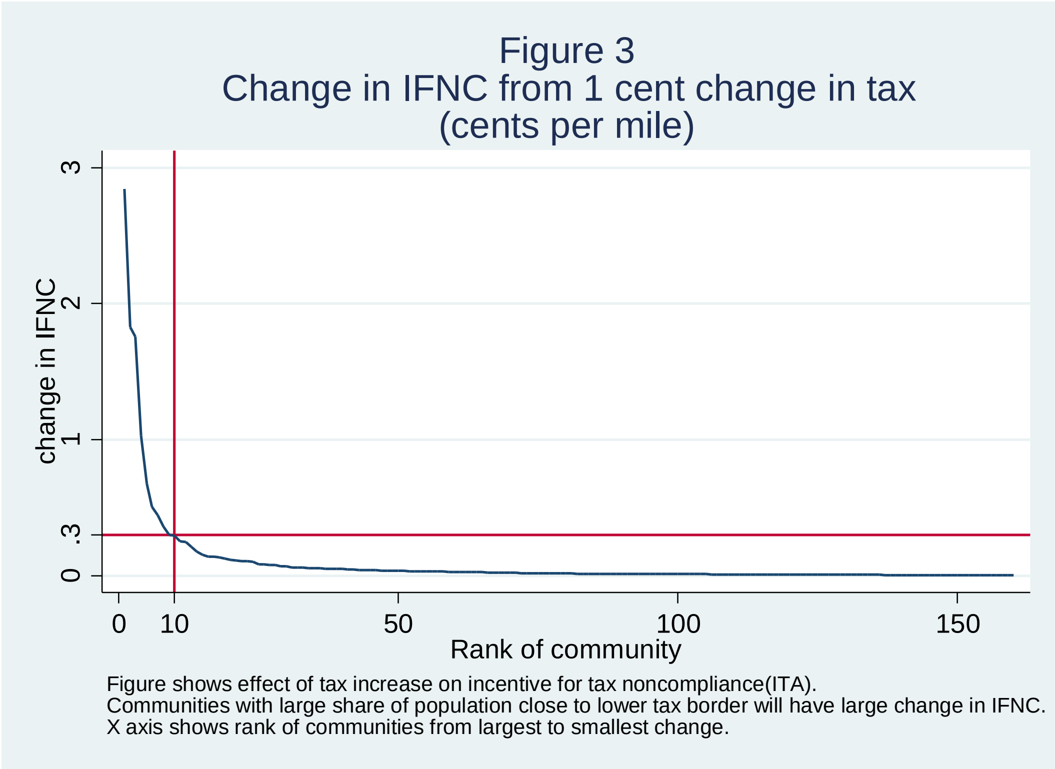


Figure 3: Change in IFNC from One-Cent Change in Tax (Cents per Mile)



Notes: Table shows simulated share of non-compliant littered packs at several levels of IFNC. Simulations derived from coefficients estimated using maximum likelihood binomial logit procedure that relates number of packs that either had no stamp or a stamp from a lower tax jurisdiction to the independent variables used in Table 2 column 7.

Table 1: Number of packs by tax differential and distance between purchase origin and sales destination

(Estimates unweighted)

	Tax differential (TD)							
	TD<0	**TD=0**	**0<TD<5c**	**5c <TD <25c**	**25c <TD < 50c**	**50c < TD < $1**	**TD>$1**	**Total**
d=0	0	1,719	0	0	0	0	0	**1,719**
0<d<=20	19	0	4	4	1	22	5	**55**
20<d<=50	1	0	0	0	1	4	7	**13**
50<d<=100	2	0	0	0	1	0	4	**7**
100<d<=150	0	0	1	1	0	0	4	**6**
d>150	11	0	1	6	3	4	47	**72**
No stamp(1)	0	0	0	0	31	51	101	**183**
Unidentifiable, possibly foreign(2)	Unknown							**61**
Total	**33**	**1,682**	**35**	**13**	**40**	**88**	**225**	**2,116**
	1.56%	79.49%	1.65%	0.61%	1.89%	4.16%	10.63%	

TD is the difference in cigarette tax rate between the jurisdiction where a pack was found and the jurisdiction that issued the stamp on the pack. TD equals to zero if the pack had the appropriate stamp that is consistent with the location it was found. A positive value of TD indicates the cents of tax saved through noncompliance. A negative tax differential indicates that the consumer paid higher cigarette taxes, and the difference in two locations is due to reasons such as travel. We do not count these packs as tax noncompliance.

