St. John Fisher University

Fisher Digital Publications

Economics Faculty/Staff Publications

Economics

8-2017

Retail Bottle Pricing at the Border: Evidence of Cross-Border Shopping, Fraudulent Redemptions, and Use Tax Evasion

Ben J. Niu

St. John Fisher University, bniu@sjf.edu

Follow this and additional works at: https://fisherpub.sjf.edu/economics_facpub

Part of the [Economics Commons](#)

[How has open access to Fisher Digital Publications benefited you?](#)

Publication Information

Niu, Ben J. (2017). "Retail Bottle Pricing at the Border: Evidence of Cross-Border Shopping, Fraudulent Redemptions, and Use Tax Evasion." *Journal of Economic Geography* 18.6, 1253-1283.

Please note that the Publication Information provides general citation information and may not be appropriate for your discipline. To receive help in creating a citation based on your discipline, please visit <http://libguides.sjfc.edu/citations>.

This document is posted at https://fisherpub.sjf.edu/economics_facpub/10 and is brought to you for free and open access by Fisher Digital Publications at . For more information, please contact fisherpub@sjf.edu.

Retail Bottle Pricing at the Border: Evidence of Cross-Border Shopping, Fraudulent Redemptions, and Use Tax Evasion

Abstract

This paper examines the pattern of retail prices for deposit eligible goods near Michigan's borders. Michigan's unique bottle redemption system and lower sales tax generate incentives for various potentially illegal household responses. Such incentives and behavior should be capitalized in the prices of affected goods. I empirically quantify the spatial price effects and find patterns consistent with theoretical predictions. Michigan's border prices are higher (lower) for goods with higher (lower) per unit costs by up to 38%. Price-distance trends reflect the waning of these effects away from the border.

Keywords

fsc2017

Disciplines

Economics

Comments

This is a pre-copyedited, author-produced version of an article accepted for publication in *Journal of Economic Geography* following peer review. The version of record [Ben J. Niu; Retail bottle pricing at the border: evidence of cross-border shopping, fraudulent redemptions, and use tax evasion, *Journal of Economic Geography*, Volume 18, Issue 6, 1 November 2018, Pages 1253–1283] is available online at: <https://doi.org/10.1093/jeg/lbx025>

Retail Bottle Pricing at the Border: Evidence of Cross-Border Shopping, Fraudulent Redemptions, and Use Tax Evasion

Journal of Economic Geography Submission
May 2017

Abstract

This paper examines the pattern of retail prices for deposit eligible goods near Michigan's borders. Michigan's unique bottle redemption system and lower sales tax generate incentives for various potentially illegal household responses. Such incentives and behavior should be capitalized in the prices of affected goods. I empirically quantify the spatial price effects and find patterns consistent with theoretical predictions. Michigan's border prices are higher (lower) for goods with higher (lower) per unit costs by up to 38%. Price-distance trends reflect the waning of these border effects away from the border.

Keywords: Bottle Deposits, Cross-Border Shopping, Tax Evasion, Spatial Pricing
JEL Codes: H71; R12; H26

1 Introduction

This paper focuses on the areas of Michigan and its two neighboring states of Indiana and Ohio. With nearly 200 miles of unobstructed land borders, the region is particularly susceptible to two prominent policy discontinuities. First, Michigan has a state level deposit-redemption system. This levies an additional per-bottle deposit cost on purchases of sodas and beer in Michigan. For those consumers that paid into the system, they can recoup their deposits by redeeming the empty containers. Conversely, Indiana and Ohio do not have such a policy. All else being equal, this makes purchases in Michigan more costly. Moreover, the inability to differentiate between Michigan bottles (paid the deposit) versus Indiana and Ohio bottles (did not pay the deposit) makes Michigan susceptible to fraudulent redemptions, i.e. payments made for returns of deposit-less, non-Michigan bottles. This concept was made famous in a 1996 episode of *Seinfeld* whereby a scheme was concocted to arbitrage between New York's five cent deposit and Michigan's ten cent deposit. By law, households can only recoup money that was paid into the system. As such, each fraudulent redemption constitutes a net drain of ten cents from Michigan's coffers. The policy discontinuity therefore creates a deposit wedge that pushes purchases to the Indiana and Ohio sides of the border. The second discontinuity arises due to differences in state sales tax rates. Michigan's 6% rate is lower than the 7% and 6.25% - 7.75% rates in Indiana and Ohio, respectively. Conversely, this pushes purchases towards the Michigan side of the border. The effect is particular strong for soda purchases. While all three states exempt groceries from sales taxation, Michigan is unique in including soda within this category. The tax wedge therefore promotes cross-border shopping and use tax evasion.

Despite the implied behavioral incentives, no direct data exist to confirm and/or quantify these illegal activities. This paper attempts to indirectly identify these effects by analyzing price patterns for deposit eligible goods near the borders of these three states. If arbitrage incentives do have a significant and real impact on household behavior, then such effects should be capitalized in the form of a sharp price discontinuity at the border. A simple, cross-border model suggests that the tax wedge is stronger than the deposit wedge when per unit prices are high. In these situations, Michigan's border prices should be higher than those on the other side. Conversely, Michigan's border prices should be lower when per unit prices are low and the wedges are reversed. Because the strength of these wedges vary with proximity to the border, this also suggests the presence of a second price effect, a price-distance trend, within each state.

The empirical analysis therefore focuses on identifying these two pricing effects. Using an original data set of retail prices for four deposit eligible goods near the borders, I find that actual price patterns

coincide for the most part with the theoretical predictions. The high per unit price good, 2 liter Cola, exhibits higher prices on the Michigan side whereas the lower per unit price good, 12 packs of Cola, exhibits lower prices on the Michigan side. Specifically, I estimate a border gap ranging between \$0.37 and \$0.66 for the case of 2 liter Colas (mean price of \$1.72), and -\$1.20 to -\$1.72 for 12 packs of Cola (mean price of \$4.68). A lesser Michigan border price is also expected in the case of 24 packs of Bud Lite. At the Ohio border, Michigan prices are indeed \$1.11 to \$1.23 cheaper at the Ohio border. At the Indiana border, however, the opposite is true where they are \$1.67 more expensive (mean price of \$17.52). In terms of the price-distance trends, estimates of these patterns indicate that the advantages/disadvantages of the wedges do dissipate as distance from the border increases. For example, 2 liter Colas are \$0.22 more expensive and \$0.27 cheaper when purchased at a retailer located 30 minutes from the border in Indiana and Michigan, respectively. Relative to the mean price, this constitutes a difference of roughly 13% and 16%. In the majority of cases where significant price effects are found, patterns are consistent with the theory and provide evidence in support of the wedges.

The remainder of the paper proceeds as follows. Section 2 provides background on the differing policies and literature. Section 3 presents a theoretical model of household and firm behavior that incorporates these two wedges. Section 4 discusses the data and empirical strategy. Section 5 presents and discusses the results. Finally, Section 6 discusses issues in the analysis.

2 Background

2.1 State-Level Policies

Figure 1: Map of the Region

Geographically, the length of the border and the distribution of population in the three states makes the region particularly susceptible to cross-border activity. In Michigan, the majority of the denser population centers are located in the middle of the state. The border areas are relatively rural with only two, relatively small metropolitan statistical areas. The Monroe MSA benefits by being the intermediate point between the Detroit-Livonia-Wayne (MI) and Toledo (OH) MSAs. Likewise, the Niles/Benton Harbor MSA benefits by being part of the Michigan-Indiana-Illinois corridor, a manufacturing and shipping route with Chicago, as well as leaning on Lake Michigan tourism. In Indiana and Ohio, the South Bend-Mishawaka (IN) and Toledo (OH) MSAs are very urban with significant suburban offshoots. In comparison to those on the Michigan border, these MSAs are larger. The location and proximity of these MSAs are

relevant when considering the two policy discontinuities of interest. These differences are summarized in Table 1.¹

Table 1: Policy Differences

	Bottle Deposit/ Redemption	Sales Tax on Beer	Sales Tax on Soda
Indiana	N/A	7%	7%
Michigan	10 cents	6%	0%
Ohio	N/A	6.25% - 7.75%	6.25% - 7.75%

Colloquially known as bottle bills, state-level deposit-redemption systems impose a per-bottle fee when consumers make purchases of sodas and beers in their original containers. This fee is then returned to the consumer when the empty bottles are redeemed typically via reverse vending machines (RVMs) located at the points of purchase, or at redemption centers. The usage of the word bottle refers generically to beverage containers made from aluminum, plastic, or glass. In 1971, Oregon became the first state to implement a bottle bill. It was originally intended as a method to address littering as cheap gas, highway expansions, and changes in product packaging trends increased the amount of empty containers being thrown onto roadways. Within ten years of implementation, the Oregon Department of Environmental Quality found that the percentage of roadside litter attributable to bottles decreased from 40% to 9%. Since then, California, Connecticut, Hawaii, Iowa, Maine, Massachusetts, Michigan, New York, and Vermont have all adopted similar systems.²

The theory behind using a deposit and redemption system is analyzed in Dobbs (1991). Post-consumption activities can generate added social costs coming not only from littering (eyesore and clean-up costs) but also from proper disposal (landfill, recycling process costs). Therefore, two Pigouvian instruments, a consumption/disposal tax (deposit) and a user fee (redemption), are necessary. A single redemption subsidy is therefore insufficient. As with current bottle bills, the paper also finds that the optimal deposit and redemption amounts are identical and offset. Ignoring administrative costs, this makes the policy cost and revenue neutral for both consumers and governments. Other papers, e.g. Eggert and Weichenrieder (2004); Fullerton and Wu (1998); Fullerton and Wolverton (2000), address this optimality question in the presence of additional factors such as escheat (unclaimed deposits) division and monopoly power.

More recently, the focus of bottle bills has shifted from litter reduction to increased recycling. Using national survey data, Viscusi et al. (2013) find that most households follow an all or nothing recycling pattern. Absent any policy, 45% always recycle while 25% never recycle. With a five cent deposit, the same sharp dichotomy was 62% and 8% in regards to redeeming and not redeeming, respectively. From their survey data, they also find that the effect of a bottle deposit is strongest for those with lower incomes. This is confirmed by Ashenmiller (2011) who finds that low income individuals in Santa Barbara (CA) derive a non-trivial portion of their income from redemptions. The author estimates that households earning less than \$10,000 generated \$340 annually from redemptions. Intuitively, the bottle bill acts as a mechanism that sorts labor such that those with low wages redeem both their own and other consumers' bottles. This added income was also found to potentially have a beneficial impact on crime rates. Using differences in the timing of bottle bill roll-outs, Ashenmiller (2010) estimates that crime rates were 11% lower on average in states with deposit-redemption systems.

¹ States also differ in their alcohol pricing laws. There are legislated minimum mark-up formulas for retailers. The extent to which firms actually price to the lowest mark-up is unclear but it is likely to depend on the degree of firm competition. Moreover, there are loopholes that allow retailers to skirt such regulations.

