# Getting into the Weeds of Tax Invariance

Benjamin Hansen,* University of Oregon, NBER, IZA
Kendall Houghton,* University of Oregon
Keaton Miller,* University of Oregon
Caroline Weber,* University of Kentucky

December 2020

**Abstract**

We provide the first general empirical test of tax invariance (TIV). When a 25 percent gross receipts tax was eliminated on cannabis manufacturers in Washington state and the retail cannabis excise tax was simultaneously increased from 25 to 37 percent—a shift intended to be revenue-neutral—TIV did not hold. Manufacturers kept two-thirds of their tax savings, instead of passing all their savings through to retail firms via lower prices. One-third of the retail tax increase was passed on to consumers via higher retail prices, leaving retail firms to bear the remaining third of the retail tax increase.

*JEL Codes*: H20, H30, H70.

*Keywords*: Tax Invariance, Natural Experiment, Excise Taxes, Cannabis, Tax Incidence.

*Hansen: bchansen@uoregon.edu; Houghton: khoughto@uoregon.edu; Miller: keatonm@uoregon.edu; Weber: caroline.weber@uky.edu. The authors would like to thank David Agrawal, Nathan Anderson, Youssef Benzarti, David Evans, Michael Grossman, Bill Hoyt, Donald Kenkel, Michael Kuhn, Nathan Seegert, and Joel Slemrod for helpful comments. We appreciate comments and feedback from participants at seminars at Case Western, Columbia, Cornell, University of Kentucky, Portland State, Rutgers, and conference participants at IHEA, WEAI, and the IIOC meetings, as well as industry participants and Cannabis Science and Policy Summit attendees. Many thanks to David Shi for excellent research assistance. This paper previously circulated as part of "The Taxation of Recreational Marijuana: Evidence from Washington State" and many of our thanks are for comments provided on that earlier draft.

# 1 Introduction

Tax invariance (TIV) is the principle that who remits taxes does not influence who ultimately bears the burden of the tax. This principle is a key component of tax policy analysis: TIV allows taxes to be collected at any point in the supply chain without concern as to the ultimate incidence. Recent empirical work suggests that TIV can fail under specific circumstances—when tax evasion opportunities vary along the supply chain (Brockmeyer and Hernandez, 2016; Kopczuk et al., 2016; Slemrod, 2008), when there are price rigidities (Lehmann et al., 2013; Saez et al., 2012; Muysken et al., 1999), or if tax salience is different for consumers and firms (Chetty et al., 2009; Finkelstein, 2009).

There has not been a general empirical test of TIV in a setting without these features. This is due in part to the scarcity of tax reforms that move the point of tax collection in industries with detailed data. We provide this test using the cannabis market in Washington state. The frequently-audited comprehensive reporting system in the Washington market—the source of our data—makes tax evasion difficult. Prices both increase and decrease often; price rigidities are unlikely. Regulatory requirements ensure that owners are highly-skilled and well-capitalized—thus taxes are likely salient. Furthermore, the market is a closed system—each gram of cannabis purchased in Washington was grown in Washington, and vice versa. Neighboring states did not have legal cannabis markets at the time, which means tax leakage and competition are not relevant.

The tax system was reformed in an ideal manner for examining TIV. Prior to July 1, 2015, a 25% gross receipts tax was assessed at each transfer of cannabis—i.e., between cultivators and manufacturers, manufacturers and retailers, and retailers and consumers.[1] Afterward, the only tax collected was a 37% excise tax at retail. Crucially, this change was unexpected by market participants. The reform was passed on June 27, 2015, and signed by the Governor on June 30. Contemporaneous media reports suggest that the industry was aware of and supported the reform but did not expect passage and market participants did not have confidence in their price forecasts at the time of the change (La Corte, 2015).

We begin with a brief theoretical motivation. We ask if TIV holds for a percent-based tax—the relevant tax in our setting—for perfect competition and for a monopolist retailer and a monopolist manufacturer. TIV holds in perfect competition and we derive a modified version of TIV for the monopolist setting. Under modified TIV, moving tax assessment from manufacturing to retail still causes the manufacturer to pass along their entire savings. However, retailers respond to their new demand curve by cutting prices and expanding output to maximize profit under the new tax system. Our setting lies between these extremes. Cannabis retailers have substantial market power (Hollenbeck and Uetake, 2019; Mace et al., 2020). Manufacturers are likely more competitive but retain some power (see Table 1). We use these results to set benchmarks against which we can compare outcomes in the data.

We then empirically examine how the prices that manufacturers charge when selling to retailers change post-reform. Our framework predicts that manufacturers' prices should decrease 28.7% from pre-reform levels. Given that per-gram tax revenue would fall in that scenario,[2] we also consider a second benchmark: the amount manufacturers needed to pass-

[1]Cultivators grow cannabis plants, manufacturers transform raw plant material into 'usable marijuana', and retailers sell products to end-users. We ignore cultivators as roughly 95% of cultivators were integrated with manufacturers at the time of the reform and no excise tax was owed on intra-firm transfers (Hansen et al., 2019). Manufacturer-retailer integration was banned.

[2]It would have remained roughly constant if we took the naive view that tax-exclusive prices would remain constant—i.e. retailers had passed along their entire tax increase to consumers.

through to leave retailers' per-gram profits and consumer-facing tax-inclusive prices constant post-reform (17.7%). We find that manufacturers reduce their prices by only 7.2%; we reject the null hypothesis of TIV based on either benchmark at the 0.1 percent level.

Finally, we examine retail behavior. Our framework predicts that retailers should either leave their tax-inclusive prices constant or decrease them. Instead, we find tax-inclusive retail prices *increased* by an average of 2.5%. Retailers pass through one-third of the tax increase to consumers. Another roughly one-third is borne by manufacturers, leaving retailers to bear about one-third of the increase. We find evidence that retailers adjusted prices to maintain constant tax-exclusive markups.

Our findings relate to recent work on tax incidence by Benzarti et al. (2020). They find that value-added tax increases are passed-through to consumers at twice the rate of tax decreases. We find this asymmetry persists in the presence of simultaneous tax increases and decreases that were intended to be roughly revenue- and welfare- neutral or even enhancing. Moreover, unlike the VAT context, where consumers may be less attentive to price changes, our setting features two sets of firms who are highly aware of the tax and engaged in long relationships with repeated transactions.

Our work also contributes to the growing literature exploring violations of traditional profit-maximization models of firm behavior. DellaVigna and Gentzkow (2019) document that national retail chains charge uniform prices across geographies even in the face of large cross-store demand differences. Butters et al. (2019) study retailers' responses to changes in liquor, cigarette, and soda excise tax rates and find that firms respond similarly to local and national cost shocks. In contrast to these papers, which focus exclusively on the prices charged by retailers to consumers, we additionally observe interactions between manufacturers and retailers in a fragmented market where uniform pricing is impossible and study how the full supply chain responds to a change in the point of tax collection instead of the rate alone.

## 2 Background

Our analysis focuses on the adult-use cannabis market in Washington state, which opened in July 2014 after cannabis was legalized by ballot initiative in November 2012. We have written elsewhere about the history of this market (Miller and Seo, 2018; Hansen et al., 2020,a)—here we focus on features of the market and the reform that are key to our analysis.

Washington's cannabis market consists of three types of firms: cultivators, who grow and harvest cannabis plants, manufacturers, who transform harvested plant material into marijuana products and sell them wholesale, and retailers, who sell products they obtain from manufacturers to consumers at up to three locations. Applicants for licenses of any type have to pass background checks and undergo a lengthy application and zoning process requiring substantial capital investment before entry. Cultivators face capacity constraints—the largest firms may cultivate 30,000 ft.$^2$ of plant canopy and may not merge to increase capacity. While retailers must be financially independent from other firms, cultivators and manufacturers may be vertically integrated. When the reform was implemented, approximately 95% (by weight) of usable marijuana—dried and cured cannabis flowers—was produced through a vertically-integrated process (Hansen et al., 2019). Thus, we focus our analyses on two types of firms, "manufacturers" and "retailers".[3]

[3] State law calls cultivators "producers" and manufacturers "processors"—we choose nomenclature to represent functional equivalents in other markets.

