I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15525

Tax Elasticity of Border Sales: A Meta-Analysis

Dat Huynh
Anna Sokolova
Mehmet S. Tosun

AUGUST 2022

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15525

Tax Elasticity of Border Sales: A Meta-Analysis

Dat Huynh
University of Nevada Reno

Anna Sokolova
University of Nevada Reno and HSE University

Mehmet S. Tosun
University of Nevada Reno and IZA

AUGUST 2022

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Tax Elasticity of Border Sales: A Meta-Analysis*

When regions in close proximity have different tax rates, residents may engage in cross-border shopping and take advantage of tax differentials. The extent of this activity can be captured by the tax elasticity of border sales (TEBS). We collect 749 estimates of TEBS reported in 60 studies, and conduct the first meta-analysis of this literature. We show that the literature is prone to selective reporting: positive estimates of TEBS are systematically discarded—this biases the mean reported estimate away from the 'true' underlying effect. Reported estimates also vary widely; we construct 29 control variables that capture empirical strategies used to obtain them, and employ Bayesian Model Averaging to pin down the sources of this variation. We find that sales of food, retail and fuel are more elastic compared to sales of tobacco and other individual 'sin' products; that while the cross-border shopping is prominent in the US, it is much less prevalent in Europe and other countries.

JEL Classification: H71, H73, H26, H22, R12, H31

Keywords: cross-border shopping, taxation, tax elasticity, meta-analysis, Bayesian Model Averaging

Corresponding author:
Mehmet S. Tosun
Department of Economics
University of Nevada-Reno
Mail Stop 0030
Reno, Nevada 89557
USA

E-mail: tosun@unr.edu

* Data and code supporting results in this study are available from the authors upon request. Mehmet Tosun acknowledges support from a grant from the Donald W. Reynolds Foundation, in honor of Barbara Smith Campbell. Anna Sokolova acknowledges support from the Basic Research Program at the HSE University.

1 Introduction

In countries with decentralized fiscal systems, jurisdictions have authority to determine their unique tax policies, adjusting tax rates to raise tax revenues or discourage consumption of certain products (e.g. cigarettes). But the desired outcomes of tax policy changes may be jeopardized when individuals respond by traveling to shop at a neighboring jurisdiction, taking advantage of tax differentials. When such cross-border shopping is prominent, an increase in, for example, a cigarette tax in a jurisdiction can result in more tax avoidance while having little effect on cigarette consumption or tax revenues.

There is a considerable literature within public finance that studies the link between such cross-border sales and taxation, with many studies reporting estimates that can capture the elasticity of border sales to changes in tax rates, which we refer to as 'tax elasticity of border sales' (TEBS, for brevity). Examples of studies on cross-border sales and taxation are Tosun & Skidmore (2007), Goel & Nelson (2012), Chernick & Merriman (2013), Jansen & Jonker (2018), Hansen *et al.* (2020).[1] However, the reported estimates vary widely; they are also context specific, and it is not clear to what extent estimates pertaining to a tax change in one region can be used to understand the response to a different tax change in another region, at a different point in time. Yet, knowing TEBS is crucial for addressing a variety of theoretical and empirical questions in public finance, from the behavioral response of individuals to taxation, to tax incidence and the impact of taxation on individuals' welfare.

We conduct the first meta-analysis of the literature estimating tax elasticity of border sales. Unlike previous studies in this field, we address and reconcile the diversity of results reported by the empirical literature estimating TEBS. First, we examine the sources of variation in TEBS estimates. We collect 749 estimates reported by 60 studies in the field. For each estimate we construct a comprehensive set of explanatory variables detailing features of the underlying data and methodology employed. We then use meta-regressions to pin down sources of systematic variation in estimates, addressing model uncertainty with Bayesian Model Averaging. Meta-regressions have previously been used to detect systematic patterns in estimates produced by a number of empirical literatures, examining the effects of minimum wage on employment (Card & Krueger 1995, Doucouliagos & Stanley 2009), of distance on trade (Disdier & Head 2008), of IMF programs on growth

[1] Tosun & Skidmore (2007) use data on food sales and sales tax rates for West Virginia counties to estimate TEBS over the 1988-1991 period. Chernick & Merriman (2013) study the response of cross-border cigarette sales in New York City before and after New York State raised its cigarette tax rate in June 2008. Jansen & Jonker (2018) examine the responses of Dutch drivers to the variation in cross-border fuel price differences. Hansen *et al.* (2020) study the cross-border sales of marijuana along the Washington-Oregon border.

(Balima & Sokolova 2021); literatures that study elasticities of labor supply (e.g. Chetty *et al.* 2013), monopsony in labor markets (Sokolova & Sorensen 2021) the effects of active labor market programs (Card *et al.* 2017), etc.

Second, we investigate publication bias in the literature estimating TEBS, the bias that has previously been found to be prominent in many fields of economics research, see e.g. Card & Krueger (1995), Stanley (2001), Brodeur *et al.* (2016), Ioannidis *et al.* (2017), Sokolova & Sorensen (2021). Third, we use the results of the meta-regression analysis to construct fitted estimates of TEBS for different categories of goods and geographic regions, corrected for the selective reporting in the literature.

The mean TEBS reported in the literature is -0.57, suggesting that cross-border shopping should have a non-negligible effect on sales at the border. However, we show that, like many other literatures in economics, the literature estimating TEBS is prone to selective reporting of the results. Specifically, studies tend to underreport results with a positive sign: such results are counter-intuitive as they are not in line with the economic theory. Yet, given the randomness of the data, counter-intuitive results should still appear from time to time. This underreporting leads to the asymmetry in the distribution of reported TEBS estimates, with the mean reported estimate being biased away from the 'true' effect. We construct a corrected mean effect—a task that can only be accomplished within the framework of a meta-analysis. We find that the corrected effect is somewhat less prominent: most of our specifications find it to be between -0.4 and -0.5, some indicate it may be even smaller in magnitude.

We also find that TEBS estimates vary systematically with the category of goods considered. Cross-border sales of food and retail, which have a broad tax base, are found to be much more elastic compared to sales of tobacco: the absolute value of TEBS for food and retail exceed that of tobacco by about 2. Food and retail comprise a large fraction of household budget, and increases in sales taxes may strongly affect overall household expenses—more so, compared to an increase in an excise tax on one product.[2] Thus, traveling to save on food and retail may be more economically justified compared to traveling to save on tobacco. Compared to tobacco, fuel sales are also found to be more elastic: with TEBS greater in absolute value by about 0.7-0.8. Gas stations tend to display fuel prices at their entrances in a way that eases cross-jurisdiction comparisons by travelers passing by, thus possibly encouraging cross-border shopping. By contrast, TEBS of tobacco does not appear to differ from TEBS of other individual 'sin' products, such as marijuana, alcohol and soda. Finally, while cross-border sales appear very elastic

[2]Of course, we acknowledge the fact that sales of food (specifically grocery sales) are exempt from taxation in a majority of states in the US. According to Federation of Tax Administrators (FTA), there are 16 states that tax food to some extent, see taxadmin.org. We should also note that there is no exemption for retail sales aside from short periods of sales tax holidays.

for the US where 'car culture' is prevalent, the cross-border effect is much less prominent in Europe and other countries.

Our study is also related to the literature examining the question of "who (ultimately) bears the burden of taxation in a jurisdiction?" To the extent that there is cross-border shopping, some of the tax burden will be shared between individuals in different jurisdictions—this shifting of the tax burden from one jurisdiction to another is addressed by the literature on tax incidence (e.g. McLure 1967, Hogan & Shelton 1973, Mutti & Morgan 1983, Fujii *et al.* 1985, Gade & Adkins 1990, Pollock 1991, Morgan *et al.* 1996, Murray 2006, Khan *et al.* 2020, Minnesota Department of Revenue Tax Research Division 2021). Cross-border sales of different commodities have been seen as a direct way in which tax exporting occurs, and knowing TEBS is pertinent for assessing the magnitude of the corresponding effect. In a similar vein, the assumptions made about the elasticity of cross-border sales affect the results reported by a broader set of studies in public finance (e.g. Trandel 1992, Nechyba 1996, Poterba 1996, Beard *et al.* 1997, Besley & Rosen 1999, Ohsawa 1999, and Nelson 2002).[3] Our fitted estimates of TEBS can serve as a reference for these literatures.

The fitted estimates of TEBS reported in this study can also be helpful to state and local policymakers as they contemplate tax changes in their jurisdiction. On the one hand, having accurate TEBS estimates is essential when assessing the stability and adequacy of tax revenues—a crucial task for jurisdictions with balanced budget requirements. For example, taxes on cigarettes and other tobacco products are usually earmarked for health and/or education programs. Cross-border sales can jeopardize the stability of those tax revenues and may in turn affect the expenditures financed by earmarked revenues. On the other hand, aside from the tax revenue concerns, cross-border shopping can undermine the outcomes of Pigouvian taxation. Excise taxes on cigarettes, alcohol, gasoline and marijuana are partially aimed at correcting the negative consumption externalities. If TEBS is high, an increase in cigarette tax does not necessarily mean that more people quit smoking, as they may continue to buy cigarettes from a lower-tax jurisdiction nearby. Interjurisdictional tax rate differences, including those that are due to recent sales tax holidays which aim at providing relief to families impacted by price increases, continue to generate incentives for cross-border shopping.[4] Knowing TEBS is important when evaluating the effects of such policies; the estimates provided here can help policymakers gauge the magnitude of TEBS suggested by the relevant literature.

[3]Additionally, there are also policy oriented (mostly non-academic) studies from the U.S. that provide economic and fiscal impact of tax policy changes with a focus on state and local governments.

[4]Federation of Tax Administrators provide information on state sales tax holidays www.taxadmin.org/sales-tax-holidays. Tax amnesties may also create tax differences between states and lead to revenue impacts as shown by Alm *et al.* (1990), Alm & Beck (1991), Alm & Beck (1993), Luitel & Sobel (2007), and Luitel & Tosun (2014) among others.

2 Data Collection Strategy

2.1 Approaches to estimating TEBS

Studies examining cross-border sales typically estimate the responsiveness of sales at the border in one jurisdiction to changes in relative price between the jurisdiction under consideration and its neighbor—where the neighbor may have a different tax rate or even a different tax structure. The estimates are usually presented in the form of tax elasticity, or other formats from which the elasticity can be deduced. Due to lack of a name that is commonly used in this literature, we term such estimates 'tax elasticity of border sales', or TEBS.

Studies can estimate TEBS with a version of the following regression model:

$$\ln Sales_{it} = \epsilon \cdot \ln P_{it} + \sum_{l=1}^{M} \beta_l Z_{l,it} + C_i + T_t + u_{ij}, \tag{1}$$

where $Sales_{it}$ are per capita sales of taxable items (e.g. general sales, food) at the border in region i (e.g. county) at time t; $Z_{l,it}$ are controls that capture the overall demand in region i (e.g. log per capita income); C_i and T_t are region- and time-specific effects and u_{ij} is the disturbance term (see e.g. Tosun & Skidmore 2007). The variable P_{it} captures the after-tax price of taxable goods of interest in region i relative to the neighboring region:

$$P_{it} = \frac{P_{it}^H(1+\tau_{it}^H)}{P_{it}^N(1+\tau_{it}^N)}, \tag{2}$$

where P_{it}^H and τ_{it}^H are the before-tax price and tax rate in the home region i, and P_{it}^N and τ_{it}^N are the before-tax price and tax rate in the neighboring region (see e.g. Walsh & Jones 1988, Wooster & Lehner 2010). The parameter ϵ in (1) captures a percentage change in sales that would occur if the relative price P_{it} increases by 1%—i.e. TEBS.

Studies differ in their choices over model specifications used to estimate ϵ. A number of studies make a simplifying assumption that the before-tax prices at home and in the neighboring region are the same, i.e. $P_{it}^H = P_{it}^N$ (e.g. Goolsbee *et al.* 2010). Under this assumption it is enough to consider the difference between tax rates at home and in the neighboring region to capture ϵ, TEBS. Studies estimating regression (1) make an implicit assumption that border sales respond symmetrically to a 1% increase in after-tax prices at home and a 1% decrease in prices of the neighbor; in other studies this assumption is relaxed, and the coefficient on the log of after-tax home price is allowed to differ from that on the after-tax neighbor price, as $\ln[P_{it}^H(1+\tau_{it}^H)]$ and $\ln[P_{it}^N(1+\tau_{it}^N)]$ enter separately into regression (1), e.g. Goel (2008), Connelly *et al.* (2009), and Goel & Nelson (2012).

In such regression models, the coefficient pertaining to the after-tax neighbor price can be interpreted similarly to ϵ under the assumption that the after-tax price in the home region is constant. We collect estimates coming from all specifications discussed above, but introduce controls to capture these subtle differences in their interpretation.