² Delaware had a bottle bill but replaced it in 2010 with a state-wide curbside recycling policy. The program is partially financed via a non-refundable deposit.

While the deposit system is nominally price neutral, there is a real effort cost to redeeming. Not all households do so given the hassle associated with storage and transportation. This generates a real behavioral effect on consumption decisions as purchases in Michigan are relatively more costly, all else being equal, than purchases in the bottle bill-free states of Indiana and Ohio. Michigan's added cost is distinct even among other bottle bill states. It is the only one with a ten cent standard deposit as compared to the five cent standard in other states. Even if redeeming were costless, the bottle bill would still generate incentives for cross-border shopping because of the possibility for fraudulent redemptions. Legally, consumers are only able to recoup money that they have paid into the system - hence the usage of the term bottle *deposit*. Given that product labels are not state-specific, RVMs are unable to make the Michigan/non-Michigan distinction. Therefore, individuals can redeem deposit-free, Indiana/Ohio bottles and illegally claim money. If caught, however, such acts are punishable by fines and jail time.

Evidence of the pervasiveness of this behavior became particularly glaring in 1992 when the redemption-to-deposit rate rose to 100.41%; there were 15 million more redemption claims than deposits. Michigan consistently has higher redemptions rates (90% range) than other bottle bill states (70% - 80% range) because of the higher deposit. However, this magnitude of excess cannot be solely attributed to a five cent difference.³ Consumer demand for bottled goods is therefore pushed south of the border by this policy.

Concurrently, a tax wedge acts in the opposite direction. Michigan levies a 6% sales tax while Indiana and Ohio have tax rates of 7% and 6.25% - 7.75%, respectively. In the case of Ohio, the base state-wide rate of 5.5% is augmented by varying local taxes. Because Michigan's sales tax is lower, the tax wedge makes purchases in Michigan more attractive. This tax differential is even larger for soda purchases. All three states exempt non-prepared foods but only Michigan includes soda in this category. Whereas the deposit wedge pushes purchases from Michigan to the Indiana and Ohio sides, this tax wedge pushes purchases into Michigan. Legally, tax savings from these cross-border purchases should be remitted to the home states. In practice, however, most households do not pay this difference and are, knowingly or not, committing use tax evasion.⁴

The analysis of this paper focuses on the capitalization of this consumer behavior in the prices of retailers located near the borders. This indirect approach is necessary as direct evidence is lacking. To my knowledge, this is the first academic paper to study the effects of bottle bills and tax differences in this context.⁵ However, it does tie into a rich and varied set of literature analyzing the economic consequences of border/policy discontinuities. For example, Lovenheim (2008) and Merriman (2010) examined cross-border shopping caused by differences in cigarettes taxes and prices. Using US Current Population Survey data, Lovenheim estimates that up to 25% of smokers cross-border shopped or purchased in border locations because of price differences. Merriman focuses more specifically on the highly tax stratified Chicago area. He finds that a one mile increase in distance to a low tax border increased the probability of finding a littered "home" cigarette pack by one percent. This change in consumer behavior can generate reactions in other economic actors as well. On the firm side, the seminal Engel and Rogers (1996) paper finds that price variation within countries is far lower than the price variation across countries, even in the case of similar and close neighbors. The US-Canadian border in particular generates an estimated price differential

³ A 2009 Michigan law, the RVM Antifraud Act, aimed at curbing fraudulent redemptions mandated that bottlers put an identifying Michigan mark on bottles sold within the state. In 2011, the law was in the early stages of implementation when the US Sixth Circuit Court of Appeals struck it down in December. In the case of American Beverage Association v. Snyder et al, they found that forcing an identifying mark on bottles sold in Michigan also created an extraterritorial effect on out-of-state bottlers.

⁴ Internet sales are also an area with significant use tax evasion. As a result, a number of states have been making a concerted push for increased support of the Streamlined Sales and Use Tax Agreement.

⁵ One report in the late 1990's, Analysis of Foreign Containers in the Michigan Deposit Stream, was produced by a consulting group for the Michigan Beer and Wine Wholesalers Association in their push to enact stricter anti-fraud legislation. This report used surveys of retailers to estimate that roughly \$16 million in redemptions came via fraud. A separate, state commissioned report in 2000 suggests that this number is conflated due to biased sampling in a small region of the state. It put the estimate at closer to \$10 million.

equivalent to 75,000 miles of within-country distance. In the context of federalism, papers such as Kanbur and Keen (1993) and Agrawal (2012) analyze the impact of state sales taxes on local sales taxes. There is some evidence that local tax rates smooth out state level differences.

3 Theoretical Framework

3.1 Setup

Figure 2: Hotelling Line

For simplicity, I will analyze the theoretical framework underlying the tax and deposit wedges using a four city, two state Hotelling pricing model. Assume that the two border cities, 2 and 3, are zero distance apart while the outer cities are all located one unit away from the border. Aside from their locations, the cities are identical. Households derive benefit V_i from the consumption of a homogenous, single-bottle good. This product is sold by identical, zero cost firms located in each city. The homogeneity of the good and costs implies that firms are competing under a Bertrand pricing scenario with spatial competition between cities. Thus, gross prices in this model are city specific: p_j for $j = \{1, 2, 3, 4\}$. This assumes collusion within but not across cities.⁶

Households make two decisions. First, they choose the city from which to purchase assuming that $p_j \leq V_i$ for at least some households. Households can cross-city and/or cross-border shop subject to a marginal (round trip) travel cost c_i^T . In addition to the gross price, this purchase decision is also affected by the tax rate at each location. In the absence of local taxes, assume that $\tau_1 = \tau_2 \equiv \tau_{IN}$ and $\tau_3 = \tau_4 \equiv \tau_{MI}$.⁷ Additionally, the presence of the bottle bill in Michigan implies that purchases made in Cities 3 and 4 also incur the deposit D . Second, households must choose whether or not to redeem the good. In order to recoup D , households incur a hassle cost c_i^R as well as any additional travel costs associated with bringing the bottle back to Michigan.

3.2 Border Gap

There are two main points of interest in the analysis: the price discontinuity at the border, and the within-state price trend moving away from the border. Regarding the first aspect, define the *border gap* as the Michigan border city price less the Indiana border city price, i.e. $p_3 - p_2$. While demand from the outer cities can effect these prices, the border gap is primarily determined by the behavior of households in Cities 2 and 3.

Table 2: City 2 Matrix

	Purchase in City 2	Purchase in City 3
Not Redeem	$V_i - (1 + \tau_{IN})p_2$	$V_i - (1 + \tau_{MI})p_3 - D$
Redeem	$V_i - (1 + \tau_{IN})p_2 + D - c_i^R$	$V_i - (1 + \tau_{MI})p_3 - c_i^R$

Table 2 shows the decision matrix for an Indiana household in City 2.⁸ Notice that the redemption and purchase location decisions are independent of each other. Households that choose to redeem, or not to

⁶ Prices and cost parameters in the theoretical model are denoted in cents.

⁷ Because bottled goods are only a small fraction of the overall set of taxed goods, I argue that states do not adjust the tax in response to this particular behavior. They are assumed to be exogenous in the model.

⁸ I have assumed that their purchase choice is limited to these two locations for simplicity.

redeem, will do so regardless of where they purchase the good. As such, the effective prices in Cities 2 and 3 must be equal, independent of the redemption decision. This is shown in Eq. 1. For beer products, the tax rates are 7% and 6% for Indiana and Michigan, respectively. For soda, the Michigan rate is now 0% because of the exemption. Conducting the same analysis from the perspective of City 3 generates identical border gap equations.⁹ As expected, a larger Indiana-Michigan tax differential (tax wedge) increases the competitiveness of Michigan. This allows it to set a higher price and increase the border gap. The bottle deposit (wedge) has the opposite effect as it hurts the competitiveness of the Michigan side.

$$p_3 - p_2 = \frac{(\tau_{IN} - \tau_{MI})p_2 - D}{1 + \tau_{MI}} \quad (1)$$

$$p_3 - p_2 = \begin{cases} \frac{0.01p_2 - 10}{1.06}, & \tau_{IN} = 0.07 \quad \tau_{MI} = 0.06 \text{ [Beer]} \\ 0.07p_2 - 10, & \tau_{IN} = 0.07 \quad \tau_{MI} = 0.00 \text{ [Soda]} \end{cases} \quad (2)$$

Figure 3: MI – IN/OH Border Gap Graph

Given the ten cent deposit, Fig. 3 plots the border gap as the City 2 price varies under the beer (1% tax differential) and soda (7% tax differential) cases. Intuitively, the potential savings from the lower sales tax in Michigan is small when the price is low whereas the cost of the added deposit is constant. The deposit wedge dominates the tax wedge at lower prices and the border gap is negative, i.e. Michigan must charge a lower price than Indiana to remain competitive. Conversely, a higher price generates a higher potential tax savings. The tax wedge is now dominant and the border gap is positive, i.e. Michigan can charge a higher price than Indiana. For beer, the border gap switches from negative to positive when p_2 is approximately 1000 or \$10 per bottle. All commonly found beer products are priced below this threshold (on a per unit basis). It is therefore likely that Indiana border prices will be higher than Michigan border prices in this category. For soda, however, the significantly larger tax gradient allows for a much lower switching threshold of 143 or \$1.43. While still higher than the per unit price of most soda products, this threshold is within the price range of 2 liter bottles of soda. Thus, it is most likely that a higher Michigan border price exists for this case.

⁹ The assumption that Cities 2 and 3 are located zero distance away from each other allows for symmetry in these results. Alternatively, assuming that the two are one unit apart adds an additional travel cost term to the numerator. The choice of City 2's firms or City 3's firms now creates a differential border gap equation. This differing assumption generates more significant implications for the behavior of households and firms in the outer cities.

3.3 Price-Distance Trend

The second point of interest focuses on the intrastate price patterns, i.e. how prices react moving away from the border. Define the *price-distance trends* as $p_1 - p_2$ in Indiana and $p_4 - p_3$ in Michigan. A positive (negative) price-distance trend therefore implies that prices increase (decrease) with distance. The impact of the tax and deposit wedges are strongest at the border and weakest in the interior. The price-distance trends reflect this tapering effect.

Figure 4: Price Effects at Various Travel Cost and Interior Price Levels

For a fixed V_i , the main determining factor of the price-distance trend is the degree of household mobility, i.e. the moving cost c_i^T . Consider a city located far from the border. Absent the effects of the two wedges, firms will simply set the monopoly or autarky price. However, this is only possible when moving costs are high such that households do not price compare. With lower moving costs, greater inter-city competition puts downward pressure on prices. Thus, the moving cost determines the prevailing, interior price levels in these cities. This is depicted by the dotted line in Fig. 4.