The distance is measured from the centroid of the community where a pack was found the nearest border of the state that issued the tax stamp on the pack.

(1) For the packs with no stamp, distance from purchase origin to destination is unknown. We calculated tax differential based on the assumption that no state or local taxes are paid for these packs. However, it is possible that these are packs from North Carolina, South Carolina, or North Dakota, where state cigarette taxes are levied but no tax stamps are affixed on the pack.

(2) There are 61 packs that have stamps unidentifiable. These stamps can possibly be foreign stamps. Tax differential is thus unknown because we don't know where the packs were bought.

Table 2: Summary Statistics

Panel 1: all communities

	N	Mean	Median	SD	Min	Max
Tax noncompliance (share of packs)	132	0.21	0.11	0.28	0	1
IFNC (cents per mile)	132	7.43	0.63	30.38	0	312.51
Total tax rate ($s per pack)	132	1.60	1.34	1.22	0.17	5.85
Distance to tax border (miles)	132	155.42	87.16	213.99	0.53	1365.58
Number of packs per community	132	16	9	18	1	83
Median income ($s)	132	61,832	57,580	24,464	25,000	157,690
Share of households with cars	132	0.93	0.95	0.07	0.56	1
Population density (people per sq. mile)	132	2,613	1,015	4,931	21	31,145
share of retail/service land use	132	0.32	0.32	0.15	0.04	0.77
Poverty rate	132	0.12	0.11	0.08	0.03	0.41
Far West	132	0.14	0	0.35	0	1
Great Lakes	132	0.17	0	0.37	0	1
Midwest	132	0.16	0	0.37	0	1
New England	132	0.05	0	0.23	0	1
Plains	132	0.06	0	0.24	0	1
Rocky Mountain	132	0.02	0	0.15	0	1
Southeast	132	0.27	0	0.44	0	1
Southwest	132	0.13	0	0.34	0	1

Panel 2: only communities within 38 miles from the state border

	N	Mean	Median	SD	Min	Max
Tax noncompliance (share of packs)	67	0.28	0.15	0.34	0	1
IFNC (cents per mile)	67	13.78	0.86	41.81	0	312.51
Total tax rate ($s per pack)	67	1.883	1.600	1.484	0.170	5.850
Distance to tax border (miles)	67	16.24	14.21	10.65	0.20	35.01
Number of packs per community	67	17	11	17	1	76
Median income ($s)	67	61,440	53,842	27,596	25,000	157,690
share of households with cars	67	0.92	0.95	0.08	0.56	0.99
Population density (people per sq. mile)	67	2,660	665	5,901	24	31,145
Share of retail/service land use	67	0.33	0.31	0.14	0.07	0.77
Poverty rate	67	0.12	0.12	0.07	0.03	0.34
Far West	67	0.05	0	0.21	0	1
Great Lakes	67	0.16	0	0.37	0	1
Midwest	67	0.28	0	0.45	0	1
New England	67	0.10	0	0.31	0	1
Plains	67	0.06	0	0.24	0	1
Rocky Mountain	67	0.02	0	0.12	0	1
Southeast	67	0.31	0	0.47	0	1
Southwest	67	0.02	0	0.12	0	1

To construct tax noncompliance and IFNC, we collected littered packs from each SEZ, and compared the jurisdiction that issued the tax stamp on each pack with the community location. Distance to tax border is an average of straight line as-the-crow-flies distance from the block centroid to the nearest lower-rate state border weighted by population in each block that constitutes a community. Both the geographical distance and population were collected from US Census. Total tax rates were collected from government web sites. Median income, share of households with cars, population density, share of retail/service land use, and poverty rate are collected from US Census.

Table 3: Determinants for the Share of Packs with Tax Noncompliance

Panel 1: All 132 communities included

	(1)	(2)	(3)	(4)	(5)	(6)
IFNC	0.0032***	0.0033***	0.0034***	0.0035***	0.0034***	0.0029***
	(0.0056)	(0.0062)	(0.0073)	(0.0066)	(0.0068)	(0.0074)
Ln(Median_income)		-0.1032	-0.1296*	-0.0278	-0.059	-0.1709**
		(0.0639)	(0.0732)	(0.0707)	(0.0740)	(0.0851)
share of households with cars			0.363	-0.2531	-0.5549	0.0605
			(0.3550)	(0.3537)	(0.3976)	-0.4713
Ln(Population density)				-0.0379**	-0.0248	-0.0281
				(0.0160)	(0.0151)	(0.0175)
share of retail/service land use					-0.3707*	-0.2338
					(0.2068)	(0.1907)
Constant	0.1876***	1.3183*	1.2689*	0.9725	1.6302**	2.3153***
	(0.0235)	(0.7093)	(0.6835)	(0.6341)	(0.8008)	(0.8262)
Region dummies	No	No	No	No	No	Yes
R squared	0.107	0.118	0.118	0.148	0.159	0.184