Finally, a number of studies construct experiments to directly measure changes in border sales in response to tax innovations by e.g. collecting discarded cigarette packs in a given location and calculating the percentages of packs purchased locally before and after a tax change (see Merriman 2010, Chernick & Merriman 2013). For these studies we approximate ϵ by 1) collecting the provided estimate of the percentage change in sales, $\Delta Sales/Sales$, 2) computing the percentage change in the relative prices $\Delta P/P$ from the information provided in the study and 3) setting $\epsilon = (\Delta Sales/Sales)/(\Delta P/P)$ and approximating the standard error with the delta method.

2.2 Data collection and inclusion criteria

We use Google Scholar to search for the studies estimating TEBS. We implement a search query containing the words 'relative price', 'tax difference elasticity estimate border', 'cross-border', 'neighboring sales', 'tax', 'shopping', 'purchases', and 'estimates' and save the search results on April 30th, 2021. We read through abstracts of papers appearing on the first 50 pages, download and save papers that seem to report relevant results. We then read through the downloaded papers making sure they contain estimates of TEBS, or results that could be used to derive the implied TEBS estimates, implementing additional conversions discussed in the previous subsection. We also require the studies to report some measures of statistical significance of each estimate (i.e. standard errors, t-statistics or p-values). We also examine the references of the articles we find to locate any research papers that our search procedure may have missed.

We include papers published in peer-reviewed journals: these papers would have gone through the peer-review process and therefore have passed a quality check. We also include working papers that came out after the year 2018, as for these more recent works the lack of publication does not imply anything about the papers' quality or relevance.

This search strategy yielded 60 studies providing 749 estimates.[5] For each estimate, we record its corresponding standard error and detail the context in which the estimate was obtained (e.g. product type, empirical methodology, geographic location, etc.). These estimates cover a diverse set of geographic regions and product groups. We report the sample statistics in Table 1.

The mean estimate of TEBS is -0.57, but estimates vary widely. Estimates of TEBS

[5]The full list of studies in our sample can be found in Appendix D.

Table 1: Estimates of TEBS

	Mean	Median	5%	95%	N	N Stud.
All	-0.57	-0.19	-2.20	0.23	749	60
Product groups						
Food & Retail	-1.79	-0.67	-7.77	0.12	94	10
Fuel	-0.26	-0.06	-1.50	0.40	88	6
Tobacco	-0.43	-0.20	-1.55	0.16	485	35
Other 'sin' products	-0.33	-0.04	-2.20	0.70	82	10
Regions						
U.S.	-0.59	-0.19	-2.33	0.20	584	50
E.U. & other countries	-0.48	-0.21	-2.08	0.23	165	12

Notes: 5% and 95% refer to the corresponding percentiles.

seem to differ depending on the types of sales considered, with the most prominent mean elasticity of -1.79 observed for food and retail purchases, and somewhat smaller results for fuel (-0.26), tobacco (-0.43) and other 'sin' products, such as marijuana, alcohol and soda (-0.33 overall). US sales appear to be more responsive to tax changes compared to sales in Europe and other countries (with mean elasticity of -0.59 compared to -0.48).

The observed variation in estimates may be due to fundamental differences between the cross border effects for different groups of goods. At the same time, it is possible that, coincidentally, studies estimating cross-border effects e.g. for food and retail tend to use empirical methods and data that are markedly different from those employed to estimate TEBS for other goods. To disentangle sources of variation in estimates we therefore need to carefully consider features of each studies' design. Furthermore, it is possible that mean reported estimates are affected by preferences of the authors working in TEBS literature—in other words, that the literature is prone to some selective reporting of the results. In the subsequent sections we will evaluate these alternative explanations for observed variation in TEBS and distinguish between its sources.

3 Publication bias

Estimates of TEBS reported by the literature may be affected not only by fundamental features of the underlying data and studies' design, but also by practices surrounding the publication process and preferences of the profession. Brodeur *et al.* (2016) examine 50,000 tests for empirical significance published in top three journals in economics and conclude that authors tend to under-report results with p-values between 0.25 and 0.10, possibly due to them engaging in specification searches to obtain results that are statistically significant. Card & Krueger (1995) and Doucouliagos & Stanley (2009) document publication selection bias in the literature studying the effect of minimum wage increases on employment; they find that, even though a large fraction of the literature reports a strong negative effect, this effect disappears once publication bias is accounted for. In a similar vein, Havranek & Sokolova (2020) find that the mean reported estimates of shares of rule-of-thumb consumers exhibit strong upward bias due to systematic discarding of negative results. Sokolova & Sorensen (2021) document a similar bias in the literature estimating monopsony power on the labor market. Ioannidis *et al.* (2017) argue that about 80% of the findings reported in economics literatures are exaggerated and single out meta-analysis as a tool for correcting such biases.

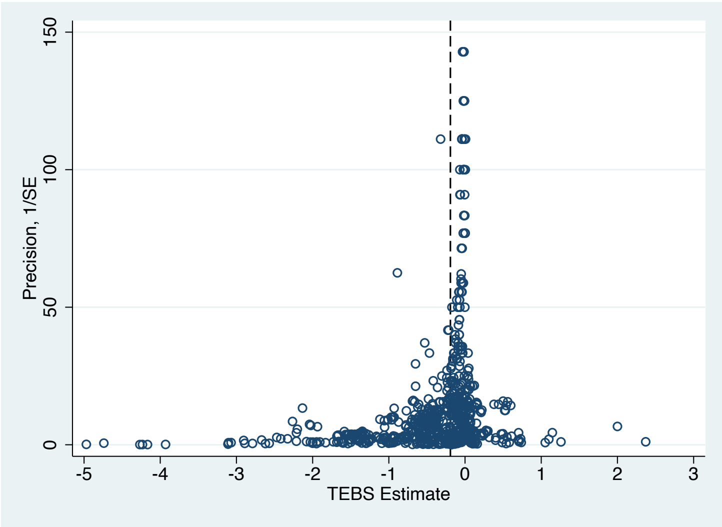
The reported estimates of TEBS may also be prone to publication selection bias. Consider the following thought experiment. Suppose that in line with the economic theory the underlying 'true' elasticity is negative, but that it is also relatively close to zero, i.e. the magnitude of the effect is small. Estimating this parameter on noisy data using standard unconstrained estimators should yield a symmetric distribution of estimates that is centered around the 'true' underlying parameter.

Because (by assumption) the 'true' effect is relatively close to zero, a fraction of estimates with low precision ends up positive. Suppose that because a positive TEBS has no good theoretical interpretations researchers obtaining positive results engage in further specification searches until the estimates become negative. If a large fraction of positive estimates is thus discarded, the distribution of estimates that are being reported by researchers would become skewed. This type of selective reporting would make it difficult for researchers to assess the underlying 'true' TEBS, as the mean estimate reported by the literature would exhibit a downward bias.

We now turn to investigate whether the literature on TEBS is affected by selective reporting of the estimates. The intuition behind testing for publication bias can be summarized as follows. Absent selective reporting, estimates found in empirical literature should be distributed symmetrically around the 'true' underlying effect—provided that the estimates can be thought of as assessing the same underlying parameter. A skewed distribution of reported estimates would signal that certain results are being systematically discarded.

Figure 1 depicts TEBS estimates in our sample; the estimates are plotted against their precision (1/*Standard error*). In the absence of selective reporting, this funnel plot should be symmetrical, with more precise estimates clustering around the underlying 'true' effect (see Egger *et al.* 1997). In our sample, the most precise estimates lie close to zero, singling that the 'true' underlying effect may be small in magnitude. At the same time, the left tail of the funnel appears to be much more prominent compared to the right tail. This

Figure 1: Funnel plot



Notes: The figure plots TEBS estimates from our dataset against the reported precision; the dashed line indicates sample mean. The figure is truncated at $1/SE = 150$.

could mean that the positive estimates are being systematically discarded, which would imply that the mean estimate reported in Table 1 exhibits a downward bias.

We will now conduct a formal funnel asymmetry test (see Stanley 2005). Estimation methods commonly used by researchers assume the ratio of the estimate to its standard error to be symmetrically distributed (e.g. t-distributed). If this assumption holds, the estimates should not be correlated with their standard errors. We will test this by regressing the estimates in our sample on their standard errors:

$$\hat{\epsilon}_{ij} = \epsilon_0 + \gamma \cdot SE(\hat{\epsilon}_{ij}) + u_{ij}, \tag{3}$$

where $\hat{\epsilon}_{ij}$ is the i-th estimate of TEBS reported by j-th study, $SE(\hat{\epsilon}_{ij})$ is its standard error, and u_{ij} is noise. Absent publication bias, the coefficient γ should be zero; if authors systematically discard positive results, this coefficient would be negative. Importantly, the estimate of the constant term ϵ_0 can be interpreted as an approximation of the unbiased effect (Stanley 2008 shows this approximation to perform well in Monte Carlo simulations). We will use estimates of ϵ_0 to construct a correction for the effect of publication bias.

The first column of Table 2 shows OLS estimates of model (3). We cluster the standard errors at the study level to address possible within-study correlation in estimates. The coefficient on the standard error is negative and statistically significant: as discussed above, this suggests selective reporting of the negative estimates. The constant term

ϵ_0 is estimated at -0.4, giving the approximate 'true' unbiased effect. This effect is less pronounced when compared with the sample mean of -0.57, suggesting that the presence of publication selection bias leads to an exaggerated mean reported TEBS—in absolute value. Thus, once the publication bias is accounted for, the cross-border trade effect reported by the literature becomes somewhat more modest.

Next, it is possible that estimates are affected by the unobserved study-level characteristics; we therefore evaluate a specification featuring study-level fixed effects, as is done in e.g. Havranek 2015, Sokolova & Sorensen 2021—see Table 2, column 'FE'. Studies often report multitudes of estimates, some of which may be inferior to others. We next evaluate a specification in which we estimate γ and ϵ_0 using only the variation across studies (see e.g. Havranek & Sokolova 2020); we report the results under 'BE' in Table 2. To address the potential heteroscedasticity and give more weight to more precise estimates, we also evaluate a specification in which each estimate is weighted by its precision, as recommended in Stanley & Doucouliagos (2015) (see 'Precision' in Table 2). Finally, some studies report many more estimates compared to others. To give studies roughly equal weight, we assess a specification in which each estimate is weighted by the inverse of the number of estimates reported in the associated study (as is done in e.g. Gunby *et al.* 2017, Havranek & Sokolova 2020); these results appear under 'Study' in Table 2.

Table 2: Testing for publication bias

	OLS	FE	BE	Precision	Study
SE	-0.315	-0.306	-0.687	-0.998	-0.730
	(0.113)	(0.083)	(0.179)	(0.333)	(0.330)
Constant	-0.400	-0.405	-0.504	-0.034	-0.511
	(0.095)	(0.045)	(0.258)	(0.015)	(0.211)
Studies	60	60	60	60	60
Observations	749	749	60	749	749

Notes: Standard errors appearing in parenthesis are clustered at the study level. *FE*=specification with study-level fixed effects; *BE*=specification that uses median estimates and standard errors for each study; *Precision*=specification in which each estimate is weighted by its precision; *Study*=specification in which each estimate is weighted by the inverse of the number of estimates reported in the associated study.

In all specifications listed above there is a statistically significant correlation between the reported estimates and their standard errors—this, again, suggests that the literature on TEBS is prone to selective reporting of the results. The constant term (that can be interpreted as the bias-corrected effect) is negative and statistically different from zero; its magnitude is smaller compared to the sample mean of -0.57, particularly in the specification that employs precision weights. Thus, there is evidence of a statistically significant TEBS even when publication bias is accounted for, albeit the bias-corrected effect is less prominent compared to the mean estimate of -0.57 reported by the literature.

While the evidence of selective reporting uncovered thus far appears strong, it is possible that some of it is driven by outliers in the data: a small group of very imprecise negative estimates could contribute to the observed negative association between estimates and their standard errors. We investigate this possibility by repeating the exercise of Table 2 for samples of estimates subject to a number of outlier treatments. Specifically, we consider the samples with outliers winsorized at 1% and 5% (.5% and 2.5% in each tail), and with 1% and 5% of outliers dropped—see Table B1. The results appear largely unaffected by the treatment of outliers.

Another caveat of the strategy outlined above is that it does not, technically, allow to distinguish between forms of selective reporting (e.g. selection for the 'right sign', or selection for statistical significance). In a recent work, Andrews & Kasy (2019) propose an alternative strategy for testing for publication bias, which involves explicit modeling of the publication selection process. The approach developed by Andrews & Kasy (2019) models probability of a result being reported as a step function of the associated Z-score; the reporting probabilities may differ between negative significant results (Z-score below -1.96), negative insignificant results (Z-score between -1.96 and 0), positive insignificant results (Z-score between 0 and 1.96) and positive significant results (Z-score over 1.96). Furthermore, the approach developed by Andrews & Kasy (2019) allows calculating the unbiased underlying mean effect. Previously, this method has been implemented in a similar context by Sokolova & Sorensen (2021).