This is important for two reasons. Border city prices asymptote towards the baseline, interior price levels. The moving cost therefore determines the tail end of the price trend. Moreover, it also determines the beginning of the price trend, i.e. the border gaps. When c^T is low (left diagram of Fig. 4), the low interior price implies that the deposit wedge dominates the tax wedge. Michigan's border city has to set a price that is lower than the interior price level. The high degree of household mobility therefore pulls Indiana's border city down as well. In this scenario, there is a negative border gap with positive price-distance trends in both states. Conversely, when c^T is high (right diagram of Fig. 4), the high interior price implies that the tax wedge dominates the deposit wedge. Michigan's border city has a strong comparative advantage and can set a higher price. This also pulls Indiana's border city above the interior price. In this scenario, there is a positive border gap with negative price-distance trends in both states. The center diagram of Fig. 4 highlights the most likely scenario when price-distance trends are different in the two states. The interior price level is high enough such that the tax wedge dominates. Michigan's border city is able to charge a higher price (positive border gap) but not enough to pull Indiana's border price above the baseline. Thus, Michigan has a negative price-distance trend while Indiana has a positive price-distance trend.¹⁰ In the absence of information on the degree of household mobility, the price level acts as a rough proxy. Thus, the three cases in Fig. 4 also apply to differing per unit prices. Positive price-distance trends are more likely to be seen in cases with low per unit price (higher quantity products). Conversely, negative price-distance trends are more likely to be seen in cases with high per unit prices (low quantity products). A positive/negative (IN/MI) trend is also possible in this scenario.

¹⁰ In comparing the low versus high c^T cases, notice that both border prices are increasing. This intermediate case assumes that the increase in border prices is monotonic and relatively larger for the Michigan city.

4. Empirical Setup

4.1 Price and Retailer Data

Figure 5: Map of Retailers Sampled

As is common in the literature, direct economic data detailing illegal behavior is often unavailable or inaccurate. This paper therefore indirectly analyzes the effect of the deposit and tax wedges by investigating the retail prices of affected bottled goods near the border. Geographically, the data set spans 177 ZIP code zones, 142 cities, and 66 counties from the Michigan, Indiana, and Ohio border regions. In total, 346 retailers ranging from gas stations and pharmacies to liquor stores and supermarkets were included. These retailers were identified using a combination of Google Maps and the Yellow Pages website with an emphasis placed on types that were likely to have greater bottles sales.¹¹ Google Maps was also used to calculate distance from the border. As opposed to a straight line, as the crow flies measurement, calculations for the fastest and shortest driving routes to the border are a more effective measure of travel. Greater emphasis was also placed on areas closer to the border. Overall, distance from the border ranged between 0 and 150 miles, and 0 and 165 minutes.

For each retailer, I collected pricing data on four bottled goods between December 2012 and April 2013: 2 liter bottles of Coca Cola, 12 packs of Coca Cola in cans, 6 packs of Budweiser in bottles, and 24 packs of Bud Lite in cans. Summary statistics in Table 3 reflect the gross, pre-tax prices for the cheapest available versions at each location.¹² Additionally, I noted if each observation was identified as being a sale price. Notice that there is a fair degree of price variation both within and across the states. These differences can potentially be explained by supply side factors such as cost differences between supermarkets and non-supermarkets, and demand side factors such as different income levels. The central hypothesis of this paper, however, is that some of this variation is determined by the two wedges. The choice of these four specific goods was purposely made to address this possibility.¹³ The differences in quantity and beverage type generate significant variations in per unit prices and tax differentials. Given the implications of the model, this provides testable predictions in regards to empirical estimates of the border gaps and price-distance trends.

¹¹ In contrast, existing UPC/scanner data focus almost exclusively on supermarket prices. See Appendix A.

¹² Some supermarkets offer lower prices for enrolling in their membership clubs. Additionally, a number of retailers charged higher prices for refrigerated versions. In the presence of quantity deals, I used the average price.

¹³ The price of an 11 ounce bag of Doritos was originally collected for usage as a control good. It is dropped due to a lack of variation in prices. Nearly all retailers followed the suggested retail price printed on each bag. In cases where the price diverged, this coincided almost perfectly with sale or retail type. Other possible candidates included gum and bottled water but neither proved feasible. Indiana state laws prohibit the sale of gum in liquor stores. Bottled water, not subject to the deposit at that time, is problematic because it also a substitute and/or complementary good.

Table 3: Summary of Variables - Mean (Standard Deviation in Parentheses)

Variable	Indiana	Michigan	Ohio	All States
<i>Retailer Level</i>				
2 Liter Cola	1.66 (0.36)	1.78 (0.37)	1.73 (0.30)	1.72
Range	[1, 2.49]	[1, 2.49]	[1.18, 2.29]	[1, 2.49]
Avg. Unit Price	1.66	1.78	1.73	1.72
N / N _{sale}	99 / 29	108 / 30	86 / 20	293 / 79
12 Pack Cola	4.66 (0.91)	4.81 (0.94)	4.57 (0.98)	4.68
Range	[2.50, 5.99]	[3.33, 6.99]	[3, 6.98]	[2.50, 6.99]
Avg. Unit Price	0.39	0.40	0.38	0.39
N / N _{sale}	92 / 26	91 / 31	81 / 22	264 / 79
6 Pack Bud	6.12 (0.51)	6.46 (0.58)	6.23 (0.48)	6.30
Range	[5.47, 7.50]	[5.19, 8.99]	[5.19, 9.10]	[5.19, 9.10]
Avg. Unit Price	1.02	1.08	1.04	1.05
N / N _{sale}	49 / 2	97 / 0	77 / 0	223 / 2
24 Pack BudLite	16.91 (2.02)	17.53 (1.74)	18.07 (0.58)	17.52
Range	[14.99, 24.91]	[14.99, 22.99]	[14.99, 19.99]	[14.99, 24.91]
Avg. Unit Price	0.70	0.73	0.75	0.73
N / N _{sale}	70 / 25	105 / 24	91 / 12	266 / 61
Minutes	30.0 (39.5)	31.4 (28.0)	44.3 (53.9)	35.3 (42.1)
Distance	26.1 (40.0)	31.7 (32.4)	42.4 (57.0)	33.5 (44.7)
Sales Tax Rate	7 (0)	6 (0)	6.75 (0.20)	6.57 (0.45)
Supermarket	0.31	0.25	0.30	0.29
N _{retailers}	110	126	110	346
<i>Zip Code Level</i>				
Tot. Population	9,079 (10048)	13,856 (14654)	13,360 (13166)	12,186 (12987)
Med. HH Income	49,776 (11980)	49,904 (13031)	47,459 (13076)	49,046 (12740)
Med. Home Value	129,034 (27889)	141,375 (47078)	124,886 (39280)	131,979 (39759)
Pop. Density	383 (886)	652 (1262)	1,518 (3598)	857 (2310)
Retail Density	3.10 (3.27)	3.15 (4.42)	7.57 (9.82)	4.61 (6.83)

When looking at the unconditional means, 2 liter Colas have the highest average per unit price at \$1.72. Because sodas are exempt from sales taxation, it also has the highest tax differentials at 7% and 6.25% - 7.75% in Indiana and Ohio, respectively. Given the theoretical threshold of \$1.43, the model predicts that a positive border gap is most likely for this case. At this average price, purchases in Michigan generate a potential tax savings of \$0.11 - \$0.13 whereas purchases in Indiana or Ohio allow for avoidance and/or fraudulent redemption of only the ten cent deposit. Michigan border prices should therefore be higher. Moreover, it is likely that the price-distance trends should be negative on both sides, i.e. prices fall with distance from the border, given the high per unit price.

Conversely, a negative border gap is predicted for the other three goods. The 12 pack of Cola has the lowest per unit price at \$0.39 per unit followed by the 24 pack of Bud Lite at \$0.73 per unit. While the 24 pack of Bud Lite is a bit more expensive, it has the smaller tax differential. Regardless, the potential tax savings are significantly lower than the deposit in both cases. Michigan border prices are most likely to be lower for these products. Price-distance trends are also likely to be positive such that prices are increasing

with distance from the border. These three products will therefore be the focus of the analysis. Table 4 highlights the predicted signs on the price effects. Additionally, I use the average price levels from Table 3 to create an approximate estimate for the specific, price equalizing border gaps.

Table 4: Predicted Price Effects (Estimates from Avg. Price Levels)

	Border Gap		Border Gap	Price-Distance Trends
Good	IN-MI	OH-MI	General	General
2 Liter Cola*	\$0.016	\$0.017	+	-/- or +/-
12 Pack Cola	-\$0.874	-\$0.892	-	+/+
6 Pack Budweiser*	-\$0.539	-\$0.553	-	+/+
24 Pack Bud Lite*	-\$2.231	-\$2.264	-	+/+

From a data perspective, I also selected these major brands in commonly sold quantities to mitigate issues with missing observations. However, this still occurred to some extent as evidenced by the fact that there are fewer price observations than retailers. Unavailability was due to the fact that some retailers did not stock the larger quantities or did not sell alcohol.¹⁴ Table 3 also presents summary statistics for the other variables of interest used in the empirical identification. These were obtained from the American Community Survey (2011 - 5 Year) and the County Business Patterns (2010) data sets. All census variables are specified at the standardized ZIP code level (ZCTA5). Additional data at the city, town, and township levels used in the robustness analyses were taken from www.city-data.com.

4.2 Regression Specifications

The Indiana-(Western) Michigan border sub-sample includes 478 price observations spread across 22 Indiana counties and 9 Michigan counties. The Ohio-(Eastern) Michigan border sub-sample includes 568 price observations spread across 21 Ohio counties and 14 Michigan counties. This separation is necessary because of the differences in sales tax rates between Indiana and Ohio. Similarly, the model suggests that differences in rates and price levels between the four goods requires separation as well. To estimate these border and good specific price effects, I utilize two main specifications: the multiple and the pooled specifications.

The multiple regression specification is given by Eq. 3. This is the preferred specification as it estimates the price effects using separate, good-specific regressions. Given the two borders and four products, this generates eight main regressions. The dependent variable, p_{ijk} , is the pre-tax, pre-deposit retail price in dollars of bottled good j at retailer i in border sample k (IN-MI versus OH-MI). The two main independent variables of interest are the distance and state terms. As shown in Table 5, these two variables allow for the identification of the border gap and price-distance trends. The $\ln minutes$ term denotes the drive time from the border for a given retailer. It therefore captures the distance effect of moving further in-state, i.e. the price-distance trend. The $mich$ dummy captures the difference in state prices at the border, i.e. the border gap.¹⁵ The interaction between the two is necessary as it allows for the separate identification of price-distance trends on either side of the border.

$$\begin{aligned}
 p_{ijk} = & \beta_0 + \beta_1 spmkt_{ijk} + \beta_2 \ln minutes_{ijk} + \beta_3 mich_{ijk} + \beta_4 spmkt_{ijk} \\
 & * mich_{ijk} \\
 & + \beta_5 spmkt_{ijk} * \ln minutes_{ijk} + \beta_6 mich_{ijk} * \ln minutes_{ijk} \\
 & + \beta_7 spmkt_{ijk} * mich_{ijk} * \ln minutes_{ijk} + \beta_8 X_{ijk}
 \end{aligned} \tag{3}$$

¹⁴ The price of the Pepsi and Coors equivalent was used in the case of five and two retailers, respectively.