Panel 2: Only communities within 38 miles from the state border

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IFNC	0.0027***	0.0053*	0.0062**	0.0091***	0.0089***	0.0088***	0.0082**
	(0.0004)	(0.0027)	(0.0024)	(0.0027)	(0.0024)	(0.0025)	(0.0032)
Ln(Median_income)			-0.1701*	-0.2404**	-0.0552	-0.0780	-0.1850
			(0.0968)	(0.1024)	(0.0998)	(0.1028)	(0.1492)
% of households with cars				1.091**	0.0320	(0.5039)	0.0650
				(0.4173)	(0.4897)	(0.5111)	(0.7636)
Ln(Population density)					-0.0736**	-0.0537*	-0.0518
					(0.0284)	(0.0288)	(0.0369)
% of retail/service land use						-0.6588*	-0.3894
						(0.3414)	(0.3540)
Constant	0.2432***	0.2261***	2.0781*	1.8247*	1.2454	2.0765**	2.7612**
	(0.0404)	(0.0432)	(1.0708)	(1.0010)	(0.8780)	(0.9944)	(1.0625)
Region dummies	No	No	No	No	No	No	Yes
R squared	0.103	0.041	0.069	0.112	0.179	0.201	0.170
N	67	65	65	65	65	65	65

Robust standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01
Region dummy variables: reference group="Far West"
Two observations omitted in Panel 2.
One omitted community has an IFNC of 113. The other omitted community has an IFNC of 312.

Table 4: Robustness Check with Alternative Predictors

Panel 1: All communities included & communities with number of packs >=10

	(1)	(2)	(3)	(4)	(5)	(6)
IFNC	0.0030***			0.0025***	0.0054	0.0061**
	(0.0006)			(0.0004)	(0.0032)	(0.0020)
Total tax rate		0.0800**	0.0736*	0.0590*	0.0413	
		(0.0286)	(0.0305)	(0.0287)	(0.0286)	
Distance to border		-0.0002*	-0.0002*	-0.0002	-0.0002	
		(0.0001)	(0.0001)	(0.0001)	(0.0001)	
Ln(Median_income)		-0.2382*				-0.1078
		(0.0931)				(0.1163)
Poverty rate	0.2533		0.5124	0.5459	0.6182	
	(0.3435)		(0.3973)	(0.3756)	(0.3883)	
% of households with cars	-0.2829	0.6480	0.2596	0.1728	0.3347	0.2772
	(0.4490)	(0.5065)	(0.5157)	(0.4464)	(0.4966)	(0.6037)
Ln(Population density)	-0.0417*	-0.0188	-0.0350*	-0.0388*	-0.0365*	-0.0155
	(0.0176)	(0.0184)	(0.0174)	(0.0168)	(0.0170)	(0.0219)
% of retail/service land use	-0.2088	-0.2577	-0.2413	-0.2412	-0.2472	0.1721
	(0.1878)	(0.1986)	(0.1968)	(0.1931)	(0.1924)	(0.2215)
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes
R squared	0.1536	0.1661	0.1324	0.1958	0.1433	0.2177
N	132	132	132	132	130	56

Panel 2: Only communities within 38 miles from the state border

	(1)	(2)	(3)	(4)
IFNC	0.0059			0.0064
	(0.0036)			(0.0038)
Total tax rate		0.0958*	0.0620	0.0199
		(0.0384)	(0.0380)	(0.0365)
Distance to border		0.0032	0.0033	0.0053
		(0.0038)	(0.0036)	(0.0037)
Ln(Median_income)		-0.2078		
		(0.1565)		
Poverty rate	-0.8401		-0.7490	-0.5836
	(0.7343)		(0.8118)	(0.7937)
% of households with cars	-1.1581	-0.0885	-1.3062	-0.9899
	(0.7915)	(0.8132)	(0.7863)	(0.8046)
Ln(Population density)	-0.0861*	-0.0504	-0.0850*	-0.0877*
	(0.0369)	(0.0392)	(0.0382)	(0.0379)
% of retail/service land use	-0.4018	-0.4610	-0.4457	-0.3538
	(0.3479)	(0.3316)	(0.3555)	(0.3461)
Region dummies	Yes	Yes	Yes	Yes
R squared	0.1635	0.159	0.145	0.1640
N	65	65	65	65

Robust standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01
In column 5 &6 of Panel 1, and all columns in Panel 2, two observations omitted in Panel 2. One omitted community has an IFNC of 113. The other omitted community has an IFNC of 312. Column 6 of Panel 1 sample includes only communities in which 10 or more useable littered packs were found.