The market is a closed supply system: all cannabis sold by retailers is grown in the state, and every ounce grown legally within the state is sold at a Washington retailer. These rules are enforced through the state's "seed-to-sale" traceability system, which tracks each plant from cultivation through processing and retail. This system was implemented to respond to the informal federal regulations created by the "Cole Memo" (Cole, 2013). The system provides information that can be used to check for tax evasion: retailers cannot sell cannabis without manufacturing records, which forces manufacturers to report accurately.[4] Reporting is enforced through frequent audits—firms face an average of 7 visits per year—backed by significant penalties for non-compliance (Hansen et al., 2018). Non-Washington residents may purchase cannabis; significant demand came from inter-state shoppers during the period we study (Hansen et al., 2020b). No neighboring state had legal cannabis retailers.

Washington's initial tax regime consisted of a 25% tax assessed at every transfer of cannabis. Vertically-integrated manufacturers did not have to pay taxes on intra-firm transfers. The reform we analyze eliminated the 25% excise taxes within the supply chain and increased the retail excise tax rate to 37%. The excise tax applied to the state-and-local-sales tax-inclusive price pre-reform and the sales-tax-exclusive price post-reform. Accounting for changes to the base and rate of the retail tax, the reform changed the retail tax rate by 6.93%.[5] This change was designed to be revenue neutral under the assumption that tax-*exclusive* prices remained constant (whereas TIV predicts constant tax-*inclusive* prices). We account for both the change in the rate and the base of the retail excise tax in our analyses.

[4]Retailers can under-report their sales, but such behavior is detectable as retail sales can be compared to purchases from manufacturers. Under-reporting is uncommon. Our estimates are unaffected by dropping the few retailers that engage in significant under-reporting.

[5]The average total sales tax rate during this period was 8.9%, thus $\log\left(\frac{1.25(1+.089)}{1.37+.089}\right) = -0.0693$

Other regulations concerning cannabis production, distribution, and sales were unaffected. The posted prices faced by consumers include all taxes—these tax changes are thus fully salient to consumer decisions.

Our identification assumes that the policy change was unanticipated and thus the bill's history is relevant. The bill originated and was passed in the Washington House during the 2015 Regular Session, but stalled in the Senate. The bill was reintroduced in the First Special Session; again it stalled in the Senate. Finally, on June 27, the last day of the Second Special Session, the bill passed both chambers. The Governor signed it on June 30 and the law went into effect the next day. Contemporaneous media reporting portrayed the industry as unprepared. According to one retail store manager, "this is supposed to happen tomorrow. You have a few hours to change an entire market's pricing structure. It is an exceptionally short window for such a tremendous change" (La Corte, 2015).

# 3 A Framework for Tax Invariance

To motivate our empirical analyses, we introduce stylized models of manufacturers, retailers, and consumers. We assume a constant manufacturing marginal cost of $mc$ and a constant demand elasticity. For tax-inclusive retail price $p$, the inverse demand curve is $q(p) = kp^{\varepsilon}$ with $\varepsilon < -1$. We evaluate two extremes: (1) perfectly competitive manufacturers and perfectly competitive retailers and (2) a monopolist manufacturer and monopolist retailer. We expect our empirical setting lies between the two (see Section 4).

Let $p_m^i$ be the price charged by manufacturers to retailers including all taxes. Let $p_m^e$ be the price charged by manufacturers exclusive of taxes and let. Given a manufacturing tax rate of $\tau_m$, $p^e_m = p^i_m(1-\tau_m)$. Similarly, let $p^i_r$ be the tax-inclusive retail price and let $p^e_m$ be the tax-exclusive retail price. Given a retail tax rate of $\tau_r$, $p^i_r = p^e_r(1+\tau_r)$. While these definitions are not parallel, they match our empirical setting and the way in which gross receipts taxes are commonly implemented.

**Perfect Competition-Perfect Competition** In perfect competition, the tax-exclusive price earned by manufacturers is equal to their marginal cost, and so the tax-inclusive price is $p^i_m = \frac{mc}{1-\tau_m}$. Perfectly competitive retailers face this price as their marginal cost, and so the tax-inclusive retail price is $p^i_r = mc\frac{1+\tau_r}{1-\tau_m}$. The total tax revenue collected is $TR = k\left(\frac{mc(1+\tau_r)}{1-\tau_m}\right)^{\epsilon}\frac{mc}{1-\tau_m}(\tau_r+\tau_m)$. To see TIV holds, define $\tau = \frac{1+\tau_r}{1-\tau_m}$. Then $p^i_r = mc\cdot\tau$ and $TR = k(mc)^{\epsilon+1}\tau^{\epsilon}(\tau-1)$. Given some $\tau$, a policy maker can freely move one of $\tau_r$ or $\tau_m$, solve for the other, and hold $p^i_r$ and $TR$ constant.

**Monopolist-Monopolist** Under monopoly, the retailer's profit maximization problem is $\max_{p^i_m}\left(\frac{p^i_r}{1+\tau_r} - p^i_m\right)k(p^i_r)^{\epsilon}$ which implies $p^i_r = \frac{\epsilon}{1+\epsilon}p^i_m(1+\tau_r)$. Note that the retailer's tax-exclusive price is a constant markup over their marginal cost $p^i_m$. The quantity is $q = b(p^i_m)^{\epsilon}$ where $b \equiv k\left[\frac{\epsilon}{1+\epsilon}(1+\tau_r)\right]^{\epsilon}$. The wholesaler's problem is $\max_{p^i_m}[p^i_m(1-\tau_m)-mc]b(p^i_m)^{\epsilon}$ which implies $p^i_m = \frac{\epsilon}{1+\epsilon}\frac{mc}{1-\tau_m}$; the tax-inclusive price charged by the manufacturer is independent of the retail tax. Thus $p^r_i = \left(\frac{\epsilon}{\epsilon+1}\right)^2\frac{1+\tau_r}{1-\tau_m}mc$ and $TR = k\left(\frac{\epsilon}{\epsilon+1}\right)^{2\epsilon+2}\left(mc\frac{1+\tau_r}{1-\tau_m}\right)^{\epsilon+1}\left[\frac{\tau_r}{1+\tau_r}+\frac{\tau_m}{1+\tau_r}\frac{\epsilon+1}{\epsilon}\right]$. Mechanically, if $\tau$ is defined as above, the term in brackets cannot be simplified to be a function of $\tau$ alone. Given some $p_r$, if a policy maker shifts $\tau_r$ and $\tau_m$ to hold $p_r$ constant, $TR$ must change. Thus, TIV fails. Intuitively, the percentage taxes act as demand shifters, but the wholesaler does not internalize the retailer's response to retail percentage taxes because its effective demand elasticity is unchanged.

Given TIV generally does not hold in this monopolist-monopolist case, we want to understand the effect of a movement from a manufacturing to a retail tax on outcomes. First, suppose that the policy $\omega_1 = \{\tau_r = 0, \tau_m = \tau\}$ is replaced with $\omega_2 = \{\frac{\tau}{1-\tau}, 0\}$. From the equations above, it is clear that $p_r^i$ (and thus quantities) remains constant. The manufacturer passes through all of its tax savings, and earns identical profits. However, the retailer's profits decrease because the $\tau$ savings on the manufacturer's price is more than offset by the $\frac{\tau}{1-\tau}$ tax on their price. By the same logic, $TR$ increases as $\frac{\varepsilon+1}{\varepsilon} < 1$. Now consider the policy $\omega_3 = \{\tau', 0\}$ where $\tau' = \frac{p_m^i(\omega_1)\tau}{p_r^i(\omega_1) - p_m^i(\omega_1)\tau}$ is "naive-revenue-neutral": it would raise the same total revenue *if* the tax-inclusive retail price was the same after the reform. In this case, since $\tau' < \frac{\tau}{1-\tau}$, $p_r^i(\omega_3) < p_r^i(\omega_1)$ and $q(\omega_3) > q(\omega_1)$. Since $\varepsilon < -1$, profits for both firms and total tax revenues increase. Finally, suppose $\omega_1$ is replaced with $\omega_4 = \{\tau'', 0\}$ where $\tau''$ is chosen to be "true-revenue-neutral": $TR(\omega_1) = TR(\omega_4)$. Since $TR(\omega_3) > TR(\omega_1)$, $\tau'' < \tau'$ and thus $\omega_4$ increases profits for both the retailer and the manufacturer beyond $\omega_3$.