¹⁵ Retailers located 0 minutes from the border are recoded to be 0.1 for the $\ln minutes$ transformation.

p_{ijk} : pre-tax, pre-deposit price of good j
 $\ln minutes_{ijk}$: log driving time

$spmkt_{ijk}$: supermarket dummy
 X_{ijk} : vector of controls

Similarly, this specification separates the supermarket versus non-supermarket analysis via the $spmkt$ dummy.¹⁶ This is a point of focus as the geographically diverse but more centralized pricing decisions of supermarkets could make them less responsive. In fact, an analysis of the data shows that this separation is very necessary. The main supermarket chains, Kroger, Meijer, and Wal-Mart, have numerous stores both within and across states. Within-state price variation is nearly zero when controlling for the week of observation.¹⁷ Even interstate variation is limited and concentrated in a few specific cases. Conversely, non-supermarkets have more independent price setting abilities. Greater significance is expected in such cases. As such, separately identifying effects for supermarkets versus non-supermarkets is important. Failure to do so has the potential to dampen estimates for non-supermarkets.

The triple, distance-state-supermarket interaction combination, while necessary, does generate complexity in regards to the interpretation. With the baseline unit being a non-supermarket retailer in the non-Michigan state, Table 5 relates the coefficients of interest to each price effect. Border gaps are estimated as the dollar difference in price between the Michigan and non-Michigan border retailers. Because of the level-log specification, price-distance trends estimates are interpreted as the change in gross prices (in cents) for a 1% increase in driving time.

Table 5: Coefficients of Interest in Eq. 3

	Indiana or Ohio	Michigan
Non-Supermarkets		
Price-Distance Trend	β_2	$\beta_2 + \beta_6$
Border Gap	β_3	
Supermarkets		
Price-Distance Trend	$\beta_2 + \beta_5$	$\beta_2 + \beta_5 + \beta_6 + \beta_7$
Border Gap	$\beta_3 + \beta_4$	

To account for other supply and demand factors that could be affecting the price levels, X is a vector of logged demographic and economic variables. These additional controls, detailed in Table 3, include logged population, median household income, median home value, population density, and retail establishment density. Some towns near the Michigan border are smaller, have fewer stores, and exhibit greater monopoly power. All else being equal, this allows for a higher price independent of the two wedges. Not accounting for this fact would negatively bias estimates of the distance coefficient. Retail establishment density proxies for the level of economic activity and firm competition. I also include controls for the month in which the price was observed, and whether the price was marked on sale. Results are presented for regressions both with and without the *sale* dummy because of the uncertain nature of the designation. While some sales do reflect below trend prices, there are cases where discounts are common and expected.¹⁸ If being on sale is more of a spurious categorization, then including the *sale* term would underestimate the border gap, and vice versa for the opposite case. The true estimate is likely to be somewhere in between the two regressions.¹⁹ Lastly, I include a combined state and local sales tax variable in the Ohio-Michigan

¹⁶ To be classified as a supermarket as opposed to a grocery or convenience store, retailers need to satisfy two requirements. First, they need to sell an expanded selection of foods, e.g., bakery, deli, seafood, etc. Second, the retailer must have at least five chain locations.

¹⁷ See Fig. 8 in Section 5.

¹⁸ Consider the example of the home goods retailer Bed, Bath, and Beyond that regularly sends out 20% discount coupons. A majority of consumers therefore consistently pay 80% of the listed retail prices even though they are technically paying a discounted amount.

¹⁹ Table 14 in Appendix B shows the results of a probit regression examining the likelihood of being on sale.

regressions. This is not necessary in the Indiana-Michigan sample because neither state allows for local taxation. Including the sales tax rate is pointless as it is perfectly collinear with the *mich* dummy.

$$\begin{aligned}
 p_{ijk} = & \beta_0 + \beta_1\{int.set\} + \beta_2 pack_{ijk}\{int.set\} \\
 & + \beta_3 bud_{ijk}\{int.set\} \\
 & + \beta_4 budlite_{ijk}\{int.set\} + \beta_5 X_{ijk}
 \end{aligned} \tag{4}$$

int.set: set of spmkt-distance-state interactions ***bud*_{ijk}**: dummy for 6 packs of Budweiser
***pack*_{ijk}**: dummy for 12 packs of Cola ***budlite*_{ijk}**: dummy for 24 packs of Bud Lite

The second specification, Eq. 4, differs in that it combines the four goods into one pooled regression. It includes the same set of triple interactions from Eq. 3, $\{int.set\}$, but with an additional, fourth interaction level. Specifically, three new variables act as goods dummies for the 12 packs of Cola, 6 packs of Budweiser, and 24 packs of Bud Lite. The baseline in the pooled specification is therefore the 2 liter Cola from a non-supermarket in the non-Michigan state. Note that this pooled specification is only partially interacted. There is no interaction between the goods dummies and the control variables. This assumes that the effects of the demographic and economic controls are the same for all products. Under the multiple regression approach, controls are specific to each good. Alternatively, a fully interacted specification would add these terms. In this case, coefficients from the multiple and pooled regressions would be identical. This full specification is not reported in this paper because it generates only marginal differences in the significance. As such, only the partial specification is presented. Further discussion of the pros and cons are in Section 5.2.

5 Results

5.1 Multiple Regressions Approach

The full regression tables from the multiple approach, Eq. 3, are shown in Tables 6 and 7 for the Indiana-Michigan and Ohio-Michigan sub-samples, respectively. Recall that the baseline in these regressions is a non-supermarket retailer in the non-Michigan state. The coefficient of the *mich* dummy is therefore the border gap for non-supermarkets. Similarly, the coefficient of the *lnminutes* term is price-distance trend for non-supermarkets in the non-Michigan state. The other price effects, as calculated according to the linear combinations described in Table 5, are shown at the bottom of the two tables. Given the number of different estimates, estimated price effects from Tables 6 and 7 are collected and summarized in Tables 8 (non-supermarkets) and 9 (supermarkets). Also shown are estimated price effects from regressions excluding the *sale* term as a comparison.

Table 6: Gross Prices in Indiana-Michigan (Multiple Regressions)

VARIABLES	(1) 2LCola-InMi	(2) 12Cola-InMi	(3) 6Bud-InMi	(4) 24BudLite-InMi
spmkt	0.02194 (0.113)	0.80711* (0.416)	-0.35894 (0.212)	0.09092 (1.001)
lnminutes	0.06660*** (0.020)	0.02793 (0.077)	0.11660** (0.050)	0.08247 (0.101)
mich	0.37408*** (0.131)	0.67210 (0.690)	0.24698 (0.280)	1.56917 (1.045)
spmkt*mich	-0.45488* (0.264)	-2.39050*** (0.795)	-0.52733** (0.247)	-3.55587* (1.774)
spmkt*lnminutes	-0.06593* (0.036)	-0.32758** (0.147)	-0.14025* (0.072)	-0.28855 (0.209)
mich*lnminutes	-0.11853*** (0.037)	-0.08264 (0.293)	-0.11108 (0.086)	-0.45229** (0.200)
spmkt*mich*lnminutes	0.15646* (0.086)	0.58946* (0.302)	0.29489*** (0.091)	1.02279** (0.476)
Log Total Population	-0.02645 (0.025)	-0.11550 (0.081)	0.01678 (0.079)	-0.18698 (0.314)
Log Med. HH Income	0.11507 (0.137)	0.57455 (0.344)	0.36270 (0.216)	-0.34417 (0.535)
Log Med. Home Value	0.02283 (0.073)	-0.75206 (0.452)	-0.45631** (0.166)	-0.55721 (0.582)
Log Density	0.04581** (0.019)	0.14539* (0.074)	0.05711 (0.069)	0.21334 (0.334)
Log Retail Density	-0.03298*** (0.009)	-0.10360* (0.054)	-0.05437 (0.068)	-0.21088 (0.286)
Sale	-0.47936*** (0.071)	-0.99242*** (0.118)	0.14170 (0.372)	-1.50599** (0.685)
Constant	0.20006 (1.430)	8.02685 (5.225)	7.24679** (3.282)	28.29351*** (6.473)
<i>N</i>	144	126	93	115
<i>R</i> ²	0.594	0.578	0.471	0.429
MI NonSpmkt Trend	-0.052	-0.055	0.006	-0.370**
P-Value	0.110	0.854	0.927	0.031
Spmkt Border Gap	-0.081	-1.718***	-0.280	-1.987
P-Value	0.729	0.003	0.155	0.413
IN Spmkt Trend	0.001	-0.300**	-0.024	-0.206
P-Value	0.983	0.015	0.662	0.396
MI Spmkt Trend	0.039	0.207**	0.160**	0.364
P-Value	0.466	0.040	0.014	0.437
Robust standard errors are in parentheses and clustered by county. Baseline is an Indiana non-supermarket in February. Month controls are included but not shown.				
*** p<0.01, ** p<0.05, * p<0.1				

Table 7: Gross Prices in Ohio-Michigan (Multiple Regressions)

VARIABLES	(1)	(2)	(3)	(4)
	2LCola-OhMi	12Cola-OhMi	6Bud-OhMi	24BudLite-OhMi
spmkt	-0.25251*** (0.060)	-0.58590*** (0.146)	-0.22039 (0.172)	0.56787*** (0.200)
lnminutes	0.06645*** (0.021)	0.10991 (0.070)	-0.05377 (0.048)	0.22430*** (0.060)
mich	-0.91990 (0.930)	1.40505 (2.670)	0.76233 (0.642)	-1.22901* (0.712)
spmkt*mich	0.06619 (0.083)	-1.32857*** (0.248)	-1.14665*** (0.414)	0.11825 (0.498)
spmkt*lnminutes	-0.03587 (0.024)	-0.10146 (0.066)	-0.06335 (0.113)	-0.07216 (0.076)
mich*lnminutes	-0.07498** (0.030)	0.00633 (0.096)	-0.07171 (0.210)	0.43994* (0.223)
spmkt*mich*lnminutes	-0.00891 (0.036)	0.33484*** (0.121)	0.25771 (0.160)	-0.56584*** (0.196)
Log Total Population	-0.04613 (0.028)	-0.09070 (0.079)	0.13733 (0.094)	-0.05957 (0.126)
Log Med. HH Income	-0.00373 (0.140)	0.13583 (0.477)	0.53923 (0.460)	-0.72422* (0.401)
Log Med. Home Value	-0.01899 (0.122)	-0.02646 (0.451)	-0.49788 (0.361)	0.49257 (0.546)
Log Density	0.03393 (0.025)	0.06186 (0.064)	-0.10833 (0.093)	0.10581 (0.111)
Log Retail Density	-0.01727 (0.023)	0.00339 (0.057)	0.05943 (0.062)	-0.05124 (0.102)
Sales Tax Rate	-0.17542 (0.134)	0.16605 (0.397)	0.32425 (0.260)	0.40906 (0.326)
Sale	-0.30415*** (0.065)	-0.85626*** (0.094)		-0.76052** (0.327)
Constant	3.22066*** (1.111)	3.43614 (3.618)	3.32620* (1.930)	16.34004*** (5.010)
<i>N</i>	149	138	130	151
<i>R</i> ²	0.588	0.636	0.229	0.366
MI NonSpmkt Trend	-0.009	0.116	-0.125	0.664***
P-Value	0.783	0.204	0.562	0.003
Spmkt Border Gap	-0.854	-1.329**	-0.384	-1.111**
P-Value	0.375	0.000	0.443	0.027
OH Spmkt Trend	0.031	0.008**	-0.117	0.152**
P-Value	0.286	0.024	0.443	0.027
MI Spmkt Trend	-0.053	0.350***	0.069	0.026
P-Value	0.149	0.000	0.783	0.847
Robust standard errors are in parentheses and clustered by county. Baseline is an Ohio non-supermarket in February. Month controls are included but not shown.				
*** p<0.01, ** p<0.05, * p<0.1				