Table 5: Binomial Logit on the Share of Packs with Tax Noncompliance

	All communities (1)	Only communities within 38 miles from a state border (2)
IFNC	0.0447***	0.0417***
	(0.0083)	(0.0088)
Ln(Median_income)	-0.7048	-0.0858
	(0.5681)	(0.6775)
% of households with cars	2.2842	-0.7499
	(2.7460)	(3.9570)
Ln(Population density)	-0.1632*	-0.2081
	(0.0990)	(0.1789)
% of retail/service land use	0.5508	-0.3615
	(1.1281)	(2.9620)
Region dummies	Yes	Yes
N	130	65

Robust standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01
Two observations omitted. One community has an IFNC of 113
the other omitted community has an IFNC of 312

Table 6: Simulations of the relationship between five cent tax increase, predicted noncompliance and revenue at various levels of IFNC

(1) Level of IFNC (cents per mile)	(2) Share of packs that do not comply	(3) Standard error	(4) Share of tax paid packs (1-column (2))	(5) percentage decline in quantity due to increased noncompliance	(6) highest tax at which 5 cent tax increase will increase revenue[1]
15	0.153	0.049	0.847		
16	0.159	0.051	0.841	-0.65%	$7.72
17	0.164	0.052	0.836	-0.67%	$7.45
18	0.170	0.054	0.830	-0.69%	$7.20
19	0.176	0.056	0.824	-0.72%	$6.95
20	0.182	0.057	0.818	-0.74%	$6.72
21	0.188	0.059	0.812	-0.77%	$6.49
22	0.195	0.061	0.805	-0.80%	$6.28
23	0.202	0.063	0.798	-0.82%	$6.07
24	0.208	0.065	0.792	-0.85%	$5.87
25	0.215	0.067	0.785	-0.88%	$5.68
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
45	0.490	0.133	0.510		
46	0.500	0.134	0.500	-2.04%	$2.45
47	0.510	0.136	0.490	-2.09%	$2.40
48	0.521	0.137	0.479	-2.13%	$2.35
49	0.531	0.138	0.469	-2.17%	$2.30
50	0.542	0.140	0.458	-2.21%	$2.26
51	0.552	0.141	0.448	-2.26%	$2.22
52	0.562	0.142	0.438	-2.30%	$2.18
53	0.573	0.142	0.427	-2.34%	$2.14
54	0.583	0.143	0.417	-2.38%	$2.10
55	0.593	0.144	0.407	-2.42%	$2.06

[1]Calculations assume five cent change in tax increases IFNC by 1 cent. This is equivalent to assuming that the average resident lives five miles from the lower tax border. Simulations use parameter estimates reported in column 2 of Table 4. All variables except IFNC are held at their mean level in the close to the border sample. See appendix 1 for derivation of calculations in column (6). We assume no change in adjacent areas tax rates and no change in consumption. Appendix 1 demonstrates that the assumption of no change in consumption is inconsequential.

DISCLOSURES

This research is supported by the National Cancer Institute of the National Institutes of Health under the State and Community Tobacco Control Initiative, grant number 5U01-CA154248 and the Robert Wood Johnson Foundation, under the Bridging the Gap: Research Informing Practice and Policy for Healthy Youth Behavior Initiative, grant numbers 64702 and 70175. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors have no financial arrangements that might give rise to conflicts of interest with respect to the research reported in this paper.

REFERENCES

Bachman, Jerald G.; Lloyd D. Johnston; Patrick M. O'Malley and John E. Schulenberg. 2011. "The Monitoring the Future Project after Thirty-Seven Years: Design and Procedures," *Monitoring the Future Paper Series.* Ann Arbor, MI: University of Michigan, Institute for Social Research.

Baltagi, Badi H. and Dan Levin. 1986. "Estimating Dynamic Demand for Cigarettes Using Panel Data: The Effects of Bootlegging, Taxation and Advertising Reconsidered." *The Review of Economics and Statistics* 68(1), 148-55.