We discuss this technique in Appendix B.3 and apply it to our estimates. We find prominent evidence of selection for the 'right' sign, as positive estimates tend to be 18 times less likely to be reported compared to estimates that are negative and significant. The probability of reporting of negative insignificant estimates is about 7 times smaller compared to that of negative significant results. The bias-corrected effect is found to be very close to zero.

One remaining caveat pertaining to both the funnel asymmetry test and the Andrews & Kasy (2019) approach is that they rely on the assumption of independence between the estimates produced by latent studies and their standard errors. However, it is possible that some choices that researchers make that determine studies' design would affect the estimate and the standard error in the same direction. We remedy this using two strategies. First, we consider relatively homogenous subsamples of estimates that are more likely to pertain to the same 'true' underlying values of TEBS: the subsample of estimates characterizing TEBS for tobacco purchases only, and the subsample of estimates obtained using OLS. We repeat the exercises of Figure 1 and Table 2 for these two subsamples and report the results in Figure B1 and Table B2.[6] The evidence of publication selection bias remains strong for both subsamples. The second, alternative, strategy we employ is presented in the following section, in which we account for the observed heterogeneity in estimates by including a number of studies' characteristics as additional controls in regression (3).

4 Variation in TEBS estimates

Estimates may vary due to variation in the underlying 'true' elasticity. They may also vary with empirical methods employed by researchers and other features of studies' design. In this section we construct a comprehensive set of explanatory variables to capture key factors likely contributing to the variation in estimates and use meta-regression analysis to determine which study characteristics have systematic effects on the reported TEBS.

4.1 Explanatory variables

We describe below the key explanatory variables that capture features of data and techniques used by researches estimating TEBS. The full list of the 29 recorded variables and their descriptive statistics are available in Table A1.

Data characteristics

Tax elasticity of border sales can be estimated with different kinds of data. The majority of studies in our sample use data at the annual frequency. At the same time, a number of studies employ daily data (e.g. Khan *et al.* 2020), weekly data (e.g. Hansen *et al.* 2020), monthly data (e.g. Ye & Kerr 2016), and as well as data at quarterly frequency (e.g. Baranzini & Weber 2013). Similarly, while most studies estimate TEBS with state-level data, some use data at the household or county level (e.g. Lesley & Erich 2008), as well as data on border sales associated with travel across countries (e.g. Friberg *et al.* 2022). Finally, some studies investigating TEBS in the US obtain data from local sources (e.g. Merriman 2010), while others use statistics collected at the federal level (e.g. Stehr 2007). We construct a set of controls to reflect the above differences. We additionally record the number of observations used to obtain each estimate, and the average year of data. As before, we include the standard errors associated with the estimates to capture publication bias.

Economic, geographic & demographic controls and characteristics

Studies evaluating TEBS typically estimate how sales in the home region are affected

[6] We also implement the test proposed by Andrews & Kasy (2019) for these two subsamples—see Appendix B.3. The results are very similar to those obtained for the full sample.

by the discrepancy between home and neighbor prices. But this difference is not the only factor that could potentially affect sales. The volume of sales in a given region is likely to depend on the economic wellbeing of the underlying population, as well as its demographic composition. We construct a set of variables reflecting whether the authors control for such factors in their analysis. The 'true' TEBS itself may vary across geographic regions. Car travel is much more prevalent and culturally accepted in the US compared to Europe. We construct controls to distinguish between TEBS estimates for the US and other countries.

Consumption data features

Consumers may exhibit different propensities to travel depending on commodities being taxed—the underlying 'true' value of TEBS may thus be different for different commodities. A large number of studies in our sample focuses exclusively on tobacco sales (e.g. Becker *et al.* 1994, Goel & Nelson 2012, and Lakhdar *et al.* 2016); other studies calculate TEBS for fuel (e.g.Leal *et al.* 2009, and Jansen & Jonker 2018), food and retail (Fox 1986, and Wooster & Lehner 2010) and other 'sin' products, such as marijuana, alcohol and soda (e.g. Ye & Kerr 2016, Hansen *et al.* 2020, and Seiler *et al.* 2021). For each TEBS estimate, we record the type of consumer goods it pertains to.

Estimation approaches

The vast majority of studies in our sample estimate TEBS—or effects from which TEBS can be imputed—with OLS (e.g. Apergis *et al.* 2014, and Friberg *et al.* 2022). A number of studies employ alternative methods, such as IV (e.g. Baltagi & Griffin 1995), GLS (e.g. Banfi *et al.* 2005), difference in differences (e.g. Leal *et al.* 2009), probit (e.g. Chernick & Merriman 2013) or other methods (e.g. spatial regression in Tosun & Skidmore 2007, seemingly unrelated regression in Asplund *et al.* 2007, ARIMA in Ye & Kerr 2016). Many studies also employ state and time fixed effects (e.g. Farrelly *et al.* 2003, Seiler *et al.* 2021). We construct a set of controls to account for these differences.

Specification

While many studies in our sample estimate versions of regression (1), their precise definitions of the left- and right-hand side variables may differ. For example, the dependent variable may come in the form of dollar sales (e.g. Ballard & Lee 2007) or sales expressed in quantities (e.g. number of packs of cigarettes purchased, Connelly *et al.* 2009, or volume of gasoline, Banfi *et al.* 2005). A number of studies weight the prices of various neighbors by the size of the population at the respective borders (e.g. Farrelly *et al.* 2003). While some studies in our sample estimate how sales respond to changes in the log price index (as in 2), many measure tax elasticity by estimating the coefficient on the log of the after-tax neighbor price. There are therefore some differences in the interpretation of TEBS estimates produced by the studies—we capture them by constructing corresponding controls.

For some estimates we perform conversions to make them comparable with our baseline specification (i.e. specification 1). A number of studies estimate interactions between sales and distances from the border, obtaining distance-specific TEBS (e.g. Lovenheim 2008). To make these estimates comparable with our baseline that does not include distance, we use information provided by these studies to construct estimates of TEBS conditional on mean distance considered in each paper. Some papers employ log-linear specifications (e.g. DeCicca *et al.* 2013, Lovenheim 2008). We likewise use descriptive statistics included in the papers to convert the provided estimates to a format comparable with our baseline specification. We construct controls to reflect these differences.

Publication characteristics

It is possible that studies in our sample are of differing quality. Even though we control for a number of observable characteristics of the studies and the underlying datasets, it is possible that there are some more subtle features related to study quality that we are missing. To account for the potential unobservable variation in study quality, we construct two additional controls. First, we record whether the study was published in one of the top five general interest journals in economics or in the top field journal for public finance (similar to Sokolova & Sorensen 2021). Second, we record the average yearly number of citations that the study received since its first appearance on Google Scholar (as in e.g. Havranek *et al.* 2015).

4.2 Variation in TEBS estimate and model uncertainty

In the previous section we pointed out a large number of study features that could potentially contribute to the observed variation in TEBS estimates. We accordingly constructed 29 associated control variables. We will now employ these controls to pin down the sources of variation in TEBS estimates with meta-regression analysis. We will estimate the following regression model:

$$\hat{\epsilon}_{ij} = \alpha_0 + \sum_{l=1}^{29} \beta_l X_{l,ij} + u_{ij}, \tag{4}$$

where $\hat{\epsilon}_{ij}$ is the i-th estimate of TEBS reported by j-th study, $X_{l,ij}$ are the explanatory variables capturing variation in estimates, and u_{ij} is noise. The estimates of β_l would reflect the contribution of each factor from the set of 29 predictors to variation in elasticity estimates.

One important caveat to consider while estimating regression (4) is model uncertainty. While it is likely that a subset of the 29 variables we constructed is part of the 'true' data generating process for TEBS estimates, it is probably not true that *every one* of these variables contributes to variation in TEBS estimates in a meaningful way. We thus face the following problem. On the one hand, including all 29 variables in regression (4) would likely render the model misspecified. On the other hand, choosing one smaller subset of the 29 variables would not account for the possibility that some of the remaining $2^{29} - 1$ variable combinations could do better at capturing the data generating process for estimates of TEBS.[7]

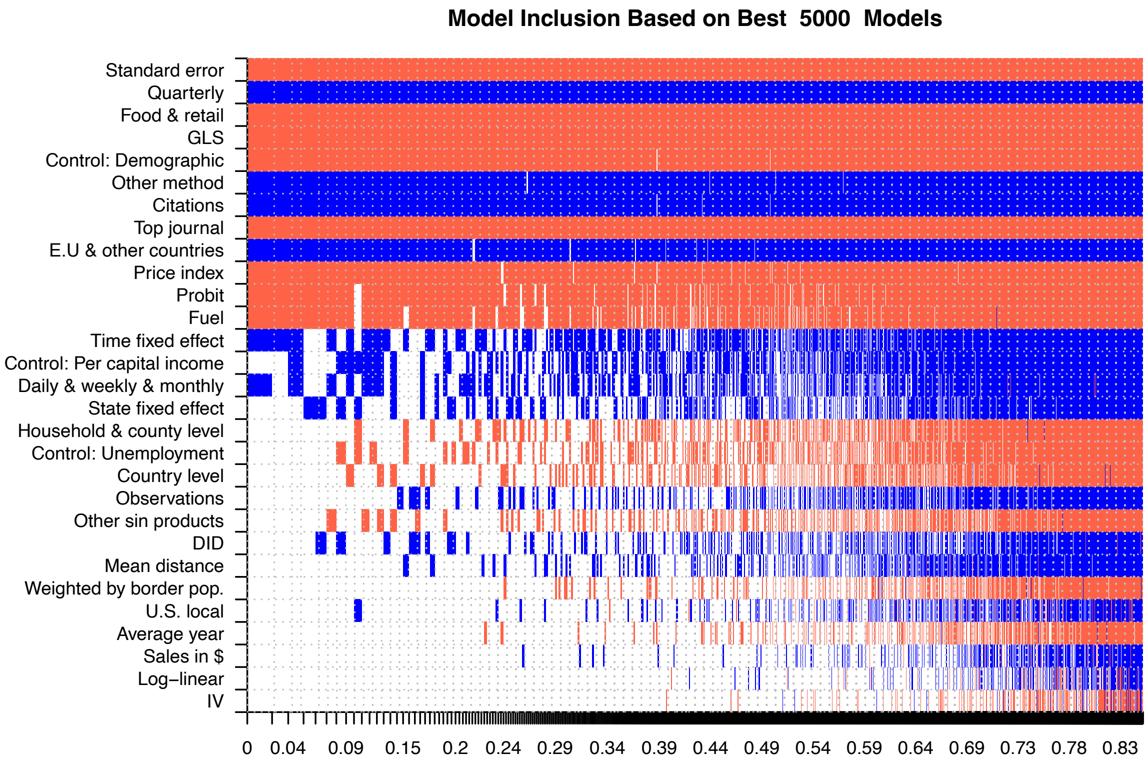
Fernandez *et al.* (2001) point out the importance of model uncertainty in cross-country growth regressions and argue in favor of a more rigorous approach—the Bayesian Model Averaging (BMA, see Raftery *et al.* 1997). BMA evaluates and compares all of the 2^{29} possible combinations of the explanatory variables, assigning each a metric (posterior model probability, PMP) that measures how likely each model is to represent the underlying data generating process. BMA then constructs inference by averaging results across all evaluated models weighted by the corresponding PMP. BMA has been used in other meta-analyses in contexts similar to ours (e.g. Havranek *et al.* 2017, Balima & Sokolova 2021). Following these studies, we estimate model (4) with BMA.

Figure 2 provides graphical representation of the BMA estimation results. Each column on the graph represents one of the 2^{29} combinations of the explanatory variables, i.e. a model. For each model, white coloring means that a given variable is not included, while red (blue) indicates that the control is included and the corresponding estimated effect is negative (positive). The horizontal axis lists cumulative posterior model probabilities for models depicted, while vertical axis ranks the 29 control variables by their posterior inclusion probabilities (PIP). For each variable, a PIP reflects the likelihood of the variable belonging to the 'true' data generating process for TEBS estimates; it is computed by adding up all PMPs of models in which the variable is present.

The variables listed at the top of Figure 2 are the most likely to be part of the 'true' data generating process for TEBS estimates. The signs of the effects associated with these controls are very stable across different models. One such variable is the standard error associated with the TEBS estimates: it is present in all 'good' models depicted on Figure 2. We document a strong negative relationship between TEBS estimates $\hat{\epsilon}_{ij}$ and their respective standard errors, $SE(\hat{\epsilon}_{ij})$,—corroborating the findings of Section 3. We conclude that the evidence of selective reporting of estimates with the 'right sign' remains intact even after we control for other potential sources of observable variation in TEBS.