Table 8: Non-Supermarket Price Effects (Multiple Regressions)

<i>Sales</i>	Indiana-Michigan				Ohio-Michigan			
	Border Gap	IN Trend	W. MI Trend	<i>N/R</i> ²	Border Gap	OH Trend	E. MI Trend	<i>N/R</i> ²
2L Cola	0.374***	0.067***	-0.052	144/0.594	-0.920	0.066***	-0.009	149/0.588
12 Cola	0.672	0.028	-0.055	126/0.578	1.405	0.110	0.116	138/0.636
6 Bud	0.247	0.117**	0.006	93/0.471	0.762	-0.054	-0.125	130/0.229
24 BudLite	1.569	0.082	-0.371**	115/0.429	-1.229*	0.224***	0.664***	151/0.366
<i>No Sales</i>	Border Gap	IN Trend	W. MI Trend	<i>N/R</i> ²	Border Gap	OH Trend	E. MI Trend	<i>N/R</i> ²
2L Cola	0.447***	0.063**	-0.080***	144/0.385	-0.185	0.054**	-0.016	149/0.495
12 Cola	0.580	0.053	-0.009	126/0.420	1.038	0.153*	0.130	138/0.530
6 Bud	0.230	0.115**	0.006	93/0.470	0.762	-0.054	-0.125	130/0.229
24 BudLite	1.667*	0.096	-0.412**	115/0.362	-1.038	0.159***	0.560**	151/0.330

Table 9: Supermarket Price Effects (Multiple Regressions)

<i>Sales</i>	Indiana-Michigan				Ohio-Michigan			
	Border Gap	IN Trend	W. MI Trend	<i>N/R</i> ²	Border Gap	OH Trend	E. MI Trend	<i>N/R</i> ²
2L Cola	-0.081	0.001	0.039	144/0.594	-0.854	0.031	-0.053	149/0.588
12 Cola	-1.718***	-0.300**	0.207**	126/0.578	-1.329**	0.008	0.350***	138/0.636
6 Bud	-0.280	-0.024	0.160***	93/0.471	-0.384	-0.117	0.069	130/0.229
24 BudLite	-1.987	-0.206	0.364	115/0.429	-1.111**	0.152	0.026	151/0.366
<i>No Sales</i>	Border Gap	IN Trend	W. MI Trend	<i>N/R</i> ²	Border Gap	OH Trend	E. MI Trend	<i>N/R</i> ²
2L Cola	0.659***	0.036	-0.120**	144/0.385	-0.115	0.035	-0.073**	149/0.495
12 Cola	-1.677**	-0.506***	-0.006	126/0.420	-1.198***	-0.085	0.177***	138/0.530
6 Bud	-0.290	-0.024	0.162***	93/0.470	-0.384	-0.117	0.069	130/0.229
24 BudLite	-0.854	-0.252	-0.007	115/0.362	-0.746	0.083	-0.194	151/0.330

2 Liter Cola

Recall from the model and Table 5 that a positive border gap is most likely to be seen for the 2 liter Cola case as it has the highest per unit price. It also has the larger sales tax differential because sodas are exempt in Michigan. As such, the tax wedge should dominate the deposit wedge. This gives Michigan's border retailers an advantage. The high price level also highlights the likelihood of negative price-distance trends on either side of the border, i.e. prices should be falling with distance. This is consistent with the right diagram in Fig. 4. Alternatively, if the price level is in the intermediate level, the trends could be positive in Indiana and Ohio but negative in Michigan as shown in the middle diagram of Fig. 4.

The prediction of a positive border gap is borne out by the empirical estimates, particularly for the Indiana-Michigan border sample. Estimated border gaps are positive in a majority of the specifications with significance in all but one case. Michigan's non-supermarket prices are \$0.374 to \$0.447 higher than those on the Indiana border, with supermarket prices being \$0.659 higher, all else being equal. Given an unconditional mean price of \$1.72, this border gap constitutes a 22% to 38% price difference. However, there is no significant result in the Ohio-Michigan border sample.

In regards to the price-distance trends, Michigan prices are estimated to be decreasing with distance from the border. The negative coefficients are significant in three out of eight cases. This coincides with

the higher Michigan border price dissipating further in-state. On the Indiana and Ohio sides, the price-distance trends flip. Coefficients are positive and significant especially in the non-supermarket cases. Estimates of the price-distance effect therefore match up with the intermediate case shown in Fig. 4. Prices are high enough where the border gap is positive but not so high that the Michigan border price can pull the non-Michigan price above the interior level. Overall, this provides evidence in support of the theoretical model even though significance is not found in every case.

Figure 6: Price Effects for Non-Supermarket 2 Liter Cola (IN-MI Border, w/o *Sale*)

Fig. 6 plots the estimated price effects for the non-supermarket prices in the Indiana-Michigan border sample. This particular case of 2 liter Cola prices finds significant estimates of \$0.447 for the border gap and price-distance trends of 0.063 and -0.08 for the Indiana and Michigan sides, respectively. Because of the log-level specification, the price-distance coefficients implies the price change in cents for a 1% increase in minutes from the border. In a better context, the 0.067 estimate for Indiana implies that a retailer located 30 minute from the border would have a \$0.22 higher price. For Michigan, the -0.08 estimate implies that a retailer located 30 minutes from the border would have a \$0.27 lower price.

12 Pack Cola

On the other end of the spectrum, 12 packs of Cola have the lowest per unit price at \$0.39. This suggests that the border gap should be negative. Even with a higher soda tax differential, the deposit wedge should still dominate the tax wedge putting Michigan at a disadvantage. Given the low price and negative border gap, prices should increase in the interior. Thus, the model predicts positive price-distance trends for both sides of the border. The regressions confirm the presence of a statistically significant, negative border gap. At the Indiana border, Michigan prices are \$1.677 to \$1.718 lower. At the Ohio border, Michigan prices are \$1.198 to \$1.329 lower. Given the unconditional mean price of \$4.68, this is a 26% to 37% price difference across the two borders. However, this significance is only found on the supermarket side.

Similarly, significant price-distance trend estimates are also found predominantly in the supermarket samples. On the Michigan side, estimates range between 0.177 and 0.35. This is consistent with the model. For the Indiana/Ohio side, there is only significance in Indiana. Moreover, the price-distance trend is negative. Ranging between -0.3 and -0.506, this estimate suggests that supermarket prices are decreasing by upwards of half a cent for every 1% increase in drive time. The pattern of price-distance trends contradicts the predicted pattern of the theoretical model.

Fig. 7 plots the combined supermarket price effects for the Indiana-Michigan border sample. All three price effects are significant when including the *sale* variable. The border gap is estimated to be \$1.718

with price-distance trends of -0.3 and 0.207 for the Indiana and Michigan sides, respectively. This implies that prices 30 minutes from the border are \$1.02 lower in Indiana and \$0.70 higher in Michigan.

Figure 7: Supermarket 12 Pack Cola (IN-MI Border, w/*Sale*)

This pattern of price effects is not predicted in the model. Given a negative border gap, both price-distance trends should be positive. The most plausible reason why this result does not line up with the model has to do with the spatial composition of regional versus larger supermarkets in the Indiana sample. Recall that supermarkets and non-supermarkets are separately identified in the specifications because of inherent differences in how prices are set. Supermarket chains are more centralized and have low within-state price variation. For the big three retailers of Kroger, Meijer, and Wal-Mart in particular, weekly advertisements are nearly identical across regions. This hinders the ability of individual stores to tailor pricing decisions.²⁰

Figure 8: Comparing Regional Supermarket Prices in Indiana (12 Pack Cola)

²⁰ Using a *chain* retailer dummy instead of the *spkmt* dummy decreases the significance of most estimates.

The same is true for the smaller, regional supermarket chains (Martin's and Harding's Supermarkets). Tables 8 and 9 bear out this point; there are five fewer significant price-distance estimates for supermarkets than non-supermarkets. Moreover, supermarket prices may also be less responsive because they are more competitive, i.e. prices are closer to the lower bound. This is also supported by the regressions.²¹

Thus, this negative price-distance trend arises because of the differences *across* supermarket chains. Fig. 8 shows a (truncated) scatter plot of prices in Indiana. The big three supermarkets are present throughout the state whereas the regional supermarkets are concentrated in two areas. The square points represent the observations from the big three supermarkets. The pattern is fairly flat with little variation. On the other hand, the circular points show the price observations from the regional supermarkets. It is evident that the negative trend is driven by these retailers as they are charging far higher prices at the border. Re-running the specification when recoding these regional supermarkets as non-supermarkets does indeed produce more muted price-distance effects.²² Thus, the result in Fig. 7 does not necessarily contradict the theoretical predictions.

24 Pack Bud Lite

The good with the second lowest per unit price is the 24 pack of Bud Lite. At \$0.73 per bottle, this is more expensive than the 12 pack of Cola. However, it has the advantage of facing a smaller tax differential. This implies that the deposit wedge should dominate the tax wedge. As such, it is also likely to have a negative border gap and positive price-distance trends.

In the Ohio-Michigan border sample, all four border gap estimates are negative with two of them being significant. In the non-supermarket case, the border gap is -\$1.229. In the supermarket case, the border gap is -\$1.111. Given the unconditional mean price of \$17.52, this represents a 6% to 7% difference. Price-distance trends also coincide with the model's predictions. Estimates are positive in all but one case as well as being positive and significant in all four of the non-supermarket cases.

Fig. 9 plots the combined non-supermarket price effects for the Ohio-Michigan border sample. All three price effects are significant when including the *sale* variable. The border gap is estimated at -\$1.229

²¹ Coefficients for the supermarket dummy in Tables 7 are significant and negative. This is not true in Table 6 because the specification is controlling for sales (supermarkets are more likely to offer sale prices). When excluding the *sale* term, coefficients for the supermarket dummies are consistently and significantly negative.