To summarize: The combination of market power and percent taxes leads traditional TIV to fail. However, revenue-neutral policies (whether "naive" or "true") that shift taxes from manufacture to retail lead to full pass-through from the manufacturer to the retailer and a decrease in tax-inclusive retail prices. Firms and consumers benefit from the change. We refer to this as modified TIV.

# 4 Data and Methods

Our data consist of administrative records from the "traceability" (or seed-to-sale) system maintained by the Washington State Liquor and Cannabis Board (WSLCB). The system tracks the entire legal marijuana supply chain, enabling regulators to collect taxes and prevent black-market diversion. Compliance is enforced through random audits, backed by civil and criminal penalties. The end result is data that tracks the planting, harvest, and production of cannabis plants into usable goods, the sale of those goods to a retailer, and final retail sale of marijuana products.

We use data on all plants, products, and sales. Harvested flowers and other plant material are converted into an "inventory lot" that is assigned a unique identifier (ID). Products or material within a single inventory lot are assumed to be homogeneous. Large inventory lots of finished product are split into smaller "retail" lots for sale to retailers. Each retail lot consists of multiple sealed packages of a specific weight of cannabis (e.g. 1 gram, 3.5 grams, etc) which are considered identical. When lots are sold to retailers, the system records the date, quantity, and price, and assigns a new lot ID. Thus, retail lot IDs uniquely identify the retailer, manufacturer, cultivator, product, and package size.[6] We observe each retail sale and link the price, quantity, and transaction time to the relevant inventory lots.

We restrict our analysis to the "usable marijuana" product category—74.5% of the total transactions observed in our data. Within this category, products are differentiated by "strain" (analogous to fruit cultivars), potency, and whether the marijuana is loose or pre-rolled into a joint. These characteristics are generally subsumed by our fixed effects. We

[6]A small number of lots have multiple package sizes, which we identify and correct for—thus our "inventory lot" fixed effects can be thought of as "inventory-lot-package-weight" fixed effects.

aggregate by inventory-lot-week. We exclude firms with less than two months of pre- and post-reform data. The reform also changed certain technical reporting requirements which affect the price data. We clean the price data for each firm to reflect the prices faced by consumers using an algorithm based on rounding behavior verified by spot checks of historical menus.[7] See Appendix B for full details.

Table 1 reports summary statistics for retail inventory lots for the six weeks pre-reform (the basis for our analyses in Section 5). The average tax-inclusive retail price was \$13.03 per gram and the tax-exclusive price was \$9.57 per gram. Retailer tax-exclusive prices are more than double manufacturer tax-inclusive prices. Both manufacturer and retail prices change week-over-week by more than one percent almost 40 percent of the time, split fairly evenly between price increases and price decreases, suggesting prices are not rigid.

The average market share of retailers in the 10-mile radius around their location was 31%, suggesting that there is substantial market power at retail, consistent with Hollenbeck and Uetake (2019) and Mace et al. (2020). The manufacturer market is effectively state-wide and the average market share is 1.4%. No manufacturer has more than 7% of the total market share. However, manufacturers may exert market power on individual retailers—on average, about seven manufacturers supply 75% of a retailer's inventory.

These data allow us to construct model-informed expected post-reform price changes. Given the details of the reform, in both the perfect competition and monopoly settings, manufacturers should pass through all of their tax savings to retailers, a $\log(1 - 0.25) = -28.7\%$ decrease in manufacturer tax-inclusive prices. This is true regardless of whether

[7]Retail firms in this market do not have access to traditional financial services and so choose to set tax-inclusive prices that are round numbers (generally "to-the-quarter" e.g. \$6.75 or \$10.25) to lower cash-handling costs. While this represents a potential friction, the effective minimum price change is smaller than the effects we estimate.

the reform is revenue neutral. This reform was not quite revenue neutral – rates were designed to be revenue-neutral on the assumption that tax-*exclusive* prices remain constant, but TIV predicts that tax-*inclusive* prices will remain constant. Given this, it may be possible to construct alternative models which both rationalize the price changes we observe in response to Washington's reform and which feature a TIV result. To rule out this concern, we construct a second benchmark for manufacturer price changes: given pre-reform prices, to maintain both a constant tax-inclusive retail price and constant per-gram retail profits (and therefore to satisfy TIV), manufacturers would have to decrease their prices by an average of 64 cents, or 17.7%.[8]

Under a revenue-neutral reform, TIV predicts that retailers would reduce their tax-*exclusive* prices by 6.93% (the amount of the change in the retail tax rate) and maintain constant tax-*inclusive* prices. Under modified TIV, we predict tax-inclusive prices will decline. As we calculate the reform is slightly revenue-decreasing,[9] our framework suggests retailers should reduce tax-inclusive prices further.

To summarize: if we estimate a decrease in average manufacturer tax-inclusive prices of less than 28.7% (in a statistically significant sense, e.g. a decrease of 15%), we will reject our model and reject TIV indirectly. If we estimate a decrease in average manufacturer tax-inclusive prices of less than 17.7%, we can reject TIV directly. If we estimate any *increase* in retailer tax-inclusive prices, we can reject TIV directly.

Figure 1 plots the panel of retail tax-exclusive prices normalized to the reform week. For each week, we take inventory lots in their first week of sale and match them with the

[8] 13.03/(1.37+0.089)-13.03/(1.25*(1+0.089)) = 64 cents.

[9] If prices had remained constant, the reform would have decreased the average total tax revenue per gram from $4.49 to $4.10.

price paid to the manufacturer, restricting observations to those for which the first retail sale and manufacturer sale both happened pre- or post-reform; thus, this illustrates the relation between retailer per-gram revenue and variable costs. The two series move roughly in lock-step with one another through the entire pre- and post-reform period (including the period around 4/20, an industry promotional event). This implies a constant markup of the retail tax-exclusive price over variable costs (the manufacturer price) that may be preserved in response to the tax reform. We return to this point in Section 5.2.

We estimate responses to the tax reform using an interrupted time-series model in first differences:

$$\Delta log(p_{it}) = \alpha_0 + \alpha_1 \Delta TaxReform_t + \alpha_2 FE_i + u_{it}, \tag{1}$$

where $p$ is the wholesale or retail price per gram, $TaxReform$ is an indicator variable that is one after July 1, 2015 and zero before, and FE are fixed effects.[10] Our analysis window spans six weeks before and after the reform—we examine the robustness of our estimates to this bandwidth. We cluster standard errors two-ways on manufacturer and retail location.[11]

For the manufacturer analysis, we aggregate inventory lot sales to the manufacturer-retailer-strain-week level, so that $i$ is a manufacturer-retailer-strain observation and $t$ indicates the week, and then take first differences.[12] Each manufacturer-retailer-strain tuple does not sell every week. We thus calculate varying length differences and include difference-length fixed effects.[13] The maximum difference-length allowed is 4 weeks. We are thus

[10]Without fixed effects, this regression is equivalent to an interrupted time series regression in levels with fixed effects at the level of our first differences and a control for distance in weeks from the reform.

[11]Firm clusters or two-way clusters on firm and week yield similar standard errors.

[12]Aggregation beyond the inventory lot is required because each lot is sold only once. The other possible aggregation is parent-lot-retailer-week, which produces similar estimates with lower power (though statistical significance remains).