[7] In such cases researchers often employ sequential t-testing; however, a sequential exclusion of insignificant regressors could lead to an accidental omission of variables that belong to the data generating process—thus, it does not adequately address the issue of model uncertainty.

Figure 2: Bayesian Model Averaging: Models



Notes: The figure depicts the results of BMA estimation of regression (4). Each column gives one model; the models are ranked by their posterior model probability (cumulative PMPs are marked on the horizontal axis). The vertical axis lists explanatory variables ranked by their PIPs. White cell means that the variable on the left-hand-side is not included in the given model; red (blue) means the variable is included and the estimated coefficient is negative (positive). Variable description is available in Table A1. The figure was obtained with the BMS package for R developed by Zeugner & Feldkircher (2015). We use a combination of the Unit Information Prior for model parameters and the Uniform prior for model space; alternative priors are considered on Figure C1.

Among other variables with high PIPs are controls reflecting the types of products the elasticity is estimated for and their geographic region. It appears that, compared to tobacco, the sales of fuel, food and retail are more elastic with respect to changes in relative price. This means that consumers are more likely to take advantage of cross-border price differences when shopping for these goods—relative to tobacco. At the same time, TEBS estimates of other 'sin' products do not seem to differ very much from those for tobacco. Furthermore, US consumers appear to be more engaged in cross-border shopping compared to consumers living in Europe and other countries. This may be due to the prevalence of 'car culture' in the US and its overall greater reliance on personal transportation.

Table 3: Heterogeneity and model uncertainty

	BMA			OLS	
Response variable:	Post. Mean	Post. SD	PIP	Coef.	SE
Data characteristics					
Standard error	-0.256	0.038	1.000	-0.296	0.066
Observations	0.015	0.030	0.257		
Average year	-0.016	0.062	0.096		
Daily & weekly & monthly	0.250	0.307	0.473		
Quarterly	2.047	0.371	1.000	1.979	0.690
Household & county level	-0.181	0.294	0.336		
Country level	-0.139	0.265	0.265		
U.S local	0.061	0.208	0.122		
Economic, geographic & demographic controls					
Control: Per capita income	0.175	0.200	0.504	0.130	0.170
Control: Unemployment	-0.129	0.210	0.325		
Control: Demographic	-0.585	0.130	0.992	-0.687	0.281
E.U. & other countries	0.773	0.274	0.962	0.936	0.353
Consumption data features					
Food & retail	-1.993	0.273	1.000	-1.973	0.944
Fuel	-0.706	0.411	0.825	-0.805	0.454
Other sin products	-0.107	0.209	0.256		
Estimation approaches					
GLS	-0.623	0.186	0.996	-0.620	0.498
DID	0.122	0.249	0.237		
IV	-0.002	0.024	0.026		
Probit	-0.883	0.418	0.904	-0.860	0.745
Other method	0.723	0.219	0.988	0.601	0.427
State fixed effect	0.102	0.161	0.340		
Time fixed effect	0.228	0.215	0.609	0.318	0.169
Specifications					
Sales in $	0.015	0.085	0.057		
Weighted by border pop.	-0.041	0.127	0.128		
Price index	-0.589	0.220	0.944	-0.592	0.453
Mean distance	0.098	0.213	0.218		
Log-linear	0.003	0.045	0.030		
Publication characteristics					
Top journal	-1.109	0.346	0.981	-1.035	0.508
Citations	0.192	0.052	0.984	0.195	0.090
Constant	-0.349		1.000	-0.182	0.195
Studies	60			60	
Observations	749			749	

Notes: The left panel of the table reports numerical results of BMA estimation corresponding to Figure 2. 'SD' is standard deviation; 'PIP' is posterior inclusion probability; 'SE' is standard error. The right panel reports results from an OLS estimation of a specification in which we only include variables with PIP higher than 50%. Standard errors are clustered at the study level.

We also observe that sales appear more elastic when TEBS is estimated on annual data—particularly compared to datasets of quarterly frequency. This difference may reflect the fact that it takes time for consumers to notice the price difference between regions and change their shopping behavior. The difference between annual and high frequency data (monthly or higher) has a similar direction, though it appears less statistically prominent—possibly due to limitations of our data.

Finally, we document some differences in TEBS estimates that arise due to differences in studies' design, such as the use of GLS and some other estimation methods. We also find that studies published in top journals tend to report more prominent TEBS estimates compared to studies published in other outlets; however, studies that accumulate greater numbers of citations per year tend to report more modest effects.

The left panel of Table 3 summarizes the numerical results from the BMA estimation. The right panel of Table 3 reports results from the 'frequentist check', in which we run an OLS with only the variables that showed PIP higher than 50% in the BMA—that is, the variables that were deemed likely to be part of the 'true' data generating process for TEBS estimates.

Sales of food and retail appear substantially more elastic than those of tobacco: when the price difference increases by one percentage point, these sales fall by about 2 percentage points more compared to tobacco sales. Unlike purchases of individual 'sin' products, spending on food and retail typically amounts to a large fraction of household budget—potential savings from shopping at a jurisdiction with a lower sales tax may thus justify the travel costs. In a similar vein, sales of fuel are more responsive to changes in regional price differences compared to sales of tobacco. A one percentage point increase in the price difference leads to a 0.7-0.8 percentage point decrease in fuel sales. Unlike all other goods, fuel consumption is directly related to the act of traveling by personal transport. Furthermore, fuel prices are typically very clearly displayed at gas station entrances, and are easily noticed and compared by those shopping for gas. All of these considerations may make consumers more responsive to changes in the prices of fuel. At the same time, the difference between TEBS for tobacco and other sin goods (alcohol, soda, marijuana) is not prominent, nor statistically significant. Thus, we can conclude that the sales elasticities of individual 'sin' goods (excluding fuel) are relatively similar.

Aside from differences across goods, we document a prominent discrepancy between TEBS estimates across geographic regions. TEBS estimates for the US are lower than those for Europe and other countries by about 0.8-0.9 percentage points, making US sales considerably more elastic.

The negative coefficient on the standard errors associated with TEBS estimates remains statistically significant—even after we control for an extensive list of features of study design that may cause estimates to vary systematically across studies. This corroborates the conclusions of the previous section, suggesting that publication selection bias is indeed prominent in the literature on TEBS.

We next examine the extent to which these results may be driven by outliers in the data. As in Section 3, we consider four outlier treatments: samples with outliers in estimates and standard errors winsorized at 1% and 5%, as well as samples in which 1%(5%) of outliers are dropped—see Figure C1. None of the outlier treatments affect the inference made about the standard error, suggesting that the evidence of selective reporting we find is not driven by the outliers in the data. At the same time, the difference between TEBS for tobacco on the one hand, and food and retail as well as fuel on the other, seems to be affected by the strong outlier treatments (winsorization at 5% and dropping 5% of outliers). Similarly, the strong outlier treatments affect the inference about TEBS for Europe and other countries.

We turn to study the extent to which our BMA results may be driven by the prior assumptions made about the model parameters and the model space. In our baseline specification, we employ the unit information prior (UIP) for parameters—that is, a prior that communicates the amount of information similar to that of one observation; for model space, we use the uniform prior. This combination of priors was found to work well for predictive estimations (Eicher *et al.* 2011). Figure C1 compares our baseline results to those obtained under alternative combinations of priors. For parameters, we consider benchmark g-priors ('BRIC' on Figure C1, see Fernández *et al.* 2001) and the flexible data-dependent priors ('HyperBRIC' on Figure C1, see Liang *et al.* 2008). For model space, we introduce the beta-binomial model prior ('Random' on Figure C1, see Ley & Steel 2009). We find that these alternative prior assumptions do not have much of an effect on the posterior means—albeit they tend to result in higher posterior inclusion probabilities for most of our explanatory variables.

Finally, we employ LASSO—an alternative approach to addressing model uncertainty (see Tibshirani 1996). This procedure runs an OLS minimization subject to a constraint on the sum of the absolute values of model coefficients. Because of the constraint, the minimization yields corner solutions, setting some of the model coefficients to exact zeros. For our dataset, LASSO does not exclude any of the variables (see Table C1). The post-LASSO OLS estimation indicates that the standard error is still statistically significant, even despite the presence of 28 other explanatory variables in the regression. The other variables that show significance in the BMA—such as the effects associated with fuel, food retail—have coefficients that in magnitude and direction are similar to those obtained before, albeit not statistically significant.

4.3 Bias-corrected estimates

We will now consider the implications of these results for the magnitudes of TEBS estimates. Using the OLS specification from Table 3, we construct fitted estimates of TEBS for different groups of consumer goods and geographic regions—conditional on correcting for publication bias. Panel A of Table 4 reports bias-corrected point estimates of TEBS, obtained by substituting sample means for all variables except the standard error and the corresponding group variable. The overall point estimate of -0.411 has similar interpretation as the constant term in Table 2: it is the overall bias-corrected estimate of TEBS; as before, this estimate is smaller in magnitude compared to the mean TEBS reported by the literature—because it is corrected for selective reporting.

Table 4: Fitted Estimates by Group

	Point Estimate	95% interval
Panel A: Fitted estimates with bias correction		
All	-0.411	[-0.534 ; -0.287]
Product groups		
Food & retail	-2.042	[-3.628; -0.455]
Fuel	-0.873	[-1.532; -0.215]
Tobacco and other sin products	-0.068	[-0.359; 0.223]
Regions		
U.S.	-0.617	[-0.848; -0.385]
E.U. & other countries	0.319	[-0.183; 0.821]
Panel B: 'Best practice' estimates with bias correction		
	Point Estimate	95% interval
All	-0.978	[-1.971; 0.015]
Product groups		
Food & retail	-2.609	[-4.858; -0.360]
Fuel	-1.441	[-2.817; -0.065]
Tobacco and other sin products	-0.636	[-1.484; 0.212]
Regions		
U.S.	-1.184	[-2.212; -0.156]
E.U. & other countries	-0.249	[-1.302; 0.805]

Notes: Here we report fitted values of TEBS obtained using estimation results of the frequentist check reported in the right panel of Table 3. For both panels, to compute the estimates and the confidence intervals, we substitute zeros for standard errors—to correct for publication bias. In panel B we additionally substitute the values of the 90th percentile for the citations and one for 'top journal'.

At the same time, fitted TEBS estimates differ depending on the consumption group and the geographic region considered. Sales of food and retail appear to be the most elastic, with TEBS of about -2; sales of fuel are elastic as well (TEBS of -0.87). Tobacco sales and sales of other 'sin' products appear much less elastic: the point estimate is small in magnitude and its confidence interval includes zero. There are also prominent geographic differences in overall estimated effects. While for the US, TEBS is negative and prominent, it is not so for EU and other countries: for these regions, once we account for selective reporting, the effect becomes statistically indistinguishable from zero.

We additionally compute TEBS estimates that are conditional on 'best practice' in the literature: for this exercise, we substitute 1 for *Top journal* and the value of the 90th percentile for *Citations* (as well as, again, correcting for publication bias). This is done in an attempt to give higher weight to studies of potentially higher quality. As we show in Panel B of Table 4, TEBS estimates that are widely cited and published in top journals tend to be greater in magnitude compared to the estimates reported by the literature overall—at least according to the corresponding point estimates.

Overall, even when the publication bias is accounted for, the literature still suggests that consumers engage in the cross-border shopping behavior—particularly where fuel, food and retail purchases are concerned. At the same time, sales of tobacco and other individual 'sin' products appear less elastic: with the exception of fuel, consumers may be less likely to travel across jurisdictions to take advantage of the price differences for individual 'sin' goods.

5 Conclusion

We conduct the first meta-analysis of the literature estimating tax elasticity of border sales, collecting 749 estimates from 60 studies in this field. We demonstrate that the literature is prone to systematic underreporting of positive estimates, which gives rise to a bias in the mean reported TEBS estimate that exaggerates the elasticity of sales at the border. We provide appropriate bias corrections.

We find that cross-border sales of food, retail and fuel are much more elastic to changes in tax rates compared to sales of tobacco and other 'sin' products. Sales tax has a broader base compared to excise taxes. Spending on food and retail constitutes a large fraction of household budget—changes in the local sales tax may prompt households to seek opportunities to save by shopping across the border. Compared to food and retail, household spending on fuel may be smaller. However, unlike other goods, fuel consumption is itself associated with travel; furthermore, gas stations tend to prominently display the associated fuel prices, facilitating cross-border comparisons for travelers—these considerations may explain the associated high elasticity of border sales for fuel.

Our meta-analysis shows that there is cross-border shopping behavior that may affect the stability of profits from sales and excise taxes, as well as the intended use of the latter as Pigouvian instruments aimed at reducing consumption of the 'sin' products. In particular, we find that cross-border shopping is prevalent in the US, possibly due to the wide-spread emphasis on personal transportation. At the same time, once publication selection bias is accounted for, TEBS for Europe and other countries becomes statistically indistinguishable from zero, and TEBS for the US becomes somewhat less prominent.