²² This recoding also dampens the estimated border gaps. For supermarkets, estimates regarding 12 packs of cola are now no longer significant.

with price-distance trends of 0.224 and 0.664 for the Ohio and Michigan sides, respectively. A retailer located 30 minutes from the Ohio border would have a \$0.76 higher price. The analogous Michigan retailer would have a \$2.26 higher price. Fig. 8 therefore coincides exactly with the left, low price scenario from Fig. 4.

However, the opposite result is found in the Indiana-Michigan border sample. The estimated border gap for the non-supermarket case is actually positive and equal to \$1.667. The Michigan price-distance trend is also negative, i.e. prices are decreasing with distance from the border. While both these estimates run counter to the theoretical predictions, it is possible that this outcome stems from Indiana's alcohol laws. Specifically, Indiana prohibits the sale of any alcohol on Sundays. Retailers are forced to close or, at a minimum, cordon off the affected aisles on these days. It is therefore plausible that Michigan border retailers can set higher effective prices because of this policy. Anecdotally, there is at least one store on the Michigan side of the border that specifically advertises its availability on Sundays. Thus, it is plausible that this positive border gap arises because retailers choose to leverage and focus on Sunday-specific, cross-border beer purchases. The close proximity of the Notre Dame and its students to the border implies that there is likely to be a high proportion of beer consumers.

6 Pack Budweiser

As with the previous two products, the model predicts that 6 packs of Budweiser should have a negative border gap with positive price-distance trends. However, its relatively higher per unit price situates the product in an awkward middle. It has lower price than the 2 liter Cola but a higher price than the 24 pack of Bud Lite. Thus, there is a lower chance of identifying the corresponding price effects.

The regression results show that border gap coefficients are positive for non-supermarket and negative for supermarkets. While there is a discrepancy, none of the estimates are significant. In regards to the price-distance trends, positive and significant estimates are found in the Indiana-Michigan border samples. However, these results are inconsistent as a significant trend is found in Michigan's supermarkets but Indiana's non-supermarkets. The weak results for the 6 pack of Budweiser do not coincide with any theoretical predictions. However, this noisiness actually provides more evidence in favor of the model. Given the intermediate price level, the model suggests that price effects are least likely to be found for this product. Because price effects were much more prevalent in the other three cases, this lack of a concrete identification is partially expected and consistent.

5.2 Pooled Regression Approach

There are pros and cons to using the pooled approach over the multiple regression approach. One obvious benefit is that it reduces the number of regressions from eight to two. However, this comes at the cost of having to deal with a much larger set of variables. Calculating the estimated price effects now requires upwards of eight coefficients under the pooled approach. However, the added complexity does tend to produce weakly smaller standard errors. Pooled regressions are usually more efficient because they assume that the variances of the residuals are equal across all goods. This is not the case with the multiple regression approach.

Similarly, the pooled specification in Eq. 4 assumes that the effect of the control variables are equal across goods. In general, this is not necessarily true. Consider the multiple regression results for the Indiana-Michigan border shown in Table 6. The estimated impact of logged median home value on is positive and insignificant for the 2 liter Cola case. However, the impact is negative and significant for the 6 pack of Budweiser case. Therefore, estimates of the price effects under the pooled regression are likely to be biased to some degree as it forces these effects to be identical. While average R^2 values under the pooled and multiple regressions are 0.98 and 0.49, respectively, the multiple regression specification is the preferred approach for this particular reason. Less potential bias in the coefficients is preferable to marginal gains in significance.

Because there are over 40 total variables in the pooled specification, full regression tables are not shown. Analogous to the summary tables from the multiple regressions approach, summaries of the

estimated price effects are presented in Tables 10 (non-supermarkets) and Table 11 (supermarkets). Comparatively, the pooled estimates tend to have less significance in the soda cases but greater significance in the beer cases. However, this pattern is not always consistent. Overall, ten price effects are no longer significant when switching to the pooled specification. At the same time, seven price effects become significant. A similar lack of consistency is also true in regards to the magnitude of the price effects. This points again to the difference in assumptions between the two specifications, especially in regards to the impact of the demographic and economic controls.

Table 10: Non-Supermarket Price Effects (Pooled Regressions Summary)

<i>Sales</i>	Indiana-Michigan			Ohio-Michigan		
	Border Gap	IN Trend	W. MI Trend	Border Gap	OH Trend	E. MI Trend
2L Cola	0.214	0.056*	0.008	0.904	0.086***	0.038
12 Cola	0.355	0.005	0.041	0.947	0.106	0.126**
6 Bud	0.043	0.082*	0.043	0.402	0.031	0.063
24 BudLite	1.935**	0.127	-0.450***	-1.222**	0.142***	0.455***
<i>N</i> and <i>R</i> ²	478/0.979			568/0.990		
<i>No Sales</i>	Border Gap	IN Trend	W. MI Trend	Border Gap	OH Trend	E. MI Trend
2L Cola	0.327	0.061**	-0.032	1.539	0.068**	0.033
12 Cola	0.196	0.017	0.084	1.640	0.123	0.123*
6 Bud	0.133	0.095**	0.030	0.502	0.010	0.022
24 BudLite	2.087**	0.142	-0.476***	-1.153**	0.110**	0.410**
<i>N</i> and <i>R</i> ²	478/0.976			568/0.989		

Table 11: Supermarket Price Effects (Pooled Regressions Summary)

<i>Sales</i>	Indiana-Michigan			Ohio-Michigan		
	Border Gap	IN Trend	W. MI Trend	Border Gap	OH Trend	E. MI Trend
2L Cola	-0.949	-0.070	0.188	1.104	0.064	-0.028
12 Cola	-1.769**	-0.350***	0.152	-0.227	-0.029	0.262***
6 Bud	-0.398	0.150*	0.309***	-0.622*	0.003	0.165
24 BudLite	-1.301	-0.141	0.289	-1.636***	0.019	-0.004
<i>N</i> and <i>R</i> ²	478/0.979			568/0.990		
<i>No Sales</i>	Border Gap	IN Trend	W. MI Trend	Border Gap	OH Trend	E. MI Trend
2L Cola	0.426	0.031	-0.096	1.752	0.070	-0.074*
12 Cola	-1.812**	-0.575***	-0.045	0.559	-0.105	0.130*
6 Bud	-0.451	0.062	0.262***	-0.543	-0.035	0.128
24 BudLite	-0.396	-0.147	0.034	-1.352**	-0.012	-0.167
<i>N</i> and <i>R</i> ²	478/0.976			568/0.989		

Specific comparisons of the results from the three highlighted cases in Figs. 6, 7, and 9 show that price effects are usually less significant and smaller in magnitude. For the IN-MI, non-supermarket 2 liter

Cola case shown in Fig. 6, there is no longer significance in the border gap and Michigan trends. The same drop is also found in the IN-MI, supermarket 12 pack of Cola case shown in Fig. 7. However, the Indiana price-distance trend is actually more significant. Magnitude wise, the price effects are larger except in Michigan. Finally, the price effects from the OH-MI, non-supermarket 24 pack of Bud Lite case shown in Fig. 8 are all smaller in magnitude. Thus, gains in significance as a result of using the pooled specification are found almost exclusively in other cases.

5.3 Additional Robustness

Different Price Measures

The dependent variable in the previous regressions is the gross, retail price of the bottled goods. Alternatively, I can analyze the net-of-tax price (net price) or the net-of-tax, net-of-deposit price (effective price). For retailers in Michigan and Indiana, changing from gross to net prices is a simple shift. Applying a 6% or 7%, or 0% in the case of soda taxation in Michigan, to all retailers within a state simply moves the baseline trend up by the amount of the tax. For retailers in Ohio, however, the net price is potentially different because of its local tax rates. For the effective price, I add on the value of the bottle deposit to the net price but only for the Michigan observations. For the most part, re-running the regressions under these new prices has a predictable effect. Estimates of the border gaps simply adjust by the tax and/or deposit value. This naturally makes the border gap estimates more significant. In regards to the price-distance trends, there should only be an effect in the Ohio sample. Recall that this price effect measures within-state variation at various locations. Using the gross versus net versus effective price does not make a difference in the Indiana or Michigan samples because *state level* deposit and sales tax amounts are constant. Ohio is the only sample where the price-distance trend could potentially be different because it has *local level* differences in sales tax rates. However, the estimated trends are nearly identical in both magnitude and significance to those from the gross price regressions. This is likely a result of including the sales tax rate as a control in the original regressions. Aside from the expected rescaling of the border gap estimates, the choice of price therefore has little substantive impact on the results.

Different Distance Measures

In regards to the right hand side variables, using logged distance in miles as opposed to minutes has a relatively small impact on the magnitude of the price effects. This is due to the numerical similarity between the two variables. However, it does decrease the significance of some estimates. In cases where significance was already marginal, the switch to distance led to insignificance. This supports the notion that driving time as opposed to driving distance is a more effective measure of the effort cost associated with travel and cross-border shopping. Switching from the logged measure to the level measure, however, drastically decreases the significance of the price-distance estimates. This highlights the non-linearity of the travel cost. Specifically, the added psychological cost of a 20 minute drive versus a 10 minute drive is higher than that of a 70 minute drive versus a 60 minute drive. The log transformation is also more appropriate as it smooths out the wide range and variation of driving time in the data.

I also consider different definitions of the border. For the previous regressions, the border region is technically defined as those observations where logged minutes equals zero. However, it is possible that the border is a thick construct. Intuitively, a retailer located one minute away is effectively the same as one located three minutes away, i.e. both are considered border retailers by consumers. As such, I re-run the regressions from the multiple approach in Eq. 3 for two additional definitions of the border. Specifically, I consider the three and five minute thick borders. For each border specification, *lnminutes* is recoded to equal zero for all retailers located within the thick border. Retailers located further away are then re-scaled according to the new border definition. Summaries of the price effects, in comparison to the original results from Tables 6 and 7, are shown in Tables 12 and 13 for the Indiana-Michigan and Ohio-Michigan samples, respectively.