[13]If a one week difference is not available, we use a two week difference, and so on. The difference length

estimating the magnitude of price changes in response to the reform within a specific firm-product pairing. When we add retailer-manufacturer-strain fixed effects, we allow each retailer-manufacturer-strain to have a separate trend over time.

For our main retail analysis, our data are aggregated by inventory-lot-week so that $i$ is an inventory lot.[14] Retail sales from an inventory lot are frequent, so we construct one-week differences. We are thus estimating the change in the retail price of an inventory lot in response to the tax reform and therefore holding all possible product and firm variation constant. Sales of retail inventory lots typically last multiple weeks, so we include fixed effects for the week since the first week a particular inventory lot sold. When we add inventory lot fixed effects, we allow prices in each inventory lot to have a separate time trend.

We separately examine the first week of retail sales for each inventory lot and include only those that were purchased from manufacturers in the same week. Similar to our manufacturer analysis, we aggregate by retailer-manufacturer-strain and take varied differences. We include difference-length fixed effects. In these regressions, we are asking how prices for *new inventory lots purchased post-reform* change relative to *pre-reform lots of the same strain from the same manufacturer*. This allows us to examine whether prices change more or less if the inventory was purchased post-reform relative to inventory that had already been purchased and was selling pre-reform.

fixed effects are not significant and our estimates are similar if we only use one-week differences, but we gain power by using this varied-length difference strategy.

[14] We are able to work at this level because retailers repeatedly sell out of a single inventory lot.

# 5 Results

## 5.1 Manufacturer Price Response

Table 2 reports estimates of Equation (1) for manufacturers. The estimate in Column (1), which includes no fixed effects, implies that prices changed by -6.5% in response to the tax reform (statistically significant at the 0.1% level). When we include manufacturer-retailer-strain fixed effects in Column (2)—our baseline specification—the point estimate becomes -7.2% (significant at the 0.1% level). This is roughly one-third of the 17.7% price decrease needed to preserve retailer per-gram profits (and therefore to minimially satisfy TIV), and one-quarter of the 28.7% decrease predicted by our framework. We can reject the null hypothesis that our estimate is consistent with TIV at the 0.1% level. Column (3) repeats Column (2) for the price in levels instead of logs—we find that the reform decreased manufacturer prices by 23 cents, about one-third of the 64 cent bound. Across specifications, the observed price adjustment was greater than 1% for more than 75% of our observations—another indication that firms were aware of this reform and prices are not rigid. Even if we rescaled our estimate assuming that any observation without a price adjustment was because of one of these frictions, the data would still reject the null hypothesis of TIV.

The bottom panel of Table 2 repeats the specification of each column for a placebo reform dated one year later. The estimates are near zero across all four columns, providing evidence that our regression specifications are valid. The top panel of Figure A.1 considers bandwidths from 2 to 8 weeks and confirms that our estimates are not sensitive to the bandwidth chosen. We provide an event study of these results in Section 5.3.

## 5.2 Retail Price Response

Table 3 reports estimates of Equation (1) for retailers. The estimate in Column (1), which include no fixed effects, implies that the reform decreased the tax-exclusive price by 4.5% (significant at the 0.1% level). We include inventory lot fixed effects in Column (2)—our baseline specification. The estimates are very similar; the coefficient in Column (2) implies that the reform reduced tax-exclusive retail prices by 4.4% (significant at the 0.1% level). Combined with the rate change, this implies that tax-*inclusive* prices increased by 2.5%; retailers passed through roughly one-third of the tax to consumers. We find that we can reject the null hypothesis of TIV-consistent pricing behavior at the 0.1 percent level.

As firms might have taken time to adjust (and the Independence Day holiday may have generated temporary price adjustments), Column (3) repeats Column (2) for two week differences and drops the first week after the reform, so that the effect of the reform is identified from the difference between the week before and the week after the reform. The estimates are approximately the same, indicating that neither of these concerns play a large role. We will return to a broader discussion of timing in Section 5.3.

Table 3 Column (4) repeats Column (2) with the dependent variable in levels—we estimate that average retail tax-exclusive prices fell by 41 cents per gram. This implies that retailers are an average of 41 cents per gram worse off on existing inventory as a result of the reform. On fresh inventory, firms were roughly 18 cents per gram worse off (41 less the estimated 23 cent decrease in manufacturer prices estimated in Table 2). In other words, under TIV this reform should have caused manufacturer and retail tax-exclusive prices to fall by 64 cents leaving profit and consumers unaffected. Instead, it caused smaller manufacturer price cuts leaving both retailers and consumers worse off.

Table 3 Columns (5) and (6) take an alternative approach to identification examining inventory lots only in their first week and only if retailers purchased the inventory lot from the manufacturer in that week. For this, we create a panel of retail-processor-strain-weight group-weeks. The estimates are quite similar—a 4.9% decrease in Column (5) versus a 4.4% decrease in Column (2), suggesting that retailers' price responses are largely unaffected by whether they are still selling inventory lots purchased pre-reform or selling new inventory lots purchased post-reform. Column (6) adds the first-differenced log manufacturer price. When included, the coefficient on the wholesale price is not statistically different from one and the coefficient on $\Delta TaxReform$ is now approximately zero. This suggests that retailers largely maintained a constant tax-exclusive markup. This is consistent with the pricing rule derived in Section 3. In other words, while retail behavior *as a whole* is inconsistent with TIV, after conditioning on the pass-through from manufacturers, retailers behaved, on average, in a way consistent with marginal-cost pricing (and therefore potentially consistent with TIV).

The bottom panel of Table 3 repeats each column in the top panel for a one-year-later placebo reform. If our estimates are valid, the coefficient on $\Delta Placebo$ should be roughly zero—exactly what we find. Even in Column (1) where the estimate is marginally significant, the coefficient is very close to zero. The bottom panel of Figure A.1 considers bandwidths from 2 to 8 weeks and confirms that our estimates are not sensitive to the bandwidth chosen.

## 5.3 Event Studies

The analyses above indicates that prices changed at the time of the reform—yet it is possible that these changes were part of the ongoing evolution of the market and that the placebo test results one year later reflect a different market structure (Doraszelski et al., 2018). Moreover, the estimates above do not indicate whether there is additional adjustment towards TIV beyond the first week. To address these issues, we conduct event studies for both the manufacturer and retailer responses using our baseline specifications (i.e. including inventory-lot fixed effects) from Tables 2 and 3. For manufacturers, we do not drop the $t-1$ tax reform coefficient due to our varied difference lengths.[15] Figure 2 plots the relevant coefficients and confidence intervals. In both event studies there is no clear trend in prices pre-reform. The entire response happens in period $t$, the reform date. Given the varied difference lengths for manufacturers, this implies that manufacturers adjust their prices the first time they sell a particular retail-strain pair post-reform. The fact that the entire response occurs immediately implies that the changes were unlikely to be part of the ongoing market evolution and are instead a true response to the reform. Furthermore, this suggests that it is very unlikely that through learning (Doraszelski et al., 2018; Huang et al., 2018), or some other mechanism, TIV will be achieved in the long-run.

# 6 Conclusion

TIV is a bedrock principle of tax design—it allows policymakers to focus their efforts on minimizing administrative costs and evasion without worrying about the welfare effects of

[15] E.g., for a two week difference that spans t-1 to t+1, both the t and t+1 coefficients are relevant.

alternative tax collection strategies. Indeed, TIV is routinely taught in "Principles of Economics" courses (e.g. McConnell et al., 2018; Mankiw, 2020). The literature has documented specific cases in which TIV fails, yet it is unclear whether TIV simply does not hold, or just that it cannot be applied in particular circumstances. We answer this question and show that it generally fails—instead of leaving the welfare of manufacturers, retailers and consumers unchanged or improved as predicted, the reform we study increased profits of manufacturers at the expense of retailers and consumers. The fact that a reform intended to be welfare-neutral or even welfare-enhancing had negative welfare consequences for both retailers and consumers (at least in the short run) has wide-ranging implications for tax policy.[16]

First, *de novo* tax designers may face welfare tradeoffs when choosing where in a supply chain to locate a tax. Both efficiency and equity considerations arise. When considering efficiency, variation in elasticities or competitive structures across the market may affect optimal tax placement. In terms of equity, if a policy goal is to ensure all market participants bear portions of the tax, it may be necessary to impose taxes on these different groups directly rather than relying on pass-through.