Our meta-analysis provides future researchers with a comprehensive framework for comparing their estimates to those reported in the existing literature. Our findings may be particularly beneficial for the discussions that emerged in recent years, of the new types of excise taxes—taxes on sugar, marijuana and fat—and of the interjurisdictional tax differences coming from sales tax holidays and other tax relief measures by state and local governments. In the future, many states may expand their sales tax base to include more services, which could also create differences in the tax systems of neighboring jurisdictions and trigger more cross-border shopping.

We believe that there will be many more studies of TEBS in the coming years: studies that will focus on the new types of taxes and address the sharp changes in cross-border shopping patterns prompted by the COVID-19 pandemic.

References

Alm, J. & W. Beck (1991): "Wiping the Slate Clean: Individual Response to State Tax Amnesties." *Southern Economic Journal* **57(4)**: pp. 1043–1053.

Alm, J. & W. Beck (1993): "Tax Amnesties and Compliance in the Long Run: A Time Series Analysis." *National Tax Journal* **46(1)**: pp. 53–60.

Alm, J., M. J. McKee, & W. Beck (1990): "Amazing Grace: Tax Amnesties and Compliance." *National Tax Journal* **43(1)**: pp. 23–37.

Andrews, I. & M. Kasy (2019): "Identification of and Correction for Publication Bias." *American Economic Review* **109(8)**: pp. 2766–94.

Apergis, N., R. Goel, & J. Payne (2014): "Dynamics of U.S. State Cigarette Consumption: Evidence from Panel Error Correction Modeling." *Atlantic Economic Journal* **42(1)**: pp. 3–20.

Asplund, M., R. Friberg, & F. Wilander (2007): "Demand and Distance: Evidence on Cross-border Shopping." *Journal of public Economics* **91(1-2)**: pp. 141–157.

Balima, H. W. & A. Sokolova (2021): "IMF Programs and Economic Growth: A meta-analysis." *Journal of Development Economics* **153**: p. 102741.

Ballard, C. L. & J. Lee (2007): "Internet Purchases, Cross-Border Shopping, and Sales Taxes." *National Tax Journal* **60(4)**: pp. 711–725.

Baltagi, B. H. & J. M. Griffin (1995): "A Dynamic Demand Model for Liquor: the Case for Pooling." *The Review of Economics and Statistics* **77(3)**: pp. 545–554.

Banfi, S., M. Filippini, & L. C. Hunt (2005): "Fuel Tourism in Border Regions: the Case of Switzerland." *Energy Economics* **27(5)**: pp. 689–707.

Baranzini, A. & S. Weber (2013): "Elasticities of Gasoline Demand in Switzerland." *Energy Policy* **63(C)**: pp. 674–680.

Beard, T. R., P. A. Gant, & R. P. Saba (1997): "Border-Crossing Sales, Tax Avoidance, and State Tax Policies: An Application to Alcohol." *Southern Economic Journal* **64(1)**: pp. 293–306.

Becker, G. S., M. Grossman, & K. M. Murphy (1994): "An Empirical Analysis of Cigarette

Addiction." *American Economic Review* **84(3)**: pp. 396–418.

Besley, T. J. & H. S. Rosen (1999): "Sales Taxes and Prices: An Empirical Analysis." *National Tax Journal* **52(2)**: pp. 157–178.

Brodeur, A., M. Le, M. Sangnier, & Y. Zylberberg (2016): "Star Wars: The Empirics Strike Back." *American Economic Journal: Applied Economics* **8(1)**: pp. 1–32.

Card, D., J. Kluve, & A. Weber (2017): "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations." *Journal of the European Economic Association* **16(3)**: pp. 894–931.

Card, D. & A. B. Krueger (1995): "Time-Series Minimum-Wage Studies: A Meta-Analysis." *American Economic Review* **85(2)**: pp. 238–43.

Chernick, H. & D. Merriman (2013): "Using Littered Pack Data to Estimate Cigarette Tax Avoidance in NYC." *National Tax Journal* **66(3)**: pp. 635–668.

Chetty, R., A. Guren, D. Manoli, & A. Weber (2013): "Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities." *NBER Macroeconomics Annual* **27(1)**: pp. 1–56.

Connelly, R., R. Goel, & R. Ram (2009): "Demand for Cigarettes in the United States: Effects of Prices in Bordering States and Contiguity with Mexico and Canada." *Applied Economics* **41(18)**: pp. 2255–2260.

DeCicca, P., D. Kenkel, & F. Liu (2013): "Excise Tax Avoidance: the Case of State Cigarette Taxes." *Journal of Health Economics* **32(6)**: pp. 1130–1141.

Disdier, A.-C. & K. Head (2008): "The Puzzling Persistence of the Distance Effect on Bilateral Trade." *The Review of Economics and Statistics* **90(1)**: pp. 37–48.

Doucouliagos, H. & T. D. Stanley (2009): "Publication Selection Bias in Minimum-Wage Research? A Meta-Regression Analysis." *British Journal of Industrial Relations* **47(2)**: pp. 406–428.

Egger, M., G. D. Smith, M. Scheider, & C. Minder (1997): "Bias in Meta-Analysis Detected by a Simple, Graphical Test." *British Medical Journal* **316**: pp. 629–634.

Eicher, T. S., C. Papageorgiou, & A. E. Raftery (2011): "Default Priors and Predictive Performance in Bayesian Model Averaging, with Application to Growth Determinants." *Journal of Applied Econometrics* **26(1)**: pp. 30–55.

Farrelly, M. C., T. F. Pechacek, & F. J. Chaloupka (2003): "The Impact of Tobacco Control Program Expenditures on Aggregate Cigarette Sales: 1981-2000." *Journal of Health Economics* **22(5)**: pp. 843–859.

Fernández, C., E. Ley, & M. F. J. Steel (2001): "Benchmark Priors for Bayesian Model Averaging." *Journal of Econometrics* **100(2)**: pp. 381–427.

Fernandez, C., E. Ley, & M. F. J. Steel (2001): "Model Uncertainty in Cross-Country Growth Regressions." *Journal of Applied Econometrics* **16(5)**: pp. 563–576.

Fox, W. F. (1986): "Tax Structure And The Location Of Economic Activity Along State Borders." *National Tax Journal* **39(4)**: pp. 387–401.

Friberg, R., F. Steen, & S. A. Ulsaker (2022): "Hump-shaped Cross-price Effects and the Extensive Margin in Cross-border Shopping." *American Economic Journal: Microeconomics* **14(2)**: pp. 408–38.

Fujii, E., M. Khaled, & J. Mak (1985): "The Exportability of Hotel Occupancy and Other Tourist Taxes." *National Tax Journal* **38(2)**: pp. 169–177.

Gade, M. N. & L. C. Adkins (1990): "Tax Exporting and State Revenue Tax Structures." *National Tax Journal* **43(1)**: pp. 39–52.

Goel, R. K. (2008): "Cigarette Smuggling: Price vs. Nonprice Incentives." *Applied Economics Letters* **15(8)**: pp. 587–592.

Goel, R. K. & M. A. Nelson (2012): "Cigarette Demand and Effectiveness of U.S. Smoking Control Policies: State-Level Evidence for More than Half a Century." *Empirical Economics* **42(3)**: pp. 1079–1095.

Goolsbee, A., M. F. Lovenheim, & J. Slemrod (2010): “Playing with Fire: Cigarettes, Taxes, and Competition from the Internet.” *American Economic Journal: Economic Policy* **2(1)**: pp. 131–54.

Gunby, P., Y. Jin, & W. Robert Reed (2017): “Did FDI Really Cause Chinese Economic Growth? A Meta-Analysis.” *World Development* **90(C)**: pp. 242–255.

Hansen, B., K. Miller, & C. Weber (2020): “Federalism, Partial Prohibition, and Cross-border Sales: Evidence from Recreational Marijuana.” *Journal of Public Economics* **187(C)**.

Havranek, T. (2015): “Measuring Intertemporal Substitution: The Importance Of Method Choices And Selective Reporting.” *Journal of the European Economic Association* **13(6)**: pp. 1180–1204.

Havranek, T., R. Horvath, Z. Irsova, & M. Rusnak (2015): “Cross-Country Heterogeneity in Intertemporal Substitution.” *Journal of International Economics* **96(1)**: pp. 100–118.

Havranek, T., M. Rusnak, & A. Sokolova (2017): “Habit Formation in Consumption: A Meta-Analysis.” *European Economic Review* **95**: pp. 142 – 167.

Havranek, T. & A. Sokolova (2020): “Do Consumers Really Follow a Rule of Thumb? Three Thousand Estimates from 144 Studies Say 'Probably Not'.” *Review of Economic Dynamics* **35**: pp. 97–122.

Hogan, T. D. & R. B. Shelton (1973): “Interstate Tax Exportation and States' Fiscal Structure.” *National Tax Journal* **26(4)**: pp. 553–564.

Ioannidis, J. P. A., T. D. Stanley, & H. Doucouliagos (2017): “The Power of Bias in Economics Research.” *Economic Journal* **127(605)**: pp. 236–265.

Jansen, D.-J. & N. Jonker (2018): “Fuel Tourism in Dutch Border Regions: Are Only Salient Price Differentials Relevant?” *Energy Economics* **74(C)**: pp. 143–153.

Khan, M. S., P. N. Thompson, & V. J. Tremblay (2020): “Marijuana Tax Incidence, Stockpiling, and Cross-border Substitution.” *International Tax and Public Finance* **27(1)**: pp. 103–127.

Lakhdar, C. B., N. G. Vaillant, & F.-C. Wolff (2016): “Does Smoke Cross the Border? Cigarette Tax Avoidance in France.” *The European Journal of Health Economics* **17(9)**: pp. 1073–1089.

Leal, A., J. López-Laborda, & F. Rodrigo (2009): “Prices, Taxes and Automotive Fuel Cross-border Shopping.” *Energy Economics* **31(2)**: pp. 225–234.

Lesley, C. & M. Erich (2008): “Crossing the Line: Direct Estimation of Cross-Border Cigarette Sales and the Effect on Tax Revenue.” *The B.E. Journal of Economic Analysis & Policy* **8(1)**: pp. 1–41.

Ley, E. & M. F. Steel (2009): “On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regressions.” *Journal of Applied Econometrics* **24(4)**: pp. 651–674.

Liang, F., R. Paulo, G. Molina, M. A. Clyde, & J. O. Berger (2008): “Mixtures of G Priors for Bayesian Variable Selection.” *Journal of the American Statistical Association* **103**: pp. 410–423.

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

Luitel, H. & M. Tosun (2014): “A Reexamination of State Fiscal Health and Amnesty Enactment.” *International Tax and Public Finance* **21(5)**: pp. 874–893.

Luitel, H. S. & R. S. Sobel (2007): “The Revenue Impact of Repeated Tax Amnesties.” *Public Budgeting & Finance* **27(3)**: pp. 19–38.

McLure, C. E. (1967): “The Interstate Exporting of State and Local Taxes: Estimates for 1962.” *National Tax Journal* **20(1)**: pp. 49–77.

Merriman, D. (2010): “The Micro-geography of Tax Avoidance: Evidence from Littered Cigarette Packs in Chicago.” *American Economic Journal: Economic Policy* **2(2)**: pp. 61–84.

Minnesota Department of Revenue Tax Research Division (2021): “2021 Minnesota Tax Incidence Study: An Analysis of Minnesota's Household and Business Taxes.”

Morgan, W., J. Mutti, & D. Rickman (1996): "Tax Exporting, Regional Economic Growth, and Welfare." *Journal of Urban Economics* **39(2)**: pp. 131–159.

Murray, M. (2006): "Exporting State and Local Taxes: An Application to the State of Maine." Brookings Institution Background Paper.

Mutti, J. H. & W. E. Morgan (1983): "The Exportation of State and Local Taxes in a Multilateral Framework: The Case of Household Type Taxes." *National Tax Journal* **36(4)**: pp. 459–475.

Nechyba, T. (1996): "Fiscal Federalism and Local Public Finance: A Computable General Equilibrium (CGE) Framework." *International Tax and Public Finance* **3(2)**: pp. 215–231.

Nelson, M. A. (2002): "Using Excise Taxes to Finance State Government: Do Neighboring State Taxation Policy and Cross-Border Markets Matter?" *Journal of Regional Science* **42(4)**: pp. 731–752.

Ohsawa, Y. (1999): "Cross-Border Shopping and Commodity Tax Competition Among Governments." *Regional Science and Urban Economics* **29(1)**: pp. 33–51.

Pollock, S. H. (1991): "Mechanisms for Exporting the State Sales Tax Burden in the Absence of Federal Deductibility." *National Tax Journal* **44(3)**: pp. 297–310.

Poterba, J. M. (1996): "Retail Price Reactions to Changes in State and Local Sales Taxes." *National Tax Journal* **49(2)**: pp. 165–176.