Barker, Dianne; Shu Wang; David Merriman; Andrew Crosby; Elissa Resnick and Frank Chaloupka. Forthcoming. "Estimating Cigarette Tax Avoidance and Evasion: Evidence from a National Sample of Littered Packs." *Tobacco Control*.

Becker, Gary. S.; Michael Grossman and Kevin M. Murphy. 1991. "Rational Addiction and the Effect of Price on Consumption." *American Economic Review*, 81(2), 237-41.

Lakhdar, Ben C. 2008. "Quantitative and Qualitative Estimates of Cross-Border Tobacco Shopping and Tobacco Smuggling in France." *Tobacco Control,* 17(1), 12-16.

Chaloupka, F. 1991. "Rational Addictive Behavior and Cigarette-Smoking." *Journal of Political Economy*, 99(4), 722-42.

Chernick, Howard and David Merriman. 2011. "Using Littered Pack Data to Estimate Cigarette Tax Avoidance in NYC." *National Tax Journal*, 66 (3), 635-68.

Chiou, Lesley and Erich Muehlegger. 2008. "Crossing the Line: Direct Estimation of Cross-Border Cigarette Sales and the Effect on Tax Revenue." *B.E. journal of economic analysis and policy: Contribution to Economic Analysis & Policy,* 8(1), 1-39.

Crosby, Andrew; David Merriman and Shu Wang. 2014. *How Much Cigarette Tax Avoidance Is There in the Us? Different Estimates from Smoker Surveys and Physical Pack Collection Methods*. Unpublished Manuscript. University of Illinois at Chicago.

Davis, Kevin C.; Victoria Grimshaw; David Merriman; Matthew C. Farrelly; Howard Chernick; Micaela H. Coady; Kelsey Campbell and Susan M. Kansagra. 2013. "Cigarette Trafficking in

Five Northeastern Us Cities." *Tobacco Control*, 1-7.

DeCicca, Philip; Donald Kenkel and Feng Liu. 2013. "Excise Tax Avoidance: The Case of State Cigarette Taxes." *Journal of Health Economics*, 32(6), 1130-41.

Fix, Brian. V.; Andrew Hyland; Richard. J. O'Connor; K Michael Cummings; Geoffrey. T. Fong; Frank. J. Chaloupka and Andrea. S. Licht. 2014. "A Novel Approach to Estimating the Prevalence of Untaxed Cigarettes in the USA: Findings from the 2009 and 2010 International Tobacco Control Surveys." *Tobacco Control*, 23 Suppl 1, i61-66.

Galbraith, John W. and Murray Kaiserman. 1997. "Taxation, Smuggling and Demand for Cigarettes in Canada: Evidence from Time-Series Data." *Journal of Health Economics*, 16(3), 287-301.

Harding, Matthew; Ephraim Leibtag and Michael F. Lovenheim. 2012. "The Heterogeneous Geographic and Socioeconomic Incidence of Cigarette Taxes: Evidence from Nielsen Homescan Data." *American Economic Journal: Economic Policy*, 4(4), 169-98.

Institute of Medicine, National Research Council. 2015. *Understanding the U.S. Illicit Tobacco Market: Characteristics, Policy Context, and Lessons from International Experience*. Washington DC: The National Academies Press.

Kurti, Marin. K.; Klaus. von Lampe and Douglas. E. Thompkins. 2013. "The Illegal Cigarette Market in a Socioeconomically Deprived Inner-City Area: The Case of the South Bronx." *Tobacco Control*, 22(2), 138-40.

Lovenheim, Michael F. 2008. "How Far to the Border?: The Extent and Impact of Cross–Border Casual Cigarette Smuggling." *National Tax Journal*, 61(1), 7-33.

Merriman, David. 2010. "The Micro-Geography of Tax Avoidance: Evidence from Littered Cigarette Packs in Chicago." *American Economic Journal: Economic Policy*, 2(2), 61-84.

Pesko, Michael F.; John A. Tauras; Jidong Huang and Frank J. Chaloupka. 2016. "The Influence of Geography and Measurement in Estimating Cigarette Price Responsiveness, National Bureau of Economic Research Working Paper, w22296

Saba, Richard R.; T. Randolph Beard; Robert B. Ekelund and Rand W. Ressler. 1995. "The Demand for Cigarette Smuggling." *Economic Inquiry*, 33(2), 189-202.

Thursby, Jerry G. and Marie C. Thursby. 2000. "Interstate Cigarette Bootlegging: Extent, Revenue Losses, and Effects of Federal Intervention." *National Tax Journal* 53(1), 59-77.