Second, policymakers considering changes to existing tax policy face greater consequences for doing so. While it may be possible to implement revenue-neutral reforms, restructuring will create clear winners and losers. In this case, manufacturers benefited—despite being in an arguably more competitive market—while retailers and consumers were harmed.

Taken together, these concerns point to broader political economy issues surrounding tax policy (Winer and Hettich, 2006; von Hagen, 2008). Political systems may be designed

[16] In other work we find that, over time, the reform increased trade between cultivators and manufacturers which could have positive effects on all market participants in the long run that offset the short-run negative effects on retailers and consumers (Hansen et al., 2019).

to limit the ability of policymakers to enact tax reforms and thus rational actors may unknowingly design systems which have additional inefficiencies as described here. Indeed, in Washington state, the legislature may not reform measures passed by ballot initiative for two years after passage. Though local government officials knew from the moment of passage that the gross receipts tax was likely to have negative consequences on the market, their hands were tied. Flexibility in political and policy systems may help avoid these concerns—though at the cost of volatility.

Finally, these results demonstrate a need for further experimental and modelling work. Modern models of competition, growth, trade, inflation, and the business cycle generally make assumptions about taxes which are appealing from a tractability standpoint. This assumptions generally imply TIV (see, e.g. Judd, 2002; Melitz, 2003; Galí, 2015). Instead of failures of TIV being the exception, our work provides evidence that TIV simply may not hold in practice. In the absence of TIV, it may be necessary to conduct experiments which examine the way in which firms and consumers respond to tax policy and construct models which more accurately capture these responses.

# References


Benzarti, Y., D. Carloni, J. Harju, and T. Kosonen (2020). What goes up may not come down: Asymmetric incidence of value-added taxes. *Journal of Political Economy*, forthcoming.

Brockmeyer, A. and M. Hernandez (2016). *Taxation, information, and withholding: evidence from Costa Rica*. The World Bank.

Butters, A., D. Sacks, and B. Seo (2019). How do national firms respond to local shocks? evidence from excise taxes. *Kelley School of Business Research Paper*.

Chetty, R., A. Looney, and K. Kroft (2009, September). Salience and taxation: Theory and evidence. *American Economic Review 99*(4), 1145–77.

Cole, J. M. (2013). Memorandum for all United States attorneys. `https://www.justice.gov/iso/opa/resources/3052013829132756857467.pdf`. Accessed: 2017-07-20.

DellaVigna, S. and M. Gentzkow (2019). Uniform pricing in us retail chains. *The Quarterly Journal of Economics 134*(4), 2011–2084.

Doraszelski, U., G. Lewis, and A. Pakes (2018). Just starting out: Learning and equilibrium in a new market. *American Economic Review 108*(3), 565–615.

Finkelstein, A. (2009, 08). E-ztax: Tax Salience and Tax Rates*. *The Quarterly Journal of Economics 124*(3), 969–1010.

Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.

Hansen, B., K. Miller, B. Seo, and C. Weber (2020). Taxing the potency of sin goods: Evidence from recreational cannabis and liquor markets. *National Tax Journal*.

Hansen, B., K. Miller, and C. Weber (2018). Auditing and enforcement in the recreational marijuana industry. In *111th Annual Conference on Taxation*. NTA.

Hansen, B., K. Miller, and C. Weber (2019). The welfare effects of a gross receipts tax. Working paper.

Hansen, B., K. Miller, and C. Weber (2020a). Early evidence on recreational marijuana legalization and traffic fatalities. *Economic inquiry 58*(2), 547–568.

Hansen, B., K. Miller, and C. Weber (2020b). Federalism, partial prohibition, and cross-border sales: Evidence from recreational marijuana. *Journal of Public Economics 187*, 104–159.

Hollenbeck, B. and K. Uetake (2019). Taxation and market power in the legal marijuana industry. *Available at SSRN 3237729*.

Huang, Y., P. B. Ellickson, and M. J. Lovett (2018). Learning to set prices in the washington state liquor market. *Manuscript. University of Rochester. Simon Business School*.

Judd, K. L. (2002). Capital-income taxation with imperfect competition. *American Economic Review 92*(2), 417–421.

Kopczuk, W., J. Marion, E. Muehlegger, and J. Slemrod (2016, May). Does tax-collection invariance hold? evasion and the pass-through of state diesel taxes. *American Economic Journal: Economic Policy 8*(2), 251–86.

La Corte, R. (2015, June 30). Gov. inslee signs recreational marijuana reform bill. *Associated Press*.

Lehmann, E., F. Marical, and L. Rioux (2013). Labor income responds differently to income-tax and payroll-tax reforms. *Journal of Public Economics 99*, 66 – 84.

Mace, C., E. Patel, and N. Seegert (2020). Marijuana taxation and imperfect competition. *National Tax Journal 73*(2), 545–592.

Mankiw, N. G. (2020). *Principles of economics*. Cengage Learning.

McConnell, C. R., S. L. Brue, and S. Flynn (2018). *Economics: Principles, problems, and policies*. McGraw-Hill.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *econometrica 71*(6), 1695–1725.

Miller, K. and B. Seo (2018). Tax revenues when substances substitute: Marijuana, alcohol, and tobacco. *Kelley School of Business Research Paper* (18-27).

Muysken, J., T. V. Veen, and E. D. Regt (1999). Does a shift in the tax burden create employment? *Applied Economics 31*(10), 1195–1205.

Saez, E., M. Matsaganis, and P. Tsakloglou (2012). Earnings determination and taxes: Evidence from a cohort-based payroll tax reform in greece. *The Quarterly Journal of Economics 127*(1), 493–533.

Slemrod, J. (2008). Does it matter who writes the check to the government? the economics of tax remittance. *National Tax Journal 61*(2), 251–275.

von Hagen, J. (2008). Political economy of fiscal institutions. In D. A. Wittman and B. R. Weingast (Eds.), *The Oxford Handbook of Political Economy*.

Winer, S. L. and W. Hettich (2006). Structure and coherence in the political economy of public finance. *The Oxford handbook of political economy 7*, 441.

# Tables

### Table 1: Pre-Reform Retail Summary Statistics

| Variable | Obs. | Mean | Std. Dev. |
|---|---|---|---|
| *Prices and Taxes* | | | |
| Tax-Inclusive Retail Price (\$/g) | 63,668 | 13.033 | 3.798 |
| Tax-Exclusive Retail Price (\$/g) | 63,668 | 9.570 | 2.783 |
| Probability of $>$ 1% Retail Price Increase | 63,668 | 0.17 | 0.375 |
| Probability of $>$ 1% Retail Price Decrease | 63,668 | 0.204 | 0.403 |
| Manufacturer Price (\$/g) | 63,668 | 4.103 | 1.309 |
| Probability of $>$ 1% Manufacturer Price Increase† | 7,954 | 0.177 | 0.382 |
| Probability of $>$ 1% Manufacturer Price Decrease† | 7,954 | 0.196 | 0.397 |
| Retail State + Local Sales Tax Rate | 63,668 | 1.089 | 0.006 |
| Tax Revenue Pre-Reform (\$/g) | 63,668 | 4.489 | 1.246 |
| *Competition* | | | |
| Market Share of Retailer in 10 Mile Radius | 63,668 | 0.313 | 0.282 |
| Market-level Manufacturer Market Share | 63,668 | 0.014 | 0.016 |
| Retail-Level Manufacturer Concentration Index | 63,668 | 6.997 | 2.691 |
| *Benchmarks Assuming TIV* | | | |
| Expected Tax Revenue Post-Reform (\$/g) | 63,668 | 4.104 | 1.200 |
| Manufacturer Pass-Through Cents | 63,668 | −0.640 | 0.185 |
| Manufacturer Pass-Through Percent Change | 63,668 | −0.177 | 0.058 |

An observation is an inventory-lot-week pre-reform. The data come from our retail analysis set and cover the six weeks prior to the tax reform. Tax revenue is calculated using both excise and state and local sales taxes. The retail-level manufacturer concentration index is calculated as follows: for a given retailer, we sort their suppliers by the weight of inventory sold, and count the number needed to comprise at least 75% of total sales. The "benchmarks assuming TIV" account for changes in the base and rate of the retail excise tax. The "manufacturer pass-through" statistics assume constant tax-inclusive retail prices and indicate the post-reform changes to manufacturer prices that would have left retailer variable-profit-per-gram constant.