Raftery, A. E., D. Madigan, & J. A. Hoeting (1997): "Bayesian Model Averaging for Linear Regression Models." *Journal of the American Statistical Association* **92(437)**: pp. 179–191.

Seiler, S., A. Tuchman, & S. Yao (2021): "The Impact of Soda Taxes: Pass-Through, Tax Avoidance, and Nutritional Effects." *Journal of Marketing Research* **58(1)**: pp. 22–49.

Sokolova, A. & T. Sorensen (2021): "Monopsony in Labor Markets: A Meta-Analysis." *ILR Review* **74(1)**: pp. 27–55.

Stanley, T. D. (2001): "Wheat from Chaff: Meta-analysis as Quantitative Literature Review." *Journal of Economic Perspectives* **15(3)**: pp. 131–150.

Stanley, T. D. (2005): "Beyond Publication Bias." *Journal of Economic Surveys* **19(3)**: pp. 309–345.

Stanley, T. D. (2008): "Meta-Regression Methods for Detecting and Estimating Empirical Effects in the Presence of Publication Selection." *Oxford Bulletin of Economics and Statistics* **70(1)**: pp. 103–127.

Stanley, T. D. & H. Doucouliagos (2015): "Neither Fixed nor Random: Weighted Least Squares Meta-Analysis." *Statistics in Medicine* **34(13)**: pp. 2116–27.

Stehr, M. (2007): "The Effect of Sunday Sales Bans and Excise Taxes on Drinking and Cross–Border Shopping for Alcoholic Beverages." *National Tax Journal* **60(1)**: pp. 85–105.

Tibshirani, R. (1996): "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society. Series B (Methodological)* **58(1)**: pp. 267–288.

Tosun, M. S. & M. Skidmore (2007): "Cross-Border Shopping and the Sales Tax: An Examination of Food Purchases in West Virginia." *The B.E. Journal of Economic Analysis & Policy* **7(1)**: pp. 1–20.

Trandel, G. A. (1992): "Evading the Use Tax on Cross-Border Sales : Pricing and Welfare Effects." *Journal of Public Economics* **49(3)**: pp. 313–331.

Walsh, M. J. & J. D. Jones (1988): "More Evidence on the "Border Tax" Effect: The Case of West Virginia, 1979-84." *National Tax Journal* **41(2)**: pp. 261–265.

Wooster, R. B. & J. W. Lehner (2010): "Reexamining The Border Tax Effect: A Case Study Of Washington State." *Contemporary Economic Policy* **28(4)**: pp. 511–523.

Ye, Y. & W. C. Kerr (2016): "Estimated Increase in Cross-border Purchases by Washington Residents Following Liquor Privatization and Implications for Alcohol Consumption Trends." *Addiction* **111(11)**: pp. 1948–1953.

Zeugner, S. & M. Feldkircher (2015): "Bayesian Model Averaging Employing Fixed and Flexible Priors – The BMS Package for R." *Journal of Statistical Software* **68(4)**: pp. 1–37.

Appendix A Description of Variables

Table A1: Definitions and summary statistics of explanatory variables.

Variable	Description	Mean (all)	Std. dev (all)
Data characteristics			
Standard error	= standard error of the TEBS estimate.	0.536	1.264
Observations	= log number of observations in the dataset.	7.318	2.019
Average year	= the log of midyear of data used minus 1961, i.e. the earliest midyear in our sample.	3.288	0.746
Daily & weekly & monthly	= 1 if the frequency of the data used is daily, weekly or monthly (baseline: frequency is annual).	0.210	0.407
Quarterly	= 1 if the frequency of the data used is quarterly (baseline: frequency is annual).	0.075	0.263
Household & county level	= 1 if estimated at a household or county level, or at a region level of size comparable to a county (baseline: state level).	0.144	0.352
Country level	= 1 if estimated at a country level (baseline: state level).	0.117	0.322
U.S. local	= 1 if data is taken from a U.S. state or local source.	0.148	0.356
Economic, geographic & demographic controls			
Control: Per capita income	= 1 if per capita personal income is included in regression.	0.712	0.453
Control: Unemployment	= 1 if unemployment rate is included in regression.	0.089	0.286
Control: Demographic	= 1 if age and/or race are controlled for.	0.272	0.445
E.U. & other countries	= 1 if data comes from Europe, or other countries excluding the US (baseline: data from the US).	0.220	0.415
Consumption data features			
Food & retail	= 1 if estimate corresponds to food or retail (baseline: tobacco).	0.126	0.332
Fuel	= 1 if estimate corresponds to fuel (baseline: tobacco).	0.117	0.322
Other sin products	= 1 if estimate corresponds to other 'sin' products, i.e. marijuana, alcohol or soda (baseline: tobacco).	0.109	0.312
Estimation approaches			
GLS	= 1 if GLS or FGLS is employed (baseline: OLS).	0.077	0.267
DID	= 1 if Difference in Differences method is employed (baseline: OLS).	0.063	0.243
IV	= 1 if IV method is employed (baseline: OLS).	0.097	0.297
Probit	= 1 if Probit is employed (baseline: OLS).	0.049	0.217
Other method	= 1 if other method is employed, e.g. Spatial, Semi-structual, ARIMA (baseline: OLS).	0.056	0.230
State fixed effect	= 1 if state fixed effects are included.	0.497	0.500
Time fixed effect	= 1 if time fixed effects are included.	0.247	0.432

Continued on next page

Table A1: Definitions and summary statistics of explanatory variables (continued)

Variable	Description	Mean (all)	SD (all)
Specifications			
Sales in $	= 1 if the dependent variable is sales measured in dollars/local currency (baseline: sales measured as quantity).	0.203	0.402
Weighted by border pop.	= 1 if neighbor prices are weighted by the size of the respective border population.	0.077	0.267
Price index	= 1 if elasticity reflects the response to changes in the relative price index as in (2) (baseline: elasticity reflects response to changes in neighbor price).	0.414	0.493
Mean distance	= 1 if regression includes interaction of neighbor price (price index) with distance. In such cases we obtain the estimate comparable to the rest of the sample by calculating elasticity assuming mean distance considered in the paper.	0.134	0.340
Log-linear	= 1 if the model is log-linear.	0.072	0.259
Publication characteristics			
Top journal	= 1 if the study was published in one of the top five general interest journals in economics or the top field journal in public fiance.	0.025	0.157
Citations	= log value of citations per year since the paper first appeared on Google Scholar.	0.534	1.635
Numbers of studies	60		
Observations	749		

Notes: We report means and standard deviations for the full sample of 749 observations. When indicator variables form groups, we state the reference category.

Appendix B Publication Bias: Additional Results

Appendix B.1 Outlier tratments

Table B1: Testing for publication bias

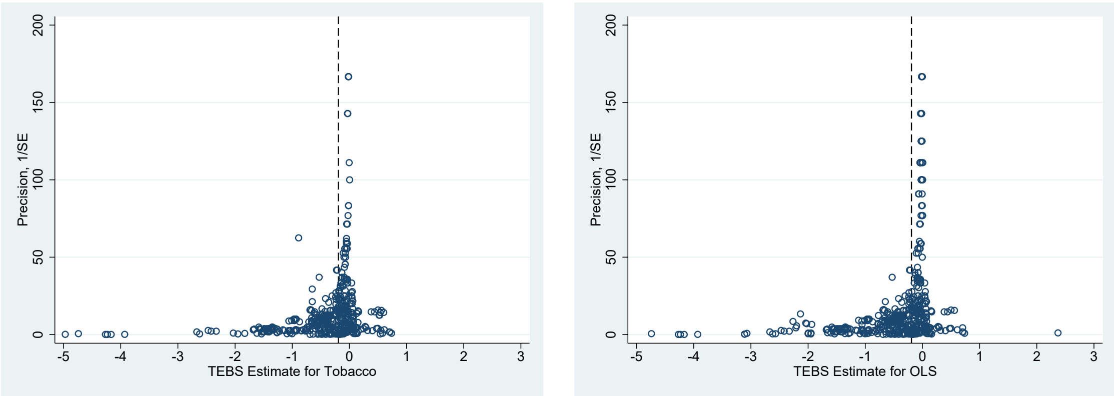
	OLS	FE	BE	Precision	Study
Winsorized at 1%					
SE dummy	-0.330	-0.314	-0.681	-1.004	-0.782
	(0.121)	(0.092)	(0.174)	(0.331)	(0.332)
Constant	-0.395	-0.403	-0.509	-0.038	-0.485
	(0.093)	(0.049)	(0.251)	(0.017)	(0.204)
Studies	60	60	60	60	60
Observations	749	749	60	749	749
Winsorized at 5%					
SE dummy	-0.483	-0.503	-0.942	-1.028	-0.986
	(0.204)	(0.166)	(0.122)	(0.306)	(0.128)
Constant	-0.291	-0.282	-0.198	-0.045	-0.209
	(0.083)	(0.075)	(0.132)	(0.019)	(0.083)
Studies	60	60	60	60	60
Observations	749	749	60	749	749
Dropped 1%					
SE dummy	-0.294	-0.307	-0.679	-0.955	-0.733
	(0.096)	(0.083)	(0.169)	(0.315)	(0.331)
Constant	-0.386	-0.379	-0.510	-0.035	-0.516
	(0.087)	(0.044)	(0.243)	(0.015)	(0.203)
Studies	60	60	60	60	60
Observations	743	743	60	743	743
Dropped 5%					
SE dummy	-0.195	-0.239	-0.190	-0.831	-0.329
	(0.059)	(0.053)	(0.071)	(0.299)	(0.122)
Constant	-0.345	-0.324	-0.351	-0.036	-0.337
	(0.074)	(0.026)	(0.089)	(0.016)	(0.066)
Studies	54	54	54	54	54
Observations	714	714	54	714	714

Notes: This table checks the robustness of the results presented in Table 2 against the treatment of outliers. *Winsorized at 1%*= sample with outliers in each tail winsorized at 0.5%; *Winsorized at 5%*= sample with outliers in each tail winsorized at 2.5%; *Dropped 1%*= sample with 1% of outliers dropped; *Dropped 5%*= sample with 5% of outliers dropped. See also notes for Table 2.

Appendix B.2 Homogenous subsamples

Figure B1: Funnel plots.
Homogenous subsamples

(a) *Tobacco* only (b) *OLS* only



Notes: Figure B1(a) plots TEBS estimates for Tobacco against their precision; Figure B1(b) plots TEBS estimates obtained using OLS against their precision; the dashed lines indicate respective sample means. The figures are truncated at $1/SE = 200$.

Table B2: Testing for publication bias.
Homogenous subsamples

	OLS	FE	BE	Precision	Study
Panel A: Tobacco only					
SE	-0.204	-0.287	-0.030	-0.694	-0.233
	(0.057)	(0.080)	(0.050)	(0.329)	(0.026)
Constant	-0.318	-0.272	-0.331	-0.045	-0.313
	(0.099)	(0.045)	(0.070)	(0.025)	(0.063)
Studies	35	35	35	35	35
Observations	485	485	35	485	485
Panel B: OLS only					
	OLS	FE	BE	Precision	Study
SE	-0.244	-0.252	-0.607	-0.773	-0.670
	(0.072)	(0.052)	(0.139)	(0.331)	(0.310)
Constant	-0.366	-0.361	-0.329	-0.045	-0.386
	(0.101)	(0.032)	(0.224)	(0.027)	(0.171)
Studies	46	46	46	46	46
Observations	492	492	46	492	492

Notes: Panel A considers a subset of TEBS estimates that refer to Tobacco. Panel B considers a subset of TEBS estimates obtained using OLS. See also notes for Table 2.

Appendix B.3 Publication bias using Andrews & Kasy 2019

In this section we use the method designed by Andrews & Kasy 2019 (AK2019) to model selective reporting, estimate publication probabilities depending on the sign and significance of the results, and obtain the corrected 'unbiased' estimate implied by this model.

The AK2019 model we use assumes the following structure. The 'true' underlying elasticities vary across studies—for each latent study j, the corresponding underlying effect ϵ_j is drawn from a distribution (e.g. a t-distribution); the mean of this distribution, $\bar{\epsilon}$, is unknown, but will be estimated later—the estimate of $\bar{\epsilon}$ will be interpreted as the bias-corrected 'true' elasticity. The study j then generates estimates ϵ_{ij} that are drawn from a normal distribution with ϵ_j as a mean. Some of these estimates will be reported; the probability of reporting, p(Z), depends on the value of the estimates divided by their standard errors, Z.