Table 12: Different Borders Definitions (Indiana-Michigan)

Price Effect	(1) 2LCoke	(2) 2LCoke-3	(3) 2LCoke-5	(4) 12Coke	(5) 12Coke-3	(6) 12Coke-5	(7) 6Bud	(8) 6Bud-3	(9) 6Bud-5	(10) 24BudLite	(11) 24BudLite-3	(12) 24BudLite-5
Non-Sp Border Gap	0.37408*** 0.008	0.2903* 0.063	0.22059* 0.063	0.67210 0.339	0.55537 0.296	0.55898* 0.087	0.24698 0.389	0.07195 0.870	-0.05570 0.887	1.56917 0.145	1.86625 0.236	
Non-Sp Trend (IN)	0.06660*** 0.003	0.03920 0.118	0.02911 0.188	0.02793 0.718	-0.01932 0.825	-0.02625 0.733	0.11660** 0.031	0.09116 0.410	0.03872 0.704	0.08247 0.422	-0.03788 0.868	-0.08275 0.708
Non-Sp Trend (MI)	-0.052 0.110	-0.046 0.294	-0.033 0.337	-0.055 0.854	-0.057 0.810	-0.069 0.671	0.006 0.927	0.055 0.498	0.057 0.423	-0.370** 0.031	-0.596** 0.028	-0.555** 0.023
Sp Border Gap	-0.081 0.729	-0.141 0.449	-0.097 0.544	-1.718*** 0.003	-1.467*** 0.000	-1.245*** 0.001	-0.280 0.155	-0.156 0.343	-0.140 0.359	-1.987 0.413	-1.732 0.374	-1.525 0.415
Sp Trend (IN)	0.001 0.983	-0.005 0.828	-0.008 0.740	-0.300** 0.015	-0.272*** 0.005	-0.233** 0.012	-0.024 0.662	0.005 0.905	-0.001 0.975	-0.206 0.396	-0.168 0.346	-0.176 0.322
Sp Trend (MI)	0.039 0.466	0.054 0.246	0.039 0.320	0.207** 0.040	0.184** 0.012	0.170** 0.011	0.160** 0.014	0.148*** 0.006	0.138*** 0.003	0.364 0.437	0.370 0.301	0.330 0.335
N	144	144	144	126	126	126	93	93	93	115	115	115
R ²	0.594	0.576	0.571	0.578	0.583	0.578	0.471	0.456	0.448	0.429	0.438	0.439
Non-Sp Border Gap	0.44657*** 0.004	0.41133** 0.024	0.32181** 0.026	0.57961 0.511	0.49862 0.456	0.52937 0.205	0.22969 0.422	0.04579 0.916	-0.08207 0.830	1.66691* 0.089	2.11543 0.140	
Non-Sp Trend (IN)	0.06345** 0.015	0.03833 0.242	0.02597 0.402	0.05275 0.492	0.02414 0.800	0.01800 0.826	0.11506** 0.037	0.08618 0.435	0.03213 0.747	0.09574 0.306	0.02768 0.884	-0.03550 0.853
Non-Sp Trend (MI)	-0.080*** 0.005	-0.092** 0.031	-0.076** 0.023	-0.009 0.980	-0.010 0.975	-0.028 0.898	0.006 0.917	0.056 0.483	0.058 0.409	-0.412** 0.016	-0.651** 0.023	-0.601** 0.025
Sp Border Gap	0.659*** 0.002	0.470*** 0.007	0.475*** 0.001	-1.677** 0.024	-1.435*** 0.009	-1.193** 0.020	-0.290 0.135	-0.165 0.301	-0.148 0.313	-0.854 0.640	-0.727 0.613	-2.343** 0.021
Sp Trend (IN)	0.036 0.290	0.028 0.275	0.024 0.298	-0.506*** 0.002	-0.441*** 0.001	-0.380*** 0.004	-0.024 0.655	0.005 0.913	-0.002 0.965	-0.252 0.390	-0.183 0.377	-0.194 0.367
Sp Trend (MI)	-0.120** 0.044	-0.074 0.171	-0.083* 0.073	-0.006 0.965	0.010 0.912	0.012 0.894	0.162*** 0.010	0.150*** 0.004	0.139*** 0.002	-0.007 0.983	0.055 0.816	0.030 0.892
N	144	144	144	126	126	126	93	93	93	115	115	115
R ²	0.385	0.362	0.360	0.420	0.424	0.411	0.470	0.454	0.446	0.362	0.368	0.369

Results in the upper (lower) portion are from regressions including (excluding) the sale term. All regressions include the standard set of controls shown in Table 6. P values are given below the estimates. The non-supermarket border gap under the 5 minute border is absent due to a collinearity issue with the data at that sample split.

Table 13: Different Border Definitions (Ohio-Michigan)

Price Effects	(1) 2LCoke	(2) 2LCoke-3	(3) 2LCoke-5	(4) 12Coke	(5) 12Coke-3	(6) 12Coke-5	(7) 6Bud	(8) 6Bud-3	(9) 6Bud-5	(10) 24BudLite	(11) 24BudLite-3	(12) 24BudLite-5
Non-Sp Border Gap	-0.91990 0.330	-0.91823 0.343	-0.93346 0.323	1.40505 0.603	1.40304 0.596	1.09766 0.673	0.76233 0.245	0.65622 0.195	0.68120 0.125	-1.22901* 0.094	-0.88921 0.119	-0.77732 0.132
Non-Sp Trend (OH)	0.06645*** 0.003	0.05424*** 0.006	0.05353*** 0.002	0.10991 0.129	0.08233 0.137	0.06379 0.225	-0.05377 0.272	-0.03322 0.419	-0.03269 0.395	0.22430*** 0.001	0.15691*** 0.008	0.14422** 0.015
Non-Sp Trend (MI)	-0.009 0.783	-0.008 0.703	-0.006 0.765	0.116 0.204	0.073 0.334	0.059 0.370	-0.125 0.562	-0.071 0.680	-0.088 0.573	0.664*** 0.003	0.514*** 0.004	0.480*** 0.004
Sp Border Gap	-0.854 0.375	-0.863 0.374	-0.860 0.363	-1.329*** 0.000	0.378 0.888	0.022 0.993	-0.384 0.443	-0.295 0.342	-0.264 0.408	-1.111** 0.027	-1.172*** 0.002	-1.237*** 0.001
Sp Trend (OH)	0.031 0.155	0.023 0.207	0.030* 0.099	0.008 0.895	-0.004 0.941	-0.037 0.362	-0.117 0.299	-0.094 0.331	-0.104 0.272	0.152 0.102	0.110 0.173	0.099 0.221
Sp Trend (MI)	-0.053 0.149	-0.046** 0.064	-0.042* 0.094	0.350*** 0.000	0.245*** 0.000	0.249*** 0.000	0.069 0.783	0.081 0.648	0.060 0.732	0.026 0.847	0.006 0.954	0.014 0.896
N	149	149	149	138	138	138	130	130	130	151	151	151
R ²	0.588	0.587	0.588	0.636	0.633	0.634	0.229	0.226	0.229	0.366	0.363	0.360
Non-Sp Border Gap	-0.18535 0.861	-0.18798 0.859	-0.19989 0.849	1.03824 0.748	0.97505 0.762	0.56394 0.862	0.76233 0.245	0.65622 0.195	0.68120 0.125	-1.03840 0.160	-0.74763 0.197	-0.64500 0.218
Non-Sp Trend (OH)	0.05360** 0.027	0.04529** 0.030	0.04302** 0.024	0.15267* 0.091	0.11235* 0.092	0.08230 0.176	-0.05377 0.272	-0.03322 0.419	-0.03269 0.395	0.15917*** 0.007	0.10534** 0.041	0.09185* 0.081
Non-Sp Trend (MI)	-0.016 0.648	-0.010 0.666	-0.010 0.672	0.130 0.182	0.085 0.277	0.057 0.394	-0.125 0.562	-0.071 0.680	-0.088 0.573	0.560** 0.015	0.437** 0.019	0.404** 0.020
Sp Border Gap	-0.115 0.915	-0.146 0.890	-0.147 0.888	-1.198*** 0.000	-0.018 0.996	-0.468 0.887	-0.384 0.443	-0.295 0.342	-0.264 0.408	-0.746 0.130	-0.936** 0.012	-1.016*** 0.008
Sp Trend (OH)	0.035 0.175	0.034 0.132	0.036* 0.093	-0.085 0.232	-0.076 0.174	-0.118** 0.014	-0.117 0.299	-0.094 0.331	-0.104 0.272	0.083 0.321	0.057 0.435	0.044 0.547
Sp Trend (MI)	-0.073** 0.012	-0.055*** 0.005	-0.054*** 0.008	0.177*** 0.005	0.136*** 0.001	0.130*** 0.001	0.069 0.783	0.081 0.648	0.060 0.732	-0.194 0.140	-0.165* 0.072	-0.155 0.106
N	149	149	149	138	138	138	130	130	130	151	151	151
R ²	0.495	0.494	0.494	0.530	0.527	0.529	0.229	0.226	0.229	0.330	0.329	0.326

Results in the upper (lower) portion are from regressions including (excluding) the sale term. All regressions include the standard set of controls shown in Table 7. P values are given below the estimates.

In general, switching to a thicker border has a depressive effect on the estimates. The magnitudes of the border gaps and price-distance trends both shift closer to zero such that prices are less reactive both across and within states. The coefficients are also less significant. Intuitively, this result makes sense as the price impact of the tax and deposit wedges are less pronounced when averaged out over a larger distance. However there are a number of cases, most notably and consistently for 24 packs of Bud Lite, where the opposite occurs. For example, the non-supermarket trend in the Indiana-Michigan sample becomes more pronounced when increasing the border designation. Likewise, two of the supermarket estimates switch from being insignificant to significant. This may be caused by the aforementioned regional supermarkets. Despite these changes, most results are still fairly consistent and persist across the different border definitions. Magnitudes for both trends and p values, while changing, are all relatively close.

Different Geographic Levels

Using city level controls instead of zip code level controls has a similar, weakly depressive effect. For the case of larger locales where a city includes numerous zip codes, using city level controls explain less of the variation because it aggregates at a larger geographic scale. Zip code level controls are therefore more precise in these cases. In smaller locales, there is little difference as the two levels are usually identical, i.e. a given city/town only has one zip code. In a few isolated, extreme cases, the zip code level is actually bigger than the city/town borders. Thus, switching to city level controls does not improve the identification of the price effects. Moreover, it explains less of the variation as R^2 values are lower. While it may be preferable to include both zip code and city level controls, the limited sample size is an issue.

In the previous regressions, standard errors are clustered at the county level. Clustering at the city or zip code levels is not appropriate given the fact some regions only have one price observation. Typically, clustering improves standard errors. However, it may actually be detrimental if there are fewer than 50 clusters. While the total number of counties was 66 in the data, the number in each border sub-sample was roughly half of the total. I therefore re-calculated the standard errors from the main regressions using wild bootstraps as suggested by Cameron et al. (2008). Significance did decrease in some of the estimates but overall results are still consistent.²³

6 Conclusion and Discussion

This paper looked broadly at the effects of two policy discontinuities present at the borders between Michigan and its two neighboring states of Indiana and Ohio. These two wedges have the potential to incentivize cross-border shopping, fraudulent redemptions, and use tax evasion by households. Such incentives and behavior, if they do exist, should be capitalized in the retail prices of affected goods. Importantly, these predicted patterns should vary across products and borders. High per unit price goods such as 2 liter Colas should exhibit a positive border gap while low per unit priced goods such as 24 packs of Bud Lite should exhibit a negative border gap. For the most part, empirical estimates of the price effects coincide with the theory. In cases where such effects did not line up, there were plausible justifications.