† These probabilities are calculated for the subset of retail-processor-strain-weight group-weeks when the inventory lot changes (and thus a new purchase from a manufacturer has occurred).

**Table 2: Manufacturer Price Response**

| | (1) $\Delta$log(Price) | (2) $\Delta$log(Price) | (3) $\Delta$Price |
|---|---|---|---|
| | | Tax Reform | |
| $\Delta$Tax Reform | $-0.065^{***}$ | $-0.072^{***}$ | $-0.228^{***}$ |
| | (0.015) | (0.018) | (0.068) |
| Observations | 12,087 | 12,087 | 12,087 |
| Manufacturer Firms | 199 | 199 | 199 |
| P-Value for Test of TIV-Predicted Pass-Through | 0.000 | 0.000 | 0.000 |
| | | Placebo | |
| $\Delta$Placebo | 0.001 | 0.000 | 0.014 |
| | (0.012) | (0.014) | (0.040) |
| Observations | 21,288 | 21,288 | 21,288 |
| Manufacturer Firms | 180 | 180 | 180 |
| Bandwidth | 6 weeks | 6 weeks | 6 weeks |
| MRS FE? | No | Yes | Yes |

This table reports estimates of Equation (1) – other variables in that equation are included, but not reported. An observation is a manufacturer-retailer-strain-week. The outcome is the change in the log of the price per gram charged by the manufacturer to the retailer (except for in column (3) which is the same outcome, but not logged). MRS stands for manufacturer-retailer-strain fixed effects. The estimates are weighted by the total grams sold across the two weeks of the difference. The P-value tests the null hypothesis that the estimated pass-through is equal to that predicted by TIV. For the placebo regressions, we repeat the analysis one year later. These regressions are estimated with `reghdfe` in Stata. Standard errors twoway-clustered by manufacturer and retailer are in parentheses. $^{*}$5% significance level. $^{**}$1% significance level. $^{***}$0.1% significance level.

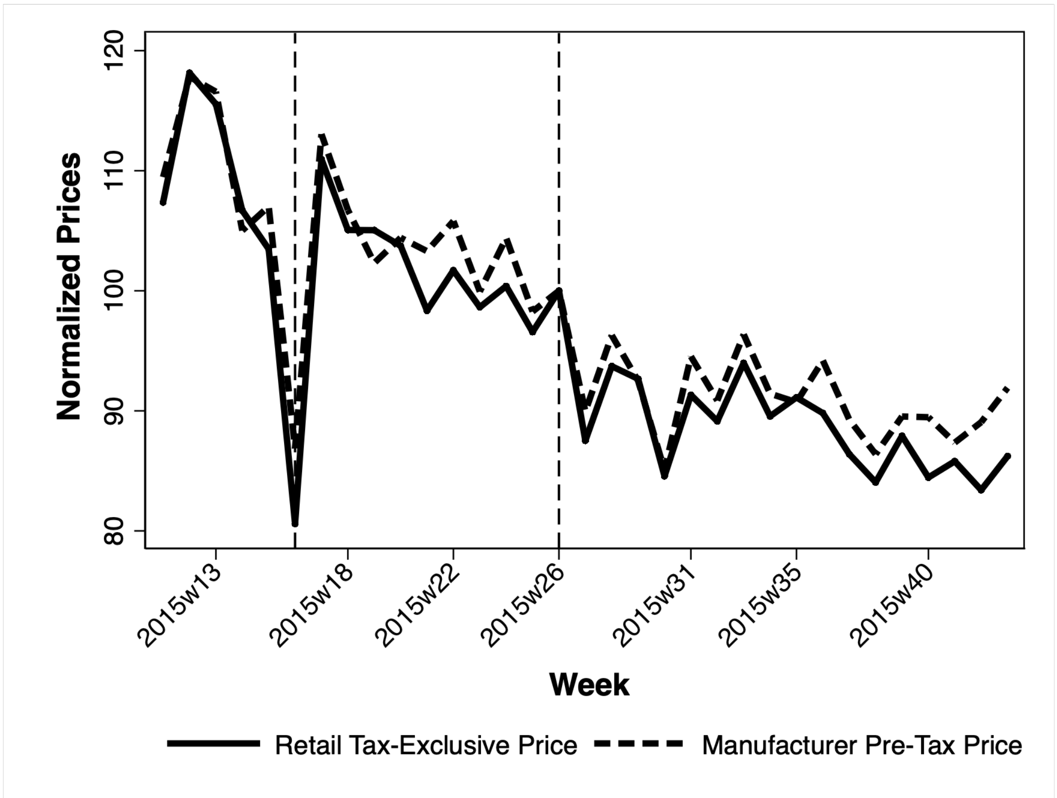
## Table 3: Retail Tax-Exclusive Price Response

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | $\Delta$log(Price) | $\Delta$log(Price) | $\Delta$log(Price) | $\Delta$Price | $\Delta$log(Price) | $\Delta$log(Price) |
| | | | Tax Reform | | | |
| $\Delta$Tax Reform | $-0.045^{***}$ | $-0.044^{***}$ | $-0.046^{***}$ | $-0.413^{***}$ | $-0.049^{**}$ | 0.011 |
| | (0.006) | (0.007) | (0.006) | (0.065) | (0.018) | (0.017) |
| $\Delta$log(Manufacturer Price) | | | | | | $0.887^{***}$ |
| | | | | | | (0.084) |
| Observations | 145,357 | 145,357 | 145,357 | 145,357 | 11,265 | 11,265 |
| Retail Firms | 110 | 110 | 110 | 110 | 110 | 110 |
| Implied Tax-Inclusive Price Change | 0.024 | 0.025 | 0.023 | 0.230 | 0.020 | 0.080 |
| P-Value for Test of Constant Tax-Inclusive Price | 0.000 | 0.000 | 0.000 | 0.000 | 0.270 | 0.000 |
| | | | Placebo | | | |
| $\Delta$Placebo | $-0.006^{*}$ | $-0.004$ | 0.001 | $-0.029$ | $-0.016$ | $-0.004$ |
| | (0.003) | (0.003) | (0.002) | (0.017) | (0.012) | (0.009) |
| $\Delta$log(Manufacturer Price) | | | | | | $0.642^{***}$ |
| | | | | | | (0.053) |
| Observations | 253,123 | 253,123 | 253,123 | 253,123 | 11,534 | 11,534 |
| Retail Firms | 106 | 106 | 106 | 106 | 105 | 105 |
| Bandwidth | 6 weeks | 6 weeks | 6 weeks | 6 weeks | 6 weeks | 6 weeks |
| MRS FE? | No | No | No | No | Yes | Yes |
| Inventory Lot FE? | No | Yes | Yes | Yes | No | No |
| Difference Length | 1 week | 1 week | 2 weeks | 1 week | 1-4 weeks | 1-4 weeks |
| Restricted to First Week Sales? | No | No | No | No | Yes | Yes |

This table reports estimates of Equation (1) – other variables in that equation are included but not reported. An observation is an inventory-lot-week. The outcome is the log of the tax-exclusive price per gram charged by the retailer to consumers (except for in column (4) which is the same outcome, but not logged). MRS stands for manufacturer-retailer-strain fixed effects. The estimates are weighted by the total grams sold in the first week of the difference. The P-value tests the null hypothesis that the tax-inclusive price remained constant as predicted by TIV. For the placebo regressions, we repeat the analysis one year later. These regressions are estimated with reghdfe in Stata. In the last two columns we only include observations in their first week of being sold at retail and only if the cannabis was also purchased from the processor in that same week. Standard errors twoway-clustered by manufacturer and retailer are in parentheses. $^{*}$5% significance level. $^{**}$1% significance level. $^{***}$0.1% significance level.