In this paper we estimate the version of the AK2019 model discussed in Section IIIC of the body of their paper. Following the authors, we assume that the 'true' elasticity ϵ_j for each latent study j comes from a t-distribution with mean $\bar{\epsilon}$, $\bar{v}$ degrees of freedom and a scale parameter $\tilde{\tau}$. For each result, the probability of it being reported is encompassed by the following step function:

$$p(Z) \propto \begin{cases} \beta_{\mathrm{p},1} \text{ if Z } > 1.96, \\ \beta_{\mathrm{p},2} \text{ if } 0 < \text{ Z} \leq 1.96, \\ \beta_{\mathrm{p},3} \text{ if } -1.96 < \text{ Z} \leq 0, \\ 1 \text{ if Z } \leq -1.96. \end{cases} \tag{5}$$

Here, we normalize probability of reporting a negative result significant at the 5% level to 1 and use maximum likelihood to estimate the relative reporting probabilities for negative and positive insignificant results ($\beta_{\mathrm{p},3}$ and $\beta_{\mathrm{p},2}$), and for positive results significant at the 5% level ($\beta_{\mathrm{p},1}$).

The estimation results are reported in Table B3. We find the probability of reporting a positive result to be dramatically lower compared to a result that is negative and significant: a positive estimate is about 18 times less likely to be reported. A negative insignificant result is about seven times less likely to be reported compared to a result with a Z-score below -1.96. The estimate of $\bar{\epsilon}$, the mean 'true' elasticity, is very close to zero and positive. Thus, overall, the AK2019 test indicates even more selective reporting compared to the tests examined in Section 3.

Table B3: Testing for publication bias using Andrews & Kasy 2019

$\bar{\epsilon}$	$\tilde{\tau}$	$\tilde{\nu}$	$\beta_{p,1}$	$\beta_{p,2}$	$\beta_{p,3}$
0.014	0.066	1.561	0.055	0.055	0.144
(0.006)	(0.020)	(0.135)	(0.025)	(0.018)	(0.041)

Notes: Standard errors appearing in parenthesis are clustered at the study level. The results are obtained using the `Matlab` code for AK2019 that replicates their Table 3.

As discussed in the body of the paper, one caveat associated with the AK2019 model is that the estimates examined need to be relatively homogenous—otherwise the assumptions about the distributions of 'true' elasticities and the estimates produced by studies might not be appropriate. To check the robustness of our findings, we consider two subsamples of our data that are more homogenous: a subsample of estimates pertaining to sales of tobacco, and a subsample of estimates obtained through OLS. The results are reported in Table B4. The results are similar to those for the baseline sample, albeit the underlying 'true' effect is found to be negative for the OLS estimates. For the subsample of OLS estimates, the TEBS of -0.004 is statistically different from zero—but not for tobacco.

Table B4: Testing for publication bias using Andrews & Kasy (2019).
Homogenous subsamples

Panel A: Tobacco only					
$\bar{\epsilon}$	$\tilde{\tau}$	$\tilde{\nu}$	$\beta_{p,1}$	$\beta_{p,2}$	$\beta_{p,3}$
0.005	0.091	1.218	0.046	0.077	0.191
(0.009)	(0.056)	(0.310)	(0.027)	(0.031)	(0.065)
Panel B: OLS only					
$\bar{\epsilon}$	$\tilde{\tau}$	$\tilde{\nu}$	$\beta_{p,1}$	$\beta_{p,2}$	$\beta_{p,3}$
-0.004	0.038	0.750	0.048	0.072	0.214
(0.001)	(0.020)	(0.119)	(0.029)	(0.034)	(0.082)

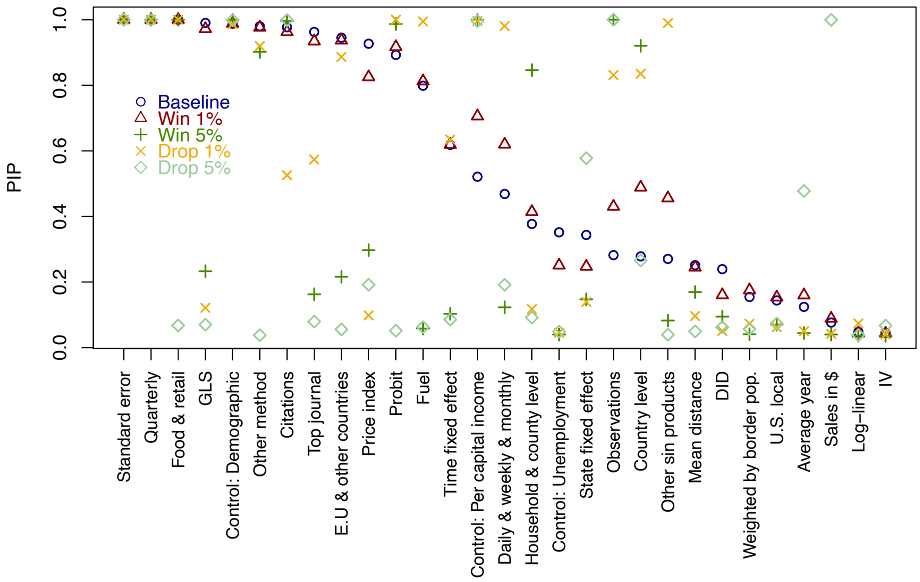
Notes: Standard errors appearing in parenthesis are clustered at the study level. The results are obtained using the `Matlab` code for AK2019 that replicates their Table 3. Panel A considers a subset of TEBS estimates that refer to Tobacco. Panel B considers a subset of TEBS estimates obtained using OLS

Appendix C Why do estimates vary? Additional results

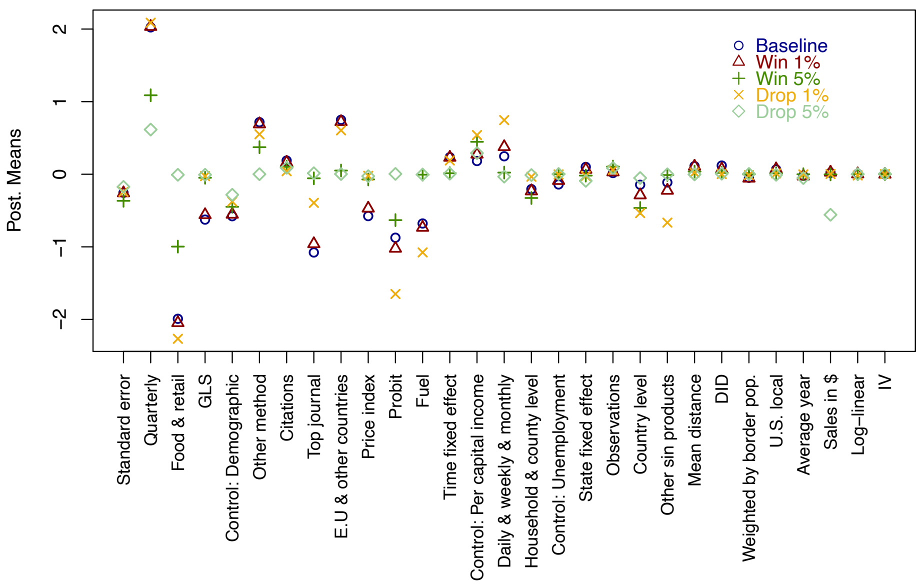
Appendix C.1 Robustness

Figure C1: Outlier treatments

(a) PIP



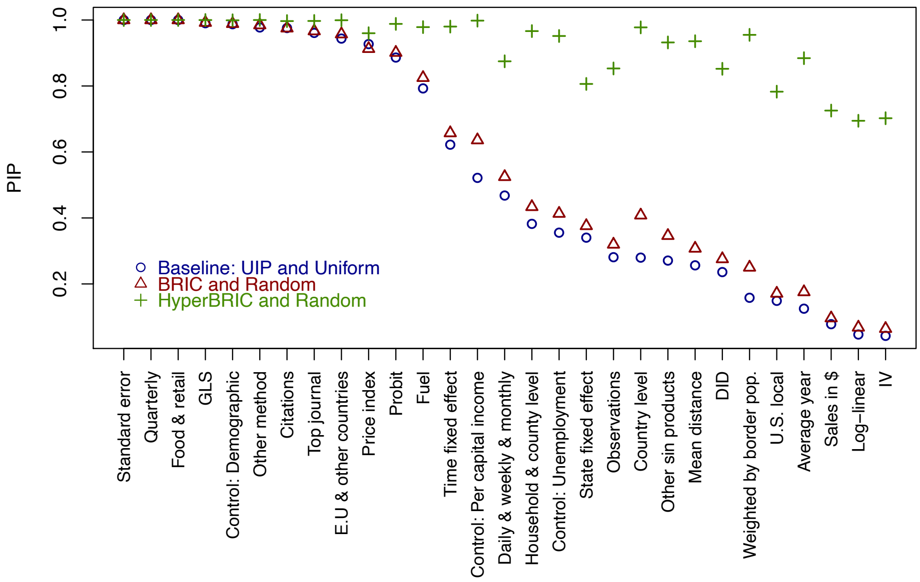
(b) Posterior means



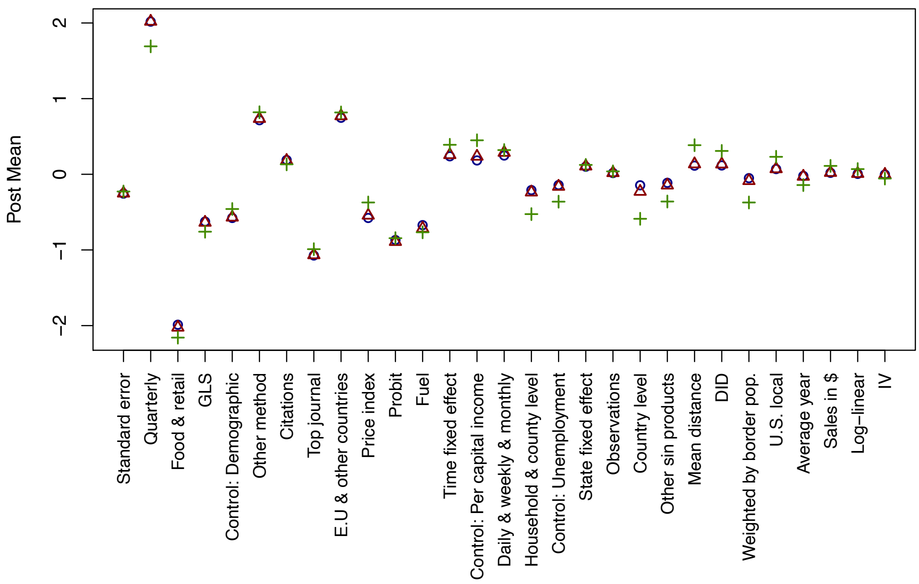
Notes: This figure compares BMA results under alternative outlier treatments. *Winsorized at 1%*= sample with outliers in each tail winsorized at 0.5%; *Winsorized at 5%*= sample with outliers in each tail winsorized at 2.5%; *Dropped 1%*= sample with 1% of outliers dropped; *Dropped 5%*= sample with 5% of outliers dropped.

Figure C1: Priors

(c) PIP



(d) Posterior means



Notes: This figure compares BMA results under alternative assumptions about priors. *UIP*=Unit information prior for model parameters; *Uniform*=Uniform prior for model space; *'BRIC'* = Benchmark g-priors for parameters, see Fernández *et al.* (2001); *'HyperBRIC'*=Flexible data-dependent priors for parameters, see Liang *et al.* (2008); *'Random'*=Beta-binomial prior for model space, see Ley & Steel (2009).

Table C1: Why do estimates of supply elasticity vary? LASSO

	LASSO	OLS using selected variables	
Response variable	Coef.	Coef.	SE
Data characteristics			
Standard error	-0.235	-0.236	0.067
Observations	0.044	0.046	0.070
Average year	-0.163	-0.174	0.224
Daily & weekly & monthly	0.292	0.315	0.696
Quarterly	1.662	1.696	0.968
Household & county level	-0.584	-0.581	0.542
Country level	-0.566	-0.610	0.470
U.S local	0.279	0.307	0.523
Economic, geographic & demographic controls			
Control: Per capita income	0.460	0.482	0.202
Control: Unemployment	-0.412	-0.412	0.507
Control: Demographic	-0.456	-0.462	0.236
E.U. & other countries	0.793	0.854	0.357
Consumption data features			
Food & Retail	-2.226	-2.297	1.178
Fuel	-0.705	-0.783	0.664
Other sin products	-0.371	-0.414	0.617
Estimation approaches			
GLS	-0.802	-0.828	0.684
DID	0.331	0.365	0.443
IV	-0.077	-0.085	0.163
Probit	-0.843	-0.876	1.164
Other method	0.836	0.866	0.452
State fixed effect	0.128	0.137	0.333
Time fixed effect	0.399	0.400	0.240
Specifications			
Sale in $	0.130	0.156	0.701
Weighted border pop.	-0.385	-0.406	0.331
Price index	-0.384	-0.389	0.362
Mean distance	0.426	0.422	0.361
Log-linear	0.105	0.096	0.504
Publication characteristics			
Top journal	-0.994	-1.050	0.740
Citations	0.134	0.135	0.105
Const.	-0.313	-0.302	0.867
Observations	749	749	.