There are a number of potential issues and critiques that should be addressed. Most notably, there is limited significance. This is especially true for 6 packs of Budweiser where none of the border gap estimates are significant. Even with the other three products, price effects are not always identifiable in every retailer-border-sale combination. One obvious culprit is the sample size. This directly weakens the statistical power. It also has an indirect effect because it limits the number of controls that can be added. There are, however, two reasons why this issue may not be as significant of a concern. First, note that there is both an absolute and a practical upper bound in the number of retailer observations. More pricing data is certainly available from the larger cities where there are many retailers. For the more rural locations, this

²³ Adding county dummies is problematic. Note that the main specifications includes price observations from both sides of the border. Given the distance measure, the border gap is fully captured by the Michigan dummy. When both sets of dummies are included, the effect of the border gap cannot be uniquely identified because of collinearity between the Michigan counties and the Michigan dummy. The coefficient on the Michigan dummy will necessarily include the border gap *plus* some additional county-specific effect.

number is much smaller and may only be one or two in some cases. The data set already includes most if not all available retailers in such locations. Unfortunately, these are the exact areas where more observations would have the most benefit. Thus, a larger sample size will come mostly from an increase in large city prices. The potential benefit this brings is small especially given the spatial distribution of population around the borders. Second, the lack of significance in some cases may not be an error. It is possible that there is no actual price effect in some situations. For the supermarket cases, it is reasonable that price-distance trends are less likely to be found. Additionally, there are non-modeled behavioral factors that could be at play. For example, consider a household that is price shopping for beer. Buying multiple 24 packs of Bud Lite is far more reasonable than buying sets of four separate 6 packs of Budweiser. This may be another reason why 6 packs of Budweiser have little significance. The absence of such considerations in the model points to the possibility of other unaccounted for factors.

This relates to a more general issue regarding omitted variable bias that is of particular concern given the static nature of the setup. Because deposit values and tax rates are constant during the data period, the regression must focus on identifying the spatial pattern of prices caused solely by the two wedges. In such a scenario, omitted variable bias is a concern because all other factors that could impact the prices must be accounted for and excluded. Unfortunately, the small sample size prevents the inclusion of a lot of controls. However, while the main regressions only include the standard demographic and economic variables, the degree of bias is not necessarily high. Most demand aspects are likely to be correlated with income, e.g. demand for soda and beer, knowledge of policy differences and prices, and etc. Similarly, most supply aspects are likely to be correlated with retail density, type, and state. Moreover, omitted factors only bias the estimates if such factors are also correlated with distance from the border. For example, assume that variable X strongly predicts a genetic propensity to cross-border shop. If X is randomly distributed, then such effects will average out among different locales at each distance level. The price trends will still be correct in the aggregate. However, if people that are inherently more prone to cross-border shop also choose to live near the border because of this tendency, then not controlling for X will bias the estimates. The extent to which is an issue is unclear.

Finally, a comparison between the predicted border gap values from Table 4 and the estimated border gaps from Tables 8 and 9 suggests that retailers and/or consumers are over-compensating for the 2 liter Cola and 12 pack of Cola cases. In the 2 liter Cola case, the estimated difference is upwards of \$0.30 more or 23 times larger than predicted. For the 12 pack of Cola case, this difference is approximately \$0.90 more or two times larger. Conversely, they are under-compensating for the 24 pack of Bud Lite case at the Ohio border. The border gap is approximately \$1 less negative (54% difference) than predicted. These differences appear far too large to simply be measurement error or noise in the data. Rather, it is possible that a behavioral aspect such as salience is at play.

In particular, it appears that households and/or retailers are discounting the deposit while overemphasizing the tax difference. For the soda cases, the 0% sales tax in Michigan may make it more prominent. However, this argument does not seem to work for the 24 pack of Bud Lite. In this case, it is possible that the Michigan dummy is picking up other state-level effects. As previously mentioned, Indiana does not allow for alcohol purchases on Sundays. This was a potential reason for the counter-running, positive border gap estimate. While there is no such state level mandate in Ohio, individual localities do have the authority to establish their own regulations. As such, there are some towns that do restrict Sunday purchases prior to 10AM or noon. This may be another potential reason that Michigan has less of a disadvantage. Alternatively, a simpler explanation may be that households do not generally engage in fraudulent redemptions. If the majority of them do not even consider the possibility of illegally collecting ten cents on every bottle, then the deposit wedge becomes far weaker. This would cause Michigan's advantage (disadvantage) to increase (decrease). As such, the pattern of differences between the predicted and estimated border gaps suggests that fraudulent redemptions, at least in relation to cross-border shopping and use tax evasion, is far less prevalent in the general population. This certainly highlights the need for further research.

References

- Agrawal, D. "Games Within Borders: Are Geographically Differentiated Taxes Optimal?", *Intern* 4 (2012), pp. 574-597.
- Ashenmiller, B. "Externalities from Recycling Laws: Evidence from Crime Rates", *American Law and Economics Review* 12 (2010), pp. 245-261.
- Ashenmiller, B. "The Effect of Bottle Laws on Income: New Empirical Results", *The American Economic Review* 101 (2011).
- Cameron, C., Gelbach, J., Miller, D. "Bootstrap-Based Improvements for Inference with Clustered Errors", *Review of Economics and Statistics* 90 (2008), pp. 414-427.
- Dobbs, I. "Litter and Waste Management: Disposal Versus User Charges", *The Canadian Journal of Economics* 24 (1991), pp. 221-227.
- Eggert, W., Weichenrieder, A. "On the Economics of Bottle Deposits", *CESifo Working Paper Series No. 1243* (2004).
- Engel, C., Rogers, J. "How Wide is the Border?", *The American Economic Review* 86 (1996), pp. 1112-1125.
- Fullerton, D., Wolverton, A. "Two Generalizations of a Deposit-Refund System", *The American Economic Review* 90(2) (2000), pp. 238-242.
- Fullerton, D., Wu, W. "Policies for Green Design", *Journal of Environmental Economics and Management* 36 (1998), pp. 131-148.
- Kanbur, R., Keen, M. "Jeux Sans Frontieres: Tax Competition and Tax Coordination When Countries Differ in Size", *The American Economic Review* 83 (4) (1993), pp. 877-892.
- Lovenheim, M. "How Far to the Border?: The Extent and Impact of Cross-Border Casual Cigarette Smuggling", *National Tax Journal* 61 (2008), pp. 7-33.
- Merriman, D. "The Micro-Geography of Tax Avoidance: Evidence from Littered Cigarette Packs in Chiago", *American Economic Journal: Economic Policy* 2 (2010), pp. 61-84.
- Viscusi, K., Huber, J., Bell, J. "Discontinuous Behavioral Responses to Recycling Laws and Plastic Water Bottle Deposits", *American Law and Economics Review* 15 (2013), pp. 110-155.

Appendix

A. Retail Price Data

Sampling wise, I first identified those cities within 20 miles of the Michigan border. This created a list of approximately 40 cities between the three states. For each of these border cities, I had a target goal of 15 retail observations between the different types of retailers with an emphasis on supermarkets, grocery stores, and liquor stores as well as retailers located within five miles of the border. However, I was only able to do so in less than ten of the border cities due to the fact the large majority these border cities did not have enough population to support such retail coverage. For cities located between 20 and 50 miles from the border, I had a target goal of five retail observations. Cities further than 50 miles had a target of only one or two observations. The list of retailers was collected using spatial search functions with the search terms “grocery”, “beer”, and “pharmacy”. A smaller subset of specific stores was then selected to identify those retailers from whom prices would be collected. Sampling wise, greater emphasis was placed on areas closer to the borders and on retailers with relatively greater potential bottle sales.

During the actual collection of prices, there was a significant amount of attrition. Nearly half of the retailers contacted via phone either did not answer (no longer in business, busy, not present) or refused to participate (company/store policy, busy, fear of competition). In these cases, I would re-sample and try to contact another store. I also collected pricing data in-person during the course of three trips. These trips focused heavily on the border regions near Toledo, Ohio and South Bend, Indiana given the importance of these MSAs. These trips were also used to fill in attrition gaps in smaller towns and cities. However, attrition was also present during my in-person visits. A number of retailers were no longer in business. Some retailers had missing or unknown prices despite the fact that employees were present. Additionally, some still refused to participate fearing competition.

B. Sales

Table 14 shows results from probit regressions of the *sale* dummy on the state, distance, and demographic variables for the 2 liter Cola, 12 pack of Cola, and 24 pack of Bud Lite samples. The three columns report marginal effects on the likelihood of a price observation being on sale. The baseline case is that of a non-supermarket in Indiana. Overall, supermarkets are more likely to have observations marked as on sale. This supports the lower average prices in supermarkets as compared to non-supermarkets. Additionally, retailers at the Michigan border are less likely to have a sale price than Indiana and Ohio, except for 12 packs of Cola. Moving away from the border, the probability tends to increase although this is not significant in all cases. There is less consistency among the other variables. Results are omitted for 6 packs of Budweiser as only two observations were on sale.

As discussed previously, the nature of the sales designation is uncertain. Some sale prices are not actually below trend if they are always on sale. Including both true sales and these pseudo sales has the potential to bias the price effects. As long as this phenomenon is not correlated with distance from the border, then estimates should still be acceptable. However, the fact that there are significant coefficients on the driving time terms suggests that there is potential bias.

Even if the sale price is truly below trend, it is possible for the sale designation to be inappropriate. Consider the case of a sophisticated household that knows the timing of the sales and can plan in advance. Under this scenario, the below trend sale price is the actual effective price. As such, it is arguable that the specification without sales may be more appropriate.

Table 14: Marginal Effects of a Probit Regression Examining Sales and Distance

VARIABLES	(1) 2LCola	(2) 12Cola	(3) 24BudLite
spmkt	0.38369*** (0.090)	0.48068*** (0.032)	0.41498*** (0.038)
mich	-0.70121*** (0.092)	0.14693 (0.164)	-0.72386*** (0.044)
ohio	-0.25167*** (0.026)	0.07943*** (0.029)	-0.23290*** (0.050)
lnminutes	0.00244 (0.006)	0.06594* (0.036)	0.01499 (0.011)
mich*lnminutes	0.11622*** (0.021)	-0.04601* (0.026)	0.09776*** (0.006)
ohio*lnminutes	0.01000 (0.009)	-0.05059*** (0.007)	0.02089** (0.009)
Log Total Pop	-0.00710 (0.038)	0.01110 (0.013)	0.03412 (0.037)
Log Med. HH Income	0.17112 (0.194)	-0.11658* (0.063)	0.15494*** (0.032)
Log Med. Home Value	0.02369 (0.136)	0.14430* (0.081)	-0.09538** (0.037)
Log Density	-0.00661 (0.044)	-0.04427*** (0.010)	-0.02840 (0.036)
Log Retail Density	0.01570 (0.050)	0.02987 (0.029)	0.05154*** (0.013)
Sales Tax Rate	-0.66015*** (0.063)	-0.02542 (0.064)	-0.46935*** (0.088)
N	346	346	337
R-squared	0.312	0.290	0.442

Robust standard errors are in parentheses and clustered at the county level. This regression was fewer interactions as it only looked at the likelihood of having a sale observation. As such, it was pooled across all three states. There are no results for the 6 pack of Budweiser as nearly all observations were not on sale. Baseline is an Indiana non-supermarket in February. Month coefficients are included but not shown.

*** p<0.01, ** p<0.05, * p<0.1