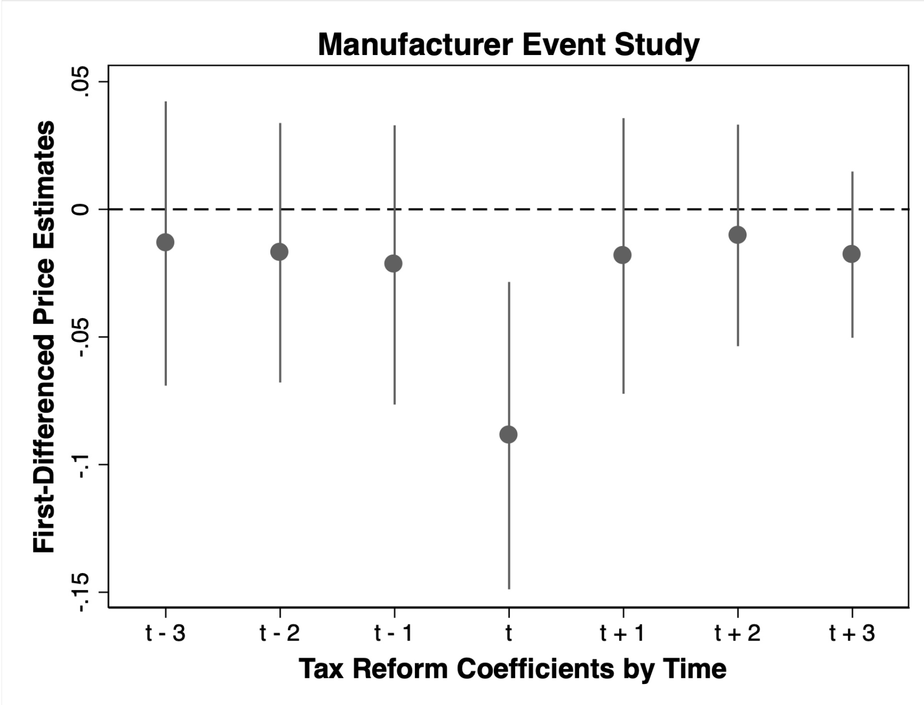
# Figures

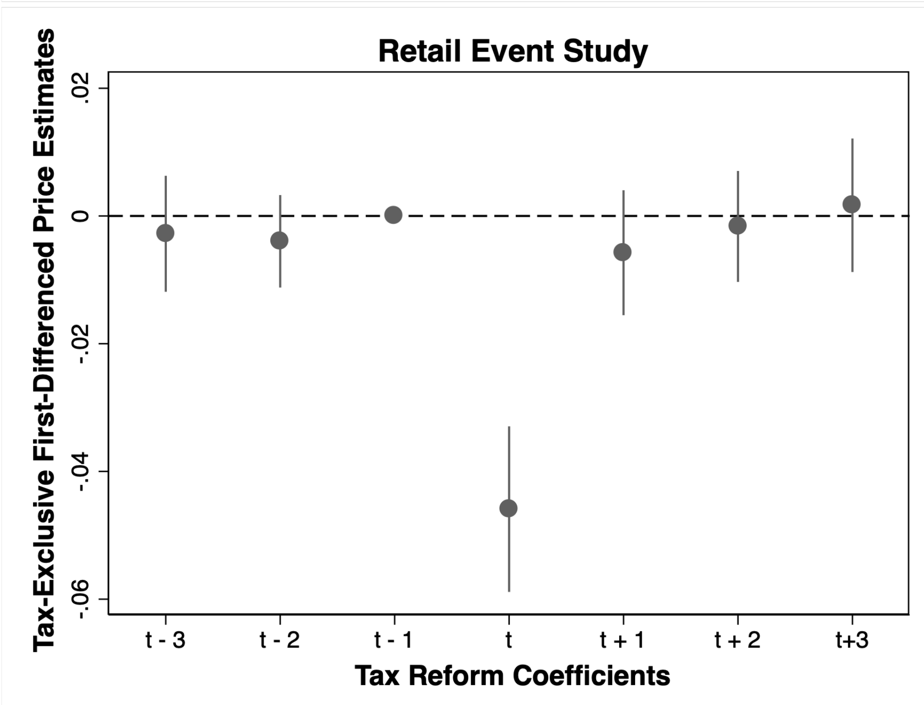
**Figure 1: Retail and Manufacturer Prices**




This figure plots average prices in Washington's cannabis industry for four months before and after the tax reform, normalized to 100 in the week before the reform. For each week, we take inventory lots in their first week of sale and match them with the price paid to the manufacturer, restricting observations to those for which the first retail sale and manufacturer sale both happened pre- or post-reform (before any applicable taxes are paid from the manufacturer to the government). This therefore illustrates the relation between retailer per-gram revenue and variable costs. The left dashed line in the figure marks 4/20 (an industry promotional period) and the right dashed line marks the week before the tax reform.

**Figure 2: Manufacturer and Retail Price Event Study**



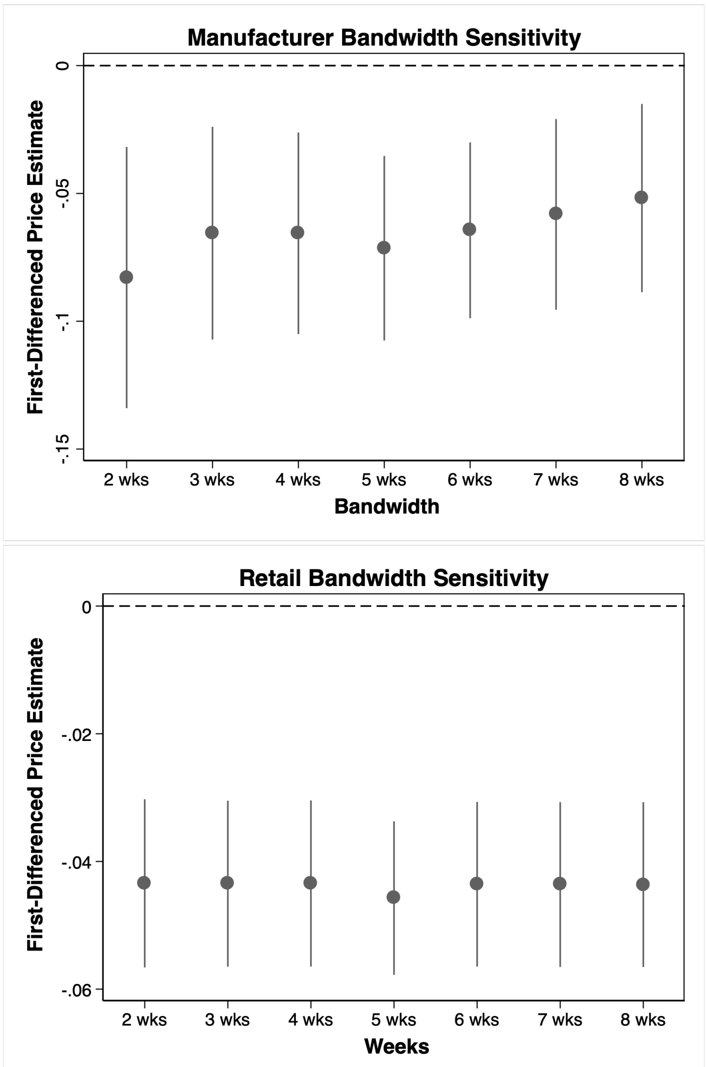




This figure plots estimates of Table 2 Column (2) (top panel) and Table 3 Column (2) (bottom panel) with additional leads and lags of $\Delta TaxReform$. The plotted coefficients are leads and lags of $\Delta TaxReform$. We include in the regression (but do not plot) leads and lags are for periods $t-4$ and before and $t+4$ and after. The dots indicate the point estimates and the lines indicate 95% confidence intervals. See the notes for Tables 2 and 3 for regression details.