Notes: The left panel presents estimates obtained using LASSO with the penalty set to minimize mean-squared prediction error with cross-validation. We implement this in `STATA` with the `cvlasso` routine. The right panel shows results of estimating OLS using the subset of variables selected by LASSO.

Appendix D Studies Used in Meta-analysis

We used the following search query to find the relevant studies:

> Our search query contains the words 'relative price', 'tax difference elasticity estimate border', 'cross-border', "cross-border", 'neighboring sales', 'tax', 'shopping', 'purchases', and 'estimates' save the Google Scholar search results on April 30th, 2021. We read through abstracts of papers appearing on the first 50 pages, download and save papers that appear to report relevant results.
>
> We then read through the downloaded papers making sure they contain estimates of TEBS, or results that could be used to derive the implied TEBS estimates, using conversions discussed in the previous section. We include papers published in peer-reviewed journals, and working papers that came out after the year 2018.
>
> This search strategy yielded 60 studies providing 749 estimates. These estimates cover a diverse set of geographic regions and types of sales.

Papers in the Study

Apergis, N., R. Goel, & J. Payne (2014): "Dynamics of U.S. State Cigarette Consumption: Evidence from Panel Error Correction Modeling." *Atlantic Economic Journal* **42(1)**: pp. 3–20.

Asplund, M., R. Friberg, & F. Wilander (2007): "Demand and Distance: Evidence on Cross-border Shopping." *Journal of public Economics* **91(1-2)**: pp. 141–157.

Ballard, C. L. & J. Lee (2007): "Internet Purchases, Cross-Border Shopping, and Sales Taxes." *National Tax Journal* **60(4)**: pp. 711–725.

Baltagi, B. H. & J. M. Griffin (1995): "A Dynamic Demand Model for Liquor: the Case for Pooling." *The Review of Economics and Statistics* **77(3)**: pp. 545–554.

Baltagi, B. H., J. M. Griffin, & W. Xiong (2000): "To Pool Or Not To Pool: Homogeneous Versus Hetergeneous Estimations Applied to Cigarette Demand." *The Review of Economics and Statistics* **82(1)**: pp. 117–126.

Baltagi, B. H. & D. Levin (1986): "Estimating dynamic demand for cigarettes using panel data: the effects of bootlegging, taxation and advertising reconsidered." *The Review of Economics and Statistics* pp. 148–155.

Baltagi, B. H. & D. Levin (1992): "Cigarette Taxation: Raising Revenues and Reducing Consumption." *Structural Change and Economic Dynamics* **3(2)**: pp. 321–335.

Banfi, S., M. Filippini, & L. C. Hunt (2005): "Fuel tourism in border regions: the case of switzerland." *Energy Economics* **27(5)**: pp. 689–707.

Baranzini, A. & S. Weber (2013): "Elasticities of Gasoline Demand in Switzerland." *Energy Policy* **63(C)**: pp. 674–680.

Beard, T. R., P. A. Gant, & R. P. Saba (1997): "Border-Crossing Sales, Tax Avoidance, and State Tax Policies: An Application to Alcohol." *Southern Economic Journal* **64(1)**: pp. 293–306.

Becker, G. S., M. Grossman, & K. M. Murphy (1994): "An Empirical Analysis of Cigarette Addiction." *American Economic Review* **84(3)**: pp. 396–418.

Chaloupka, F. J. & H. Saffer (1992): "Clean Indoor Air Laws and the Demand for Cigarettes."

Contemporary Economic Policy **10(2)**: pp. 72–83.

Chen, Y.-C. & C.-C. Lin (2012): "Information shocks and cigarette addiction: Views from dynamic panels with common structural changes." *Applied Economics* **44(35)**: pp. 4651–4660.

Chernick, H. & D. Merriman (2013): "Using Littered Pack Data to Estimate Cigarette Tax Avoidance in NYC." *National Tax Journal* **66(3)**: pp. 635–668.

Coats, R. M. (1995): "A Note on Estimating Cross-Border Effects of State Cigarette Taxes." *National Tax Journal* **48(4)**: pp. 573–584.

Connelly, R., R. Goel, & R. Ram (2009): "Demand for Cigarettes in the United States: Effects of Prices in Bordering States and Contiguity with Mexico and Canada." *Applied Economics* **41(18)**: pp. 2255–2260.

Connelly, R., R. Goel, & R. Ram (2010): "Income Inequality and Cigarette Consumption: Evidence from the United States." *Applied Economics Letters* **17(5)**: pp. 423–426.

DeCicca, P., D. Kenkel, & F. Liu (2013): "Excise Tax Avoidance: the Case of State Cigarette Taxes." *Journal of Health Economics* **32(6)**: pp. 1130–1141.

Doogan, N. J., M. E. Wewers, & M. Berman (2018): "The impact of a federal cigarette minimum pack price policy on cigarette use in the usa." *Tobacco control* **27(2)**: pp. 203–208.

Farrelly, M. C., T. F. Pechacek, & F. J. Chaloupka (2003): "The Impact of Tobacco Control Program Expenditures on Aggregate Cigarette Sales: 1981-2000." *Journal of Health Economics* **22(5)**: pp. 843–859.

Fox, W. F. (1986): "Tax Structure And The Location Of Economic Activity Along State Borders." *National Tax Journal* **39(4)**: pp. 387–401.

Friberg, R., F. Steen, & S. A. Ulsaker (2022): "Hump-shaped cross-price effects and the extensive margin in cross-border shopping." *American Economic Journal: Microeconomics* **14(2)**: pp. 408–38.

Gallet, C. A. (2003): "Advertising and Restrictions in the Cigarette Industry: Evidence of State-by-State Variation." *Contemporary Economic Policy* **21(3)**: pp. 338–348.

Ghoddusi, H., N. Rafizadeh, & M. H. Rahmati (2018): "Price Elasticity of Gasoline Smuggling: A Semi-structural Estimation Approach." *Energy Economics* **71(C)**: pp. 171–185.

Goel, R. (2004): "Cigarette Demand in Canada and the US-Canadian Cigarette Smuggling." *Applied Economics Letters* **11(9)**: pp. 537–540.

Goel, R. (2011a): "Persistence of Cigarette Advertising Across Media and Smoking Rates." *Applied Economics Letters* **18(7)**: pp. 611–619.

Goel, R. & M. Nelson (2012a): "Cigarette Demand and Effectiveness of U.S. Smoking Control Policies: State-Level Evidence for More than Half a Century." *Empirical Economics* **42(3)**: pp. 1079–1095.

Goel, R. K. (2008): "Cigarette Smuggling: Price vs. Nonprice Incentives." *Applied Economics Letters* **15(8)**: pp. 587–592.

Goel, R. K. (2011b): "Advertising Media And Cigarette Demand." *Bulletin of Economic Research* **63(4)**: pp. 404–416.

Goel, R. K. & M. A. Nelson (2012b): "Cigarette demand and effectiveness of u.s. smoking control policies: State-level evidence for more than half a century." *Empirical Economics* **42(3)**: pp. 1079–1095.

Goel, R. K. & M. A. Nelson (2012c): "Internet diffusion and implications for taxation: evidence from us cigarette sales." *Policy & Internet* **4(1)**: pp. 1–15.

Goel, R. K., J. E. Payne, & J. W. Saunoris (2016): "Spillovers From the Beer Market to U.S. Cigarette Demand." *Economic Modelling* **55(C)**: pp. 292–297.

Goel, R. K. & J. W. Saunoris (2019): "Cigarette smuggling: Using the shadow economy or creating its own?" *Journal of Economics and Finance* **43(3)**: pp. 582–593.

Goolsbee, A., M. F. Lovenheim, & J. Slemrod (2010): "Playing with Fire: Cigarettes, Taxes, and Competition from the Internet." *American Economic Journal: Economic Policy* **2(1)**: pp. 131–54.

Hansen, B., K. Miller, & C. Weber (2020): "Federalism, Partial Prohibition, and Cross-border Sales: Evidence from Recreational Marijuana." *Journal of Public Economics* **187(C)**.

Harbolt, B. (2019): "Tax avoidance through e-commerce and cross-border shopping." .

Harding, M., E. Leibtag, & M. F. Lovenheim (2012): "The Heterogeneous Geographic and Socioeconomic Incidence of Cigarette Taxes: Evidence from Nielsen Homescan Data." *American Economic Journal: Economic Policy* **4(4)**: pp. 169–198.

Huang, B.-N., C. wei Yang, & M. jeng Hwang (2004): "New Evidence on Demand for Cigarettes: A Panel Data Approach." *The International Journal of Applied Economics* **1(1)**: pp. 81–97.

Jansen, D.-J. & N. Jonker (2018): "Fuel Tourism in Dutch Border Regions: Are Only Salient Price Differentials Relevant?" *Energy Economics* **74(C)**: pp. 143–153.

Khan, M. S., P. N. Thompson, & V. J. Tremblay (2020): "Marijuana Tax Incidence, Stockpiling, and Cross-border Substitution." *International Tax and Public Finance* **27(1)**: pp. 103–127.

Lakhdar, C. B., N. G. Vaillant, & F.-C. Wolff (2016): "Does Smoke Cross the Border? Cigarette Tax Avoidance in France." *The European Journal of Health Economics* **17(9)**: pp. 1073–1089.

Lawley, C. & V. Thivierge (2018): "Refining the Evidence: British Columbias Carbon Tax and Household Gasoline Consumption." *The Energy Journal* **0(Number 2)**.

Leal, A., J. López-Laborda, & F. Rodrigo (2009): "Prices, Taxes and Automotive Fuel Cross-border Shopping." *Energy Economics* **31(2)**: pp. 225–234.

Lesley, C. & M. Erich (2008): "Crossing the Line: Direct Estimation of Cross-Border Cigarette Sales and the Effect on Tax Revenue." *The B.E. Journal of Economic Analysis & Policy* **8(1)**: pp. 1–41.

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

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

Mikesell, J. L. (1970): "Central Cities and Sales Tax Rate Differentials: the Border City Problem." *National Tax Journal* **23(2)**: pp. 206–213.

Mikesell, J. L. & C. K. Zorn (1986): "Impact of the Sales Tax Rate on Its Base: Evidence From a Small Town." *Public Finance Quarterly* **14(3)**: pp. 329–338.

Ornstein, S. I. & D. M. Hanssens (1985): "Alcohol control laws and the consumption of distilled spirits and beer." *Journal of consumer research* **12(2)**: pp. 200–213.

Seiler, S., A. Tuchman, & S. Yao (2021): "The Impact of Soda Taxes: Pass-Through, Tax Avoidance, and Nutritional Effects." *Journal of Marketing Research* **58(1)**: pp. 22–49.

Srithongrung, A. (2017): "Consumers' Behavioral Response to Sales Taxes On Food In Kansas." *Public Finance & Management* **17(2)**: p. 92.

Stehr, M. (2007): "The Effect of Sunday Sales Bans and Excise Taxes on Drinking and Cross–Border Shopping for Alcoholic Beverages." *National Tax Journal* **60(1)**: pp. 85–105.

Sung, H., T. Hu, & T. E. Keeler (1994): "Cigarette Taxation And Demand: An Empirical Model." *Contemporary Economic Policy* **12(3)**: pp. 91–100.

Thursby, J. G. & M. C. Thursby (2000): "Interstate Cigarette Bootlegging: Extent, Revenue Losses, and Effects of Federal Intervention." *National Tax Journal* **53(1)**: pp. 59–78.

Thursby, M., R. Jensen, & J. Thursby (1991): "Smuggling, camouflaging, and market structure." *The Quarterly Journal of Economics* **106(3)**: pp. 789–814.

Tosun, M. S. & M. Skidmore (2007): "Cross-Border Shopping and the Sales Tax: An Examination of Food Purchases in West Virginia." *The B.E. Journal of Economic Analysis & Policy* **7(1)**: pp. 1–20.

Walsh, M. J. & J. D. Jones (1988): "More Evidence on the "Border Tax" Effect: The Case of West Virginia, 1979-84." *National Tax Journal* **41(2)**: pp. 261–265.

Wooster, R. B. & J. W. Lehner (2010): "Reexamining The Border Tax Effect: A Case Study Of Washington State." *Contemporary Economic Policy* **28(4)**: pp. 511–523.

Ye, Y. & W. C. Kerr (2016): "Estimated Increase in Cross-border Purchases by Washington Residents Following Liquor Privatization and Implications for Alcohol Consumption Trends." *Addiction* **111(11)**: pp. 1948–1953.

Yurekli, A. A. & P. Zhang (2000): "The Impact of Clean Indoor-air Laws and Cigarette Smuggling on Demand for Cigarettes: an Empirical Model." *Health Economics* **9(2)**: pp. 159–170.