# Appendices

## A Appendix Figure

**Figure A.1: Manufacturer and Retail Price Bandwidth Choices**




This figure plots estimates of Table 2 Column (2) in the top panel and Table 3 Column (2) in the bottom panel, varying the bandwidth. The bandwidth in our baseline specifications is 6 weeks. The estimates plotted are for the coefficient on $TaxReform$. The dots indicate the point estimates and the lines indicate 95% confidence intervals. See the notes for Tables 2 and 3 for regression details.

# B Data Cleaning Appendix

In this appendix we detail our data cleaning procedure. We begin by detailing our methods for cleaning prices in the face of changing reporting requirements and tax rates. We then discuss other cleaning steps to transform the raw data into the set used in our analyses.

The retail sales prices reported by firms in the "seed-to-sale" traceability system were supposed to be all-tax-inclusive pre-reform and tax-exclusive post-reform. However, compliance varied from firm to firm and changed over time. For example, some firms reported prices with the sales tax included and some reported prices without the sales tax.

This reporting confusion means that we must infer, for each firm, how they reported their prices and therefore the true tax-inclusive and tax-exclusive prices they charged. For each firm-week, we assign a "multiplier" that reveals the relationship between the reported price and the price faced by consumers. This chosen multiplier is selected from a set of multipliers based on possible tax rates for the firm. We merge in the state and local sales tax rates for each firm in order to construct this choice set.[17] To understand the relationship between the multiplier, reported prices, and faced prices, consider the following equation:

$$Price_{Consumer} = Price_{Reported} \times Multiplier$$

We algorithmically determine which tax-based multiplier makes the prices faces by consumer's ($Price_{Consumer}$) most round for each week, where roundness is the closeness of the price to a 25 cent increment of a dollar. For each product type, $Price_{Reported}$ is the modal observed price for the week, where idiosyncratic discounts have been removed.[18]

We consider two orthogonal methods of determine the proper set of multipliers. Our results are robust to the method used. Ultimately, we find the modal firm never included

[17]For five firms, the state and local tax rates do not match the rates they are using, so we adjust these. And a few firms do not ever change their local tax rate for reporting purposes—we make that adjustment as well. This transforms these firms from very unround to very round, but otherwise has approximately no effect on the data as the difference between the statutory and reported local tax rates is very small.

[18]We determine that a price is a one-off discount if the price for that transaction is 5% to 95% (in increments of 5 percentage points) or 33%/66.67% less than the previously reported price.

the sales tax, included the excise tax pre-reform, and excluded the excise tax post-reform.

**Cash Market Identification** In order to determine how each firm reports their prices in the traceability system, we take advantage of two characteristics of retail prices. First, publicly advertised prices (or 'list' prices) are nearly universally all tax-inclusive. Second, retailers nearly always choose to set prices in whole-dollar or (rarely) quarter-dollar increments.[19] We use these two facts to determine the difference between the list prices faced by consumers and the prices reported in the traceability system.

We assign each firm a multiplier before and after the tax change. We begin by assigning the modal firm's multiplier choices to all firms—all firms' prices were adjusted by the state and local sales tax pre-reform and all firm's prices were adjusted by the excise tax plus state and local sales tax post-reform. We then make the following adjustments based on the results from our algorithm:

1. We leave prices unadjusted (i.e. a multiplier of 1) where are algorithm finds that this choice maximizes roundedness and at least 85 percent of weekly sales are round with this multiplier choice.[20] This applies to 16% of firms.

2. We adjust the multiplier post reform to account for only the excise tax when the algorithm finds that this choice maximizes roundness and at least 85 percent of weekly sales are round with this multiplier choice. This applies to one firm (out of 110).

   There are three additional firms for whom an only excise tax adjustment makes them most round, but their roundness in the immediate post period is less than 85 percent. We leave two of the firms alone because they were also left alone in the pre-reform period because of unroundness and we could either adjust them both before and after

[19] We verified this through conversations with retailers as well as using historical menus available through The Internet Archive and a full set of menus for almost all firms we took screen shots of on 7/18/2017.

[20] For the 4.5% of firms that suggest the multiplier could be 1 but are quite unround, there is too much uncertainty to confidently make an adjustment. Leaving these firms' multipliers unchanged, if wrong, will bias our estimates towards our main null hypothesis in the retail section of the paper—that firms did not adjust their prices in response to the reform.

the reform or leave them both alone with similar effects to the log price change. The third firm becomes more round a few weeks after the reform and keeps this multiplier through the end of our data (and we have confirmed the multiplier in the menu screen shots), so we make this multiplier adjustment.

3. There are two firms for whom the multiplier that makes them round post-reform is the excise tax + state and local sales taxes divided by the state and local sales tax rate. In both cases, we have clear evidence that this is because they adjusted their prices post-reform by making their prices sales-tax exclusive post-reform. One firm keeps this choice permanently and we see this in the menu screen shots at the end of our data. The other firm eventually adjusts to the modal firms' multiplier. Our assumption keeps prices roughly constant through this reporting change.

**Product Batch Price Stability** To provide additional evidence that our multiplier decisions are not systematically biasing our estimates, we consider a completely different mechanism for determining multipliers—we pick the set of multipliers that makes the tax-inclusive prices for the most number of inventory lots for a given firm the same pre- and post-reform.

There are a couple of reasons why this is a reasonable alternative to consider. A number of inventory lots did leave prices constant in response to the tax reform and the main null hypothesis in our retail analysis is that firms did not change their tax-inclusive prices—this is what we would expect if the tax reform was indeed tax invariant.

We consider two variants of this. One is to begin with the modal firms' multipliers and adjust it to another multiplier if it decreases the number of price changes by any margin. The second variant is to begin with our estimates based on roundedness and then make adjustments for firms that under the best set of multipliers leaves at least 25% of their inventory lots constant in response to the reform. The latter changes the multipliers for only four firms and three of those four leave the percent price changes quite similar. The former method decreases our baseline estimate by 0.4 percentage points and the latter decreases our baseline estimate by 0.2 percentage points. This evidence strongly supports our price cleaning methods and suggests that any remaining bias is extremely small.

**Additional Cleaning**

In addition to adjusting retail prices, we also drop some extreme outliers in the data. In particular, we drop all wholesale transactions with a usable weight above 2,500 grams[21] and all retail transactions if the usable weight was above 28.5 grams.[22] We also drop all wholesale or retail price per grams above \$42.[23] We censor the THC content data if it is zero or above 40 in both the manufacturer and retailer data.[24] We also drop wholesale prices less than \$1. This effectively drops samples from our data, which are sold well below market value. We typically see these as the first recorded sale from a parentlot.

Lastly, we drop some firms or firm-days in our data set. In particular, we require for each firm that the first sales transaction occurs two months before the tax reform and continues to have transactions through the two months after the reform (either because they had not yet opened, had closed, or because they took a long hiatus from selling any cannabis). A few retailers conduct a "soft opening" by opening briefly, closing for more than a month, and then re-open permanently. In these cases, we drop the first brief selling period and consider their first activity date the first date upon re-opening in our data. We also drop 20 retail firms for whom at some point in the 8 weeks before or after the reform report their data only once per day—this is a clear indicator of poor overall data quality and, because of this, determining the appropriate multipliers for these firms is difficult.

[21] This is about 0.025% of wholesale transactions.
[22] This is because the maximum legal sale was one ounce. This step drops 0.15% of retail transactions.
[23] This is less than 0.03% of wholesale transactions and less than 0.04% of retail transactions.
[24] This affects 0.2% of wholesale transactions and 5% of retail transactions.